EDA

Jiaheng Cai

11/16/2021

System information

```
library(tidyverse)
```

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

```
## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.3     v dplyr   1.0.7
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   2.0.0     v forcats 0.5.1
```

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

```
library(readxl)
library(gtsummary)
library(survival)
library(survminer)
```

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

```
os <- sessionInfo()$running

if (str_detect(os, "Ubuntu")) {
  path <- '~/Biostatistics_M215_Project/data/'
} else if (str_detect(os, "mac")) {
  path <- '~/Downloads/Biostat 215 Project/Biostatistics_M215_Project/Data/'
} else if(str_detect(os, "Windows")){
  path <- "C:/Users/alexc/Desktop/215/Biostats-M215/"
}
```

Data clean and recode

```
endpoint = read_excel(paste0(path, 'endpoints.xls'))
endpoint_colnames = c('ID', 'Intervention Group',
                      'Vitality Status as of 6/1/2006',
                      'Breast Cancer Status as of 6/1/2006 or last prior contact',
                      'Other Cancer (invasive, not breast) Status as of 6/1/2006',
                      'Breast Cancer Contribute to Death',
                      'Year Breast Cancer Diagnosed', 'Cancer Grade',
                      'Dummy Variable for Cancer Grade 2',
                      'Dummy Variable for Cancer Grade 3',
                      'Dummy Variable for Unknown Cancer Grade',
                      'Cancer Stage, AJCC 6th',
                      'Dummy Variable for Stage 2 AJCC 6th',
```

```r
                         'Dummy Variable for Stage 3 AJCC 6th',
                         'WHEL Clinical Site', 'Recurrence Flag',
                         'Years from Diagnosis to WHEL Study Entry',
                         'Years from Study Entry to Recurrence/New Primary, or to Censor',
                         'Years from Diagnosis to Recurrence/New Primary, or to Censor',
                         'Years from Diagnosis to Death or Censor')

colnames(endpoint) = endpoint_colnames

endpoint$`Intervention Group` = ifelse(endpoint$`Intervention Group` == 3, 'Intervention', 'Comparison')

endpoint$`Vitality Status as of 6/1/2006` = ifelse(endpoint$`Vitality Status as of 6/1/2006` == 0, 'Dead

for (i in 1:nrow(endpoint)){
if (endpoint$`Breast Cancer Status as of 6/1/2006 or last prior contact`[i] == 0){
  endpoint$`Breast Cancer Status as of 6/1/2006 or last prior contact`[i] = 'No Evidence of Recurrence'
} else if(endpoint$`Breast Cancer Status as of 6/1/2006 or last prior contact`[i] == 1){
  endpoint$`Breast Cancer Status as of 6/1/2006 or last prior contact`[i] = 'Confirmed - New Primary Bre
}else if(endpoint$`Breast Cancer Status as of 6/1/2006 or last prior contact`[i] == 2){
  endpoint$`Breast Cancer Status as of 6/1/2006 or last prior contact`[i] = 'Confirmed - Local Recurrenc
}else if(endpoint$`Breast Cancer Status as of 6/1/2006 or last prior contact`[i] == 3){
  endpoint$`Breast Cancer Status as of 6/1/2006 or last prior contact`[i] = 'Confirmed - Regional Recurr
}else{
  endpoint$`Breast Cancer Status as of 6/1/2006 or last prior contact`[i] = 'Confirmed - Distant Recurre
}

endpoint$`Other Cancer (invasive, not breast) Status as of 6/1/2006` = ifelse(endpoint$`Other Cancer (in

endpoint$`Breast Cancer Contribute to Death`[endpoint$`Breast Cancer Contribute to Death` == -1] = 'Not
endpoint$`Breast Cancer Contribute to Death`[endpoint$`Breast Cancer Contribute to Death` == 0] = 'Dead
endpoint$`Breast Cancer Contribute to Death`[endpoint$`Breast Cancer Contribute to Death` == 1] = 'Dead
endpoint$`Breast Cancer Contribute to Death`[endpoint$`Breast Cancer Contribute to Death` == 2] = 'Dead

endpoint$`Cancer Grade`[endpoint$`Cancer Grade` == 0] = 'Grade Not Applicable or Not Available'
endpoint$`Cancer Grade`[endpoint$`Cancer Grade` == 1] = 'Grade I, Well Differentiated'
endpoint$`Cancer Grade`[endpoint$`Cancer Grade` == 2] = 'Grade II, Moderately Differentiated'
endpoint$`Cancer Grade`[endpoint$`Cancer Grade` == 3] = 'Grade III, Poorly Differentiated'

endpoint$`Cancer Stage, AJCC 6th`[endpoint$`Cancer Stage, AJCC 6th` == 1] = 'Stage I'
endpoint$`Cancer Stage, AJCC 6th`[endpoint$`Cancer Stage, AJCC 6th` == 2] = 'Stage IIA'
endpoint$`Cancer Stage, AJCC 6th`[endpoint$`Cancer Stage, AJCC 6th` == 3] = 'Stage IIB'
endpoint$`Cancer Stage, AJCC 6th`[endpoint$`Cancer Stage, AJCC 6th` == 4] = 'Stage IIIA'
endpoint$`Cancer Stage, AJCC 6th`[endpoint$`Cancer Stage, AJCC 6th` == 5] = 'Stage IIIB'

endpoint$`WHEL Clinical Site`[endpoint$`WHEL Clinical Site` == 1] = 'Site A in California'
endpoint$`WHEL Clinical Site`[endpoint$`WHEL Clinical Site` == 3] = 'Site B in California'
endpoint$`WHEL Clinical Site`[endpoint$`WHEL Clinical Site` == 5] = 'Site C in California'
endpoint$`WHEL Clinical Site`[endpoint$`WHEL Clinical Site` == 7] = 'Site in Arizona'
endpoint$`WHEL Clinical Site`[endpoint$`WHEL Clinical Site` == 9] = 'Site D in California'
endpoint$`WHEL Clinical Site`[endpoint$`WHEL Clinical Site` == 11] = 'Site in Texas'
endpoint$`WHEL Clinical Site`[endpoint$`WHEL Clinical Site` == 13] = 'Site in Oregon'

endpoint$`Recurrence Flag`[endpoint$`Recurrence Flag` == 0] = 'No Invasive Breast Cancer Events'
```

```
endpoint$`Recurrence Flag`[endpoint$`Recurrence Flag` == 1] = 'Invasive Breast Cancer Event'
```

Summary table

```
endpoint %>%
  tbl_summary(by = 'Intervention Group',
              include=-c(ID, `Year Breast Cancer Diagnosed`,
                         `Dummy Variable for Cancer Grade 2`,
                         `Dummy Variable for Cancer Grade 3`,
                         `Dummy Variable for Unknown Cancer Grade`,
                         `Dummy Variable for Stage 2 AJCC 6th`,
                         `Dummy Variable for Stage 3 AJCC 6th`),
              missing_text = "Missing") %>%
  add_p()
```

```
## Table printed with `knitr::kable()`, not {gt}. Learn why at
## http://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
## To suppress this message, include `message = FALSE` in code chunk header.
```

