

```

N <- 700
Nsims <- 5000

maxFollow <- 84
cens.type <- "weibull"
##### Forest search criteria
hr.threshold <- 1.25 # Initital candidates
hr.consistency <- 1 # Candidates for many splits

pconsistency.threshold <- 0.9
stop.threshold <- 0.95

maxk <- 2
nmin.fs <- 60
pstop_futile <- 0.7

# Limit timing for forestsearch
max.minutes <- 3
m1.threshold <- Inf # Turning this off (Default)
# pconsistency.threshold<-0.70 # Minimum threshold (will choose max among
# subgroups satisfying)
fs.splits <- 400 # How many times to split for consistency
# vi is % factor is selected in cross-validation --> higher more important
vi.grf.min <- (-1) * 0.2 # This is default (to NOT exclude via VI)
# Null, turns off grf screening
d.min <- 10 # Min number of events for both arms (d0.min=d1.min=d.min)
# default=5

##### Virtual twins analysis Counter-factual difference
##### (C-E) >= vt.threshold Large values in favor of C
##### (control)

vt.threshold <- 0.175 # For VT delta

treat.threshold <- 0

maxdepth <- 2
n.min <- 60
ntree <- 1000

# GRF criteria
dmin.grf <- 12 # For GRF delta
# Note: For CRT this represents dmin.grf/2 RMS for control (-dmin.grf/2 for
# treatment)
frac.tau <- 0.6

# For forestsearch algorithm use same as GRF
frac.tau_fs <- 0.6
dmin.grf_fs <- 12
maxdepth_fs <- 2

label.analyses <- c("FS1", "GRF", "VT(24)", "VT#(24)", "VT(36)", "VT#(36)", "GRF.60")
# Classification table names
est_names <- c("$FS_{g}$", "$FS_{lg}$", "$GRF$", "$GRF_{60}$", "$VT(24)$", "{$VT}^{\{\backslash\#}(24)}$",
"$VT(36)$", "{$VT}^{\{\backslash\#}(36)}$")

```

```

outcome.name <- c("y.sim")
event.name <- c("event.sim")
id.name <- c("id")
treat.name <- c("treat")

cox.formula.sim <- as.formula(paste("Surv(y.sim,event.sim)~treat"))
cox.formula.adj.sim <- as.formula(paste("Surv(y.sim,event.sim)~treat+v1+v2+v3+v4+v5"))

get.FS <- TRUE
get.VT <- TRUE
get.GRF <- TRUE

fl_prefix <- paste0("oc_sims=", Nsims, "_")

out.loc <- paste0("results/", fl_prefix)

# m1 -censoring adjustment
muC.adj <- log(1.5)

# 0, 3, or 5
n_add_noise <- 0

mindex <- "m4a"
file.index <- "v0-4cutsTestVT"

z1_frac <- 0.25

if (mindex == "m4a") {
  k.z3 <- 1
  k.treat <- 0.9
  pH_super <- 0.125 # non-NULL re-defines z1_frac
}

if (mindex == "m4aB") {
  k.z3 <- 1
  k.treat <- 0.9
  pH_super <- 0.2 # non-NULL re-defines z1_frac
}

if (mindex == "m4b") {
  k.z3 <- 1
  k.treat <- 1.25
  pH_super <- 0.3 # non-NULL re-defines z1_frac
}

if (mindex == "m4c") {
  k.z3 <- 1
  k.treat <- 1.5
  pH_super <- 0.3 # non-NULL re-defines z1_frac
}

model.index <- paste0(mindex, "-Noise=", "")
model.index <- paste0(model.index, n_add_noise, "")
```

```

if (is.null(pH_super)) {
  # pH_check<-with(gbsg,mean(pgr<=quantile(pgr,c(z3_frac),1,0) &
  # er<=quantile(er,z1_frac)))
  pH_check <- with(gbsg, mean(meno == 0 & er <= quantile(er, z1_frac)))
  cat("Underlying pH_super", c(pH_check), "\n")
}

# pH_super specified If pH_super then override z1_frac and find z1_frac to
# yield pH_super

if (!is.null(pH_super)) {
  # Approximate Z1 quantile to yield pH proportion
  z1_q <- uniroot(propH.obj4, c(0, 1), tol = 1e-04, pH.target = pH_super)$root
  # pH_check<-with(gbsg,mean(pgr<=quantile(pgr,c(z3_frac),1,0) &
  # er<=quantile(er,z1_q)))
  pH_check <- with(gbsg, mean(meno == 0 & er <= quantile(er, z1_q)))
  cat("pH", c(pH_check), "\n")
  rel_error <- (pH_super - pH_check)/pH_super
  if (abs(rel_error) >= 0.1)
    stop("pH_super approximation relative error exceeds 10%")
  z1_frac <- z1_q
  cat("Underlying pH_super", c(pH_check), "\n")
}

## pH 0.122449
## Underlying pH_super 0.122449

# Bootstrap on log(hr) scale converted to HR (est.loghr=TRUE & est.scale='hr')
est.loghr <- TRUE
est.scale <- "hr"
t.start.all <- proc.time()[3]

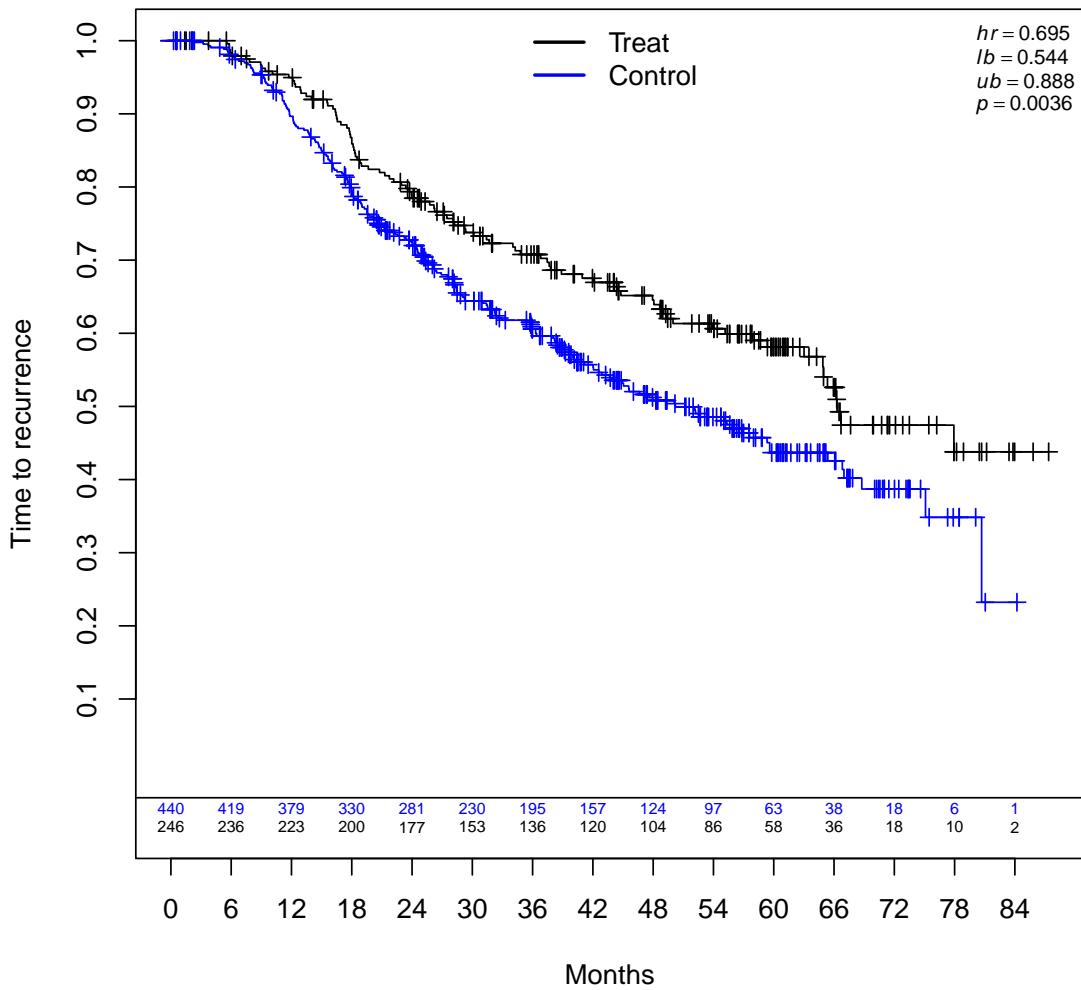
# Classification table names Note: within tab_tests (summary.VTFS) we rename so
# that denominator in ppv(hatH) is # hatH Manuscript section 3.2 will be
# updated accordingly

stat_names <- c("any(H)", "${sens}(\hat{H})$", "${sens}(\hat{H}^C)$", "${ppv}(\hat{H})$",
  "${ppv}(\hat{H}^C)$", "${avg}\sqrt{\hat{H}}\sqrt{$", "${min}\sqrt{\hat{H}}\sqrt{$",
  "${max}\sqrt{\hat{H}}\sqrt{$", "${avg}\sqrt{\hat{H}^C}\sqrt{$", "${min}\sqrt{\hat{H}^C}\sqrt{$",
  "${max}\sqrt{\hat{H}^C}\sqrt{$")

if (!get.FS) est_names <- est_names[-c(1:3)]


mod.harm <- "null"
this.dgm <- get.dgm4.OC(mod.harm = mod.harm, N = N, k.treat = k.treat, model.index = model.index,
  sol_tol = 10^-8, hrH.target = hrH.target, cens.type = cens.type, out.loc = out.loc,
  file.index = file.index, details = TRUE, parms_torand = FALSE)

