```
N <- 500
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 <- 0.225  # 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")
```

```r
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 <- "m4aB"
file.index <- "v0-4cuts"

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 0.2026239
## Underlying pH_super 0.2026239

# 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)", "${sens}(\\hat{H})$", "${sens}(\\hat{H}^{C})$", "${ppv}(\\hat{H})$",
    "${ppv}(\\hat{H}^{C})$", "${avg}\\vert \\hat{H} \\vert$", "${min}\\vert \\hat{H} \\vert$",
    "${max}\\vert \\hat{H} \\vert$", "${avg}\\vert \\hat{H}^{C} \\vert$", "${min}\\vert \\hat{H}^{C} \\v
    "${max}\\vert \\hat{H}^{C} \\vert$")

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)
```
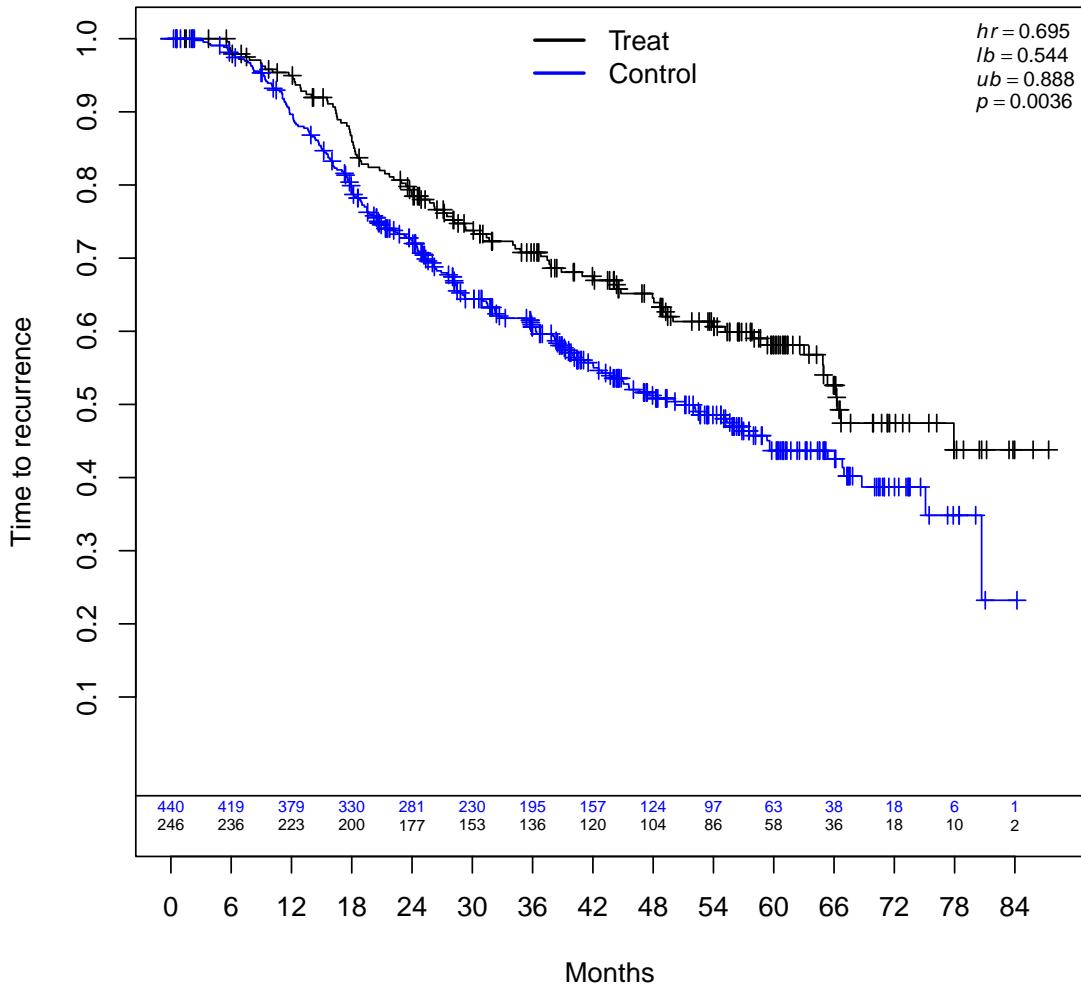
```
## Super-population empirical harm and non-harm hazard ratios= NA 0.692068
## Causal HR (empirical ITT)= 0.692068

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= 48.14834 2
```

```
##   leaf.node control.mean control.size control.se depth
## 1         2        -5.28       477.00        1.25     1
## 2         4        -6.33       396.00        1.35     2
##   leaf.node control.mean control.size control.se depth
## 1         2        -5.28       477.00        1.25     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 <= 21"   "size <= 35"   "z1"
##  [6] "z2"           "z3"           "z4"           "z5"           "grade3"
## Number of factors evaluated= 10
## Confounders per grf screening q4 q9 q3 q7 q5 q8 q10 q6 q2 q1
##         Factors Labels VI(grf)
## 4    size <= 35     q4  0.2105
## 9            z5     q9  0.1304
## 3    size <= 21     q3  0.1144
## 7            z3     q7  0.0969
## 5            z1     q5  0.0947
## 8            z4     q8  0.0921
## 10       grade3    q10  0.0815
## 6            z2     q6  0.0682
## 2    size <= 25     q2  0.0614
## 1  size <= 29.3     q1  0.0500
## 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 0.00085
## Approximately 20% of max_count met: minutes 0.001033333
## Approximately 33% of max_count met: minutes 0.001966667
## Approximately 50% of max_count met: minutes 0.003133333
## Approximately 75% of max_count met: minutes 0.00455
## Approximately 90% of max_count met: minutes 0.005516667
## # 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 22 23
## # of subgroups with sample size less than criteria 42
## # of subgroups meeting all criteria = 161
## # of subgroups fitted (Cox model estimable) = 161
## *Subgroup Searching Minutes=* 0.00595
## Number of subgroups meeting HR threshold 1
## Subgroup candidate(s) found (FS)
## # of candidate subgroups (meeting HR criteria) =  1
## SGs (1st 10) meeting screening thresholds sorted by sg_focus= hr
##     n  E d1   HR L(HR) q4.0
## 1 61 45 21 1.27   0.7    0
## Consistency 0.5275
## Subgroup Consistency Minutes= 0.0254
## NO subgroup found (FS)
## Minutes forestsearch overall= 0.04455
```

```
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 48.14834 2
##   leaf.node control.mean control.size control.se depth
## 1         2        -5.28       477.00       1.25     1
## 2         4        -6.33       396.00       1.35     2
##   leaf.node control.mean control.size control.se depth
## 1         2        -5.28       477.00       1.25     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 = 0.007637654