Characteristic	Comparison, N = 1,551	Intervention, N = 1,537	p-value
Vitality Status as of 6/1/2006			0.8
Alive	1,391 (90%)	1,381 (90%)	
Dead	160 (10%)	155 (10%)	
Unknown	0 (0%)	1 (<0.1%)	
Breast Cancer Status as of 6/1/2006 or last prior contact			0.6
Confirmed – Distant Recurrence	189 (12%)	168 (11%)	
Confirmed – Local Recurrence	28 (1.8%)	35 (2.3%)	
Confirmed – New Primary Breast Cancer	35 (2.3%)	43 (2.8%)	
Confirmed – Regional Recurrence	10 (0.6%)	10 (0.7%)	
No Evidence of Recurrence	1,289 (83%)	1,281 (83%)	
Other Cancer (invasive, not breast) Status as of 6/1/2006			0.5
Confirmed Other Cancer	60 (3.9%)	58 (3.8%)	
No evidence of Disease	1,483 (96%)	1,472 (96%)	
Reported Other Cancer (not confirmed)	4 (0.3%)	1 (<0.1%)	
Missing	4	6	
Breast Cancer Contribute to Death			0.8
Dead from a cause other than Breast Cancer	25 (1.6%)	28 (1.8%)	
Dead from Breast Cancer	135 (8.7%)	126 (8.2%)	
Dead from Cancer, not confirmed breast but likely so	0 (0%)	1 (<0.1%)	
Not Dead	1,391 (90%)	1,382 (90%)	
Cancer Grade			>0.9
Grade I, Well Differentiated	245 (16%)	239 (16%)	
Grade II, Moderately Differentiated	620 (40%)	620 (40%)	
Grade III, Poorly Differentiated	557 (36%)	551 (36%)	
Grade Not Applicable or Not Available	129 (8.3%)	127 (8.3%)	
Cancer Stage, AJCC 6th			>0.9
Stage I	607 (39%)	584 (38%)	
Stage IIA	511 (33%)	515 (34%)	
Stage IIB	190 (12%)	194 (13%)	
Stage IIIA	185 (12%)	188 (12%)	

Characteristic	Comparison, N = 1,551	Intervention, N = 1,537	p-value
Stage IIIB	58 (3.7%)	56 (3.6%)	
WHEL Clinical Site			>0.9
Site A in California	258 (17%)	272 (18%)	
Site B in California	226 (15%)	210 (14%)	
Site C in California	267 (17%)	257 (17%)	
Site D in California	260 (17%)	256 (17%)	
Site in Arizona	242 (16%)	233 (15%)	
Site in Oregon	117 (7.5%)	116 (7.5%)	
Site in Texas	181 (12%)	193 (13%)	
Recurrence Flag			0.9
Invasive Breast Cancer Event	262 (17%)	256 (17%)	
No Invasive Breast Cancer Events	1,289 (83%)	1,281 (83%)	
Years from Diagnosis to WHEL Study Entry	1.76 (1.05, 2.81)	1.84 (1.04, 2.80)	0.8
Years from Study Entry to Recurrence/New Primary, or to Censor	7.12 (5.87, 8.44)	7.11 (5.83, 8.53)	>0.9
Years from Diagnosis to Recurrence/New Primary, or to Censor	8.97 (7.36, 10.64)	9.00 (7.35, 10.67)	0.8
Years from Diagnosis to Death or Censor	9.18 (7.76, 10.87)	9.28 (7.82, 10.89)	0.6

Plots and survival analysis tools

```
endpoint.data = read_excel(paste0(path, 'endpoints.xls'))
year4.data = read_excel(paste0(path, 'healthstaty4.xls'))


endpoint.data$recur_flag = as.factor(endpoint.data$recur_flag)


endpoint.data = read_excel(paste0(path, 'endpoints.xls'))
endpoint.km.fit <- survfit(Surv(yrsdx_endr, recur_flag) ~ 1, data = endpoint.data, )
print(endpoint.km.fit)
```

```
## Call: survfit(formula = Surv(yrsdx_endr, recur_flag) ~ 1, data = endpoint.data)
##
##        n  events  median 0.95LCL 0.95UCL
##     3088     518      NA      NA      NA
```

Intervention Effects on All-Cause Mortality by Baseline Demographic and Clinical Characteristics

Preprocessing

```
demo = read_excel("../Data/demographics.xls")
phbase = read_excel("../Data/phbase.xls")
nds = read_excel("../Data/ndsfoody4.xls")
medical = read_excel("../Data/Medical.xls")


# We need the following variables:


## Survival time
SurvTime = as.numeric(endpoint$`Years from Diagnosis to Death or Censor`)
## Group and Status
Group = as_factor(endpoint$`Intervention Group`)
Group = relevel(Group, ref = "Comparison")
```

```r
a = as_factor(endpoint$`Vitality Status as of 6/1/2006`)
Status = ifelse(a == "Alive", 0, 1)
## Age at randomization, y
AgeIdx = ifelse(demo$`age at rand` < 55, "<55", ">=55") # Age indicator (<=55 or not)
## Cancer stage at diagnosis
a = endpoint$`Cancer Stage, AJCC 6th`
CancerStage = as_factor(a)
## Hormone receptor status
a = medical$`Estr Recep`
b = medical$`Prog Recep`
HormoneRecep = ifelse(a==1 & b==1, "ER+/PR+",
                      ifelse(a==1 & b==0, "ER+/PR-",
                             ifelse(a==0 & b==1, "ER-/PR+",
                                    ifelse(a==0 & b==0, "ER-/PR-", NA))))
## Time from Diag to Rand
a = endpoint$`Years from Diagnosis to WHEL Study Entry`
TimeDiagRand = as.numeric(ifelse(a <=1, 0,
                          ifelse(a <=2, 1,
                                 ifelse(a <=3, 2, 3
                          )))) # Time from diagnosis to randomization
## Tumor differentiation
a = endpoint$`Cancer Grade`
TumorDiff = as_factor(a)
## No. of positive nodes (Number axillary lymph nodes positive)
a = medical$`Node Pos`
PosNodes = ifelse(a==0, 0,
                  ifelse(a < 3, 1,
                         ifelse(a < 6, 2, 3)))
## Tumor size
a = medical$`Tumor Size`
TumorSize = ifelse(a < 2, 0,
                   ifelse(a < 3, 1,
                          ifelse(a < 4, 2,
                                 ifelse(a < 5, 3, 4))))
## Physical activity
a = phbase$`NEW METS`
PhysicalAct = ifelse(a <= 210, "<210",
                     ifelse(a <= 615, "211~615",
                            ifelse(a <= 1290, "616~1290", ">1290")))
## Energy intake
b = matrix(NA, nrow = length(a), ncol=1)
colnames(b) = "KCal"
b[endpoint$ID %in% nds$ID] = nds$Kcal
KCal = as_factor(ifelse(b <= 1430, "<1430",
                 ifelse(b <= 1680, "1430~1680",
                        ifelse(b <= 1980, "1681~1980",
                               ifelse(b > 1980, ">1980", NA)))))


##### PUT THEM TOGETHER #####
AllCauseMortalityData = data.frame(
  SurvTime, Group, Status, AgeIdx, CancerStage, HormoneRecep, TimeDiagRand,
        TumorDiff, PosNodes, TumorSize, PhysicalAct, KCal
```

```
)
##############################
```