```



```

## Super-population empirical harm and non-harm hazard ratios= NA 0.701027
## Causal HR (empirical ITT)= 0.701027

dgm <- this.dgm$dgm
output.file <- this.dgm$out.file

if (!is.null(output.file) & !grepl(mod.harm, output.file)) stop("Wrong file name for mod.harm")

# Show first simulation ans1 <- oc_analyses_m4_FS4(1)

t.start <- proc.time()[3]
res <- foreach(sim = seq_len(Nsims), .options.future = list(seed = TRUE), .combine = "rbind",
    .errorhandling = "pass") %dofuture% {
    ans <- oc_analyses_m4FourCuts_FS4(sim)
    return(ans)
}

## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 49.01796 2

```

```

##      leaf.node control.mean control.size control.se depth
## 1          2       1.67      77.00     2.98     1
## 2          3      -4.50     623.00     1.15     1
## 11         4      -5.23     179.00     2.04     2
## 21         5       1.78     171.00     1.89     2
## 4          7      -6.93     324.00     1.68     2
##      leaf.node control.mean control.size control.se depth
## 21         5       1.78     171.00     1.89     2
## GRF subgroup NOT found
## NO GRF cuts meeting delta(RMST): dmin.grf= 12
## # of continuous/categorical characteristics 1 6
## Continuous characteristics: size
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## Default cuts included (1st 20)
## Categorical: z1 z2 z3 z4 z5 grade3
## # of candidate subgroup factors= 10
## [1] "size <= 29.4" "size <= 25"    "size <= 20"    "size <= 35"    "z1"
## [6] "z2"           "z3"           "z4"           "z5"           "grade3"
## Number of factors evaluated= 10
## Confounders per grf screening q6 q8 q4 q7 q9 q5 q10 q3 q1 q2
##      Factors Labels VI(grf)
## 6          z2     q6  0.2016
## 8          z4     q8  0.1281
## 4  size <= 35   q4  0.1274
## 7          z3     q7  0.0963
## 9          z5     q9  0.0923
## 5          z1     q5  0.0856
## 10         grade3  q10 0.0730
## 3  size <= 20   q3  0.0724
## 1  size <= 29.4  q1  0.0634
## 2  size <= 25   q2  0.0598
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 210
## Approximately 5% of max_count met: minutes 0.0005833333
## Approximately 10% of max_count met: minutes 0.0009833333
## Approximately 20% of max_count met: minutes 0.0016666667
## Approximately 33% of max_count met: minutes 0.0026166667
## Approximately 50% of max_count met: minutes 0.0037333333
## Approximately 75% of max_count met: minutes 0.0055333333
## Approximately 90% of max_count met: minutes 0.0065833333
## # of subgroups evaluated based on (up to) maxk-factor combinations 210
## % of all-possible combinations (<= maxk) 100
## k.max= 2
## Events criteria for control,exp= 10 10
## # of subgroups with events less than criteria: control, experimental 21 20
## # of subgroups with sample size less than criteria 26
## # of subgroups meeting all criteria = 177
## # of subgroups fitted (Cox model estimable) = 177
## *Subgroup Searching Minutes== 0.006983333
## Number of subgroups meeting HR threshold 0
## NO subgroup candidate found (FS)
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 49.01796 2
##      leaf.node control.mean control.size control.se depth
## 1          2       1.67      77.00     2.98     1
## 2          3      -4.50     623.00     1.15     1

```

```

## 11      4      -5.23      179.00      2.04      2
## 21      5       1.78      171.00      1.89      2
## 4       7      -6.93      324.00      1.68      2
##   leaf.node control.mean control.size control.se depth
## 21      5       1.78      171.00      1.89      2
## GRF subgroup NOT found
## NO GRF cuts meeting delta(RMST): dmin.grf= 12
## # of continuous/categorical characteristics 1 6
## Continuous characteristics: size
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## CV lambda = 0.02832378
## 7 x 1 sparse Matrix of class "dgCMatrix"
##           s0
## z1      0.06829145
## z2      .
## z3      .
## z4      0.48686291
## z5     -0.73974508
## size    .
## grade3 .
## Cox-LASSO selected: z1 z4 z5
## Cox-LASSO not selected: z2 z3 size grade3
## Default cuts included from Lasso:
## Categorical after Lasso: z1 z4 z5
## # of candidate subgroup factors= 7
## [1] "size <= 29.4" "size <= 25"   "size <= 20"   "size <= 35"   "z1"
## [6] "z4"          "z5"
## Number of factors evaluated= 7
## Confounders per grf screening q6 q4 q7 q5 q3 q1 q2
##           Factors Labels VI(grf)
## 6       z4      q6  0.2034
## 4   size <= 35   q4  0.1819
## 7       z5      q7  0.1665
## 5       z1      q5  0.1391
## 3   size <= 20   q3  0.1159
## 1   size <= 29.4   q1  0.0980
## 2   size <= 25   q2  0.0951
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 105
## Approximately 5% of max_count met: minutes 0.0002333333
## Approximately 10% of max_count met: minutes 0.0004333333
## Approximately 20% of max_count met: minutes 0.0008166667
## Approximately 33% of max_count met: minutes 0.0014333333
## Approximately 50% of max_count met: minutes 0.0017833333
## Approximately 75% of max_count met: minutes 0.0027333333
## Approximately 90% of max_count met: minutes 0.0032833333
## # of subgroups evaluated based on (up to) maxk-factor combinations 105
## % of all-possible combinations (<= maxk) 100
## k.max= 2
## Events criteria for control,exp= 10 10
## # of subgroups with events less than criteria: control, experimental 15 15
## # of subgroups with sample size less than criteria 18
## # of subgroups meeting all criteria = 81
## # of subgroups fitted (Cox model estimable) = 81
## *Subgroup Searching Minutes=* 0.003383333
## Number of subgroups meeting HR threshold 0

```

```

## NO subgroup candidate found (FS)
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 48.60698 2
##   leaf.node control.mean control.size control.se depth
## 1      2      -2.96      698.00      1.06      1
## 2      4      -3.92      322.00      1.51      2
## 3      5       3.72      86.00      3.38      2
## 4      6       1.68      90.00      2.75      2
## 5      7      -6.28      202.00      1.96      2
##   leaf.node control.mean control.size control.se depth
## 3      5       3.72      86.00      3.38      2
## GRF subgroup NOT found
## NO GRF cuts meeting delta(RMST): dmin.grf= 12
## # of continuous/categorical characteristics 1 6
## Continuous characteristics: size
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## Default cuts included (1st 20)
## Categorical: z1 z2 z3 z4 z5 grade3
## # of candidate subgroup factors= 10
## [1] "size <= 29.2" "size <= 25"    "size <= 20"    "size <= 35"    "z1"
## [6] "z2"           "z3"          "z4"          "z5"          "grade3"
## Number of factors evaluated= 10
## Confounders per grf screening q5 q6 q3 q9 q10 q7 q8 q4 q2 q1
##   Factors Labels VI(grf)
## 5      z1     q5  0.1346
## 6      z2     q6  0.1193
## 3  size <= 20   q3  0.1112
## 9      z5     q9  0.1047
## 10     grade3  q10 0.1041
## 7      z3     q7  0.1003
## 8      z4     q8  0.0996
## 4  size <= 35   q4  0.0819
## 2  size <= 25   q2  0.0818
## 1  size <= 29.2  q1  0.0626
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 210
## Approximately 5% of max_count met: minutes 0.0004333333
## Approximately 10% of max_count met: minutes 0.00085
## Approximately 20% of max_count met: minutes 0.004083333
## Approximately 33% of max_count met: minutes 0.004866667
## Approximately 50% of max_count met: minutes 0.006233333
## Approximately 75% of max_count met: minutes 0.008066667
## Approximately 90% of max_count met: minutes 0.00905
## # of subgroups evaluated based on (up to) maxk-factor combinations 210
## % of all-possible combinations (<= maxk) 100
## k.max= 2
## Events criteria for control,exp= 10 10
## # of subgroups with events less than criteria: control, experimental 20 21
## # of subgroups with sample size less than criteria 28
## # of subgroups meeting all criteria = 175
## # of subgroups fitted (Cox model estimable) = 175
## *Subgroup Searching Minutes=* 0.009533333
## Number of subgroups meeting HR threshold 2
## Subgroup candidate(s) found (FS)
## # of candidate subgroups (meeting HR criteria) = 2
## SGs (1st 10) meeting screening thresholds sorted by sg_focus= hr

