

```

N <- 700
Nsims <- 1000

maxFollow <- 84
cens.type <- "weibull"
##### Forest search criteria
hr.threshold <- 1.25 # Initial 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
ml.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 <- 0.225 # 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("$FS_{g}$", "$FS_{lg}$", "$GRF$", "$GRF_{60}$", "$VT(24)$", "${VT}~{\\#}(24)$",
"$VT(36)$", "${VT}~{\\#}(36)$")

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 <- "v0B-4cuts"

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]

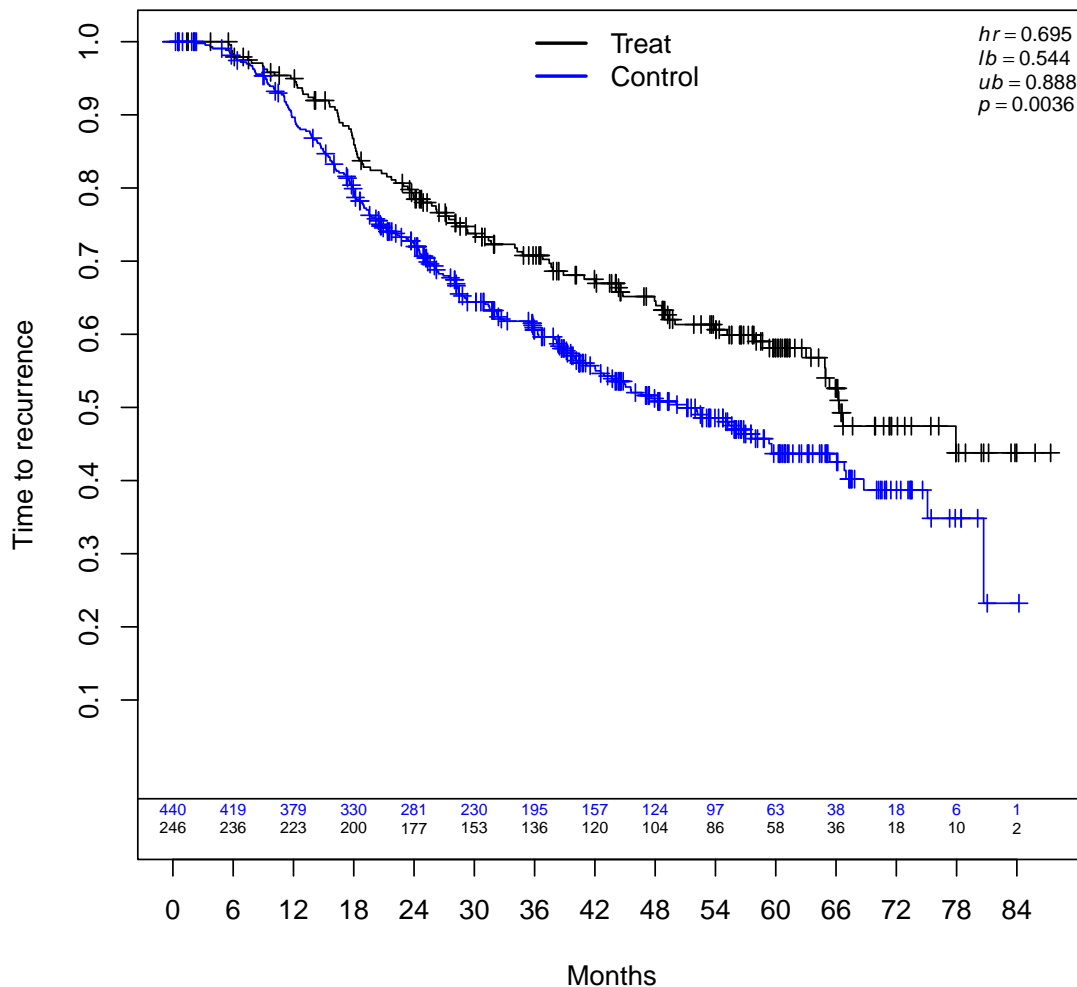
# Classification table names Note: within tab_tests (summary.VTFS) we rename so
# that denominator in ppv(hatH) is # hatH Manuscript section 3.2 will be
# updated accordingly

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

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

mod.harm <- "null"
this.dgm <- get.dgm4.OG(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 5e-04
## Approximately 10% of max_count met: minutes 0.000866667
## Approximately 20% of max_count met: minutes 0.001483333
## Approximately 33% of max_count met: minutes 0.002383333
## Approximately 50% of max_count met: minutes 0.003466667
## Approximately 75% of max_count met: minutes 0.005233333
## Approximately 90% of max_count met: minutes 0.006266667
## # 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=* 0.00665
## 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 0.0002166667
## Approximately 10% of max_count met: minutes 0.0004333333
## Approximately 20% of max_count met: minutes 8e-04
## Approximately 33% of max_count met: minutes 0.001366667
## Approximately 50% of max_count met: minutes 0.001733333
## Approximately 75% of max_count met: minutes 0.00265
## Approximately 90% of max_count met: minutes 0.003166667
## # 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=* 0.003266667
## 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= 50.12699 2
##   leaf.node control.mean control.size control.se depth
## 1         2        -2.39        657.00        1.05    1
## 2         4        -4.57        274.00        1.58    2
## 3         5         6.69        111.00        2.53    2
## 4         6        -3.91        272.00        1.64    2
##   leaf.node control.mean control.size control.se depth
## 3         5         6.69        111.00        2.53    2
## GRF subgroup found
## All splits
## [1] "size <= 27" "size <= 23" "size <= 56"
## Terminating node at max.diff (sg.harm.id)
## [1] "size <= 23"
## # 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 <= 27 size <= 23 size <= 56
## Initial GRF cuts included size <= 27 size <= 23 size <= 56
## # of candidate subgroup factors= 13
## [1] "size <= 27" "size <= 23" "size <= 56" "size <= 29.9" "size <= 25"
## [6] "size <= 21" "size <= 35" "z1" "z2" "z3"
## [11] "z4" "z5" "grade3"
## Number of factors evaluated= 13
## Confounders per grf screening q6 q2 q11 q1 q12 q10 q7 q8 q9 q13 q3 q4 q5
##   Factors Labels VI(grf)
## 6   size <= 21    q6 0.1885
## 2   size <= 23    q2 0.1511
## 11  z4          q11 0.0857
## 1   size <= 27    q1 0.0847
## 12  z5          q12 0.0835
## 10  z3          q10 0.0712
## 7   size <= 35    q7 0.0641
## 8   z1          q8 0.0624
## 9   z2          q9 0.0616
## 13  grade3      q13 0.0604
## 3   size <= 56    q3 0.0389
## 4   size <= 29.9  q4 0.0305
## 5   size <= 25    q5 0.0174
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 351
## Approximately 5% of max_count met: minutes 0.0008166667
## Approximately 10% of max_count met: minutes 0.0017
## Approximately 20% of max_count met: minutes 0.002866667
## Approximately 33% of max_count met: minutes 0.00445
## Approximately 50% of max_count met: minutes 0.006833333
## Approximately 75% of max_count met: minutes 0.009933333
## Approximately 90% of max_count met: minutes 0.01198333
## # of subgroups evaluated based on (up to) maxk-factor combinations 351
## % 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 44 51