Universal and univariable Cox regression

```
coxph(Surv(SurvTime, Status) ~ ., data=AllCauseMortalityData)
```

```
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ ., data = AllCauseMortalityData)
##
##                                            coef exp(coef) se(coef)      z
## GroupIntervention                       -0.05926   0.94246  0.26733 -0.222
## AgeIdx>=55                               1.21522   3.37104  0.29500  4.119
## CancerStageStage I                       0.20215   1.22403  0.41601  0.486
## CancerStageStage IIB                     0.01351   1.01360  0.52685  0.026
## CancerStageStage IIIA                    0.04782   1.04898  0.81198  0.059
## CancerStageStage IIIB                    0.56099   1.75240  1.11751  0.502
## HormoneRecepER-/PR+                      0.15235   1.16457  0.77955  0.195
## HormoneRecepER+/PR-                      0.03449   1.03509  0.53046  0.065
## HormoneRecepER+/PR+                      0.43595   1.54643  0.38425  1.135
## TimeDiagRand                            -0.33445   0.71573  0.14283 -2.342
## TumorDiffGrade II, Moderately Differentiated  0.53300   1.70403  0.54220  0.983
## TumorDiffGrade I, Well Differentiated   -0.23305   0.79211  0.67657 -0.344
## TumorDiffGrade III, Poorly Differentiated  0.74232   2.10080  0.56247  1.320
## PosNodes                                 0.11410   1.12086  0.31859  0.358
## TumorSize                                0.18404   1.20207  0.15033  1.224
## PhysicalAct>1290                        -1.22969   0.29238  0.46214 -2.661
## PhysicalAct211~615                      -0.16070   0.85154  0.32319 -0.497
## PhysicalAct616~1290                     -0.62413   0.53573  0.36915 -1.691
## KCal1430~1680                           -0.67682   0.50823  0.37286 -1.815
## KCal>1980                                0.07635   1.07934  0.37498  0.204
## KCal1681~1980                           -0.68036   0.50644  0.38508 -1.767
##                                              p
## GroupIntervention                       0.82457
## AgeIdx>=55                              3.8e-05
## CancerStageStage I                      0.62703
## CancerStageStage IIB                    0.97954
## CancerStageStage IIIA                   0.95304
## CancerStageStage IIIB                   0.61567
## HormoneRecepER-/PR+                     0.84505
## HormoneRecepER+/PR-                     0.94816
## HormoneRecepER+/PR+                     0.25656
## TimeDiagRand                            0.01920
## TumorDiffGrade II, Moderately Differentiated 0.32560
## TumorDiffGrade I, Well Differentiated   0.73050
## TumorDiffGrade III, Poorly Differentiated 0.18692
## PosNodes                                0.72024
## TumorSize                               0.22087
## PhysicalAct>1290                        0.00779
## PhysicalAct211~615                      0.61902
## PhysicalAct616~1290                     0.09089
## KCal1430~1680                           0.06950
## KCal>1980                               0.83867
```

```
## KCal1681~1980                                          0.07726
##
## Likelihood ratio test=47.45  on 21 df, p=0.0008157
## n= 2274, number of events= 57
##    (814 observations deleted due to missingness)
```

```r
library(survival)
library(finalfit)
dependent_os <- "Surv(SurvTime, Status)"
explanatory <- c("AgeIdx", "CancerStage", "HormoneRecep", "TimeDiagRand",
        "TumorDiff", "PosNodes", "TumorSize", "PhysicalAct", "KCal")
AllCauseMortalityData %>%
  filter(Group=="Intervention") %>%
  finalfit(dependent_os, explanatory)
```

```
##   Dependent: Surv(SurvTime, Status)
##                              AgeIdx                                <55
##                                                                   >=55
##                         CancerStage                          Stage IIA
##                                                                Stage I
##                                                              Stage IIB
##                                                             Stage IIIA
##                                                             Stage IIIB
##                        HormoneRecep                            ER-/PR-
##                                                                ER-/PR+
##                                                                ER+/PR-
##                                                                ER+/PR+
##                        TimeDiagRand                                  0
##                                                                      1
##                                                                      2
##                                                                      3
##                            TumorDiff Grade Not Applicable or Not Available
##                                      Grade II, Moderately Differentiated
##                                           Grade I, Well Differentiated
##                                         Grade III, Poorly Differentiated
##                             PosNodes                          Mean (SD)
##                            TumorSize                          Mean (SD)
##                          PhysicalAct                               <210
##                                                                  >1290
##                                                                211~615
##                                                               616~1290
##                                 KCal                              <1430
##                                                              1430~1680
##                                                                  >1980
##                                                              1681~1980
##                                 <NA>                               <NA>
##          all         HR (univariable)         HR (multivariable)
##  908 (100.0)                        -                          -
##  629 (100.0) 1.16 (0.85-1.59, p=0.348)  2.99 (1.30-6.87, p=0.010)
##  515 (100.0)                        -                          -
##  584 (100.0) 0.69 (0.43-1.09, p=0.110)  2.14 (0.61-7.46, p=0.232)
##  194 (100.0) 2.13 (1.34-3.39, p=0.001)  1.20 (0.26-5.45, p=0.813)
##  188 (100.0) 2.51 (1.60-3.93, p<0.001) 0.84 (0.06-11.14, p=0.898)
##   56 (100.0) 3.58 (1.95-6.56, p<0.001) 1.25 (0.04-37.58, p=0.899)
##  299 (100.0)                        -                          -
```

```
##    52 (100.0) 0.97 (0.46-2.05, p=0.930)  0.83 (0.10-7.27, p=0.868)
##   198 (100.0) 0.81 (0.50-1.32, p=0.398)  0.80 (0.20-3.25, p=0.759)
##   958 (100.0) 0.52 (0.36-0.75, p<0.001)  1.07 (0.33-3.42, p=0.912)
##   352 (100.0)                          -                         -
##   488 (100.0)                          -                         -
##   375 (100.0)                          -                         -
##   322 (100.0)                          -                         -
##   127 (100.0)                          -                         -
##   620 (100.0) 1.49 (0.74-3.01, p=0.263) 3.09 (0.39-24.42, p=0.284)
##   239 (100.0) 0.84 (0.36-1.96, p=0.680) 2.60 (0.28-24.05, p=0.400)
##   551 (100.0) 2.26 (1.13-4.51, p=0.021) 3.20 (0.38-27.02, p=0.286)
##    0.7 (1.0) 1.61 (1.41-1.85, p<0.001)  1.04 (0.43-2.50, p=0.932)
##    0.9 (1.2) 1.36 (1.22-1.51, p<0.001)  1.64 (1.04-2.58, p=0.034)
##   443 (100.0)                          -                         -
##   351 (100.0) 0.61 (0.38-0.99, p=0.047)  0.20 (0.04-0.94, p=0.041)
##   368 (100.0) 1.08 (0.72-1.61, p=0.717)  0.82 (0.31-2.17, p=0.694)
##   375 (100.0) 0.84 (0.55-1.29, p=0.429)  0.73 (0.28-1.94, p=0.531)
##   446 (100.0)                          -                         -
##   287 (100.0) 0.64 (0.24-1.68, p=0.365)  0.51 (0.18-1.46, p=0.209)
##   144 (100.0) 1.19 (0.45-3.13, p=0.725)  1.12 (0.41-3.07, p=0.818)
##   249 (100.0) 0.37 (0.10-1.28, p=0.116)  0.34 (0.09-1.24, p=0.102)
##          <NA> 0.70 (0.60-0.81, p<0.001)  0.64 (0.43-0.94, p=0.022)
```