```

```

##      n  E d1   HR L(HR) q5.0 q5.1
## 1  78 40 21 1.47  0.79    0    0
## 2 152 69 38 1.33  0.82    1    0
## Consistency 0.7425
## Consistency 0.7725
## Subgroup Consistency Minutes= 0.04836667
## NO subgroup found (FS)
## Minutes forestsearch overall= 0.06425
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 48.60698 2
##   leaf.node control.mean control.size control.se depth
## 1          2        -2.96     698.00     1.06     1
## 2          4        -3.92     322.00     1.51     2
## 3          5         3.72     86.00     3.38     2
## 4          6         1.68     90.00     2.75     2
## 5          7        -6.28     202.00     1.96     2
##   leaf.node control.mean control.size control.se depth
## 3          5         3.72     86.00     3.38     2
## GRF subgroup NOT found
## NO GRF cuts meeting delta(RMST): dmin.grf= 12
## # of continuous/categorical characteristics 1 6
## Continuous characteristics: size
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## CV lambda = 0.04404393
## 7 x 1 sparse Matrix of class "dgCMatrix"
##           s0
## z1 .
## z2 .
## z3 .
## z4 0.4650680
## z5 -0.7663331
## size .
## grade3 .
## Cox-LASSO selected: z4 z5
## Cox-LASSO not selected: z1 z2 z3 size grade3
## Default cuts included from Lasso:
## Categorical after Lasso: z4 z5
## # of candidate subgroup factors= 6
## [1] "size <= 29.2" "size <= 25"   "size <= 20"   "size <= 35"   "z4"
## [6] "z5"
## Number of factors evaluated= 6
## Confounders per grf screening q5 q6 q3 q4 q1 q2
##       Factors Labels VI(grf)
## 5      z4      q5  0.2108
## 6      z5      q6  0.2009
## 3 size <= 20   q3  0.1674
## 4 size <= 35   q4  0.1559
## 1 size <= 29.2 q1  0.1325
## 2 size <= 25   q2  0.1325
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 78
## Approximately 5% of max_count met: minutes 0.0001666667
## Approximately 10% of max_count met: minutes 0.0003333333
## Approximately 20% of max_count met: minutes 0.0006166667
## Approximately 33% of max_count met: minutes 0.001016667
## Approximately 50% of max_count met: minutes 0.001483333

```

```

## Approximately 75% of max_count met: minutes 0.002266667
## Approximately 90% of max_count met: minutes 0.002366667
## # of subgroups evaluated based on (up to) maxk-factor combinations 78
## % of all-possible combinations (<= maxk) 100
## k.max= 2
## Events criteria for control,exp= 10 10
## # of subgroups with events less than criteria: control, experimental 12 13
## # of subgroups with sample size less than criteria 14
## # of subgroups meeting all criteria = 58
## # of subgroups fitted (Cox model estimable) = 58
## *Subgroup Searching Minutes=** 0.0025
## Number of subgroups meeting HR threshold 0
## NO subgroup candidate found (FS)
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 49.94034 2
##     leaf.node control.mean control.size control.se depth
## 1      2       -4.26      697.00      1.10      1
## 11     4        0.91      361.00      1.52      2
## 3      6       -9.81      282.00      1.63      2
##     leaf.node control.mean control.size control.se depth
## 11     4        0.91      361.00      1.52      2
## GRF subgroup NOT found
## NO GRF cuts meeting delta(RMST): dmin.grf= 12
## # of continuous/categorical characteristics 1 6
## Continuous characteristics: size
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## Default cuts included (1st 20)
## Categorical: z1 z2 z3 z4 z5 grade3
## # of candidate subgroup factors= 10
## [1] "size <= 29.3" "size <= 25"   "size <= 20"   "size <= 35"   "z1"
## [6] "z2"           "z3"         "z4"         "z5"         "grade3"
## Number of factors evaluated= 10
## Confounders per grf screening q7 q6 q5 q9 q8 q3 q1 q4 q10 q2
##     Factors Labels VI(grf)
## 7      z3    q7  0.2459
## 6      z2    q6  0.1947
## 5      z1    q5  0.0908
## 9      z5    q9  0.0906
## 8      z4    q8  0.0895
## 3  size <= 20  q3  0.0655
## 1  size <= 29.3 q1  0.0606
## 4  size <= 35  q4  0.0585
## 10  grade3    q10 0.0523
## 2  size <= 25  q2  0.0515
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 210
## Approximately 5% of max_count met: minutes 0.00045
## Approximately 10% of max_count met: minutes 0.00095
## Approximately 20% of max_count met: minutes 0.001683333
## Approximately 33% of max_count met: minutes 0.002766667
## Approximately 50% of max_count met: minutes 0.004116667
## Approximately 75% of max_count met: minutes 0.005766667
## Approximately 90% of max_count met: minutes 0.00655
## # of subgroups evaluated based on (up to) maxk-factor combinations 210
## % of all-possible combinations (<= maxk) 100
## k.max= 2

```

```

## Events criteria for control,exp= 10 10
## # of subgroups with events less than criteria: control, experimental 18 21
## # of subgroups with sample size less than criteria 24
## # of subgroups meeting all criteria = 178
## # of subgroups fitted (Cox model estimable) = 178
## *Subgroup Searching Minutes== 0.00945
## Number of subgroups meeting HR threshold 4
## Subgroup candidate(s) found (FS)
## # of candidate subgroups (meeting HR criteria) = 4
## SGs (1st 10) meeting screening thresholds sorted by sg_focus= hr
##      n   E d1    HR L(HR) q7.0 q7.1 q6.0 q6.1
## 1 107 42 27 1.86  0.99    0    0    0    0
## 2 104 44 29 1.62  0.87    0    0    1    0
## 3 125 57 33 1.33  0.79    1    0    0    0
## 4 129 92 41 1.30  0.86    0    0    1    0
## Consistency 0.9275
## # of splits= 400
## Model, % Consistency Met= !{z4} {size <= 20} 0.9275
## Consistency 0.895
## Consistency 0.705
## Consistency 0.7725
## SG focus= hr
## Subgroup Consistency Minutes= 0.09361667
## Subgroup found (FS)
## Minutes forestsearch overall= 0.1088
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 49.94034 2
##      leaf.node control.mean control.size control.se depth
## 1       2        -4.26     697.00     1.10     1
## 11      4         0.91     361.00     1.52     2
## 3       6        -9.81     282.00     1.63     2
##      leaf.node control.mean control.size control.se depth
## 11      4         0.91     361.00     1.52     2
## GRF subgroup NOT found
## NO GRF cuts meeting delta(RMST): dmin.grf= 12
## # of continuous/categorical characteristics 1 6
## Continuous characteristics: size
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## CV lambda = 0.01324073
## 7 x 1 sparse Matrix of class "dgCMatrix"
##           s0
## z1      .
## z2      .
## z3      0.010035680
## z4      0.436760697
## z5      -1.062587139
## size    0.006726465
## grade3  0.156893502
## Cox-LASSO selected: z3 z4 z5 size grade3
## Cox-LASSO not selected: z1 z2
## Default cuts included from Lasso:
## Categorical after Lasso: z3 z4 z5 grade3
## # of candidate subgroup factors= 8
## [1] "size <= 29.3" "size <= 25"   "size <= 20"   "size <= 35"   "z3"
## [6] "z4"          "z5"          "grade3"

```

```

## Number of factors evaluated= 8
## Confounders per grf screening q5 q6 q7 q3 q1 q2 q4 q8
##          Factors Labels VI(grf)
## 5          z3      q5  0.3941
## 6          z4      q6  0.1235
## 7          z5      q7  0.1135
## 3  size <= 20      q3  0.0847
## 1  size <= 29.3     q1  0.0749
## 2  size <= 25      q2  0.0732
## 4  size <= 35      q4  0.0729
## 8  grade3      q8  0.0631
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 136
## Approximately 5% of max_count met: minutes 0.0002833333
## Approximately 10% of max_count met: minutes 0.0005833333
## Approximately 20% of max_count met: minutes 0.0011333333
## Approximately 33% of max_count met: minutes 0.0018333333
## Approximately 50% of max_count met: minutes 0.00265
## Approximately 75% of max_count met: minutes 0.0039666667
## Approximately 90% of max_count met: minutes 0.0043166667
## # of subgroups evaluated based on (up to) maxk-factor combinations 136
## % of all-possible combinations (<= maxk) 100
## k.max= 2
## Events criteria for control,exp= 10 10
## # of subgroups with events less than criteria: control, experimental 15 17
## # of subgroups with sample size less than criteria 18
## # of subgroups meeting all criteria = 111
## # of subgroups fitted (Cox model estimable) = 111
## *Subgroup Searching Minutes= 0.004666667
## Number of subgroups meeting HR threshold 2
## Subgroup candidate(s) found (FS)
## # of candidate subgroups (meeting HR criteria) = 2
## SGs (1st 10) meeting screening thresholds sorted by sg_focus= hr
##      n   E d1   HR L(HR) q5.0 q5.1
## 1 107 42 27 1.86  0.99    0    0
## 2 125 57 33 1.33  0.79    1    0
## Consistency 0.9275
## # of splits= 400
## Model, % Consistency Met= !{z4} {size <= 20} 0.9275
## Consistency 0.705
## SG focus= hr
## Subgroup Consistency Minutes= 0.04683333
## Subgroup found (FS)
## Minutes forestsearch overall= 0.05856667

t.now <- proc.time()[3]
t.min <- (t.now - t.start)/60

print(table(res$analysis))

##
##      FS1     FSlg      GRF    GRF.60    VT(24)    VT(36)    VT#(24)    VT#(36)
##      5000     5000     5000     5000     5000     5000     5000     5000

check <- c(c(table(res$analysis)) - Nsims)
if (all(check != 0)) stop("All analyses not complete")