```

```

## # of subgroups with sample size less than criteria 66
## # of subgroups meeting all criteria = 264
## # of subgroups fitted (Cox model estimable) = 264
## *Subgroup Searching Minutes=* 0.013
## 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 q2.0 q2.1 q11.0 q11.1
## 1 111  62 41 1.77  1.04    0    0    1    0    0    0
## 2  86  46 29 1.71  0.94    0    0    1    0    0    0
## 3  71  46 25 1.37  0.76    0    0    0    0    0    0
## 4 131  73 42 1.34  0.84    0    0    1    0    0    0
## 5  64  40 22 1.34  0.72    0    0    0    0    0    0
## 6 182 100 61 1.32  0.89    1    0    0    0    0    0
## Consistency 0.9775
## # of splits= 400
## Model, % Consistency Met= !{size <= 23} {size <= 27} 0.9775
## SG focus= hr
## Subgroup Consistency Minutes= 0.0279
## Subgroup found (FS)
## Minutes forestsearch overall= 0.04753333
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 50.12699 2
##   leaf.node control.mean control.size control.se depth
## 1         2         -2.39         657.00         1.05    1
## 2         4         -4.57         274.00         1.58    2
## 3         5          6.69         111.00         2.53    2
## 4         6         -3.91         272.00         1.64    2
##   leaf.node control.mean control.size control.se depth
## 3         5          6.69         111.00         2.53    2
## GRF subgroup found
## All splits
## [1] "size <= 27" "size <= 23" "size <= 56"
## Terminating node at max.diff (sg.harm.id)
## [1] "size <= 23"
## # of continuous/categorical characteristics 1 6
## Continuous characteristics: size
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## CV lambda = 0.01622571
## 7 x 1 sparse Matrix of class "dgCMatrix"
##           s0
## z1      0.106312977
## z2     -0.179096033
## z3      .
## z4      0.675389637
## z5     -0.755501652
## size    0.003042607
## grade3  .
## Cox-LASSO selected: z1 z2 z4 z5 size
## Cox-LASSO not selected: z3 grade3
## Default cuts included from Lasso:
## Categorical after Lasso: z1 z2 z4 z5
## Factors per GRF: size <= 27 size <= 23 size <= 56
## Initial GRF cuts included size <= 27 size <= 23 size <= 56

```



```

## Factors included per GRF (not in lasso) size <= 27 size <= 23 size <= 56
## # of candidate subgroup factors= 11
## [1] "size <= 27" "size <= 23" "size <= 56" "size <= 29.9" "size <= 25"
## [6] "size <= 21" "size <= 35" "z1" "z2" "z4"
## [11] "z5"
## Number of factors evaluated= 11
## Confounders per grf screening q6 q2 q10 q11 q9 q1 q8 q7 q3 q4 q5
## Factors Labels VI(grf)
## 6 size <= 21 q6 0.2002
## 2 size <= 23 q2 0.1533
## 10 z4 q10 0.1080
## 11 z5 q11 0.1026
## 9 z2 q9 0.0964
## 1 size <= 27 q1 0.0855
## 8 z1 q8 0.0745
## 7 size <= 35 q7 0.0739
## 3 size <= 56 q3 0.0448
## 4 size <= 29.9 q4 0.0385
## 5 size <= 25 q5 0.0222
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 253
## Approximately 5% of max_count met: minutes 0.0005166667
## Approximately 10% of max_count met: minutes 0.0009833333
## Approximately 20% of max_count met: minutes 0.0019
## Approximately 33% of max_count met: minutes 0.002833333
## Approximately 50% of max_count met: minutes 0.004233333
## Approximately 75% of max_count met: minutes 0.006466667
## Approximately 90% of max_count met: minutes 0.01001667
## # 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 38 42
## # of subgroups with sample size less than criteria 55
## # of subgroups meeting all criteria = 178
## # of subgroups fitted (Cox model estimable) = 178
## *Subgroup Searching Minutes=* 0.01051667
## 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) q6.0 q6.1 q2.0 q2.1
## 1 111 62 41 1.77 1.04 0 0 1 0
## 2 86 46 29 1.71 0.94 0 0 1 0
## 3 131 73 42 1.34 0.84 0 0 1 0
## 4 182 100 61 1.32 0.89 1 0 0 0
## Consistency 0.9775
## # of splits= 400
## Model, % Consistency Met= ![size <= 23] {size <= 27} 0.9775
## SG focus= hr
## Subgroup Consistency Minutes= 0.02496667
## Subgroup found (FS)
## Minutes forestsearch overall= 0.04668333
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 48.57937 2
## leaf.node control.mean control.size control.se depth