Imputation

```
library(miceRanger)
ImputedAll <- AllCauseMortalityData %>%
  miceRanger(
  m = 3,
  returnModels = TRUE,
  verbose = TRUE
)
```

```
## One or more of the specified variables to impute contains no missing values. These will remain as a
```

```
## Converting characters to factors.
```

```
##
## Process started at 2021-11-18 21:46:51
```

```
##
## Attaching package: 'data.table'
```

```
## The following objects are masked from 'package:dplyr':
##
##     between, first, last
```

```
## The following object is masked from 'package:purrr':
##
##     transpose
```

```
##
## dataset 1
## iteration 1   | HormoneRecep | PosNodes | TumorSize | KCal
## iteration 2   | HormoneRecep | PosNodes | TumorSize | KCal
## iteration 3   | HormoneRecep | PosNodes | TumorSize | KCal
```

```
## iteration 4   | HormoneRecep | PosNodes | TumorSize | KCal
## iteration 5   | HormoneRecep | PosNodes | TumorSize | KCal
##
## dataset 2
## iteration 1   | HormoneRecep | PosNodes | TumorSize | KCal
## iteration 2   | HormoneRecep | PosNodes | TumorSize | KCal
## iteration 3   | HormoneRecep | PosNodes | TumorSize | KCal
## iteration 4   | HormoneRecep | PosNodes | TumorSize | KCal
## iteration 5   | HormoneRecep | PosNodes | TumorSize | KCal
##
## dataset 3
## iteration 1   | HormoneRecep | PosNodes | TumorSize | KCal
## iteration 2   | HormoneRecep | PosNodes | TumorSize | KCal
## iteration 3   | HormoneRecep | PosNodes | TumorSize | KCal
## iteration 4   | HormoneRecep | PosNodes | TumorSize | KCal
## iteration 5   | HormoneRecep | PosNodes | TumorSize | KCal
```

Cox Regression for each variable

```
ImputedData = as.data.frame(completeData(ImputedAll)[[3]])
for (item in explanatory) {
  len = ImputedData %>%
    select(item) %>%
    unique(.) %>%
    nrow(.)
  category = ImputedData %>%
      select(item) %>%
      unique(.)
  hi = item
  for (i in 1:len) {
    cat = category[i,]
    print(paste0("Explanatory: ", hi))
    print(paste0("Category: ", cat))
    idx = ImputedData[, hi] == cat
    print(summary(coxph(Surv(SurvTime, Status) ~ Group, data=ImputedData[idx, ])))
  }
}
```

```
## Note: Using an external vector in selections is ambiguous.
## i Use `all_of(item)` instead of `item` to silence this message.
## i See <https://tidyselect.r-lib.org/reference/faq-external-vector.html>.
## This message is displayed once per session.

## [1] "Explanatory: AgeIdx"
## [1] "Category: <55"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 1825, number of events= 170
##
##                        coef exp(coef) se(coef)    z Pr(>|z|)
## GroupIntervention 0.003098  1.003103 0.153395 0.02    0.984
##
```