```

```

dgm_alt <- dgm
outres <- out.results(res = res, dgm = dgm, output.file = output.file, t.min = t.min,
                      out_analysis = "FS1")

## [1] "results/oc_sims=5000_m4a-Noise=0_N=700_null_ktreat=0.9_v0-4cutsTestVT.Rdata"
##      sim sizeH_true propH_true sizeHc_true propHc_true any.H size.H size.Hc
##      <int>     <num>     <num>     <int>     <num> <num> <num> <int>
## 1:    1       0       0       700       1       0       0     700
## 2:    1       0       0       700       1       0       0     700
## 3:    1       0       0       700       1       0       0     700
## 4:    1       0       0       700       1       0       0     700
## 5:    1       0       0       700       1       0       0     700
## 6:    1       0       0       700       1       0       0     700
##      ppv    npv specificity sensitivity found.1 found.2 found.both found.al3
##      <lgcl> <num>     <num>     <num> <int> <int> <num> <num>
## 1:     NA     1       1       NA     0     0     0     0
## 2:     NA     1       1       NA     0     0     0     0
## 3:     NA     1       1       NA     NA    NA    NA    NA
## 4:     NA     1       1       NA     NA    NA    NA    NA
## 5:     NA     1       1       NA     0     0     0     0
## 6:     NA     1       1       NA     0     0     0     0
##      hr.H.true hr.Hc.true hr.H.hat hr.Hc.hat   b1.H   b2.H b1.Hc      b2.Hc
##      <lgcl> <num>     <num>     <num> <lgcl> <lgcl> <num>     <num>
## 1:     NA 0.6155054     NA 0.6155054     NA     NA     0 -0.08552157
## 2:     NA 0.6155054     NA 0.6155054     NA     NA     0 -0.08552157
## 3:     NA 0.6155054     NA 0.6155054     NA     NA     0 -0.08552157
## 4:     NA 0.6155054     NA 0.6155054     NA     NA     0 -0.08552157
## 5:     NA 0.6155054     NA 0.6155054     NA     NA     0 -0.08552157
## 6:     NA 0.6155054     NA 0.6155054     NA     NA     0 -0.08552157
##      p.cens analysis  taumax    hr.itt    l.itt    u.itt hr.adj.itt
##      <num> <char>     <num>     <num>     <num>     <num>     <num>
## 1: 0.4642857    FS1 81.69660 0.6155054 0.5014207 0.755547 0.5906047
## 2: 0.4642857   FSlg 49.01796 0.6155054 0.5014207 0.755547 0.5906047
## 3: 0.4642857    GRF 81.69660 0.6155054 0.5014207 0.755547 0.5906047
## 4: 0.4642857    GRF 60.49.01796 0.6155054 0.5014207 0.755547 0.5906047
## 5: 0.4642857    VT(24) 81.69660 0.6155054 0.5014207 0.755547 0.5906047
## 6: 0.4642857   VT#(24) 81.69660 0.6155054 0.5014207 0.755547 0.5906047
##      l.adj.itt u.adj.itt l.H.true u.H.true l.Hc.true u.Hc.true l.H.hat u.H.hat
##      <num>     <num> <lgcl> <lgcl>     <num>     <num> <num>     <num>
## 1: 0.480604 0.7257825     NA     NA 0.5014207 0.755547     NA     NA
## 2: 0.480604 0.7257825     NA     NA 0.5014207 0.755547     NA     NA
## 3: 0.480604 0.7257825     NA     NA 0.5014207 0.755547     NA     NA
## 4: 0.480604 0.7257825     NA     NA 0.5014207 0.755547     NA     NA
## 5: 0.480604 0.7257825     NA     NA 0.5014207 0.755547     NA     NA
## 6: 0.480604 0.7257825     NA     NA 0.5014207 0.755547     NA     NA
##      l.Hc.hat u.Hc.hat
##      <num>     <num>
## 1: 0.5014207 0.755547
## 2: 0.5014207 0.755547
## 3: 0.5014207 0.755547
## 4: 0.5014207 0.755547
## 5: 0.5014207 0.755547
## 6: 0.5014207 0.755547
## Subgroup HRs: H, H^c, Causal= NA 0.701027 0.701027
## Simulations= 5000

```

```

## Avg censoring= 0.4616311
## Min,Max,Avg tau.max= 72.81026 83.99787 81.71174
## P(H) approximation at causal(Hrc), n=60, approx= 0.701027 0.03395014
## P(H) approximation at plim(Hrc), n=60, approx= 0.7056473 0.03532477
## Minutes,hours 24.84305 0.4140508
##          FS1   FS1g     GRF   GRF.60   VT(24) VT#(24)   VT(36) VT#(36)
## any.H      0.070  0.060    0.250   0.050    0.090   0.060    0.100   0.050
## sensH       NaN    NaN    NaN    NaN    NaN    NaN    NaN    NaN
## sensHc      0.990  0.990    0.970   0.990    0.990   0.990    0.990   0.990
## ppH        0.000  0.000    0.000   0.000    0.000   0.000    0.000   0.000
## ppHc       1.000  1.000    1.000   1.000    1.000   1.000    1.000   1.000
## Avg(#H)    100.000 101.000   87.000   77.000   84.000   83.000   88.000   81.000
## minH       61.000  61.000   60.000   60.000   60.000   60.000   60.000   60.000
## maxH      234.000 234.000   269.000   168.000   247.000   193.000   210.000   174.000
## Avg(#Hc)   693.000 694.000   678.000   696.000   693.000   695.000   692.000   696.000
## minHc      466.000 466.000   431.000   532.000   453.000   507.000   490.000   526.000
## maxHc      700.000 700.000   700.000   700.000   700.000   700.000   700.000   700.000
## hat(H*)     NaN    NaN    NaN    NaN    NaN    NaN    NaN    NaN
## hat(hat[H]) 1.789  1.798    1.540   1.542    1.201   1.280    1.218   1.429
## hat(Hc*)    0.781  0.783    0.738   0.736    0.729   0.737    0.738   0.750
## hat(hat[Hc]) 0.690  0.691    0.669   0.675    0.686   0.688    0.692   0.693
## hat(H*)all   NaN    NaN    NaN    NaN    NaN    NaN    NaN    NaN
## hat(Hc*)all  0.706  0.706    0.706   0.706    0.706   0.706    0.706   0.706
## hat(ITT)all  0.706  0.706    0.706   0.706    0.706   0.706    0.706   0.706
## hat(ITTadj)all 0.664  0.664    0.664   0.664    0.664   0.664    0.664   0.664

missC <- tab_tests(res = res)

pA <- as.character(round(outres$pAnyH.approx2, 4))
tabsim_missC <- get_tabsim(missC = missC, pA = pA, est_names = est_names, stat_names = stat_names,
                             mod.harm = mod.harm, Nsims = Nsims)

```

Table 1: Average classification rates:  $avg|\hat{H}|$ ,  $min|\hat{H}|$ , and  $max|\hat{H}|$ , denote the average, minimum, and maximum of the number of subjects in the estimated subgroup  $\hat{H}$  (analogously for  $\hat{H}^c$ ). Note that under the null  $sens(\hat{H})$  is undefined and  $ppv(\hat{H}) = 0$ .

	$FS_g$	$FS_{lg}$	$GRF$	$GRF_{60}$	$VT(24)$	$VT^{\#}(24)$	$VT(36)$	$VT^{\#}(36)$
Finding H								
any(H)	0.07	0.06	0.25	0.05	0.09	0.06	0.1	0.05
$sens(\hat{H})$	.	.	.	.	.	.	.	.
$sens(\hat{H}^c)$	0.99	0.99	0.97	0.99	0.99	0.99	0.99	0.99
$ppv(\hat{H})$	0	0	0	0	0	0	0	0
$ppv(\hat{H}^c)$	1	1	1	1	1	1	1	1
Size of H and H-complement								
$avg \hat{H} $	100	101	87	77	84	83	88	81
$min \hat{H} $	61	61	60	60	60	60	60	60
$max \hat{H} $	234	234	269	168	247	193	210	174
$avg \hat{H}^c $	693	694	678	696	693	695	692	696
$min \hat{H}^c $	466	466	431	532	453	507	490	526
$max \hat{H}^c $	700	700	700	700	700	700	700	700

Note: Number of simulations= 5000 .