```

```

## 1      2      -4.51      695.00      1.09      1
## 2      4      6.98      84.00      2.84      2
## 3      5     -3.16     242.00      1.90      2
## 4      6     -8.02     369.00      1.46      2
## leaf.node control.mean control.size control.se depth
## 2      4      6.98      84.00      2.84      2
## GRF subgroup found
## All splits
## [1] "z2 <= 0"      "size <= 20" "size <= 70"
## Terminating node at max.diff (sg.harm.id)
## [1] "size <= 20"
## # 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: z2 <= 0 size <= 20 size <= 70
## Initial GRF cuts included z2 <= 0 size <= 20 size <= 70
## # of candidate subgroup factors= 11
## [1] "size <= 20"      "size <= 70"      "size <= 28.8" "size <= 25"      "size <= 35"
## [6] "z1"              "z2"              "z3"              "z4"              "z5"
## [11] "grade3"
## Number of factors evaluated= 11
## Confounders per grf screening q7 q5 q1 q10 q9 q11 q6 q8 q4 q3 q2
## Factors Labels VI(grf)
## 7      z2      q7  0.3227
## 5      size <= 35      q5  0.1402
## 1      size <= 20      q1  0.1159
## 10     z5      q10 0.0793
## 9      z4      q9  0.0723
## 11     grade3      q11 0.0612
## 6      z1      q6  0.0586
## 8      z3      q8  0.0526
## 4      size <= 25      q4  0.0507
## 3      size <= 28.8      q3  0.0465
## 2      size <= 70      q2  0.0000
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 253
## Approximately 5% of max_count met: minutes 0.0006333333
## Approximately 10% of max_count met: minutes 0.001016667
## Approximately 20% of max_count met: minutes 0.001866667
## Approximately 33% of max_count met: minutes 0.00275
## Approximately 50% of max_count met: minutes 0.004416667
## Approximately 75% of max_count met: minutes 0.006333333
## Approximately 90% of max_count met: minutes 0.007533333
## # 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 43 43
## # of subgroups with sample size less than criteria 48
## # of subgroups meeting all criteria = 177
## # of subgroups fitted (Cox model estimable) = 177
## *Subgroup Searching Minutes=* 0.007883333
## 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) q7.0 q7.1
## 1  84 40 25 1.91  1.00    1    0
## 2 107 53 29 1.30  0.76    0    0
## Consistency 0.9675
## # of splits= 400
## Model, % Consistency Met= ![z2] {size <= 20} 0.9675
## SG focus= hr
## Subgroup Consistency Minutes= 0.02601667
## Subgroup found (FS)
## Minutes forestsearch overall= 0.04066667
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 48.57937 2
##      leaf.node control.mean control.size control.se depth
## 1          2          -4.51          695.00          1.09    1
## 2          4           6.98           84.00          2.84    2
## 3          5          -3.16          242.00          1.90    2
## 4          6          -8.02          369.00          1.46    2
##      leaf.node control.mean control.size control.se depth
## 2          4           6.98           84.00          2.84    2
## GRF subgroup found
## All splits
## [1] "z2 <= 0"      "size <= 20" "size <= 70"
## Terminating node at max.diff (sg.harm.id)
## [1] "size <= 20"
## # of continuous/categorical characteristics 1 6
## Continuous characteristics: size
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## CV lambda = 0.004423165
## 7 x 1 sparse Matrix of class "dgCMatrix"
##              s0
## z1      0.248645916
## z2     -0.478590998
## z3      0.428201760
## z4      0.701900511
## z5     -0.798190193
## size   -0.001898469
## grade3  0.027491341
## Cox-LASSO selected: z1 z2 z3 z4 z5 size grade3
## Cox-LASSO not selected:
## Default cuts included from Lasso:
## Categorical after Lasso: z1 z2 z3 z4 z5 grade3
## Factors per GRF: z2 <= 0 size <= 20 size <= 70
## Initial GRF cuts included z2 <= 0 size <= 20 size <= 70
## Factors included per GRF (not in lasso) size <= 20 size <= 70
## # of candidate subgroup factors= 11
## [1] "size <= 20"      "size <= 70"      "size <= 28.8"    "size <= 25"      "size <= 35"
## [6] "z1"              "z2"              "z3"              "z4"              "z5"
## [11] "grade3"
## Number of factors evaluated= 11
## Confounders per grf screening q7 q5 q1 q10 q9 q11 q6 q8 q4 q3 q2
##      Factors Labels VI(grf)
## 7          z2      q7  0.3227
## 5    size <= 35      q5  0.1402