## 7 x 1 sparse Matrix of class "dgCMatrix"
##                  s0
## z1      0.044364827
## z2     -0.183848823
## z3         .
## z4      0.373481824
## z5     -0.927551766
## size    0.005961288
## grade3  0.298993114
## Cox-LASSO selected: z1 z2 z4 z5 size grade3
## Cox-LASSO not selected: z3
## Default cuts included from Lasso:
## Categorical after Lasso: z1 z2 z4 z5 grade3
## # of candidate subgroup factors= 9
## [1] "size <= 29.3" "size <= 25"   "size <= 21"   "size <= 35"   "z1"
## [6] "z2"           "z4"           "z5"           "grade3"
## Number of factors evaluated= 9
## Confounders per grf screening q4 q8 q3 q5 q6 q7 q9 q1 q2
##        Factors Labels VI(grf)
## 4   size <= 35     q4  0.2139
## 8           z5     q8  0.1334
## 3   size <= 21     q3  0.1329
## 5           z1     q5  0.1097
## 6           z2     q6  0.0990
## 7           z4     q7  0.0953
## 9       grade3     q9  0.0946
## 1 size <= 29.3     q1  0.0613
## 2   size <= 25     q2  0.0600
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 171
## Approximately 5% of max_count met: minutes 0.0003166667
## Approximately 10% of max_count met: minutes 0.0006666667
## Approximately 20% of max_count met: minutes 0.00085
## Approximately 33% of max_count met: minutes 0.001583333
## Approximately 50% of max_count met: minutes 0.002783333
## Approximately 75% of max_count met: minutes 0.004116667
## Approximately 90% of max_count met: minutes 0.004883333
## # 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 20 21
## # of subgroups with sample size less than criteria 35
```

```
## # of subgroups meeting all criteria = 130
## # of subgroups fitted (Cox model estimable) = 130
## *Subgroup Searching Minutes=* 0.005316667
## Number of subgroups meeting HR threshold 1
## Subgroup candidate(s) found (FS)
## # of candidate subgroups (meeting HR criteria) =  1
## SGs (1st 10) meeting screening thresholds sorted by sg_focus= hr
##    n  E d1   HR L(HR) q4.0
## 1 61 45 21 1.27   0.7    0
## Consistency 0.5275
## Subgroup Consistency Minutes= 0.02328333
## NO subgroup found (FS)
## Minutes forestsearch overall= 0.035
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 47.3259 2
##   leaf.node control.mean control.size control.se depth
## 2         3        -4.98       466.00       1.33     1
## 1         4        -8.57       226.00       2.05     2
## 4         7        -2.72       213.00       1.71     2
##   leaf.node control.mean control.size control.se depth
## 4         7        -2.72       213.00       1.71     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 q9 q8 q5 q4 q3 q6 q7 q1 q10 q2
##        Factors Labels VI(grf)
## 9             z5     q9  0.1997
## 8             z4     q8  0.1206
## 5             z1     q5  0.1036
## 4    size <= 35     q4  0.1004
## 3    size <= 20     q3  0.1003
## 6             z2     q6  0.0995
## 7             z3     q7  0.0808
## 1  size <= 29.4     q1  0.0675
## 10        grade3    q10  0.0668
## 2    size <= 25     q2  0.0608
## 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 0.001016667
## Approximately 20% of max_count met: minutes 0.00185
## Approximately 33% of max_count met: minutes 0.002983333
## Approximately 50% of max_count met: minutes 0.007283333
## Approximately 75% of max_count met: minutes 0.009083333
## Approximately 90% of max_count met: minutes 0.00995
## # 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 22 20
## # of subgroups with sample size less than criteria 38
## # of subgroups meeting all criteria = 165
## # of subgroups fitted (Cox model estimable) = 165
## *Subgroup Searching Minutes=* 0.0105
## 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= 47.3259 2
##    leaf.node control.mean control.size control.se depth
## 2          3        -4.98       466.00       1.33     1
## 1          4        -8.57       226.00       2.05     2
## 4          7        -2.72       213.00       1.71     2
##    leaf.node control.mean control.size control.se depth
## 4          7        -2.72       213.00       1.71     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 = 0.01957029
## 7 x 1 sparse Matrix of class "dgCMatrix"
##                  s0
## z1        .
## z2     -0.061646715
## z3        .
## z4      0.505919924
## z5     -0.896824809
## size    0.003664031
## grade3 -0.008192052
## Cox-LASSO selected: z2 z4 z5 size grade3
## Cox-LASSO not selected: z1 z3
## Default cuts included from Lasso:
## Categorical after Lasso: z2 z4 z5 grade3
## # of candidate subgroup factors= 8
## [1] "size <= 29.4" "size <= 25"   "size <= 20"   "size <= 35"   "z2"
## [6] "z4"           "z5"           "grade3"
## Number of factors evaluated= 8
## Confounders per grf screening q7 q6 q5 q3 q4 q8 q1 q2
##        Factors Labels VI(grf)
## 7           z5     q7  0.2248
## 6           z4     q6  0.1683
## 5           z2     q5  0.1371
## 3    size <= 20     q3  0.1211
## 4    size <= 35     q4  0.1132
## 8        grade3     q8  0.0843
## 1 size <= 29.4     q1  0.0789
## 2    size <= 25     q2  0.0724
## 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 6e-04
## Approximately 20% of max_count met: minutes 0.001166667
## Approximately 33% of max_count met: minutes 0.001883333
## Approximately 50% of max_count met: minutes 0.002616667
```

```
## Approximately 75% of max_count met: minutes 0.004016667
## Approximately 90% of max_count met: minutes 0.0044
## # 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 17 16
## # of subgroups with sample size less than criteria 28
## # of subgroups meeting all criteria = 102
## # of subgroups fitted (Cox model estimable) = 102
## *Subgroup Searching Minutes=* 0.00475
## 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= 48.15369 2
##    leaf.node control.mean control.size control.se depth
## 1          2        -7.48       288.00       1.64     1
## 2          3         0.35       212.00       2.07     1
## 11         4        -8.00       278.00       1.65     2
## 3          6         5.09        90.00       3.02     2
## 4          7        -3.15       122.00       2.80     2
##   leaf.node control.mean control.size control.se depth
## 3          6         5.09        90.00       3.02     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 <= 28.7" "size <= 25"   "size <= 20"   "size <= 35"   "z1"
##  [6] "z2"           "z3"           "z4"           "z5"           "grade3"
## Number of factors evaluated= 10
## Confounders per grf screening q5 q9 q1 q3 q6 q7 q8 q10 q2 q4
##         Factors Labels VI(grf)
## 5            z1     q5  0.2629
## 9            z5     q9  0.1145
## 1   size <= 28.7    q1  0.0971
## 3     size <= 20    q3  0.0958
## 6            z2     q6  0.0860
## 7            z3     q7  0.0810
## 8            z4     q8  0.0772
## 10       grade3    q10  0.0715
## 2     size <= 25    q2  0.0573
## 4     size <= 35    q4  0.0566
## 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 0.0009333333
## Approximately 20% of max_count met: minutes 0.001716667
## Approximately 33% of max_count met: minutes 0.003016667
## Approximately 50% of max_count met: minutes 0.004183333
## Approximately 75% of max_count met: minutes 0.005933333
## Approximately 90% of max_count met: minutes 0.00715
## # 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 23
## # of subgroups with sample size less than criteria 38
## # of subgroups meeting all criteria = 165
## # of subgroups fitted (Cox model estimable) = 165
## *Subgroup Searching Minutes=* 0.00775
## 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) q5.0 q5.1 q9.0 q9.1 q1.0
## 1  90 59 25 1.54  0.91    0    1    0    0    0
## 2 113 73 30 1.48  0.92    0    1    0    0    0
## 3  87 57 26 1.37  0.81    0    1    0    0    0
## 4  61 34 16 1.32  0.67    0    1    0    0    0
## 5  71 48 24 1.31  0.74    0    0    0    0    0
## Consistency 0.9175
## # of splits= 400
## Model, % Consistency Met= {z1} !{z2} 0.9175
## Consistency 0.895
## Consistency 0.7625
## Consistency 0.5475
## SG focus= hr
## Subgroup Consistency Minutes= 0.1146833
## Subgroup found (FS)
## Minutes forestsearch overall= 0.1276333
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 48.15369 2
##    leaf.node control.mean control.size control.se depth
## 1          2        -7.48       288.00       1.64     1
## 2          3         0.35       212.00       2.07     1
## 11         4        -8.00       278.00       1.65     2
## 3          6         5.09        90.00       3.02     2
## 4          7        -3.15       122.00       2.80     2
##   leaf.node control.mean control.size control.se depth
## 3          6         5.09        90.00       3.02     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 = 0.02885842