9

```
##                  exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention   1.003      0.9969    0.7426     1.355
##
## Concordance= 0.496  (se = 0.02 )
## Likelihood ratio test= 0  on 1 df,   p=1
## Wald test            = 0  on 1 df,   p=1
## Score (logrank) test = 0  on 1 df,   p=1
##
## [1] "Explanatory: AgeIdx"
## [1] "Category: >=55"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 1263, number of events= 146
##
##                      coef exp(coef) se(coef)      z Pr(>|z|)
## GroupIntervention -0.04775   0.95338  0.16559 -0.288    0.773
##
##                  exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention   0.9534      1.049    0.6891     1.319
##
## Concordance= 0.5  (se = 0.022 )
## Likelihood ratio test= 0.08  on 1 df,   p=0.8
## Wald test            = 0.08  on 1 df,   p=0.8
## Score (logrank) test = 0.08  on 1 df,   p=0.8
##
## [1] "Explanatory: CancerStage"
## [1] "Category: Stage IIA"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 1026, number of events= 91
##
##                     coef exp(coef) se(coef)      z Pr(>|z|)
## GroupIntervention -0.2017    0.8174   0.2107 -0.957    0.339
##
##                  exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention   0.8174      1.223    0.5408     1.235
##
## Concordance= 0.527  (se = 0.027 )
## Likelihood ratio test= 0.92  on 1 df,   p=0.3
## Wald test            = 0.92  on 1 df,   p=0.3
## Score (logrank) test = 0.92  on 1 df,   p=0.3
##
## [1] "Explanatory: CancerStage"
## [1] "Category: Stage I"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 1191, number of events= 65
##
```

```
##                      coef exp(coef) se(coef)      z Pr(>|z|)
## GroupIntervention 0.07616   1.07914  0.24811 0.307    0.759
##
##                   exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention     1.079     0.9267    0.6636     1.755
##
## Concordance= 0.516  (se = 0.033 )
## Likelihood ratio test= 0.09  on 1 df,   p=0.8
## Wald test            = 0.09  on 1 df,   p=0.8
## Score (logrank) test = 0.09  on 1 df,   p=0.8
##
## [1] "Explanatory: CancerStage"
## [1] "Category: Stage IIB"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 384, number of events= 52
##
##                     coef exp(coef) se(coef)      z Pr(>|z|)
## GroupIntervention 0.4093    1.5058   0.2857 1.433    0.152
##
##                   exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention     1.506     0.6641    0.8601     2.636
##
## Concordance= 0.553  (se = 0.036 )
## Likelihood ratio test= 2.1  on 1 df,   p=0.1
## Wald test            = 2.05  on 1 df,   p=0.2
## Score (logrank) test = 2.08  on 1 df,   p=0.1
##
## [1] "Explanatory: CancerStage"
## [1] "Category: Stage IIIA"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 373, number of events= 70
##
##                      coef exp(coef) se(coef)      z Pr(>|z|)
## GroupIntervention 0.03257   1.03311  0.23917 0.136    0.892
##
##                   exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention     1.033      0.968    0.6465     1.651
##
## Concordance= 0.505  (se = 0.031 )
## Likelihood ratio test= 0.02  on 1 df,   p=0.9
## Wald test            = 0.02  on 1 df,   p=0.9
## Score (logrank) test = 0.02  on 1 df,   p=0.9
##
## [1] "Explanatory: CancerStage"
## [1] "Category: Stage IIIB"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
```

```
## 
##   n= 114, number of events= 38
## 
##                       coef exp(coef) se(coef)      z Pr(>|z|)
## GroupIntervention -0.5660    0.5678   0.3371 -1.679   0.0932 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
##                   exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention    0.5678      1.761    0.2933     1.099
## 
## Concordance= 0.567  (se = 0.041 )
## Likelihood ratio test= 2.93  on 1 df,   p=0.09
## Wald test            = 2.82  on 1 df,   p=0.09
## Score (logrank) test = 2.89  on 1 df,   p=0.09
## 
## [1] "Explanatory: HormoneRecep"
## [1] "Category: ER+/PR+"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
## 
##   n= 1949, number of events= 164
## 
##                        coef exp(coef) se(coef)      z Pr(>|z|)
## GroupIntervention -0.08661   0.91703  0.15629 -0.554    0.579
## 
##                   exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention     0.917       1.09    0.6751     1.246
## 
## Concordance= 0.509  (se = 0.02 )
## Likelihood ratio test= 0.31  on 1 df,   p=0.6
## Wald test            = 0.31  on 1 df,   p=0.6
## Score (logrank) test = 0.31  on 1 df,   p=0.6
## 
## [1] "Explanatory: HormoneRecep"
## [1] "Category: ER-/PR-"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
## 
##   n= 629, number of events= 87
## 
##                     coef exp(coef) se(coef)    z Pr(>|z|)
## GroupIntervention 0.1050    1.1107   0.2145 0.49    0.624
## 
##                   exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention     1.111     0.9003    0.7295     1.691
## 
## Concordance= 0.514  (se = 0.027 )
## Likelihood ratio test= 0.24  on 1 df,   p=0.6
## Wald test            = 0.24  on 1 df,   p=0.6
## Score (logrank) test = 0.24  on 1 df,   p=0.6
## 
```

```
## [1] "Explanatory: HormoneRecep"
## [1] "Category: ER-/PR+"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 132, number of events= 18
##
##                      coef exp(coef) se(coef)      z Pr(>|z|)
## GroupIntervention 0.2363    1.2666   0.4751 0.497    0.619
##
##                   exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention     1.267     0.7895    0.4991     3.214
##
## Concordance= 0.539  (se = 0.062 )
## Likelihood ratio test= 0.24  on 1 df,   p=0.6
## Wald test            = 0.25  on 1 df,   p=0.6
## Score (logrank) test = 0.25  on 1 df,   p=0.6
##
## [1] "Explanatory: HormoneRecep"
## [1] "Category: ER+/PR-"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 378, number of events= 47
##
##                       coef exp(coef) se(coef)      z Pr(>|z|)
## GroupIntervention -0.06474   0.93731  0.29246 -0.221    0.825
##
##                   exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention    0.9373      1.067    0.5284     1.663
##
## Concordance= 0.52  (se = 0.038 )
## Likelihood ratio test= 0.05  on 1 df,   p=0.8
## Wald test            = 0.05  on 1 df,   p=0.8
## Score (logrank) test = 0.05  on 1 df,   p=0.8
##
## [1] "Explanatory: TimeDiagRand"
## [1] "Category: 0"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 702, number of events= 93
##
##                       coef exp(coef) se(coef)      z Pr(>|z|)
## GroupIntervention -0.03665   0.96401  0.20757 -0.177     0.86
##
##                   exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention     0.964      1.037    0.6418     1.448
##
## Concordance= 0.5  (se = 0.027 )
## Likelihood ratio test= 0.03  on 1 df,   p=0.9
```

```
## Wald test            = 0.03  on 1 df,   p=0.9
## Score (logrank) test = 0.03  on 1 df,   p=0.9
##
## [1] "Explanatory: TimeDiagRand"
## [1] "Category: 1"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 996, number of events= 109
##
##                        coef exp(coef) se(coef)      z Pr(>|z|)
## GroupIntervention -0.1040    0.9012   0.1920 -0.542    0.588
##
##                    exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention    0.9012       1.11    0.6186     1.313
##
## Concordance= 0.505  (se = 0.025 )
## Likelihood ratio test= 0.29  on 1 df,   p=0.6
## Wald test            = 0.29  on 1 df,   p=0.6
## Score (logrank) test = 0.29  on 1 df,   p=0.6
##
## [1] "Explanatory: TimeDiagRand"
## [1] "Category: 3"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 643, number of events= 52
##
##                        coef exp(coef) se(coef)      z Pr(>|z|)
## GroupIntervention -0.1019    0.9031   0.2776 -0.367    0.714
##
##                    exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention    0.9031      1.107    0.5242     1.556
##
## Concordance= 0.516  (se = 0.035 )
## Likelihood ratio test= 0.13  on 1 df,   p=0.7
## Wald test            = 0.13  on 1 df,   p=0.7
## Score (logrank) test = 0.13  on 1 df,   p=0.7
##
## [1] "Explanatory: TimeDiagRand"
## [1] "Category: 2"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 747, number of events= 62
##
##                        coef exp(coef) se(coef)      z Pr(>|z|)
## GroupIntervention 0.2405    1.2719   0.2553 0.942    0.346
##
##                    exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention    1.272     0.7862    0.7711     2.098
```

```
## 
## Concordance= 0.528  (se = 0.033 )
## Likelihood ratio test= 0.89  on 1 df,   p=0.3
## Wald test            = 0.89  on 1 df,   p=0.3
## Score (logrank) test = 0.89  on 1 df,   p=0.3
## 
## [1] "Explanatory: TumorDiff"
## [1] "Category: Grade Not Applicable or Not Available"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
## 
##   n= 256, number of events= 24
## 
##                       coef exp(coef) se(coef)     z Pr(>|z|)
## GroupIntervention -0.4682    0.6261   0.4219 -1.11    0.267
## 
##                   exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention    0.6261      1.597    0.2739     1.431
## 
## Concordance= 0.554  (se = 0.051 )
## Likelihood ratio test= 1.27  on 1 df,   p=0.3
## Wald test            = 1.23  on 1 df,   p=0.3
## Score (logrank) test = 1.25  on 1 df,   p=0.3
## 
## [1] "Explanatory: TumorDiff"
## [1] "Category: Grade II, Moderately Differentiated"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
## 
##   n= 1240, number of events= 123
## 
##                        coef exp(coef) se(coef)     z Pr(>|z|)
## GroupIntervention -0.09382   0.91045  0.18049 -0.52    0.603
## 
##                   exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention    0.9104      1.098    0.6392     1.297
## 
## Concordance= 0.509  (se = 0.024 )
## Likelihood ratio test= 0.27  on 1 df,   p=0.6
## Wald test            = 0.27  on 1 df,   p=0.6
## Score (logrank) test = 0.27  on 1 df,   p=0.6
## 
## [1] "Explanatory: TumorDiff"
## [1] "Category: Grade I, Well Differentiated"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
## 
##   n= 484, number of events= 20
## 
##                      coef exp(coef) se(coef)     z Pr(>|z|)
## GroupIntervention 0.5902    1.8044   0.4691 1.258    0.208
```

```
##
##                    exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention     1.804      0.5542    0.7196     4.525
##
## Concordance= 0.538  (se = 0.06 )
## Likelihood ratio test= 1.66  on 1 df,    p=0.2
## Wald test            = 1.58  on 1 df,    p=0.2
## Score (logrank) test = 1.63  on 1 df,    p=0.2
##
## [1] "Explanatory: TumorDiff"
## [1] "Category: Grade III, Poorly Differentiated"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 1108, number of events= 149
##
##                       coef exp(coef) se(coef)      z Pr(>|z|)
## GroupIntervention 0.03348   1.03405  0.16386 0.204     0.838
##
##                    exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention     1.034      0.9671      0.75     1.426
##
## Concordance= 0.507  (se = 0.021 )
## Likelihood ratio test= 0.04  on 1 df,    p=0.8
## Wald test            = 0.04  on 1 df,    p=0.8
## Score (logrank) test = 0.04  on 1 df,    p=0.8
##
## [1] "Explanatory: PosNodes"
## [1] "Category: 1"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 732, number of events= 70
##
##                       coef exp(coef) se(coef)      z Pr(>|z|)
## GroupIntervention 0.1953    1.2157   0.2401 0.814     0.416
##
##                    exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention     1.216      0.8226    0.7594     1.946
##
## Concordance= 0.523  (se = 0.031 )
## Likelihood ratio test= 0.66  on 1 df,    p=0.4
## Wald test            = 0.66  on 1 df,    p=0.4
## Score (logrank) test = 0.66  on 1 df,    p=0.4
##
## [1] "Explanatory: PosNodes"
## [1] "Category: 0"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 1775, number of events= 125
```