```

```

## 1      size <= 20      q1  0.1159
## 10      z5      q10  0.0793
## 9      z4      q9  0.0723
## 11      grade3      q11  0.0612
## 6      z1      q6  0.0586
## 8      z3      q8  0.0526
## 4      size <= 25      q4  0.0507
## 3      size <= 28.8      q3  0.0465
## 2      size <= 70      q2  0.0000
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 253
## Approximately 5% of max_count met: minutes 0.0005333333
## Approximately 10% of max_count met: minutes 0.001
## Approximately 20% of max_count met: minutes 0.002
## Approximately 33% of max_count met: minutes 0.002783333
## Approximately 50% of max_count met: minutes 0.00425
## Approximately 75% of max_count met: minutes 0.0061
## Approximately 90% of max_count met: minutes 0.007216667
## # 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 43 43
## # of subgroups with sample size less than criteria 48
## # of subgroups meeting all criteria = 177
## # of subgroups fitted (Cox model estimable) = 177
## *Subgroup Searching Minutes=* 0.007666667
## 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) q7.0 q7.1
## 1  84 40 25 1.91  1.00    1    0
## 2 107 53 29 1.30  0.76    0    0
## Consistency 0.9675
## # of splits= 400
## Model, % Consistency Met= ![z2] {size <= 20} 0.9675
## SG focus= hr
## Subgroup Consistency Minutes= 0.02535
## Subgroup found (FS)
## Minutes forestsearch overall= 0.04348333

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)
##      1000      1000      1000      1000      1000      1000      1000      1000

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

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

```

```

## [1] "results/oc_sims=1000_m4a-Noise=0_N=700_null_ktreat=0.9_v0B-4cuts.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      FS1g 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= 1000
## Avg censoring= 0.4613314
## Min,Max,Avg tau.max= 72.81026 83.95272 81.68616
## P(H) approximation at causal(Hrc), n=60, approx= 0.701027 0.03392115
## P(H) approximation at plim(Hrc), n=60, approx= 0.7071011 0.03573494

```

```

## Minutes, hours 5.926583 0.09877639
##          FS1      FS1g      GRF      GRF.60      VT(24)      VT#(24)      VT(36)      VT#(36)
## any.H          0.070      0.050      0.270      0.060      0.030      0.010      0.030      0.010
## sensH          NaN       NaN       NaN       NaN       NaN       NaN       NaN       NaN
## sensHc         0.990      0.990      0.970      0.990      1.000      1.000      1.000      1.000
## 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)        98.000     104.000     88.000     75.000     79.000     79.000     77.000     71.000
## minH           61.000     63.000     60.000     60.000     60.000     63.000     60.000     61.000
## maxH           218.000    218.000    269.000    116.000    113.000    116.000    113.000    90.000
## Avg(#Hc)       693.000    695.000    676.000    696.000    697.000    699.000    698.000    699.000
## minHc          482.000    482.000    431.000    584.000    587.000    584.000    587.000    610.000
## 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])     1.772     1.764     1.528     1.517     1.294     1.385     1.380     1.562
## hat(Hc*)        0.779     0.791     0.736     0.732     0.731     0.710     0.720     0.734
## hat(hat[Hc])    0.687     0.696     0.665     0.673     0.684     0.653     0.668     0.679
## hat(H*)all      NaN       NaN       NaN       NaN       NaN       NaN       NaN       NaN
## hat(Hc*)all     0.707     0.707     0.707     0.707     0.707     0.707     0.707     0.707
## hat(ITT)all     0.707     0.707     0.707     0.707     0.707     0.707     0.707     0.707
## hat(ITTadj)all  0.666     0.666     0.666     0.666     0.666     0.666     0.666     0.666

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	0.05	0.27	0.06	0.03	0.01	0.03	0.01
$sens(\hat{H})$	.	.	.	.	.	.	.	.
$sens(\hat{H}^c)$	0.99	0.99	0.97	0.99	1	1	1	1
$ppv(\hat{H})$	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} $	98	104	88	75	79	79	77	71
$min \hat{H} $	61	63	60	60	60	63	60	61
$max \hat{H} $	218	218	269	116	113	116	113	90
$avg \hat{H}^c $	693	695	676	696	697	699	698	699
$min \hat{H}^c $	482	482	431	584	587	584	587	610
$max \hat{H}^c $	700	700	700	700	700	700	700	700

Note: Number of simulations= 1000 .

Note: Probability approximation= 0.0357 .

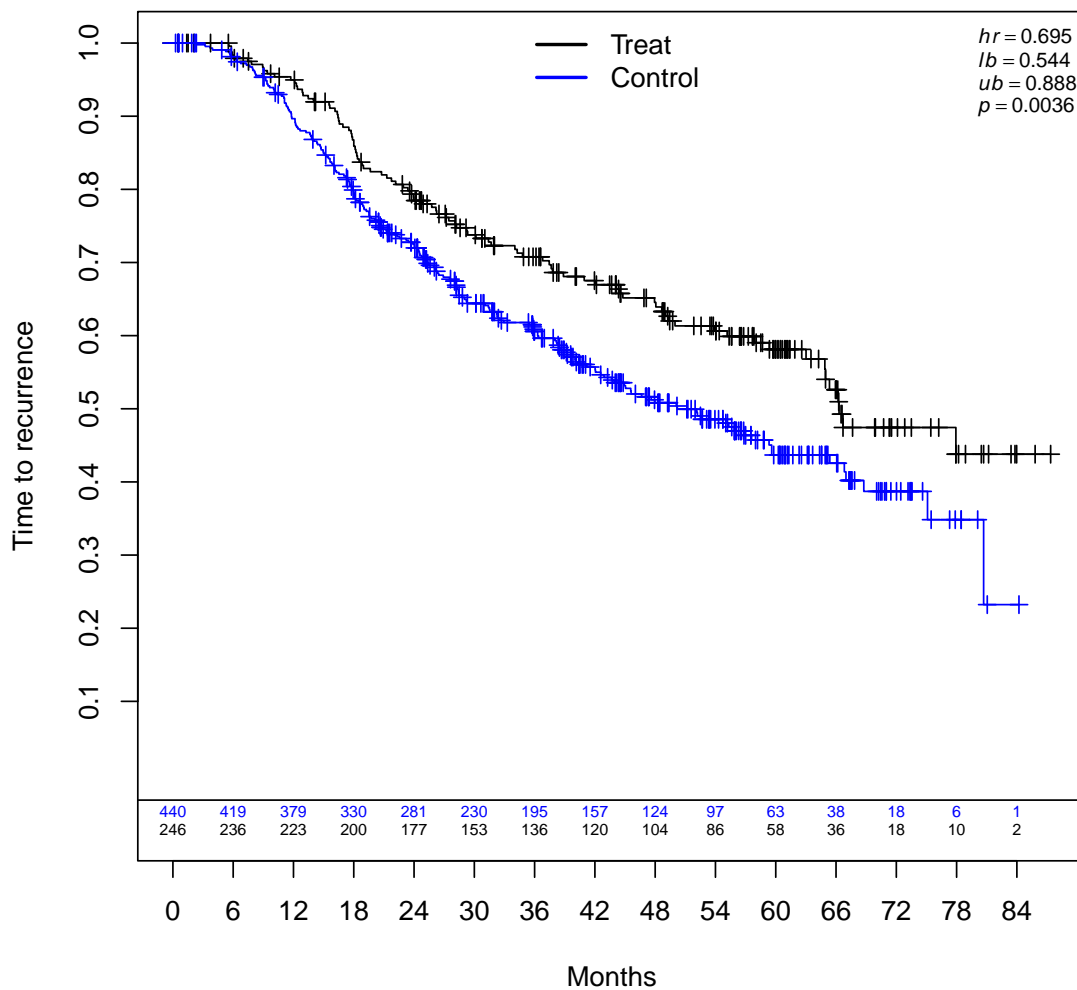
```