## 7 x 1 sparse Matrix of class "dgCMatrix"
##                 s0
## z1        .
## z2        .
## z3     -0.07206902
## z4      0.56122830
## z5     -0.62811696
## size      .
## grade3  .
## Cox-LASSO selected: z3 z4 z5
## Cox-LASSO not selected: z1 z2 size grade3
```

```
## Default cuts included from Lasso:
## Categorical after Lasso: z3 z4 z5
## # of candidate subgroup factors= 7
## [1] "size <= 28.7" "size <= 25"   "size <= 20"   "size <= 35"   "z3"
## [6] "z4"          "z5"
## Number of factors evaluated= 7
## Confounders per grf screening q7 q6 q5 q3 q1 q4 q2
##       Factors Labels VI(grf)
## 7            z5     q7  0.1776
## 6            z4     q6  0.1729
## 5            z3     q5  0.1565
## 3   size <= 20     q3  0.1492
## 1 size <= 28.7     q1  0.1451
## 4   size <= 35     q4  0.1144
## 2   size <= 25     q2  0.0843
## 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 0.00045
## Approximately 20% of max_count met: minutes 8e-04
## Approximately 33% of max_count met: minutes 0.001333333
## Approximately 50% of max_count met: minutes 0.0019
## Approximately 75% of max_count met: minutes 0.002733333
## Approximately 90% of max_count met: minutes 0.00305
## # 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 14
## # of subgroups with sample size less than criteria 20
## # of subgroups meeting all criteria = 79
## # of subgroups fitted (Cox model estimable) = 79
## *Subgroup Searching Minutes=* 0.003183333
## Number of subgroups meeting HR threshold 1
## Subgroup candidate(s) found (FS)
## # of candidate subgroups (meeting HR criteria) =  1
## SGs (1st 10) meeting screening thresholds sorted by sg_focus= hr
##    n  E d1   HR L(HR) q7.0
## 1 71 48 24 1.31  0.74    0
## Consistency 0.6425
## Subgroup Consistency Minutes= 0.0271
## NO subgroup found (FS)
## Minutes forestsearch overall= 0.03701667

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
```

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

```
## [1] "results/oc_sims=5000_m4aB-Noise=0_N=500_null_ktreat=0.9_v0-4cuts.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         500           1     0      0      500
## 2:    1         0          0         500           1     0      0      500
## 3:    1         0          0         500           1     0      0      500
## 4:    1         0          0         500           1     0      0      500
## 5:    1         0          0         500           1     0      0      500
## 6:    1         0          0         500           1     0      0      500
##      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  0.6280343       NA 0.6280343     NA     NA     0 -0.06403361
## 2:        NA  0.6280343       NA 0.6280343     NA     NA     0 -0.06403361
## 3:        NA  0.6280343       NA 0.6280343     NA     NA     0 -0.06403361
## 4:        NA  0.6280343       NA 0.6280343     NA     NA     0 -0.06403361
## 5:        NA  0.6280343       NA 0.6280343     NA     NA     0 -0.06403361
## 6:        NA  0.6280343       NA 0.6280343     NA     NA     0 -0.06403361
##    p.cens analysis    taumax     hr.itt     l.itt     u.itt hr.adj.itt l.adj.itt
##     <num>   <char>    <num>      <num>     <num>     <num>      <num>    <num>
## 1:  0.444      FSl 80.24723 0.6280343 0.4951898 0.796517  0.5866668 0.4615432
## 2:  0.444     FSlg 48.14834 0.6280343 0.4951898 0.796517  0.5866668 0.4615432
## 3:  0.444      GRF 80.24723 0.6280343 0.4951898 0.796517  0.5866668 0.4615432
## 4:  0.444   GRF.60 48.14834 0.6280343 0.4951898 0.796517  0.5866668 0.4615432
## 5:  0.444   VT(24) 80.24723 0.6280343 0.4951898 0.796517  0.5866668 0.4615432
## 6:  0.444  VT#(24) 80.24723 0.6280343 0.4951898 0.796517  0.5866668 0.4615432
##    u.adj.itt l.H.true u.H.true l.Hc.true u.Hc.true l.H.hat u.H.hat  l.Hc.hat
##        <num>  <lgcl>   <lgcl>     <num>     <num>   <num>   <num>     <num>
## 1: 0.7457113      NA       NA 0.4951898  0.796517      NA      NA 0.4951898
## 2: 0.7457113      NA       NA 0.4951898  0.796517      NA      NA 0.4951898
## 3: 0.7457113      NA       NA 0.4951898  0.796517      NA      NA 0.4951898
## 4: 0.7457113      NA       NA 0.4951898  0.796517      NA      NA 0.4951898
## 5: 0.7457113      NA       NA 0.4951898  0.796517      NA      NA 0.4951898
## 6: 0.7457113      NA       NA 0.4951898  0.796517      NA      NA 0.4951898
##    u.Hc.hat
##       <num>
## 1: 0.796517
## 2: 0.796517
## 3: 0.796517
## 4: 0.796517
## 5: 0.796517
## 6: 0.796517
## Subgroup HRs: H, H^c, Causal= NA 0.692068 0.692068
## Simulations= 5000
## Avg censoring= 0.4623128
```

```
## Min,Max,Avg tau.max= 65.82618 83.97699 80.84556
## P(H) approximation at causal(Hrc), n=60, approx= 0.692068 0.03144857
## P(H) approximation at plim(Hrc), n=60, approx= 0.6978451 0.03308917
## Minutes,hours 19.68865 0.3281442
##                    FS1     FS1g     GRF   GRF.60  VT(24) VT#(24)  VT(36) VT#(36)
## any.H             0.070   0.050   0.230   0.050   0.030   0.020   0.040   0.020
## sensH               NaN     NaN     NaN     NaN     NaN     NaN     NaN     NaN
## sensHc            0.990   0.990   0.960   0.990   0.990   1.000   0.990   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)          92.000  95.000  87.000  78.000  77.000  75.000  79.000  76.000
## minH             61.000  61.000  60.000  60.000  60.000  60.000  60.000  60.000
## maxH            267.000 267.000 283.000 147.000 136.000 135.000 154.000 182.000
## Avg(#Hc)        494.000 495.000 480.000 496.000 497.000 499.000 496.000 498.000
## minHc           233.000 233.000 217.000 353.000 364.000 365.000 346.000 318.000
## maxHc           500.000 500.000 500.000 500.000 500.000 500.000 500.000 500.000
## hat(H*)             NaN     NaN     NaN     NaN     NaN     NaN     NaN     NaN
## hat(hat[H])       1.823   1.823   1.552   1.513   1.298   1.364   1.316   1.527
## hat(Hc*)          0.803   0.808   0.753   0.753   0.751   0.746   0.742   0.761
## hat(hat[Hc])      0.679   0.682   0.651   0.666   0.689   0.680   0.675   0.679
## hat(H*)all          NaN     NaN     NaN     NaN     NaN     NaN     NaN     NaN
## hat(Hc*)all       0.698   0.698   0.698   0.698   0.698   0.698   0.698   0.698
## hat(ITT)all       0.698   0.698   0.698   0.698   0.698   0.698   0.698   0.698
## hat(ITTadj)all    0.653   0.653   0.653   0.653   0.653   0.653   0.653   0.653

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	0.23	0.05	0.03	0.02	0.04	0.02
$sens(\hat{H})$	.	.	.	.	.	.	.	.