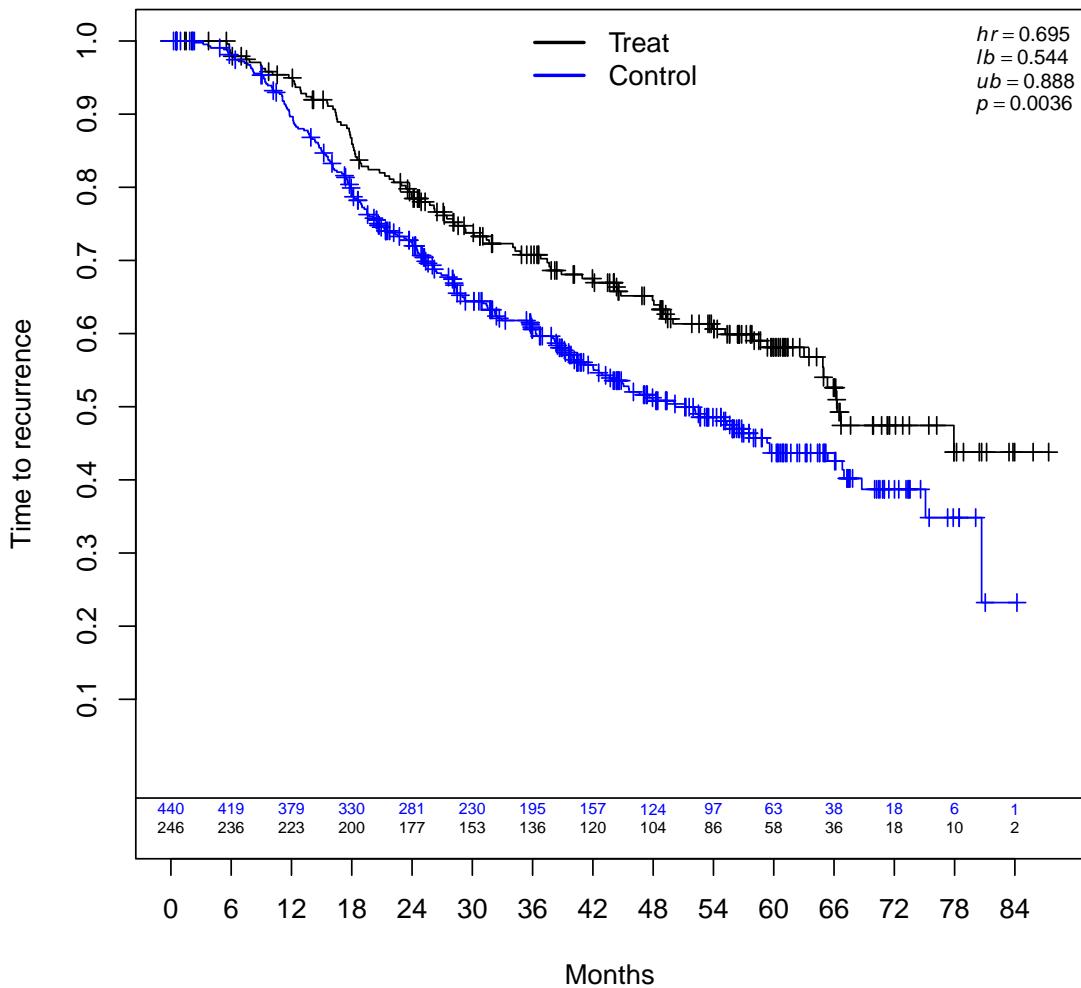
Note: Probability approximation= 0.0353 .

```

mod.harm <- "alt"
hrH.target <- 2

this.dgm <- get.dgm4.OC(mod.harm = mod.harm, N = N, k.treat = k.treat, model.index = model.index,
sol_tol = 10^-8, hrH.target = hrH.target, cens.type = cens.type, out.loc = out.loc,
file.index = file.index, details = TRUE, parms_torand = FALSE)

```



```

## Super-population empirical harm and non-harm hazard ratios= 2.000007 0.6466405
## Causal HR (empirical ITT)= 0.7057463

dgm <- this.dgm$dgm
output.file <- this.dgm$out.file

if (!is.null(output.file) & !grep1(mod.harm, output.file)) stop("Wrong file name for mod.harm")

# Show first simulation ans1 <- oc_analyses_m4_FS4(2)

t.start <- proc.time() [3]

```

```

res <- foreach(sim = seq_len(Nsims), .options.future = list(seed = TRUE), .combine = "rbind",
  .errorhandling = "pass") %dofuture% {
  ans <- oc_analyses_m4FourCuts_FS4(sim)
  return(ans)
}

## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 48.86072 2
##   leaf.node control.mean control.size control.se depth
## 1      2      -4.43     509.00     1.22     1
## 2      3       2.57     191.00     2.23     1
## 3      4      -3.89     213.00     1.97     2
## 4      5       4.14      62.00     2.97     2
## 5      6      -5.84     313.00     1.61     2
## 6      7       5.70     112.00     2.81     2
##   leaf.node control.mean control.size control.se depth
## 6      7       5.70     112.00     2.81     2
## GRF subgroup NOT found
## NO GRF cuts meeting delta(RMST): dmin.grf= 12
## # of continuous/categorical characteristics 1 6
## Continuous characteristics: size
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## Default cuts included (1st 20)
## Categorical: z1 z2 z3 z4 z5 grade3
## # of candidate subgroup factors= 10
## [1] "size <= 29.4" "size <= 25"    "size <= 20"    "size <= 35"    "z1"
## [6] "z2"           "z3"          "z4"          "z5"          "grade3"
## Number of factors evaluated= 10
## Confounders per grf screening q6 q7 q5 q8 q9 q4 q3 q10 q2 q1
##   Factors Labels VI(grf)
## 6      z2     q6  0.2715
## 7      z3     q7  0.2034
## 5      z1     q5  0.1818
## 8      z4     q8  0.0603
## 9      z5     q9  0.0596
## 4  size <= 35     q4  0.0531
## 3  size <= 20     q3  0.0472
## 10  grade3     q10 0.0426
## 2  size <= 25     q2  0.0404
## 1  size <= 29.4    q1  0.0402
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 210
## Approximately 5% of max_count met: minutes 0.00045
## Approximately 10% of max_count met: minutes 0.0008666667
## Approximately 20% of max_count met: minutes 0.00145
## Approximately 33% of max_count met: minutes 0.002433333
## Approximately 50% of max_count met: minutes 0.003766667
## Approximately 75% of max_count met: minutes 0.005416667
## Approximately 90% of max_count met: minutes 0.006283333
## # of subgroups evaluated based on (up to) maxk-factor combinations 210
## % of all-possible combinations (<= maxk) 100
## k.max= 2
## Events criteria for control,exp= 10 10
## # of subgroups with events less than criteria: control, experimental 21 20
## # of subgroups with sample size less than criteria 26
## # of subgroups meeting all criteria = 177

```

```

## # of subgroups fitted (Cox model estimable) = 177
## *Subgroup Searching Minutes= 0.006683333
## Number of subgroups meeting HR threshold 6
## Subgroup candidate(s) found (FS)
## # of candidate subgroups (meeting HR criteria) = 6
## SGs (1st 10) meeting screening thresholds sorted by sg_focus= hr
##   n  E d1    HR L(HR) q6.0 q6.1 q7.0 q7.1 q5.0 q5.1
## 1 95 78 44 1.74  1.11    0    0    0    1    0    1
## 2 122 95 52 1.48  0.98    0    1    0    0    0    1
## 3 100 80 41 1.47  0.94    0    0    0    0    0    1
## 4  77 55 30 1.41  0.83    0    0    0    0    0    1
## 5  90 64 32 1.31  0.80    0    0    0    0    0    1
## 6  67 39 21 1.30  0.69    0    0    0    1    0    0
## Consistency 0.9675
## # of splits= 400
## Model, % Consistency Met= {z3} {z1} 0.9675
## SG focus= hr
## Subgroup Consistency Minutes= 0.02251667
## Subgroup found (FS)
## Minutes forestsearch overall= 0.03696667
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 48.86072 2
##   leaf.node control.mean control.size control.se depth
## 1        2      -4.43     509.00     1.22     1
## 2        3       2.57     191.00     2.23     1
## 3        4      -3.89     213.00     1.97     2
## 4        5       4.14      62.00     2.97     2
## 5        6      -5.84     313.00     1.61     2
## 6        7       5.70     112.00     2.81     2
##   leaf.node control.mean control.size control.se depth
## 6        7       5.70     112.00     2.81     2
## GRF subgroup NOT found
## NO GRF cuts meeting delta(RMST): dmin.grf= 12
## # of continuous/categorical characteristics 1 6
## Continuous characteristics: size
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## CV lambda = 0.01355153
## 7 x 1 sparse Matrix of class "dgCMatrix"
##           s0
## z1  0.32115232
## z2 .
## z3  0.09066173
## z4  0.48299291
## z5 -0.78083587
## size .
## grade3 .
## Cox-LASSO selected: z1 z3 z4 z5
## Cox-LASSO not selected: z2 size grade3
## Default cuts included from Lasso:
## Categorical after Lasso: z1 z3 z4 z5
## # of candidate subgroup factors= 8
## [1] "size <= 29.4" "size <= 25"   "size <= 20"   "size <= 35"   "z1"
## [6] "z3"          "z4"          "z5"
## Number of factors evaluated= 8
## Confounders per grf screening q6 q5 q8 q4 q7 q3 q2 q1