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) & !grepl(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 0.001566667
## Approximately 33% of max_count met: minutes 0.002533333
## Approximately 50% of max_count met: minutes 0.00375
## Approximately 75% of max_count met: minutes 0.00545
## Approximately 90% of max_count met: minutes 0.006266667
## # 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= 0.02321667
## Subgroup found (FS)
## Minutes forestsearch overall= 0.03791667
## 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 0.0002833333
## Approximately 10% of max_count met: minutes 0.0006833333
## Approximately 20% of max_count met: minutes 0.0012
## Approximately 33% of max_count met: minutes 0.001933333
## Approximately 50% of max_count met: minutes 0.002716667
## Approximately 75% of max_count met: minutes 0.0039
## Approximately 90% of max_count met: minutes 0.004666667
## # 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=* 0.004883333
## 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= 0.02381667
## Subgroup found (FS)
## Minutes forestsearch overall= 0.03536667
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 49.01446 2
## leaf.node control.mean control.size control.se depth
## 1 2 -2.45 515.00 1.11 1
## 2 3 3.28 185.00 2.23 1
## 3 4 -5.23 203.00 1.80 2
## 4 5 3.48 301.00 1.50 2
## 5 6 -5.19 153.00 2.21 2
## leaf.node control.mean control.size control.se depth
## 4 5 3.48 301.00 1.50 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 <= 29.9" "size <= 25" "size <= 21" "size <= 35" "z1"
## [6] "z2" "z3" "z4" "z5" "grade3"
## Number of factors evaluated= 10
## Confounders per grf screening q3 q5 q7 q8 q6 q9 q4 q10 q1 q2
## Factors Labels VI(grf)
## 3 size <= 21 q3 0.2179
## 5 z1 q5 0.1985
## 7 z3 q7 0.1387
## 8 z4 q8 0.0849
## 6 z2 q6 0.0783
## 9 z5 q9 0.0762
## 4 size <= 35 q4 0.0629
## 10 grade3 q10 0.0571
## 1 size <= 29.9 q1 0.0500
## 2 size <= 25 q2 0.0355
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 210
## Approximately 5% of max_count met: minutes 0.0008833333
## Approximately 10% of max_count met: minutes 0.001483333
## Approximately 20% of max_count met: minutes 0.002683333
## Approximately 33% of max_count met: minutes 0.003583333
## Approximately 50% of max_count met: minutes 0.008166667
## Approximately 75% of max_count met: minutes 0.0104
## Approximately 90% of max_count met: minutes 0.01178333
## # 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 19 19
## # 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=* 0.01265
## Number of subgroups meeting HR threshold 21
## Subgroup candidate(s) found (FS)
## # of candidate subgroups (meeting HR criteria) = 21
## SGs (1st 10) meeting screening thresholds sorted by sg_focus= hr
## n E d1 HR L(HR) q3.0 q3.1 q5.0 q5.1 q7.0 q7.1 q8.0 q8.1 q6.0 q6.1
## 1 88 63 35 3.59 2.13 0 0 0 1 0 1 0 0 0 0
## 2 117 78 42 2.03 1.29 0 0 0 1 0 0 0 0 0 1
## 3 71 46 26 1.77 0.98 0 0 0 0 0 1 0 0 0 0
## 4 131 82 40 1.67 1.08 0 0 0 1 0 0 0 0 0 0
## 5 64 39 22 1.58 0.84 0 0 0 0 0 0 0 0 0 0
## 6 85 48 25 1.58 0.90 0 0 0 1 0 0 0 0 0 0
## 7 133 99 50 1.58 1.06 1 0 0 1 0 0 0 0 0 0
## 8 97 58 28 1.55 0.92 0 0 0 1 0 0 0 0 0 0
## 9 69 43 24 1.52 0.83 0 0 0 0 0 1 0 0 0 0
## 10 95 64 34 1.51 0.92 0 0 0 0 0 0 0 0 0 1
## Consistency 1
## # of splits= 400
## Model, % Consistency Met= {z1} {z3} 1
## SG focus= hr
## Subgroup Consistency Minutes= 0.02983333
## Subgroup found (FS)
## Minutes forestsearch overall= 0.05011667
## FS:GRF stage for cut selection with dmin,tau= 12 0.6

```