$sens(\hat{H}^C)$	0.99	0.99	0.96	0.99	0.99	1	0.99	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								
$avg|\hat{H}|$	92	95	87	78	77	75	79	76
$min|\hat{H}|$	61	61	60	60	60	60	60	60
$max|\hat{H}|$	267	267	283	147	136	135	154	182
$avg|\hat{H}^C|$	494	495	480	496	497	499	496	498
$min|\hat{H}^C|$	233	233	217	353	364	365	346	318
$max|\hat{H}^C|$	500	500	500	500	500	500	500	500

Note: Number of simulations= 5000 .
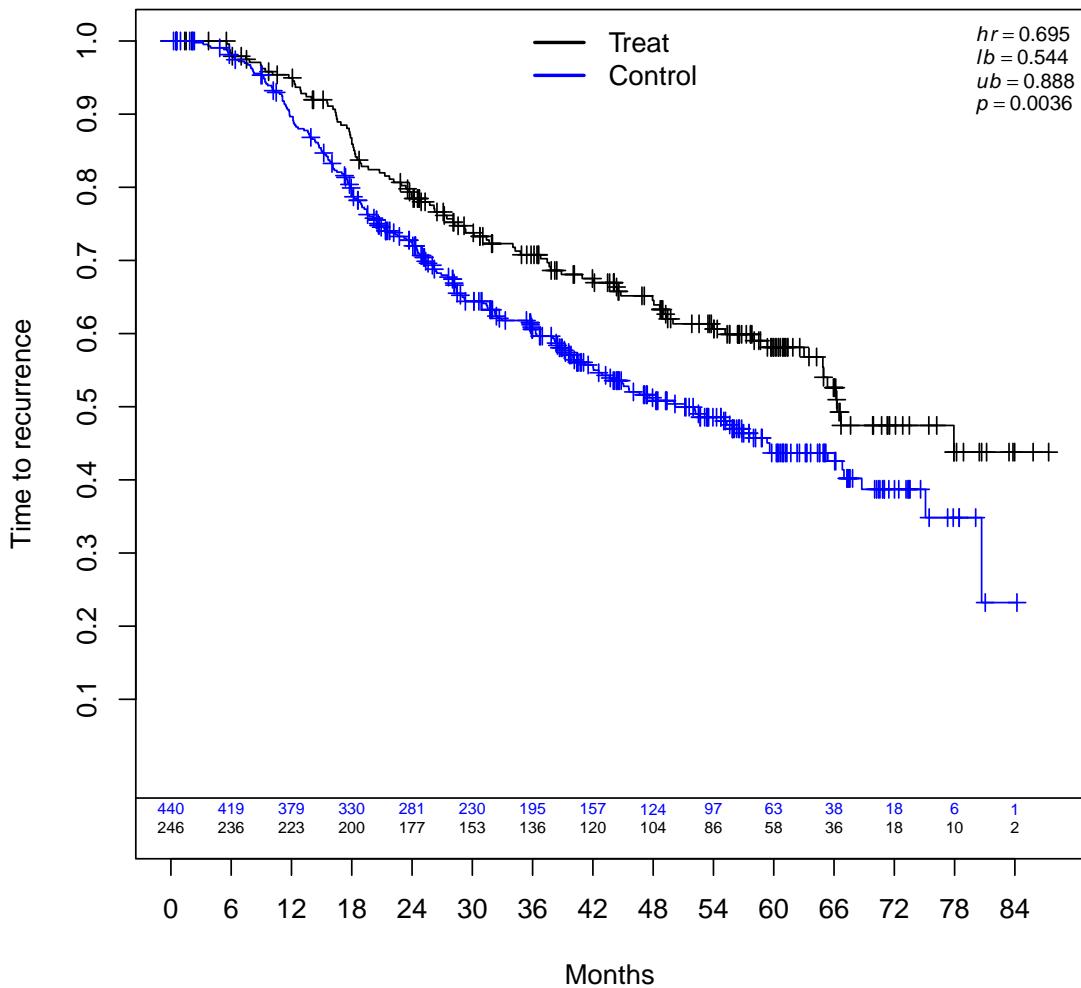
Note: Probability approximation= 0.0331 .