```

```

##      Factors Labels VI(grf)
## 6      z3     q6  0.3637
## 5      z1     q5  0.2343
## 8      z5     q8  0.0800
## 4 size <= 35    q4  0.0775
## 7      z4     q7  0.0749
## 3 size <= 20    q3  0.0638
## 2 size <= 25    q2  0.0572
## 1 size <= 29.4   q1  0.0485
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 136
## Approximately 5% of max_count met: minutes 0.0002666667
## Approximately 10% of max_count met: minutes 0.0005833333
## Approximately 20% of max_count met: minutes 0.0011
## Approximately 33% of max_count met: minutes 0.0019
## Approximately 50% of max_count met: minutes 0.002583333
## Approximately 75% of max_count met: minutes 0.003483333
## Approximately 90% of max_count met: minutes 0.004233333
## # of subgroups evaluated based on (up to) maxk-factor combinations 136
## % of all-possible combinations (<= maxk) 100
## k.max= 2
## Events criteria for control,exp= 10 10
## # of subgroups with events less than criteria: control, experimental 16 16
## # of subgroups with sample size less than criteria 19
## # of subgroups meeting all criteria = 111
## # of subgroups fitted (Cox model estimable) = 111
## *Subgroup Searching Minutes= 0.004466667
## Number of subgroups meeting HR threshold 5
## Subgroup candidate(s) found (FS)
## # of candidate subgroups (meeting HR criteria) = 5
## SGs (1st 10) meeting screening thresholds sorted by sg_focus= hr
##      n  E d1  HR L(HR) q6.0 q6.1 q5.0 q5.1 q8.0
## 1  95 78 44 1.74  1.11    0    1    0    1    0
## 2 100 80 41 1.47  0.94    0    0    0    1    1
## 3  77 55 30 1.41  0.83    0    0    0    1    0
## 4  90 64 32 1.31  0.80    0    0    0    1    0
## 5  67 39 21 1.30  0.69    0    1    0    0    0
## Consistency 0.9675
## # of splits= 400
## Model, % Consistency Met= {z3} {z1} 0.9675
## SG focus= hr
## Subgroup Consistency Minutes= 0.02393333
## Subgroup found (FS)
## Minutes forestsearch overall= 0.03475
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 48.96924 2
##      leaf.node control.mean control.size control.se depth
## 1      2        -4.64      524.00      1.16      1
## 2      3         5.87      176.00      2.51      1
## 3      4        -5.09      105.00      2.65      2
## 4      5         7.34      88.00      3.12      2
## 5      6        -6.17      372.00      1.35      2
## 6      7         5.83      135.00      2.87      2
##      leaf.node control.mean control.size control.se depth
## 4      5         7.34      88.00      3.12      2
## GRF subgroup found

```

```

## All splits
## [1] "size <= 20" "size <= 17" "z1 <= 0"
## Terminating node at max.diff (sg.harm.id)
## [1] "size <= 17"
## # of continuous/categorical characteristics 1 6
## Continuous characteristics: size
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## Default cuts included (1st 20)
## Categorical: z1 z2 z3 z4 z5 grade3
## Factors per GRF: size <= 20 size <= 17 z1 <= 0
## Initial GRF cuts included size <= 20 size <= 17 z1 <= 0
## # of candidate subgroup factors= 11
## [1] "size <= 20"    "size <= 17"    "size <= 29.2" "size <= 25"    "size <= 35"
## [6] "z1"           "z2"          "z3"          "z4"          "z5"
## [11] "grade3"
## Number of factors evaluated= 11
## Confounders per grf screening q6 q7 q2 q8 q1 q10 q9 q11 q4 q5 q3
##      Factors Labels VI(grf)
## 6      z1     q6  0.3551
## 7      z2     q7  0.1433
## 2  size <= 17   q2  0.0734
## 8      z3     q8  0.0626
## 1  size <= 20   q1  0.0611
## 10     z5     q10 0.0592
## 9      z4     q9  0.0575
## 11     grade3 q11 0.0522
## 4  size <= 25   q4  0.0477
## 5  size <= 35   q5  0.0456
## 3  size <= 29.2  q3  0.0423
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 253
## Approximately 5% of max_count met: minutes 0.00065
## Approximately 10% of max_count met: minutes 0.001166667
## Approximately 20% of max_count met: minutes 0.002333333
## Approximately 33% of max_count met: minutes 0.0038
## Approximately 50% of max_count met: minutes 0.0055
## Approximately 75% of max_count met: minutes 0.00775
## Approximately 90% of max_count met: minutes 0.009333333
## # of subgroups evaluated based on (up to) maxk-factor combinations 253
## % of all-possible combinations (<= maxk) 100
## k.max= 2
## Events criteria for control,exp= 10 10
## # of subgroups with events less than criteria: control, experimental 29 29
## # of subgroups with sample size less than criteria 40
## # of subgroups meeting all criteria = 201
## # of subgroups fitted (Cox model estimable) = 201
## *Subgroup Searching Minutes= 0.01003333
## Number of subgroups meeting HR threshold 19
## Subgroup candidate(s) found (FS)
## # of candidate subgroups (meeting HR criteria) = 19
## SGs (1st 10) meeting screening thresholds sorted by sg_focus= hr
##      n   E d1   HR L(HR) q6.0 q6.1 q7.0 q7.1 q2.0 q2.1 q8.0 q8.1 q1.0 q1.1
## 1  90  66 37 2.19  1.30   0   1   0   0   0   0   0   1   0   0
## 2 116  80 45 1.81  1.15   0   1   0   1   0   0   0   0   0   0
## 3  88  50 26 1.77  1.02   0   0   0   0   1   0   0   0   0   1
## 4 105  84 48 1.70  1.09   0   1   0   0   0   0   0   0   0   0

```

```

## 5 160 110 64 1.58 1.08 0 1 0 0 1 0 0 0 0 0 0 0
## 6 135 90 53 1.58 1.04 0 1 0 0 0 0 0 0 0 0 1 0
## 7 100 68 38 1.57 0.97 0 1 0 0 0 0 0 0 0 0 0 0 0
## 8 90 47 25 1.52 0.86 0 0 0 0 0 0 0 0 0 1 0 1
## 9 161 112 64 1.44 0.99 0 1 0 0 0 0 0 0 0 0 0 0 0
## 10 86 63 37 1.41 0.85 0 1 0 0 0 0 0 0 0 0 0 0 0 0
## Consistency 1
## # of splits= 400
## Model, % Consistency Met= {z1} {z3} 1
## SG focus= hr
## Subgroup Consistency Minutes= 0.02786667
## Subgroup found (FS)
## Minutes forestsearch overall= 0.0476
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 48.96924 2
##   leaf.node control.mean control.size control.se depth
## 1      2      -4.64     524.00      1.16    1
## 2      3       5.87     176.00      2.51    1
## 3      4      -5.09     105.00      2.65    2
## 4      5       7.34      88.00      3.12    2
## 5      6      -6.17     372.00      1.35    2
## 6      7       5.83     135.00      2.87    2
##   leaf.node control.mean control.size control.se depth
## 4      5       7.34      88.00      3.12    2
## GRF subgroup found
## All splits
## [1] "size <= 20" "size <= 17" "z1 <= 0"
## Terminating node at max.diff (sg.harm.id)
## [1] "size <= 17"
## # of continuous/categorical characteristics 1 6
## Continuous characteristics: size
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## CV lambda = 0.03111856
## 7 x 1 sparse Matrix of class "dgCMatrix"
##           s0
## z1 0.15509074
## z2 .
## z3 .
## z4 0.49617448
## z5 -0.80945129
## size .
## grade3 0.03871525
## Cox-LASSO selected: z1 z4 z5 grade3
## Cox-LASSO not selected: z2 z3 size
## Default cuts included from Lasso:
## Categorical after Lasso: z1 z4 z5 grade3
## Factors per GRF: size <= 20 size <= 17 z1 <= 0
## Initial GRF cuts included size <= 20 size <= 17 z1 <= 0
## Factors included per GRF (not in lasso) size <= 20 size <= 17
## # of candidate subgroup factors= 9
## [1] "size <= 20" "size <= 17" "size <= 29.2" "size <= 25" "size <= 35"
## [6] "z1"        "z4"        "z5"        "grade3"
## Number of factors evaluated= 9
## Confounders per grf screening q6 q2 q8 q1 q7 q4 q9 q3 q5
##          Factors Labels VI(grf)