```

## tau, maxdepth= 49.01446 2
##   leaf.node control.mean control.size control.se depth
## 1         2        -2.45        515.00        1.11    1
## 2         3         3.28        185.00        2.23    1
## 3         4        -5.23        203.00        1.80    2
## 4         5         3.48        301.00        1.50    2
## 5         6        -5.19        153.00        2.21    2
##   leaf.node control.mean control.size control.se depth
## 4         5         3.48        301.00        1.50    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 = 0.01795345
## 7 x 1 sparse Matrix of class "dgCMatrix"
##               s0
## z1      0.285277203
## z2     -0.106014514
## z3      .
## z4      0.606597879
## z5     -0.853855478
## size    0.002606491
## grade3   .
## Cox-LASSO selected: z1 z2 z4 z5 size
## Cox-LASSO not selected: z3 grade3
## Default cuts included from Lasso:
## Categorical after Lasso: z1 z2 z4 z5
## # of candidate subgroup factors= 8
## [1] "size <= 29.9" "size <= 25"  "size <= 21"  "size <= 35"  "z1"
## [6] "z2"              "z4"              "z5"
## Number of factors evaluated= 8
## Confounders per grf screening q3 q5 q6 q7 q8 q4 q1 q2
##      Factors Labels VI(grf)
## 3   size <= 21      q3 0.2576
## 5           z1      q5 0.2065
## 6           z2      q6 0.1494
## 7           z4      q7 0.1068
## 8           z5      q8 0.0953
## 4   size <= 35      q4 0.0796
## 1 size <= 29.9      q1 0.0557
## 2   size <= 25      q2 0.0491
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 136
## Approximately 5% of max_count met: minutes 0.0003666667
## Approximately 10% of max_count met: minutes 0.00085
## Approximately 20% of max_count met: minutes 0.00155
## Approximately 33% of max_count met: minutes 0.0021
## Approximately 50% of max_count met: minutes 0.003366667
## Approximately 75% of max_count met: minutes 0.0052
## Approximately 90% of max_count met: minutes 0.00625
## # 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 15

```

```

## # of subgroups with sample size less than criteria 20
## # of subgroups meeting all criteria = 110
## # of subgroups fitted (Cox model estimable) = 110
## *Subgroup Searching Minutes=* 0.00665
## Number of subgroups meeting HR threshold 11
## Subgroup candidate(s) found (FS)
## # of candidate subgroups (meeting HR criteria) = 11
## SGs (1st 10) meeting screening thresholds sorted by sg_focus= hr
##      n    E d1    HR L(HR) q3.0 q3.1 q5.0 q5.1 q6.0 q6.1 q7.0 q7.1 q8.0 q8.1
## 1  117  78 42 2.03  1.29    0    0    0    1    0    1    0    0    0    0
## 2  131  82 40 1.67  1.08    0    0    0    1    0    0    0    0    0    0
## 3   85  48 25 1.58  0.90    0    0    0    1    0    0    0    0    0    0
## 4  133  99 50 1.58  1.06    1    0    0    1    0    0    0    0    0    0
## 5   97  58 28 1.55  0.92    0    0    0    1    0    0    0    0    0    0
## 6  103  87 47 1.45  0.95    0    0    0    1    0    0    0    0    1    0
## 7  185 123 60 1.44  1.01    0    0    0    1    0    0    0    0    0    0
## 8  165 109 50 1.34  0.92    0    0    0    1    0    0    0    1    0    0
## 9  100  75 35 1.32  0.84    0    0    0    1    0    0    0    0    0    0
## 10 157  84 49 1.28  0.83    1    0    0    0    0    0    0    0    0    0
## Consistency 1
## # of splits= 400
## Model, % Consistency Met= {z1} {z2} 1
## SG focus= hr
## Subgroup Consistency Minutes= 0.03258333
## Subgroup found (FS)
## Minutes forestsearch overall= 0.04796667
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 50.23146 2
##   leaf.node control.mean control.size control.se depth
## 1         2        -5.28        512.00        1.26    1
## 2         3         0.51        188.00        2.51    1
## 3         4         3.34        107.00        2.73    2
## 4         5        -6.78        294.00        1.77    2
## 5         6        -7.81        211.00        1.92    2
## 6         7         7.70         88.00        3.59    2
##   leaf.node control.mean control.size control.se depth
## 6         7         7.70         88.00        3.59    2
## GRF subgroup found
## All splits
## [1] "z3 <= 0"      "size <= 20" "z1 <= 0"
## Terminating node at max.diff (sg.harm.id)
## [1] "z1 <= 0"
## # 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: z3 <= 0 size <= 20 z1 <= 0
## Initial GRF cuts included z3 <= 0 size <= 20 z1 <= 0
## # of candidate subgroup factors= 10
## [1] "size <= 20"      "size <= 28.8" "size <= 25"      "size <= 35"      "z1"
## [6] "z2"              "z3"            "z4"              "z5"              "grade3"
## Number of factors evaluated= 10
## Confounders per grf screening q5 q6 q4 q1 q8 q9 q10 q7 q3 q2
##      Factors Labels VI(grf)

```