```
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= 2.000002 0.6861679
## Causal HR (empirical ITT)= 0.790348

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= 48.86732 2
##    leaf.node control.mean control.size control.se depth
## 1          2        -6.67       301.00       1.66     1
## 2          3         5.53       199.00       1.98     1
## 11         4        -5.49       259.00       1.68     2
## 3          6        -8.27       113.00       2.86     2
## 4          7        14.22        96.00       2.96     2
##    leaf.node control.mean control.size control.se depth
## 4          7        14.22        96.00       2.96     2
## GRF subgroup found
## All splits
## [1] "z1 <= 0"     "grade3 <= 0" "z3 <= 0"
## Terminating node at max.diff (sg.harm.id)
## [1] "z3 <= 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: z1 <= 0 grade3 <= 0 z3 <= 0
## Initial GRF cuts included z1 <= 0 grade3 <= 0 z3 <= 0
## # of candidate subgroup factors= 10
##   [1] "size <= 29.3" "size <= 25"   "size <= 21"   "size <= 35"   "z1"
##   [6] "z2"           "z3"           "z4"           "z5"           "grade3"
## Number of factors evaluated= 10
## Confounders per grf screening q7 q5 q3 q4 q8 q6 q9 q10 q1 q2
##         Factors Labels VI(grf)
## 7            z3     q7  0.3290
## 5            z1     q5  0.1436
## 3    size <= 21     q3  0.0916
## 4    size <= 35     q4  0.0846
## 8            z4     q8  0.0788
## 6            z2     q6  0.0687
## 9            z5     q9  0.0685
## 10       grade3    q10  0.0504
## 1  size <= 29.3     q1  0.0455
## 2    size <= 25     q2  0.0393
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 210
## Approximately 5% of max_count met: minutes 0.0003666667
## Approximately 10% of max_count met: minutes 0.0008166667
## Approximately 20% of max_count met: minutes 0.00145
## Approximately 33% of max_count met: minutes 0.00225
## Approximately 50% of max_count met: minutes 0.0034
## Approximately 75% of max_count met: minutes 0.004566667
## Approximately 90% of max_count met: minutes 0.0054
## # 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 22 21
## # of subgroups with sample size less than criteria 42
## # of subgroups meeting all criteria = 161
## # of subgroups fitted (Cox model estimable) = 161
## *Subgroup Searching Minutes=* 0.005833333
## Number of subgroups meeting HR threshold 12
## Subgroup candidate(s) found (FS)
## # of candidate subgroups (meeting HR criteria) =  12
## 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 q3.0 q3.1 q4.0 q4.1 q8.0 q8.1
## 1    96 78 46 2.57  1.62    0    1    0    1    0    0    0    0    0    0
## 2    93 73 40 1.83  1.15    0    1    0    0    0    0    0    0    0    1
## 3   120 91 49 1.62  1.07    0    0    0    1    0    0    0    0    0    0
## 4    61 47 23 1.61  0.89    0    0    0    0    0    0    0    0    0    0
## 5    85 69 34 1.47  0.91    0    0    0    1    0    0    0    0    0    0
## 6   109 86 42 1.42  0.93    0    0    0    1    0    0    0    0    0    0
## 7    83 63 31 1.35  0.82    0    0    0    0    1    0    0    0    0    0
## 8    95 55 30 1.33  0.78    0    1    0    0    0    0    0    0    0    0
## 9   103 81 36 1.32  0.85    0    0    0    1    0    0    0    0    0    0
## 10   92 70 34 1.31  0.81    0    1    0    0    0    0    0    0    0    0
## Consistency 1
## # of splits= 400
## Model, % Consistency Met= {z3} {z1} 1
## SG focus= hr
## Subgroup Consistency Minutes= 0.02271667
## Subgroup found (FS)
## Minutes forestsearch overall= 0.03688333
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 48.86732 2
##    leaf.node control.mean control.size control.se depth
## 1          2        -6.67       301.00       1.66     1
## 2          3         5.53       199.00       1.98     1
## 11         4        -5.49       259.00       1.68     2
## 3          6        -8.27       113.00       2.86     2
## 4          7        14.22        96.00       2.96     2
##    leaf.node control.mean control.size control.se depth
## 4          7        14.22        96.00       2.96     2
## GRF subgroup found
## All splits
## [1] "z1 <= 0"     "grade3 <= 0" "z3 <= 0"
## Terminating node at max.diff (sg.harm.id)
## [1] "z3 <= 0"
## # of continuous/categorical characteristics 1 6
## Continuous characteristics: size
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## CV lambda = 0.01368314
## 7 x 1 sparse Matrix of class "dgCMatrix"
##                  s0
## z1      0.284127262
## z2        .
## z3      0.119025277
## z4      0.404175167
## z5     -0.879292411
## size    0.004090154
```

```
## grade3  0.213774781
## 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
## Factors per GRF: z1 <= 0 grade3 <= 0 z3 <= 0
## Initial GRF cuts included z1 <= 0 grade3 <= 0 z3 <= 0
## Factors included per GRF (not in lasso)
## # of candidate subgroup factors= 9
## [1] "size <= 29.3" "size <= 25"   "size <= 21"   "size <= 35"   "z1"
## [6] "z3"           "z4"           "z5"           "grade3"
## Number of factors evaluated= 9
## Confounders per grf screening q6 q5 q3 q4 q7 q8 q9 q1 q2
##         Factors Labels VI(grf)
## 6             z3     q6  0.3519
## 5             z1     q5  0.1460
## 3    size <= 21     q3  0.1018
## 4    size <= 35     q4  0.0893
## 7             z4     q7  0.0846
## 8             z5     q8  0.0721
## 9        grade3     q9  0.0654
## 1 size <= 29.3     q1  0.0467
## 2    size <= 25     q2  0.0422
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 171
## Approximately 5% of max_count met: minutes 0.0003333333
## Approximately 10% of max_count met: minutes 0.0006833333
## Approximately 20% of max_count met: minutes 0.001216667
## Approximately 33% of max_count met: minutes 0.001916667
## Approximately 50% of max_count met: minutes 0.00285
## Approximately 75% of max_count met: minutes 0.00385
## Approximately 90% of max_count met: minutes 0.004533333
## # 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 20 19
## # of subgroups with sample size less than criteria 35
## # of subgroups meeting all criteria = 130
## # of subgroups fitted (Cox model estimable) = 130
## *Subgroup Searching Minutes=* 0.004866667
## 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) q6.0 q6.1 q5.0 q5.1 q3.0 q3.1 q4.0 q4.1 q7.0 q7.1
## 1    96  78 46 2.57  1.62    0    1    0    1    0    0    0    0    0    0
## 2    93  73 40 1.83  1.15    0    1    0    0    0    0    0    0    0    1
## 3    61  47 23 1.61  0.89    0    0    0    0    0    0    0    0    0    0
## 4    85  69 34 1.47  0.91    0    0    0    1    0    0    0    0    0    0
## 5   109  86 42 1.42  0.93    0    0    0    1    0    0    0    0    0    0
## 6    83  63 31 1.35  0.82    0    0    0    0    1    0    0    0    0    0
## 7    95  55 30 1.33  0.78    0    1    0    0    0    0    0    0    0    0
## 8   103  81 36 1.32  0.85    0    0    0    1    0    0    0    0    0    0
## 9    92  70 34 1.31  0.81    0    1    0    0    0    0    0    0    0    0
## 10  146 112 53 1.27  0.87    0    0    0    1    1    0    0    0    0    0
```