```

```

## 6          z1      q6  0.4153
## 2  size <= 17    q2  0.0996
## 8          z5      q8  0.0849
## 1  size <= 20    q1  0.0744
## 7          z4      q7  0.0742
## 4  size <= 25    q4  0.0670
## 9          grade3  q9  0.0647
## 3 size <= 29.2   q3  0.0603
## 5  size <= 35    q5  0.0596
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 171
## Approximately 5% of max_count met: minutes 0.0004166667
## Approximately 10% of max_count met: minutes 0.0008666667
## Approximately 20% of max_count met: minutes 0.0016
## Approximately 33% of max_count met: minutes 0.002633333
## Approximately 50% of max_count met: minutes 0.003466667
## Approximately 75% of max_count met: minutes 0.004783333
## Approximately 90% of max_count met: minutes 0.00565
## # of subgroups evaluated based on (up to) maxk-factor combinations 171
## % of all-possible combinations (<= maxk) 100
## k.max= 2
## Events criteria for control,exp= 10 10
## # of subgroups with events less than criteria: control, experimental 25 25
## # of subgroups with sample size less than criteria 31
## # of subgroups meeting all criteria = 129
## # of subgroups fitted (Cox model estimable) = 129
## *Subgroup Searching Minutes= 0.00615
## Number of subgroups meeting HR threshold 13
## Subgroup candidate(s) found (FS)
## # of candidate subgroups (meeting HR criteria) = 13
## SGs (1st 10) meeting screening thresholds sorted by sg_focus= hr
##       n   E d1   HR L(HR) q6.0 q6.1 q2.0 q2.1 q8.0 q8.1 q1.0 q1.1 q7.0 q7.1
## 1  88  50 26 1.77  1.02  0    0    1    0    0    0    0    0    1    0    0
## 2 105  84 48 1.70  1.09  0    1    0    0    1    0    0    0    0    0    0
## 3 160 110 64 1.58  1.08  0    1    1    0    0    0    0    0    0    0    0
## 4 135  90 53 1.58  1.04  0    1    0    0    0    0    0    1    0    0    0
## 5 100  68 38 1.57  0.97  0    1    0    0    0    0    0    0    0    0    0
## 6 161 112 64 1.44  0.99  0    1    0    0    0    0    0    0    0    0    1
## 7  86  63 37 1.41  0.85  0    1    0    0    0    0    0    0    0    0    0
## 8  90  62 36 1.39  0.84  0    1    0    0    0    0    0    0    0    0    0
## 9  86  57 31 1.37  0.81  0    1    0    0    0    0    0    0    0    0    0
## 10 176 119 67 1.36  0.94  0    1    0    0    0    0    0    0    0    0    0
## Consistency 0.9425
## # of splits= 400
## Model, % Consistency Met= !{size <= 17} {size <= 20} 0.9425
## Consistency 0.965
## # of splits= 400
## Model, % Consistency Met= {z1} !{z5} 0.965
## SG focus= hr
## Subgroup Consistency Minutes= 0.05706667
## Subgroup found (FS)
## Minutes forestsearch overall= 0.07158333
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 48.52993 2
##     leaf.node control.mean control.size control.se depth
## 1          2        -3.06      697.00      1.06      1

```

```

## 11      4      -3.08      554.00      1.14      2
## 3       6      -8.40      93.00      3.13      2
##   leaf.node control.mean control.size control.se depth
## 1       2      -3.06     697.00      1.06      1
## GRF subgroup NOT found
## NO GRF cuts meeting delta(RMST): dmin.grf= 12
## # of continuous/categorical characteristics 1 6
## Continuous characteristics: size
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## Default cuts included (1st 20)
## Categorical: z1 z2 z3 z4 z5 grade3
## # of candidate subgroup factors= 10
## [1] "size <= 29.3" "size <= 25"    "size <= 20"    "size <= 35"    "z1"
## [6] "z2"           "z3"           "z4"           "z5"           "grade3"
## Number of factors evaluated= 10
## Confounders per grf screening q7 q5 q6 q8 q9 q1 q10 q4 q3 q2
##          Factors Labels VI(grf)
## 7       z3      q7  0.1321
## 5       z1      q5  0.1281
## 6       z2      q6  0.1188
## 8       z4      q8  0.1038
## 9       z5      q9  0.1009
## 1  size <= 29.3  q1  0.0962
## 10      grade3  q10 0.0906
## 4  size <= 35  q4  0.0805
## 3  size <= 20  q3  0.0798
## 2  size <= 25  q2  0.0692
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 210
## Approximately 5% of max_count met: minutes 0.0004666667
## Approximately 10% of max_count met: minutes 0.0008833333
## Approximately 20% of max_count met: minutes 0.0017666667
## Approximately 33% of max_count met: minutes 0.00285
## Approximately 50% of max_count met: minutes 0.0043
## Approximately 75% of max_count met: minutes 0.0065
## Approximately 90% of max_count met: minutes 0.007583333
## # of subgroups evaluated based on (up to) maxk-factor combinations 210
## % of all-possible combinations (<= maxk) 100
## k.max= 2
## Events criteria for control,exp= 10 10
## # of subgroups with events less than criteria: control, experimental 18 20
## # of subgroups with sample size less than criteria 24
## # of subgroups meeting all criteria = 179
## # of subgroups fitted (Cox model estimable) = 179
## *Subgroup Searching Minutes=* 0.00805
## Number of subgroups meeting HR threshold 4
## Subgroup candidate(s) found (FS)
## # of candidate subgroups (meeting HR criteria) = 4
## SGs (1st 10) meeting screening thresholds sorted by sg_focus= hr
##      n  E d1  HR L(HR) q7.0 q7.1 q5.0 q5.1
## 1 107 44 28 1.79  0.97    0    0    0    0
## 2 104 44 28 1.44  0.78    0    0    0    0
## 3 101 77 35 1.26  0.80    0    0    0    1
## 4 129 92 41 1.25  0.83    0    0    0    0
## Consistency 0.925
## # of splits= 400

```

```

## Model, % Consistency Met= !{z4} {size <= 20} 0.925
## Consistency 0.8075
## Consistency 0.73
## Consistency 0.7025
## SG focus= hr
## Subgroup Consistency Minutes= 0.1076167
## Subgroup found (FS)
## Minutes forestsearch overall= 0.1244
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 48.52993 2
##   leaf.node control.mean control.size control.se depth
## 1      2      -3.06     697.00      1.06      1
## 11     4      -3.08     554.00      1.14      2
## 3      6      -8.40      93.00      3.13      2
##   leaf.node control.mean control.size control.se depth
## 1      2      -3.06     697.00      1.06      1
## GRF subgroup NOT found
## NO GRF cuts meeting delta(RMST): dmin.grf= 12
## # of continuous/categorical characteristics 1 6
## Continuous characteristics: size
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## CV lambda = 0.01109449
## 7 x 1 sparse Matrix of class "dgCMatrix"
##   s0
## z1    0.221487409
## z2    .
## z3    0.161998816
## z4    0.359038687
## z5    -1.061975429
## size   0.004538426
## grade3 0.212614442
## Cox-LASSO selected: z1 z3 z4 z5 size grade3
## Cox-LASSO not selected: z2
## Default cuts included from Lasso:
## Categorical after Lasso: z1 z3 z4 z5 grade3
## # of candidate subgroup factors= 9
## [1] "size <= 29.3" "size <= 25"   "size <= 20"   "size <= 35"   "z1"
## [6] "z3"           "z4"         "z5"         "grade3"
## Number of factors evaluated= 9
## Confounders per grf screening q6 q5 q8 q1 q7 q3 q9 q4 q2
##   Factors Labels VI(grf)
## 6      z3    q6  0.1721
## 5      z1    q5  0.1481
## 8      z5    q8  0.1138
## 1 size <= 29.3  q1  0.1063
## 7      z4    q7  0.1047
## 3 size <= 20   q3  0.0967
## 9      grade3 q9  0.0962
## 4 size <= 35   q4  0.0885
## 2 size <= 25   q2  0.0735
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 171
## Approximately 5% of max_count met: minutes 0.00045
## Approximately 10% of max_count met: minutes 0.0008666667
## Approximately 20% of max_count met: minutes 0.001716667
## Approximately 33% of max_count met: minutes 0.002566667

```

```

## Approximately 50% of max_count met: minutes 0.003883333
## Approximately 75% of max_count met: minutes 0.0052
## Approximately 90% of max_count met: minutes 0.006266667
## # of subgroups evaluated based on (up to) maxk-factor combinations 171
## % of all-possible combinations (<= maxk) 100
## k.max= 2
## Events criteria for control,exp= 10 10
## # of subgroups with events less than criteria: control, experimental 16 18
## # of subgroups with sample size less than criteria 22
## # of subgroups meeting all criteria = 143
## # of subgroups fitted (Cox model estimable) = 143
## *Subgroup Searching Minutes=* 0.00665
## Number of subgroups meeting HR threshold 2
## Subgroup candidate(s) found (FS)
## # of candidate subgroups (meeting HR criteria) = 2
## SGs (1st 10) meeting screening thresholds sorted by sg_focus= hr
##      n   E d1    HR L(HR) q6.0 q6.1
## 1 107 44 28 1.79  0.97    0    0
## 2 101 77 35 1.26  0.80    0    0
## Consistency 0.925
## # of splits= 400
## Model, % Consistency Met= !{z4} {size <= 20} 0.925
## Consistency 0.73
## SG focus= hr
## Subgroup Consistency Minutes= 0.0531
## Subgroup found (FS)
## Minutes forestsearch overall= 0.06756667

t.now <- proc.time()[3]
t.min <- (t.now - t.start)/60

print(table(res$analysis))

##
##      FS1     FS1g      GRF    GRF.60    VT(24)    VT(36) VT#(24) VT#(36)
##      5000     5000     5000     5000     5000     5000     5000     5000
check <- c(c(table(res$analysis)) - Nsims)
if (all(check != 0)) stop("All analyses not complete")

dgm_alt <- dgm

outres <- out.results(res = res, dgm = dgm, output.file = output.file, t.min = t.min,
                      out_analysis = "FS1")

## [1] "results/oc_sims=5000_m4a-Noise=0_N=700_alt_ktreat=0.9_hrH=2_v0-4cutsTestVT.Rdata"
##      sim sizeH_true propH_true sizeHc_true propHc_true any.H size.H size.Hc
##      <int>      <num>      <num>      <int>      <num> <num> <num> <int>
## 1:      1       95  0.1357143      605  0.8642857     1     95    605
## 2:      1       95  0.1357143      605  0.8642857     1     95    605
## 3:      1       95  0.1357143      605  0.8642857     1     95    605
## 4:      1       95  0.1357143      605  0.8642857     0     0    700
## 5:      1       95  0.1357143      605  0.8642857     0     0    700
## 6:      1       95  0.1357143      605  0.8642857     1    100    600
##      ppv      npv specificity sensitivity found.1 found.2 found.both found.al3