```

## 5          z1      q5  0.1827
## 6          z2      q6  0.1451
## 4    size <= 35      q4  0.1151
## 1    size <= 20      q1  0.1134
## 8          z4      q8  0.1052
## 9          z5      q9  0.0891
## 10         grade3 q10  0.0826
## 7          z3      q7  0.0771
## 3    size <= 25      q3  0.0465
## 2    size <= 28.8    q2  0.0431
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 210
## Approximately 5% of max_count met: minutes 0.000566667
## Approximately 10% of max_count met: minutes 0.001166667
## Approximately 20% of max_count met: minutes 0.002216667
## Approximately 33% of max_count met: minutes 0.003366667
## Approximately 50% of max_count met: minutes 0.004666667
## Approximately 75% of max_count met: minutes 0.006916667
## Approximately 90% of max_count met: minutes 0.008366667
## # 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 21
## # 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=* 0.00915
## 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) q5.0 q5.1 q6.0 q6.1
## 1 88 68 36 1.85  1.14    0    1    0    0
## 2 84 41 25 1.78  0.95    0    0    1    0
## 3 87 65 33 1.32  0.81    0    1    0    0
## 4 76 48 25 1.29  0.73    0    1    0    0
## Consistency 0.975
## # of splits= 400
## Model, % Consistency Met= {z1} {z3} 0.975
## SG focus= hr
## Subgroup Consistency Minutes= 0.02825
## Subgroup found (FS)
## Minutes forestsearch overall= 0.0454
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 50.23146 2
##   leaf.node control.mean control.size control.se depth
## 1          2         -5.28         512.00         1.26    1
## 2          3          0.51         188.00         2.51    1
## 3          4          3.34         107.00         2.73    2
## 4          5         -6.78         294.00         1.77    2
## 5          6         -7.81         211.00         1.92    2
## 6          7          7.70          88.00         3.59    2
##   leaf.node control.mean control.size control.se depth
## 6          7          7.70          88.00         3.59    2
## GRF subgroup found

```

```

## All splits
## [1] "z3 <= 0"      "size <= 20" "z1 <= 0"
## Terminating node at max.diff (sg.harm.id)
## [1] "z1 <= 0"
## # of continuous/categorical characteristics 1 6
## Continuous characteristics: size
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## CV lambda = 0.005541341
## 7 x 1 sparse Matrix of class "dgCMatrix"
##              s0
## z1      0.4335154522
## z2     -0.5064376770
## z3      0.5355754144
## z4      0.6517793532
## z5     -0.7722640941
## size   -0.0005545369
## grade3  0.0123079851
## Cox-LASSO selected: z1 z2 z3 z4 z5 size grade3
## Cox-LASSO not selected:
## Default cuts included from Lasso:
## Categorical after Lasso: z1 z2 z3 z4 z5 grade3
## Factors per GRF: z3 <= 0 size <= 20 z1 <= 0
## Initial GRF cuts included z3 <= 0 size <= 20 z1 <= 0
## Factors included per GRF (not in lasso) size <= 20
## # of candidate subgroup factors= 10
## [1] "size <= 20"      "size <= 28.8" "size <= 25"      "size <= 35"      "z1"
## [6] "z2"              "z3"              "z4"              "z5"              "grade3"
## Number of factors evaluated= 10
## Confounders per grf screening q5 q6 q4 q1 q8 q9 q10 q7 q3 q2
##      Factors Labels VI(grf)
## 5          z1      q5 0.1827
## 6          z2      q6 0.1451
## 4    size <= 35      q4 0.1151
## 1    size <= 20      q1 0.1134
## 8          z4      q8 0.1052
## 9          z5      q9 0.0891
## 10         grade3    q10 0.0826
## 7          z3      q7 0.0771
## 3    size <= 25      q3 0.0465
## 2  size <= 28.8      q2 0.0431
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 210
## Approximately 5% of max_count met: minutes 0.0004833333
## Approximately 10% of max_count met: minutes 0.00095
## Approximately 20% of max_count met: minutes 0.001933333
## Approximately 33% of max_count met: minutes 0.003033333
## Approximately 50% of max_count met: minutes 0.004466667
## Approximately 75% of max_count met: minutes 0.006383333
## Approximately 90% of max_count met: minutes 0.007866667
## # 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 21
## # 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=* 0.008666667
## 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) q5.0 q5.1 q6.0 q6.1
## 1 88 68 36 1.85  1.14    0    1    0    0
## 2 84 41 25 1.78  0.95    0    0    1    0
## 3 87 65 33 1.32  0.81    0    1    0    0
## 4 76 48 25 1.29  0.73    0    1    0    0
## Consistency 0.975
## # of splits= 400
## Model, % Consistency Met= {z1} {z3} 0.975
## SG focus= hr
## Subgroup Consistency Minutes= 0.0316
## Subgroup found (FS)
## Minutes forestsearch overall= 0.05031667