```
##
##                      coef exp(coef) se(coef)      z Pr(>|z|)
## GroupIntervention -0.1203    0.8867   0.1794 -0.671    0.503
##
##                   exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention    0.8867      1.128    0.6239      1.26
##
## Concordance= 0.515  (se = 0.023 )
## Likelihood ratio test= 0.45  on 1 df,   p=0.5
## Wald test            = 0.45  on 1 df,   p=0.5
## Score (logrank) test = 0.45  on 1 df,   p=0.5
##
## [1] "Explanatory: PosNodes"
## [1] "Category: 2"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 333, number of events= 50
##
##                     coef exp(coef) se(coef)     z Pr(>|z|)
## GroupIntervention 0.1548    1.1675   0.2838 0.546    0.585
##
##                   exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention     1.167     0.8566    0.6694     2.036
##
## Concordance= 0.521  (se = 0.036 )
## Likelihood ratio test= 0.3  on 1 df,   p=0.6
## Wald test            = 0.3  on 1 df,   p=0.6
## Score (logrank) test = 0.3  on 1 df,   p=0.6
##
## [1] "Explanatory: PosNodes"
## [1] "Category: 3"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 248, number of events= 71
##
##                      coef exp(coef) se(coef)      z Pr(>|z|)
## GroupIntervention -0.3104    0.7331   0.2384 -1.302    0.193
##
##                   exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention    0.7331      1.364    0.4595      1.17
##
## Concordance= 0.536  (se = 0.031 )
## Likelihood ratio test= 1.7  on 1 df,   p=0.2
## Wald test            = 1.7  on 1 df,   p=0.2
## Score (logrank) test = 1.71  on 1 df,   p=0.2
##
## [1] "Explanatory: TumorSize"
## [1] "Category: 0"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
```

```
##     ])
##
##   n= 1523, number of events= 94
##
##                      coef exp(coef) se(coef)       z Pr(>|z|)
## GroupIntervention -0.1561    0.8555   0.2070 -0.754    0.451
##
##                      exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention    0.8555      1.169    0.5701     1.284
##
## Concordance= 0.509  (se = 0.027 )
## Likelihood ratio test= 0.57  on 1 df,   p=0.5
## Wald test            = 0.57  on 1 df,   p=0.5
## Score (logrank) test = 0.57  on 1 df,   p=0.5
##
## [1] "Explanatory: TumorSize"
## [1] "Category: 1"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 863, number of events= 109
##
##                     coef exp(coef) se(coef)   z Pr(>|z|)
## GroupIntervention 0.2499    1.2839   0.1923 1.3    0.194
##
##                    exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention    1.284     0.7789    0.8808     1.872
##
## Concordance= 0.53  (se = 0.025 )
## Likelihood ratio test= 1.7  on 1 df,   p=0.2
## Wald test            = 1.69  on 1 df,   p=0.2
## Score (logrank) test = 1.7  on 1 df,   p=0.2
##
## [1] "Explanatory: TumorSize"
## [1] "Category: 2"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 335, number of events= 45
##
##                      coef exp(coef) se(coef)       z Pr(>|z|)
## GroupIntervention -0.6270    0.5342   0.3075 -2.039   0.0415 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                    exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention    0.5342     1.872    0.2924    0.9761
##
## Concordance= 0.578  (se = 0.038 )
## Likelihood ratio test= 4.31  on 1 df,   p=0.04
## Wald test            = 4.16  on 1 df,   p=0.04
## Score (logrank) test = 4.29  on 1 df,   p=0.04
```

```
##
## [1] "Explanatory: TumorSize"
## [1] "Category: 3"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 152, number of events= 26
##
##                         coef exp(coef) se(coef)      z Pr(>|z|)
## GroupIntervention -0.04436   0.95661  0.39358 -0.113     0.91
##
##                     exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention    0.9566      1.045    0.4423     2.069
##
## Concordance= 0.527  (se = 0.05 )
## Likelihood ratio test= 0.01  on 1 df,   p=0.9
## Wald test            = 0.01  on 1 df,   p=0.9
## Score (logrank) test = 0.01  on 1 df,   p=0.9
##
## [1] "Explanatory: TumorSize"
## [1] "Category: 4"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 215, number of events= 42
##
##                         coef exp(coef) se(coef)      z Pr(>|z|)
## GroupIntervention 0.1728     1.1886    0.3101 0.557     0.577
##
##                     exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention    1.189      0.8413    0.6472     2.183
##
## Concordance= 0.515  (se = 0.04 )
## Likelihood ratio test= 0.31  on 1 df,   p=0.6
## Wald test            = 0.31  on 1 df,   p=0.6
## Score (logrank) test = 0.31  on 1 df,   p=0.6
##
## [1] "Explanatory: PhysicalAct"
## [1] "Category: <210"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 850, number of events= 103
##
##                         coef exp(coef) se(coef)      z Pr(>|z|)
## GroupIntervention -0.1491    0.8615    0.1972 -0.756     0.45
##
##                     exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention    0.8615      1.161    0.5853     1.268
##
## Concordance= 0.516  (se = 0.026 )
```

```
## Likelihood ratio test= 0.57  on 1 df,    p=0.4
## Wald test            = 0.57  on 1 df,    p=0.4
## Score (logrank) test = 0.57  on 1 df,    p=0.4
##
## [1] "Explanatory: PhysicalAct"
## [1] "Category: >1290"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 738, number of events= 52
##
##                       coef exp(coef) se(coef)     z Pr(>|z|)
## GroupIntervention 0.01448   1.01458  0.27767 0.052    0.958
##
##                   exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention     1.015     0.9856    0.5888     1.748
##
## Concordance= 0.512  (se = 0.036 )
## Likelihood ratio test= 0  on 1 df,    p=1
## Wald test            = 0  on 1 df,    p=1
## Score (logrank) test = 0  on 1 df,    p=1
##
## [1] "Explanatory: PhysicalAct"
## [1] "Category: 616~1290"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 749, number of events= 71
##
##                      coef exp(coef) se(coef)     z Pr(>|z|)
## GroupIntervention 0.0336    1.0342   0.2374 0.141    0.887
##
##                   exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention     1.034      0.967    0.6493     1.647
##
## Concordance= 0.491  (se = 0.031 )
## Likelihood ratio test= 0.02  on 1 df,    p=0.9
## Wald test            = 0.02  on 1 df,    p=0.9
## Score (logrank) test = 0.02  on 1 df,    p=0.9
##
## [1] "Explanatory: PhysicalAct"
## [1] "Category: 211~615"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 751, number of events= 90
##
##                       coef exp(coef) se(coef)     z Pr(>|z|)
## GroupIntervention 0.03249   1.03303  0.21086 0.154    0.878
##
##                   exp(coef) exp(-coef) lower .95 upper .95
```

```
## GroupIntervention       1.033       0.968       0.6833       1.562
##
## Concordance= 0.506   (se = 0.027 )
## Likelihood ratio test= 0.02  on 1 df,    p=0.9
## Wald test            = 0.02  on 1 df,    p=0.9
## Score (logrank) test = 0.02  on 1 df,    p=0.9
##
## [1] "Explanatory: KCal"
## [1] "Category: <1430"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 1183, number of events= 133
##
##                        coef exp(coef) se(coef)       z Pr(>|z|)
## GroupIntervention -0.02112   0.97910  0.17346 -0.122     0.903
##
##                   exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention    0.9791      1.021    0.6969     1.376
##
## Concordance= 0.502   (se = 0.023 )
## Likelihood ratio test= 0.01  on 1 df,    p=0.9
## Wald test            = 0.01  on 1 df,    p=0.9
## Score (logrank) test = 0.01  on 1 df,    p=0.9
##
## [1] "Explanatory: KCal"
## [1] "Category: 1430~1680"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 793, number of events= 77
##
##                        coef exp(coef) se(coef)       z Pr(>|z|)
## GroupIntervention 0.03448   1.03508  0.22812 0.151      0.88
##
##                   exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention    1.035     0.9661    0.6619     1.619
##
## Concordance= 0.498   (se = 0.029 )
## Likelihood ratio test= 0.02  on 1 df,    p=0.9
## Wald test            = 0.02  on 1 df,    p=0.9
## Score (logrank) test = 0.02  on 1 df,    p=0.9
##
## [1] "Explanatory: KCal"
## [1] "Category: >1980"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 423, number of events= 53
##
##                        coef exp(coef) se(coef)       z Pr(>|z|)
```

```
## GroupIntervention -0.1080     0.8976    0.2761 -0.391     0.696
##
##                     exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention     0.8976      1.114     0.5225      1.542
##
## Concordance= 0.513  (se = 0.035 )
## Likelihood ratio test= 0.15  on 1 df,    p=0.7
## Wald test            = 0.15  on 1 df,    p=0.7
## Score (logrank) test = 0.15  on 1 df,    p=0.7
##
## [1] "Explanatory: KCal"
## [1] "Category: 1681~1980"
## Call:
## coxph(formula = Surv(SurvTime, Status) ~ Group, data = ImputedData[idx,
##     ])
##
##   n= 689, number of events= 53
##
##                      coef exp(coef) se(coef)      z Pr(>|z|)
## GroupIntervention 0.01183   1.01190  0.27479 0.043     0.966
##
##                   exp(coef) exp(-coef) lower .95 upper .95
## GroupIntervention     1.012     0.9882     0.5905      1.734
##
## Concordance= 0.512  (se = 0.036 )
## Likelihood ratio test= 0  on 1 df,    p=1
## Wald test            = 0  on 1 df,    p=1
## Score (logrank) test = 0  on 1 df,    p=1
```

```
plotModelError(ImputedAll, vars = 'allNumeric')
```
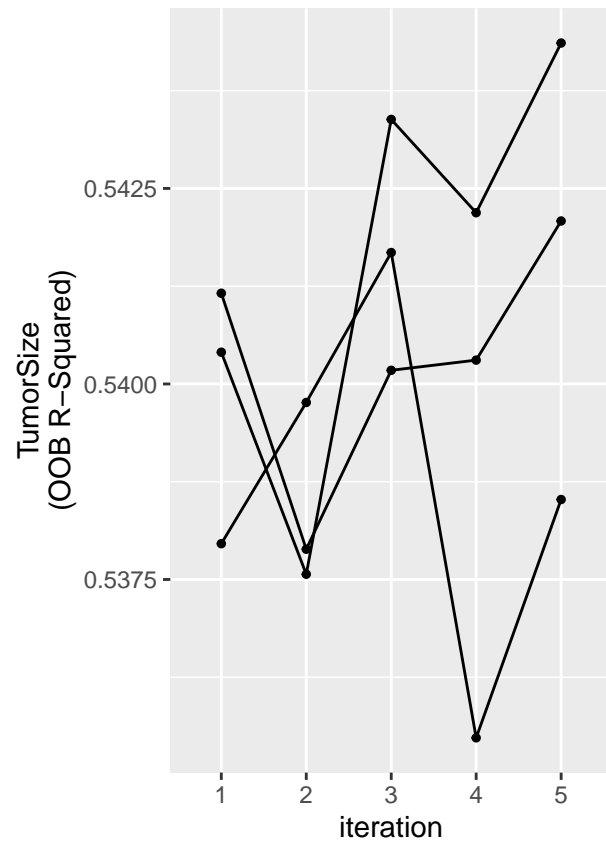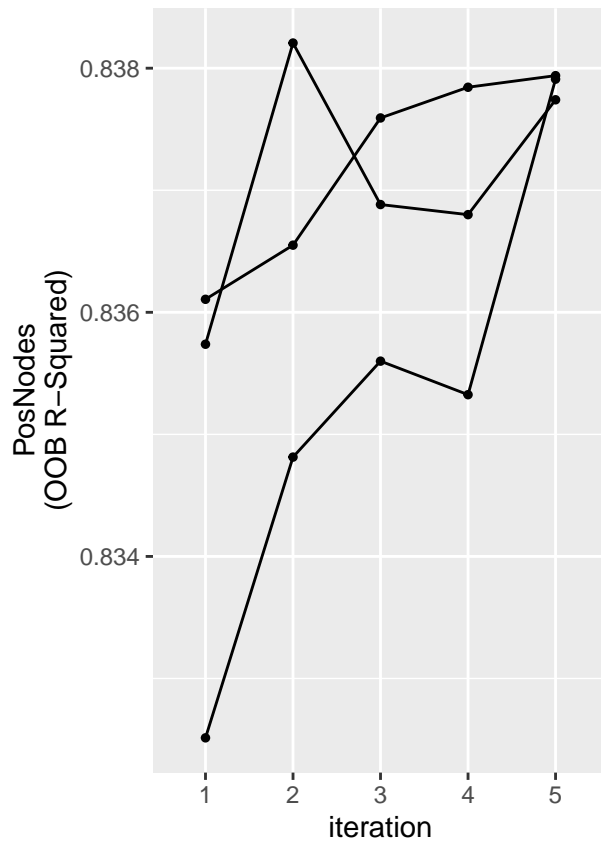
```
plotDistributions(ImputedAll, vars = 'allNumeric')
```
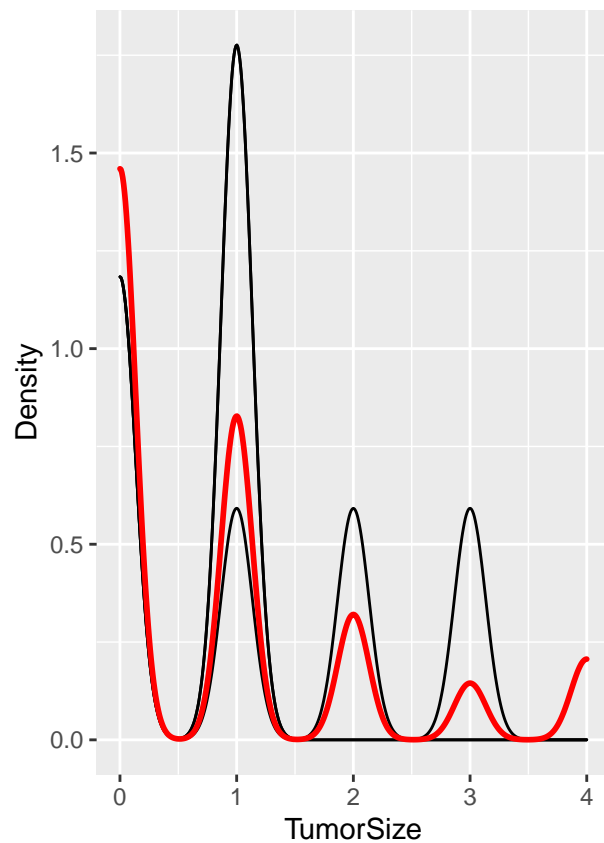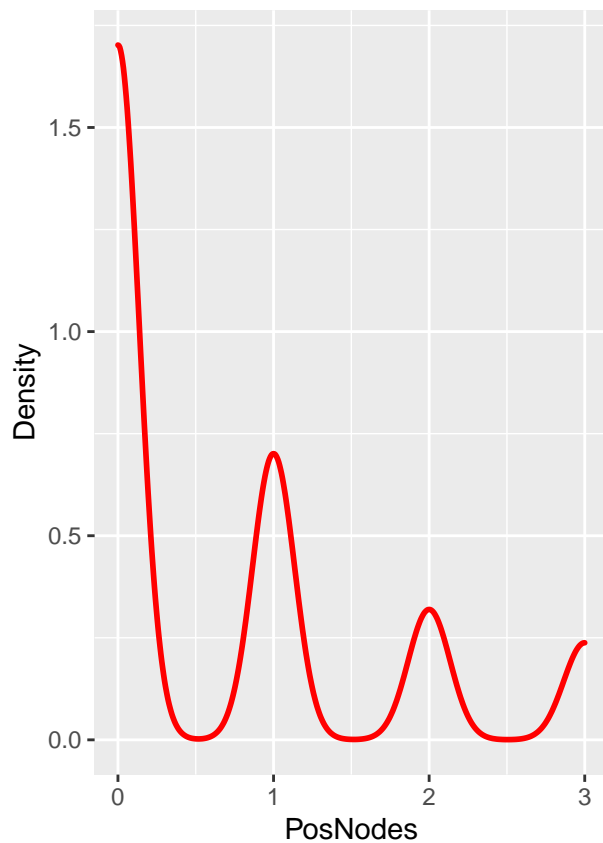
```
## Warning: Groups with fewer than two data points have been dropped.
```

```
## Warning: Groups with fewer than two data points have been dropped.
```

```
## Warning: Groups with fewer than two data points have been dropped.
```
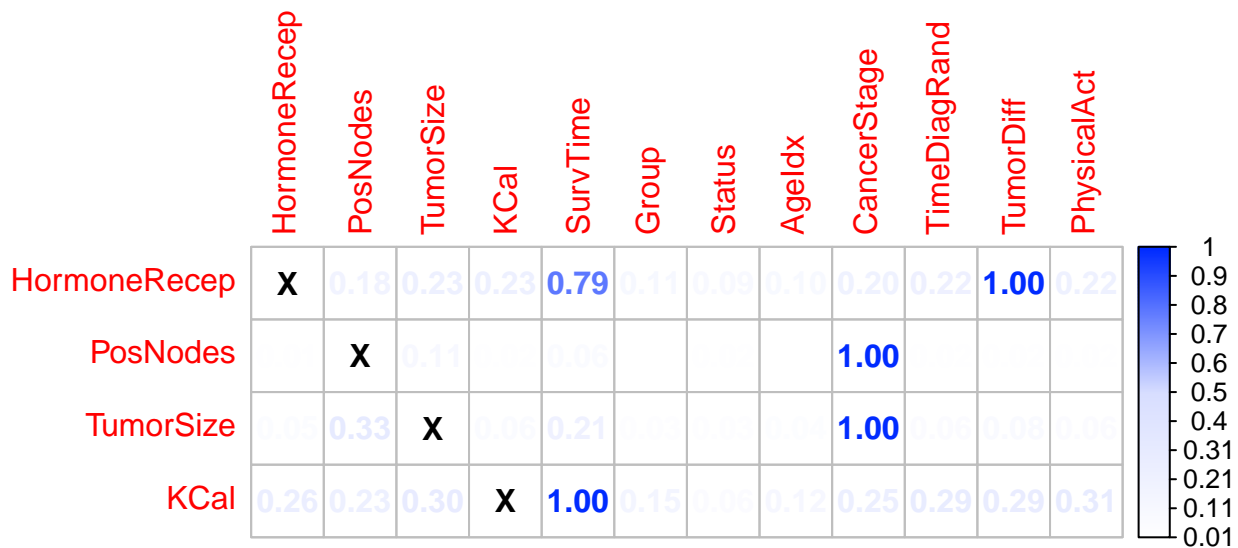
```
## Warning in max(ids, na.rm = TRUE): no non-missing arguments to max; returning
## -Inf
```

```
## Warning in max(ids, na.rm = TRUE): no non-missing arguments to max; returning
## -Inf
```

```
## Warning in max(ids, na.rm = TRUE): no non-missing arguments to max; returning
## -Inf
```

```
plotVarImportance(ImputedAll)
```