```

```

##      <num>      <num>      <num>      <num>      <num>      <num>      <num>      <num>
## 1: 1.0 1.0000000 1.0000000 1.00 0 0 0 0
## 2: 1.0 1.0000000 1.0000000 1.00 0 0 0 0
## 3: 1.0 1.0000000 1.0000000 1.00 NA NA NA NA
## 4: 0.0 1.0000000 0.8642857 0.00 NA NA NA 0
## 5: 0.0 1.0000000 0.8642857 0.00 0 0 0 0
## 6: 0.6 0.9289256 0.9366667 0.57 0 0 0 0
##      hr.H.true hr.Hc.true hr.H.hat hr.Hc.hat      b1.H      b2.H      b1.Hc
##      <num>      <num>      <num>      <num>      <num>      <num>      <num>
## 1: 1.74466 0.5577464 1.744660 0.5577464 0.0000000 -0.2553474 0.00000000
## 2: 1.74466 0.5577464 1.744660 0.5577464 0.0000000 -0.2553474 0.00000000
## 3: 1.74466 0.5577464 1.744660 0.5577464 0.0000000 -0.2553474 0.00000000
## 4: 1.74466 0.5577464 NA 0.6688550 NA NA 0.11110856
## 5: 1.74466 0.5577464 NA 0.6688550 NA NA 0.11110856
## 6: 1.74466 0.5577464 1.472201 0.5860591 -0.2724584 -0.5278057 0.02831265
##      b2.Hc p.cens analysis taumax hr.itt l.itt u.itt
##      <num>      <num>      <char>      <num>      <num>      <num>      <num>
## 1: -0.08889407 0.4514286 FS1 81.43453 0.668855 0.5464683 0.8186513
## 2: -0.08889407 0.4514286 FS1g 48.86072 0.668855 0.5464683 0.8186513
## 3: -0.08889407 0.4514286 GRF 81.43453 0.668855 0.5464683 0.8186513
## 4: 0.02221450 0.4514286 GRF.60 48.86072 0.668855 0.5464683 0.8186513
## 5: 0.02221450 0.4514286 VT(24) 81.43453 0.668855 0.5464683 0.8186513
## 6: -0.06058142 0.4514286 VT#(24) 81.43453 0.668855 0.5464683 0.8186513
##      hr.adj.itt l.adj.itt u.adj.itt l.H.true u.H.true l.Hc.true u.Hc.true
##      <num>      <num>      <num>      <num>      <num>      <num>      <num>
## 1: 0.6827343 0.5565269 0.8375626 1.105971 2.752186 0.4437854 0.7009717
## 2: 0.6827343 0.5565269 0.8375626 1.105971 2.752186 0.4437854 0.7009717
## 3: 0.6827343 0.5565269 0.8375626 1.105971 2.752186 0.4437854 0.7009717
## 4: 0.6827343 0.5565269 0.8375626 1.105971 2.752186 0.4437854 0.7009717
## 5: 0.6827343 0.5565269 0.8375626 1.105971 2.752186 0.4437854 0.7009717
## 6: 0.6827343 0.5565269 0.8375626 1.105971 2.752186 0.4437854 0.7009717
##      l.H.hat u.H.hat l.Hc.hat u.Hc.hat
##      <num>      <num>      <num>      <num>
## 1: 1.1059709 2.752186 0.4437854 0.7009717
## 2: 1.1059709 2.752186 0.4437854 0.7009717
## 3: 1.1059709 2.752186 0.4437854 0.7009717
## 4: NA NA 0.5464683 0.8186513
## 5: NA NA 0.5464683 0.8186513
## 6: 0.9431056 2.298128 0.4664573 0.7363272
## Subgroup HRs: H, H^c, Causal= 2.000007 0.6466405 0.7057463
## Simulations= 5000
## Avg censoring= 0.4500883
## Min,Max,Avg tau.max= 72.23037 83.98787 81.59443
## P(H) approximation at causal(H), n(sg)=60, approx= 2.000007 60 0.8282653
## P(H) approximation at causal(H), Avg(n(sg)), approx= 2.000007 89 0.8998486
## P(H) approximation at plim(H), Avg(n(sg)), approx= 2.115229 89 0.9264997
## Minutes,hours 28.67002 0.4778336
##      FS1   FS1g   GRF   GRF.60   VT(24)   VT#(24)   VT(36)   VT#(36)
##      any.H 0.900 0.860 0.940 0.720 0.680 0.730 0.660 0.770
##      sensH 0.840 0.800 0.850 0.650 0.620 0.680 0.570 0.730
##      sensHc 0.990 0.990 0.970 0.980 0.980 0.990 0.980 0.990
##      ppH    0.840 0.780 0.780 0.610 0.590 0.660 0.550 0.710
##      ppHc   0.980 0.970 0.980 0.960 0.950 0.960 0.940 0.970
##      Avg(#H) 89.000 91.000 102.000 100.000 94.000 93.000 94.000 92.000
##      minH   61.000 61.000 60.000 60.000 61.000 60.000 60.000 60.000

```

```

## maxH          221.000 221.000 298.000 283.000 211.000 199.000 253.000 219.000
## Avg(#Hc)     620.000 622.000 605.000 629.000 636.000 632.000 639.000 628.000
## minHc        479.000 479.000 402.000 417.000 489.000 501.000 447.000 481.000
## maxHc        700.000 700.000 700.000 700.000 700.000 700.000 700.000 700.000
## hat(H*)       2.199   2.219   2.157   2.260   2.278   2.269   2.294   2.258
## hat(hat[H])  2.234   2.211   2.023   2.069   2.133   2.171   2.129   2.189
## hat(Hc*)      0.653   0.652   0.651   0.647   0.653   0.651   0.653   0.652
## hat(hat[Hc]) 0.652   0.652   0.643   0.645   0.657   0.653   0.658   0.652
## hat(H*)all    2.115   2.115   2.115   2.115   2.115   2.115   2.115   2.115
## hat(Hc*)all   0.651   0.651   0.651   0.651   0.651   0.651   0.651   0.651
## hat(ITT)all   0.754   0.754   0.754   0.754   0.754   0.754   0.754   0.754
## hat(ITTadj)all 0.742   0.742   0.742   0.742   0.742   0.742   0.742   0.742

missC <- tab_tests(res = res)

pA <- as.character(round(outres$pAnyH.approx2, 4))

tabsim_missC <- get_tabsim(missC = missC, pA = pA, est_names = est_names, stat_names = stat_names,
                            mod.harm = mod.harm, Nsims = Nsims)

```

Table 2: Average classification rates:  $avg|\hat{H}|$ ,  $min|\hat{H}|$ , and  $max|\hat{H}|$ , denote the average, minimum, and maximum of the number of subjects in the estimated subgroup  $\hat{H}$  (analogously for  $\hat{H}^c$ ). Note that under the null  $sens(\hat{H})$  is undefined and  $ppv(\hat{H}) = 0$ .

	$FS_g$	$FS_{lg}$	$GRF$	$GRF_{60}$	$VT(24)$	$VT^{\#}(24)$	$VT(36)$	$VT^{\#}(36)$
Finding H								
any(H)	0.9	0.86	0.94	0.72	0.68	0.73	0.66	0.77
$sens(\hat{H})$	0.84	0.8	0.85	0.65	0.62	0.68	0.57	0.73
$sens(\hat{H}^c)$	0.99	0.99	0.97	0.98	0.98	0.99	0.98	0.99
$ppv(\hat{H})$	0.84	0.78	0.78	0.61	0.59	0.66	0.55	0.71
$ppv(\hat{H}^c)$	0.98	0.97	0.98	0.96	0.95	0.96	0.94	0.97
Size of H and H-complement								
$avg \hat{H} $	89	91	102	100	94	93	94	92
$min \hat{H} $	61	61	60	60	61	60	60	60
$max \hat{H} $	221	221	298	283	211	199	253	219
$avg \hat{H}^c $	620	622	605	629	636	632	639	628
$min \hat{H}^c $	479	479	402	417	489	501	447	481
$max \hat{H}^c $	700	700	700	700	700	700	700	700

Note: Number of simulations= 5000 .

Note: Probability approximation= 0.8998 .

```

t.done <- proc.time()[3]
t.min <- (t.done - t.start.all)/60
cat("Minutes and hours to finish", c(t.min, t.min/60), "\n")

## Minutes and hours to finish 53.82568 0.8970947

cat("Minutes and hours per 10,000 to finish", (10000/Nsims) * c(t.min, t.min/60),
    "\n")

## Minutes and hours per 10,000 to finish 107.6514 1.794189

# cat('Machine=',c(Sys.info()[[4]]),'\n') cat('Number of
# cores=',c(detectCores(logical = FALSE)), '\n')

```

```
require(benchmarkme)
my_system <- get_cpu()
my_ram <- get_ram()
cat("Running on system:", c(my_system$model_name), "\n")

## Running on system: AMD Ryzen Threadripper PRO 5995WX 64-Cores

cat("with number of cores and cpu/GB=", c(my_system$no_of_cores, round(c(my_ram)/10^9,
0)), "\n")

## with number of cores and cpu/GB= 128 270
```