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)
##      1000      1000      1000      1000      1000      1000      1000      1000

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

dgm_alt <- dgm

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

## [1] "results/oc_sims=1000_m4a-Noise=0_N=700_alt_ktreat=0.9_hrH=2_v0B-4cuts.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      0       0      700
##      ppv      npv specificity sensitivity found.1 found.2 found.both found.al3
##      <num> <num>      <num>      <num>      <num> <num>      <num>      <num>
## 1:      1      1 1.0000000          1      0      0          0      0
## 2:      1      1 1.0000000          1      0      0          0      0
## 3:      1      1 1.0000000          1      NA      NA          NA      NA
## 4:      0      1 0.8642857          0      NA      NA          NA      0
## 5:      0      1 0.8642857          0      0      0          0      0
## 6:      0      1 0.8642857          0      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 1.74466 0.5577464 0 -0.2553474 0.0000000
## 2: 1.74466 0.5577464 1.74466 0.5577464 0 -0.2553474 0.0000000
## 3: 1.74466 0.5577464 1.74466 0.5577464 0 -0.2553474 0.0000000
## 4: 1.74466 0.5577464 NA 0.6688550 NA NA 0.1111086
## 5: 1.74466 0.5577464 NA 0.6688550 NA NA 0.1111086
## 6: 1.74466 0.5577464 NA 0.6688550 NA NA 0.1111086
##      b2.Hc      p.cens analysis      taumax      hr.itt      l.itt      u.itt
##      <num>      <num>      <char>      <num>      <num>      <num>      <num>
## 1: -0.08889407 0.4514286      FSl 81.43453 0.668855 0.5464683 0.8186513
## 2: -0.08889407 0.4514286      FSlg 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: 0.02221450 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: 1.105971 2.752186 0.4437854 0.7009717
## 2: 1.105971 2.752186 0.4437854 0.7009717
## 3: 1.105971 2.752186 0.4437854 0.7009717
## 4:      NA      NA 0.5464683 0.8186513
## 5:      NA      NA 0.5464683 0.8186513
## 6:      NA      NA 0.5464683 0.8186513
## Subgroup HRs: H, H^c, Causal= 2.000007 0.6466405 0.7057463
## Simulations= 1000
## Avg censoring= 0.44999
## Min,Max,Avg tau.max= 74.465 83.95939 81.60594
## P(H) approximation at causal(H), n(sg)=60, approx= 2.000007 60 0.8283003
## P(H) approximation at causal(H), Avg(n(sg)), approx= 2.000007 89 0.8998777
## P(H) approximation at plim(H), Avg(n(sg)), approx= 2.127636 89 0.9289428
## Minutes,hours 6.375333 0.1062556
##      FSl      FSlg      GRF      GRF.60      VT(24)      VT#(24)      VT(36)      VT#(36)
## any.H      0.920      0.880      0.940      0.760      0.510      0.570      0.480      0.610
## sensH      0.860      0.810      0.860      0.690      0.470      0.540      0.430      0.580
## sensHc     0.990      0.990      0.970      0.980      0.990      0.990      0.990      0.990
## ppH        0.860      0.800      0.780      0.630      0.450      0.530      0.420      0.570
## ppHc       0.980      0.980      0.980      0.960      0.930      0.940      0.930      0.950
## Avg(#H)    89.000      92.000      102.000      100.000      92.000      92.000      90.000      91.000
## minH       61.000      61.000      61.000      60.000      63.000      60.000      60.000      61.000
## maxH       177.000      183.000      294.000      208.000      170.000      159.000      160.000      175.000
## Avg(#Hc)   618.000      620.000      604.000      624.000      653.000      648.000      656.000      644.000
## minHc      523.000      517.000      406.000      492.000      530.000      541.000      540.000      525.000
## maxHc      700.000      700.000      700.000      700.000      700.000      700.000      700.000      700.000
## hat(H*)    2.189      2.210      2.161      2.226      2.377      2.359      2.412      2.351
## hat(hat[H]) 2.225      2.199      2.020      2.025      2.267      2.284      2.303      2.304
## hat(Hc*)   0.653      0.652      0.649      0.648      0.652      0.653      0.652      0.653
## hat(hat[Hc]) 0.652      0.650      0.643      0.646      0.655      0.654      0.656      0.653
## hat(H*)all 2.128      2.128      2.128      2.128      2.128      2.128      2.128      2.128

```

```

## 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)	0.92	0.88	0.94	0.76	0.51	0.57	0.48	0.61
$sens(\hat{H})$	0.86	0.81	0.86	0.69	0.47	0.54	0.43	0.58
$sens(\hat{H}^c)$	0.99	0.99	0.97	0.98	0.99	0.99	0.99	0.99
$ppv(\hat{H})$	0.86	0.8	0.78	0.63	0.45	0.53	0.42	0.57
$ppv(\hat{H}^c)$	0.98	0.98	0.98	0.96	0.93	0.94	0.93	0.95
Size of H and H-complement								
$avg \hat{H} $	89	92	102	100	92	92	90	91
$min \hat{H} $	61	61	61	60	63	60	60	61
$max \hat{H} $	177	183	294	208	170	159	160	175
$avg \hat{H}^c $	618	620	604	624	653	648	656	644
$min \hat{H}^c $	523	517	406	492	530	541	540	525
$max \hat{H}^c $	700	700	700	700	700	700	700	700

Note: Number of simulations= 1000 .

Note: Probability approximation= 0.8999 .

```

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 12.60955 0.2101592

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 126.0955 2.101592

# 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

```