```
## Consistency 1
## # of splits= 400
## Model, % Consistency Met= {z3} {z1} 1
## SG focus= hr
## Subgroup Consistency Minutes= 0.02315
## Subgroup found (FS)
## Minutes forestsearch overall= 0.03333333
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 47.23728 2
##    leaf.node control.mean control.size control.se depth
## 1          2        -5.63       292.00       1.67     1
## 2          3         3.50       208.00       2.03     1
## 11         4        -3.87       275.00       1.75     2
## 3          6        -6.91       107.00       2.59     2
## 4          7         7.24       104.00       2.83     2
##   leaf.node control.mean control.size control.se depth
## 4         7         7.24       104.00       2.83     2
## GRF subgroup found
## All splits
## [1] "z1 <= 0"    "size <= 56" "z3 <= 0"
## Terminating node at max.diff (sg.harm.id)
## [1] "z3 <= 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: z1 <= 0 size <= 56 z3 <= 0
## Initial GRF cuts included z1 <= 0 size <= 56 z3 <= 0
## # of candidate subgroup factors= 11
##  [1] "size <= 56"    "size <= 29.4" "size <= 25"    "size <= 20"    "size <= 35"
##  [6] "z1"            "z2"            "z3"            "z4"            "z5"
## [11] "grade3"
## Number of factors evaluated= 11
## Confounders per grf screening q8 q7 q10 q6 q5 q9 q11 q4 q3 q2 q1
##         Factors Labels VI(grf)
## 8            z3     q8  0.2983
## 7            z2     q7  0.1366
## 10           z5    q10  0.1273
## 6            z1     q6  0.0828
## 5    size <= 35     q5  0.0815
## 9            z4     q9  0.0675
## 11       grade3    q11  0.0591
## 4    size <= 20     q4  0.0568
## 3    size <= 25     q3  0.0444
## 2  size <= 29.4     q2  0.0430
## 1    size <= 56     q1  0.0025
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 253
## Approximately 5% of max_count met: minutes 7e-04
## Approximately 10% of max_count met: minutes 0.001233333
## Approximately 20% of max_count met: minutes 0.00255
## Approximately 33% of max_count met: minutes 0.004
## Approximately 50% of max_count met: minutes 0.006483333
## Approximately 75% of max_count met: minutes 0.008966667
## Approximately 90% of max_count met: minutes 0.01036667
```

```
## # 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 36 35
## # of subgroups with sample size less than criteria 60
## # of subgroups meeting all criteria = 181
## # of subgroups fitted (Cox model estimable) = 181
## *Subgroup Searching Minutes=* 0.01078333
## Number of subgroups meeting HR threshold 9
## Subgroup candidate(s) found (FS)
## # of candidate subgroups (meeting HR criteria) =  9
## SGs (1st 10) meeting screening thresholds sorted by sg_focus= hr
##      n    E d1    HR L(HR) q8.0 q8.1 q7.0 q7.1 q10.0 q10.1 q6.0 q6.1 q5.0
## 1 104   77 39 1.81  1.14    0    1    0    0     0     0    0    1    0
## 2 100   78 39 1.71  1.08    0    1    0    0     0     0    0    0    0
## 3 111   50 25 1.49  0.85    0    1    0    0     0     1    0    0    0
## 4  67   37 18 1.45  0.76    0    1    0    0     0     0    0    0    0
## 5 129   89 44 1.39  0.92    0    0    0    1     0     0    0    1    0
## 6  83   43 22 1.39  0.76    0    0    0    1     0     0    0    0    0
## 7 165   97 48 1.32  0.88    0    1    0    0     0     0    0    0    0
## 8 134  102 52 1.28  0.86    0    0    0    1     0     0    0    0    0
## 9  90   41 22 1.25  0.68    0    0    0    0     0     1    0    0    0
## Consistency 0.9875
## # of splits= 400
## Model, % Consistency Met= {z3} {z1} 0.9875
## SG focus= hr
## Subgroup Consistency Minutes= 0.03336667
## Subgroup found (FS)
## Minutes forestsearch overall= 0.05126667
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 47.23728 2
##    leaf.node control.mean control.size control.se depth
## 1          2        -5.63       292.00       1.67     1
## 2          3         3.50       208.00       2.03     1
## 11         4        -3.87       275.00       1.75     2
## 3          6        -6.91       107.00       2.59     2
## 4          7         7.24       104.00       2.83     2
##   leaf.node control.mean control.size control.se depth
## 4         7         7.24       104.00       2.83     2
## GRF subgroup found
## All splits
## [1] "z1 <= 0"    "size <= 56" "z3 <= 0"
## Terminating node at max.diff (sg.harm.id)
## [1] "z3 <= 0"
## # of continuous/categorical characteristics 1 6
## Continuous characteristics: size
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## CV lambda = 0.01960279
## 7 x 1 sparse Matrix of class "dgCMatrix"
##                 s0
## z1      0.106432047
## z2         .
## z3      0.175971202
## z4      0.630393018
```

```
## z5      -0.867422799
## size     0.002056222
## grade3   .
## Cox-LASSO selected: z1 z3 z4 z5 size
## Cox-LASSO not selected: z2 grade3
## Default cuts included from Lasso:
## Categorical after Lasso: z1 z3 z4 z5
## Factors per GRF: z1 <= 0 size <= 56 z3 <= 0
## Initial GRF cuts included z1 <= 0 size <= 56 z3 <= 0
## Factors included per GRF (not in lasso) size <= 56
## # of candidate subgroup factors= 9
## [1] "size <= 56"   "size <= 29.4" "size <= 25"   "size <= 20"   "size <= 35"
## [6] "z1"           "z3"           "z4"           "z5"
## Number of factors evaluated= 9
## Confounders per grf screening q7 q9 q6 q5 q8 q4 q2 q3 q1
##         Factors Labels VI(grf)
## 7            z3     q7  0.3847
## 9            z5     q9  0.1467
## 6            z1     q6  0.1050
## 5   size <= 35     q5  0.0989
## 8            z4     q8  0.0823
## 4   size <= 20     q4  0.0695
## 2 size <= 29.4     q2  0.0560
## 3   size <= 25     q3  0.0522
## 1   size <= 56     q1  0.0047
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 171
## Approximately 5% of max_count met: minutes 0.0006333333
## Approximately 10% of max_count met: minutes 0.0013
## Approximately 20% of max_count met: minutes 0.002683333
## Approximately 33% of max_count met: minutes 0.004416667
## Approximately 50% of max_count met: minutes 0.006383333
## Approximately 75% of max_count met: minutes 0.008416667
## Approximately 90% of max_count met: minutes 0.01028333
## # 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 27 28
## # of subgroups with sample size less than criteria 41
## # of subgroups meeting all criteria = 119
## # of subgroups fitted (Cox model estimable) = 119
## *Subgroup Searching Minutes=* 0.01061667
## 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) q7.0 q7.1 q9.0 q9.1 q6.0 q6.1
## 1 104 77 39 1.81  1.14    0    1    0    0    0    1
## 2 100 78 39 1.71  1.08    0    1    0    0    0    0
## 3 111 50 25 1.49  0.85    0    1    0    1    0    0
## 4  67 37 18 1.45  0.76    0    1    0    0    0    0
## 5 165 97 48 1.32  0.88    0    1    0    0    0    0
## 6  90 41 22 1.25  0.68    0    0    0    1    0    0
## Consistency 0.9875
## # of splits= 400
```

```