	HormoneRecep	PosNodes	TumorSize	KCal	SurvTime	Group	Status	AgeIdx	CancerStage	TimeDiagRand	TumorDiff	PhysicalAct
HormoneRecep	**X**	0.18	0.23	0.23	**0.79**	0.11	0.09	0.10	0.20	0.22	**1.00**	0.22
PosNodes		**X**	0.11	0.12	0.00				**1.00**			
TumorSize	0.05	0.33	**X**	0.06	0.21				**1.00**		0.06	0.08	0.06
KCal	0.26	0.23	0.30	**X**	**1.00**	0.15	0.06	0.12	0.25	0.29	0.29	0.31

```
library(sjPlot)
```

```
## Install package "strengejacke" from GitHub (`devtools::install_github("strengejacke/strengejacke")`)
library(sjmisc)
```

```
## Learn more about sjmisc with 'browseVignettes("sjmisc")'.
##
```

```
## Attaching package: 'sjmisc'
## The following object is masked from 'package:purrr':
##
##      is_empty
## The following object is masked from 'package:tidyr':
##
##      replace_na
## The following object is masked from 'package:tibble':
##
##      add_case
```

```
library(sjlabelled)
```

```
##
## Attaching package: 'sjlabelled'
## The following object is masked from 'package:finalfit':
##
##      remove_labels
## The following object is masked from 'package:forcats':
##
##      as_factor
## The following object is masked from 'package:dplyr':
##
##      as_label
## The following object is masked from 'package:ggplot2':
##
##      as_label
```

```
tab_model(coxph(Surv(SurvTime, Status) ~ ., data=AllCauseMortalityData))
```

```
## Argument 'df_method' is deprecated. Please use 'ci_method' instead.
```