## Model, % Consistency Met= {z3} {z1} 0.9875
## SG focus= hr
## Subgroup Consistency Minutes= 0.03711667
## Subgroup found (FS)
## Minutes forestsearch overall= 0.05558333
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 47.39443 2
##    leaf.node control.mean control.size control.se depth
## 1          2        -7.67       288.00       1.60     1
## 2          3         8.00       212.00       2.01     1
## 11         4        -8.26       278.00       1.61     2
## 4          7         8.74       207.00       1.99     2
##   leaf.node control.mean control.size control.se depth
## 4         7         8.74       207.00       1.99     2
## GRF subgroup found
## All splits
## [1] "z1 <= 0"    "size <= 58" "size <= 11"
## Terminating node at max.diff (sg.harm.id)
## [1] "size <= 11"
## # 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: z1 <= 0 size <= 58 size <= 11
## Initial GRF cuts included z1 <= 0 size <= 58 size <= 11
## # of candidate subgroup factors= 12
##  [1] "size <= 58"   "size <= 11"   "size <= 28.7" "size <= 25"   "size <= 20"
##  [6] "size <= 35"   "z1"           "z2"           "z3"           "z4"
## [11] "z5"           "grade3"
## Number of factors evaluated= 12
## Confounders per grf screening q7 q11 q9 q10 q8 q5 q12 q3 q4 q6 q1 q2
##          Factors Labels VI(grf)
## 7             z1     q7  0.5732
## 11            z5    q11  0.0730
## 9             z3     q9  0.0624
## 10            z4    q10  0.0573
## 8             z2     q8  0.0488
## 5     size <= 20     q5  0.0411
## 12        grade3    q12  0.0404
## 3   size <= 28.7     q3  0.0384
## 4     size <= 25     q4  0.0379
## 6     size <= 35     q6  0.0275
## 1     size <= 58     q1  0.0000
## 2     size <= 11     q2  0.0000
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 300
## Approximately 5% of max_count met: minutes 0.001016667
## Approximately 10% of max_count met: minutes 0.00175
## Approximately 20% of max_count met: minutes 0.00335
## Approximately 33% of max_count met: minutes 0.006033333
## Approximately 50% of max_count met: minutes 0.008166667
## Approximately 75% of max_count met: minutes 0.0121
## Approximately 90% of max_count met: minutes 0.013
## # of subgroups evaluated based on (up to) maxk-factor combinations 300
## % 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 68 69
## # of subgroups with sample size less than criteria 85
## # of subgroups meeting all criteria = 193
## # of subgroups fitted (Cox model estimable) = 193
## *Subgroup Searching Minutes=* 0.01348333
## Number of subgroups meeting HR threshold 28
## Subgroup candidate(s) found (FS)
## # of candidate subgroups (meeting HR criteria) =  28
## SGs (1st 10) meeting screening thresholds sorted by sg_focus= hr
##       n   E d1   HR L(HR) q7.0 q7.1 q11.0 q11.1 q9.0 q9.1 q10.0 q10.1 q8.0 q8.1
## 1    87  63 32 2.96  1.73    0    1     0     0    0    0     0     0    0    0
## 2    99  74 49 2.33  1.43    0    1     0     0    0    1     0     0    0    0
## 3   122  88 54 2.19  1.42    0    1     0     0    0    0     0     0    0    1
## 4   108  72 40 2.10  1.31    0    1     0     0    0    0     0     0    0    0
## 5   124  85 49 2.02  1.31    0    1     0     0    0    0     0     0    0    0
## 6   157 111 61 1.96  1.34    0    1     0     0    0    0     0     0    0    0
## 7   169 122 64 1.96  1.36    0    1     0     0    0    0     0     1    0    0
## 8   105  65 34 1.95  1.19    0    1     0     1    0    0     0     0    0    0
## 9   107  82 45 1.94  1.24    0    1     1     0    0    0     0     0    0    0
## 10 203 142 77 1.93  1.38    0    1     0     0    0    0     0     0    0    0
## Consistency 1
## # of splits= 400
## Model, % Consistency Met= {z1} {grade3} 1
## SG focus= hr
## Subgroup Consistency Minutes= 0.03746667
## Subgroup found (FS)
## Minutes forestsearch overall= 0.05908333
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 47.39443 2
##    leaf.node control.mean control.size control.se depth
## 1          2        -7.67       288.00       1.60     1
## 2          3         8.00       212.00       2.01     1
## 11         4        -8.26       278.00       1.61     2
## 4          7         8.74       207.00       1.99     2
##   leaf.node control.mean control.size control.se depth
## 4          7         8.74       207.00       1.99     2
## GRF subgroup found
## All splits
## [1] "z1 <= 0"     "size <= 58" "size <= 11"
## Terminating node at max.diff (sg.harm.id)
## [1] "size <= 11"
## # of continuous/categorical characteristics 1 6
## Continuous characteristics: size
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## CV lambda = 0.01482319
## 7 x 1 sparse Matrix of class "dgCMatrix"
##                s0
## z1      0.2350688
## z2          .
## z3      0.1432995
## z4      0.6219367
## z5     -0.6427091
## 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
## Factors per GRF: z1 <= 0 size <= 58 size <= 11
## Initial GRF cuts included z1 <= 0 size <= 58 size <= 11
## Factors included per GRF (not in lasso) size <= 58 size <= 11
## # of candidate subgroup factors= 10
##   [1] "size <= 58"  "size <= 11"  "size <= 28.7" "size <= 25"  "size <= 20"
##   [6] "size <= 35"  "z1"          "z3"           "z4"          "z5"
## Number of factors evaluated= 10
## Confounders per grf screening q7 q8 q10 q9 q5 q3 q4 q6 q1 q2
##           Factors Labels VI(grf)
## 7              z1     q7  0.5701
## 8              z3     q8  0.0946
## 10             z5    q10  0.0834
## 9              z4     q9  0.0673
## 5     size <= 20     q5  0.0518
## 3   size <= 28.7     q3  0.0499
## 4     size <= 25     q4  0.0421
## 6     size <= 35     q6  0.0408
## 1     size <= 58     q1  0.0000
## 2     size <= 11     q2  0.0000
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 210
## Approximately 5% of max_count met: minutes 0.00065
## Approximately 10% of max_count met: minutes 0.0012
## Approximately 20% of max_count met: minutes 0.00235
## Approximately 33% of max_count met: minutes 0.003966667
## Approximately 50% of max_count met: minutes 0.005683333
## Approximately 75% of max_count met: minutes 0.008333333
## Approximately 90% of max_count met: minutes 0.00885
## # 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 55 54
## # of subgroups with sample size less than criteria 63
## # of subgroups meeting all criteria = 129
## # of subgroups fitted (Cox model estimable) = 129
## *Subgroup Searching Minutes=* 0.009083333
## 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) q7.0 q7.1 q8.0 q8.1 q10.0 q10.1 q9.0 q9.1 q5.0 q5.1
## 1   99  74 49 2.33  1.43    0    1    0    1     0     0    0    0    0    0
## 2  108  72 40 2.10  1.31    0    1    0    0     0     0    0    0    0    0
## 3  124  85 49 2.02  1.31    0    1    0    0     0     0    0    0    0    0
## 4  157 111 61 1.96  1.34    0    1    0    0     0     0    0    0    0    0
## 5  169 122 64 1.96  1.36    0    1    0    0     0     0    0    1    0    0
## 6  105  65 34 1.95  1.19    0    1    0    0     0     1    0    0    0    0
## 7  107  82 45 1.94  1.24    0    1    0    0     1     0    0    0    0    0
## 8  203 142 77 1.93  1.38    0    1    0    0     0     0    0    0    0    0
## 9  212 147 79 1.93  1.39    0    1    0    0     0     0    0    0    0    0
```