Surv(SurvTime, Status)

Predictors

Estimates

CI

p

Group [Intervention]

0.94

0.56 – 1.59

0.825

AgeIdx [>=55]

3.37

1.89 – 6.01

<0.001

CancerStage [Stage I]

1.22

$0.54 - 2.77$

0.627

CancerStage [Stage IIB]

1.01

$0.36 - 2.85$

0.980

CancerStage [Stage IIIA]

1.05

$0.21 - 5.15$

0.953

CancerStage [Stage IIIB]

1.75

$0.20 - 15.66$

0.616

HormoneRecep [ER-/PR+]

1.16

$0.25 - 5.37$

0.845

HormoneRecep [ER+/PR-]

1.04

$0.37 - 2.93$

0.948

HormoneRecep [ER+/PR+]

1.55

$0.73 - 3.28$

0.257

TimeDiagRand

0.72

$0.54 - 0.95$

0.019

TumorDiff [Grade II,ModeratelyDifferentiated]

1.70

$0.59 - 4.93$

0.326

TumorDiff [Grade I, WellDifferentiated]

0.79

0.21 − 2.98

0.730

TumorDiff [Grade III,Poorly Differentiated]

2.10

0.70 − 6.33

0.187

PosNodes

1.12

0.60 − 2.09

0.720

TumorSize

1.20

0.90 − 1.61

0.221

PhysicalAct [>1290]

0.29

0.12 − 0.72

0.008

PhysicalAct [211~615]

0.85

0.45 − 1.60

0.619

PhysicalAct [616~1290]

0.54

0.26 − 1.10

0.091

KCal [1430~1680]

0.51

0.24 − 1.06

0.069

KCal [>1980]

1.08

0.52 − 2.25

0.839

KCal [1681~1980]

0.51

0.24 – 1.08

0.077

Observations

2274

R2 Nagelkerke

0.069