```
## 10 151 111 61 1.92    1.32     0    1    0    0     0     0    0    0    1    0
## Consistency 0.9975
## # of splits= 400
## Model, % Consistency Met= {z1} {z3} 0.9975
## SG focus= hr
## Subgroup Consistency Minutes= 0.0339
## Subgroup found (FS)
## Minutes forestsearch overall= 0.05228333

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_m4aB-Noise=0_N=500_alt_ktreat=0.9_hrH=2_v0-4cuts.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         96      0.192         404       0.808     1     96     404
## 2:     1         96      0.192         404       0.808     1     96     404
## 3:     1         96      0.192         404       0.808     1     96     404
## 4:     1         96      0.192         404       0.808     1     96     404
## 5:     1         96      0.192         404       0.808     1     96     404
## 6:     1         96      0.192         404       0.808     1     96     404
##      ppv   npv specificity sensitivity found.1 found.2 found.both found.al3
##    <num> <num>       <num>       <num>   <num>   <num>      <num>     <num>
## 1:     1     1           1           1       0       0          0         0
## 2:     1     1           1           1       0       0          0         0
## 3:     1     1           1           1      NA      NA         NA        NA
## 4:     1     1           1           1      NA      NA         NA        NA
## 5:     1     1           1           1       0       0          0         0
## 6:     1     1           1           1       0       0          0         0
##    hr.H.true hr.Hc.true hr.H.hat hr.Hc.hat  b1.H     b2.H b1.Hc       b2.Hc
##        <num>      <num>    <num>     <num> <num>    <num> <num>       <num>
## 1: 2.574377  0.5911174 2.574377 0.5911174     0 0.574375     0 -0.09505049
## 2: 2.574377  0.5911174 2.574377 0.5911174     0 0.574375     0 -0.09505049
## 3: 2.574377  0.5911174 2.574377 0.5911174     0 0.574375     0 -0.09505049
## 4: 2.574377  0.5911174 2.574377 0.5911174     0 0.574375     0 -0.09505049
## 5: 2.574377  0.5911174 2.574377 0.5911174     0 0.574375     0 -0.09505049
## 6: 2.574377  0.5911174 2.574377 0.5911174     0 0.574375     0 -0.09505049
##    p.cens analysis    taumax    hr.itt    l.itt    u.itt hr.adj.itt l.adj.itt
##     <num>   <char>     <num>     <num>    <num>    <num>      <num>     <num>
## 1:  0.414      FS1 81.44553 0.7859716 0.6245689 0.9890844  0.8333264 0.6588956
## 2:  0.414     FS1g 48.86732 0.7859716 0.6245689 0.9890844  0.8333264 0.6588956
```

```
## 3:   0.414      GRF 81.44553 0.7859716 0.6245689 0.9890844   0.8333264 0.6588956
## 4:   0.414   GRF.60 48.86732 0.7859716 0.6245689 0.9890844   0.8333264 0.6588956
## 5:   0.414   VT(24) 81.44553 0.7859716 0.6245689 0.9890844   0.8333264 0.6588956
## 6:   0.414  VT#(24) 81.44553 0.7859716 0.6245689 0.9890844   0.8333264 0.6588956
##    u.adj.itt l.H.true u.H.true l.Hc.true u.Hc.true  l.H.hat  u.H.hat  l.Hc.hat
##        <num>    <num>    <num>     <num>     <num>    <num>    <num>     <num>
## 1:  1.053935 1.620097 4.090754 0.4504857 0.7756513 1.620097 4.090754 0.4504857
## 2:  1.053935 1.620097 4.090754 0.4504857 0.7756513 1.620097 4.090754 0.4504857
## 3:  1.053935 1.620097 4.090754 0.4504857 0.7756513 1.620097 4.090754 0.4504857
## 4:  1.053935 1.620097 4.090754 0.4504857 0.7756513 1.620097 4.090754 0.4504857
## 5:  1.053935 1.620097 4.090754 0.4504857 0.7756513 1.620097 4.090754 0.4504857
## 6:  1.053935 1.620097 4.090754 0.4504857 0.7756513 1.620097 4.090754 0.4504857
##     u.Hc.hat
##        <num>
## 1: 0.7756513
## 2: 0.7756513
## 3: 0.7756513
## 4: 0.7756513
## 5: 0.7756513
## 6: 0.7756513
## Subgroup HRs: H, H^c, Causal= 2.000002 0.6861679 0.790348
## Simulations= 5000
## Avg censoring= 0.4299092
## Min,Max,Avg tau.max= 65.56395 83.96284 80.6683
## P(H) approximation at causal(H), n(sg)=60, approx= 2.000002 60 0.835295
## P(H) approximation at causal(H), Avg(n(sg)), approx= 2.000002 101 0.9246477
## P(H) approximation at plim(H), Avg(n(sg)), approx= 2.310836 101 0.9716474
## Minutes,hours 23.01525 0.3835875
##                   FSl     FSlg      GRF   GRF.60  VT(24) VT#(24)  VT(36) VT#(36)
## any.H           0.980    0.970    0.980    0.830   0.660   0.710   0.630   0.770
## sensH           0.870    0.870    0.870    0.730   0.590   0.660   0.560   0.730
## sensHc          0.980    0.980    0.930    0.950   0.980   0.990   0.980   0.990
## ppH             0.880    0.870    0.790    0.670   0.590   0.660   0.550   0.730
## ppHc            0.970    0.970    0.970    0.940   0.910   0.930   0.910   0.940
## Avg(#H)        99.000  100.000  116.000  115.000 102.000 101.000 103.000 102.000
## minH           61.000   61.000   60.000   60.000  60.000  60.000  60.000  60.000
## maxH          202.000  200.000  373.000  273.000 174.000 186.000 217.000 202.000
## Avg(#Hc)      404.000  403.000  386.000  404.000 433.000 428.000 435.000 421.000
## minHc         298.000  300.000  127.000  227.000 326.000 314.000 283.000 298.000
## maxHc         500.000  500.000  500.000  500.000 500.000 500.000 500.000 500.000
## hat(H*)         2.332    2.336    2.323    2.395   2.490   2.478   2.523   2.464
## hat(hat[H])     2.398    2.381    2.123    2.116   2.372   2.406   2.408   2.433
## hat(Hc*)        0.687    0.687    0.687    0.685   0.689   0.689   0.689   0.687
## hat(hat[Hc])    0.692    0.692    0.679    0.686   0.702   0.697   0.699   0.689
## hat(H*)all      2.311    2.311    2.311    2.311   2.311   2.311   2.311   2.311
## hat(Hc*)all     0.687    0.687    0.687    0.687   0.687   0.687   0.687   0.687
## hat(ITT)all     0.871    0.871    0.871    0.871   0.871   0.871   0.871   0.871
## hat(ITTadj)all  0.895    0.895    0.895    0.895   0.895   0.895   0.895   0.895
```

```r
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)	0.98	0.97	0.98	0.83	0.66	0.71	0.63	0.77
$sens(\hat{H})$	0.87	0.87	0.87	0.73	0.59	0.66	0.56	0.73
$sens(\hat{H}^C)$	0.98	0.98	0.93	0.95	0.98	0.99	0.98	0.99
$ppv(\hat{H})$	0.88	0.87	0.79	0.67	0.59	0.66	0.55	0.73
$ppv(\hat{H}^C)$	0.97	0.97	0.97	0.94	0.91	0.93	0.91	0.94
Size of H and H-complement								
$avg|\hat{H}|$	99	100	116	115	102	101	103	102
$min|\hat{H}|$	61	61	60	60	60	60	60	60
$max|\hat{H}|$	202	200	373	273	174	186	217	202
$avg|\hat{H}^C|$	404	403	386	404	433	428	435	421
$min|\hat{H}^C|$	298	300	127	227	326	314	283	298
$max|\hat{H}^C|$	500	500	500	500	500	500	500	500

Note: Number of simulations= 5000 .

Note: Probability approximation= 0.9246 .

```r
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 43.01035 0.7168392

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 86.0207 1.433678

# 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
```