```r
N <- 700
Nsims <- 20 * 1000


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")
```

```
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 <- 3

mindex <- "m4a"
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) &
```

```r
    # 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
```

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

if (!get.FS) est_names <- est_names[-c(1:3)]
```
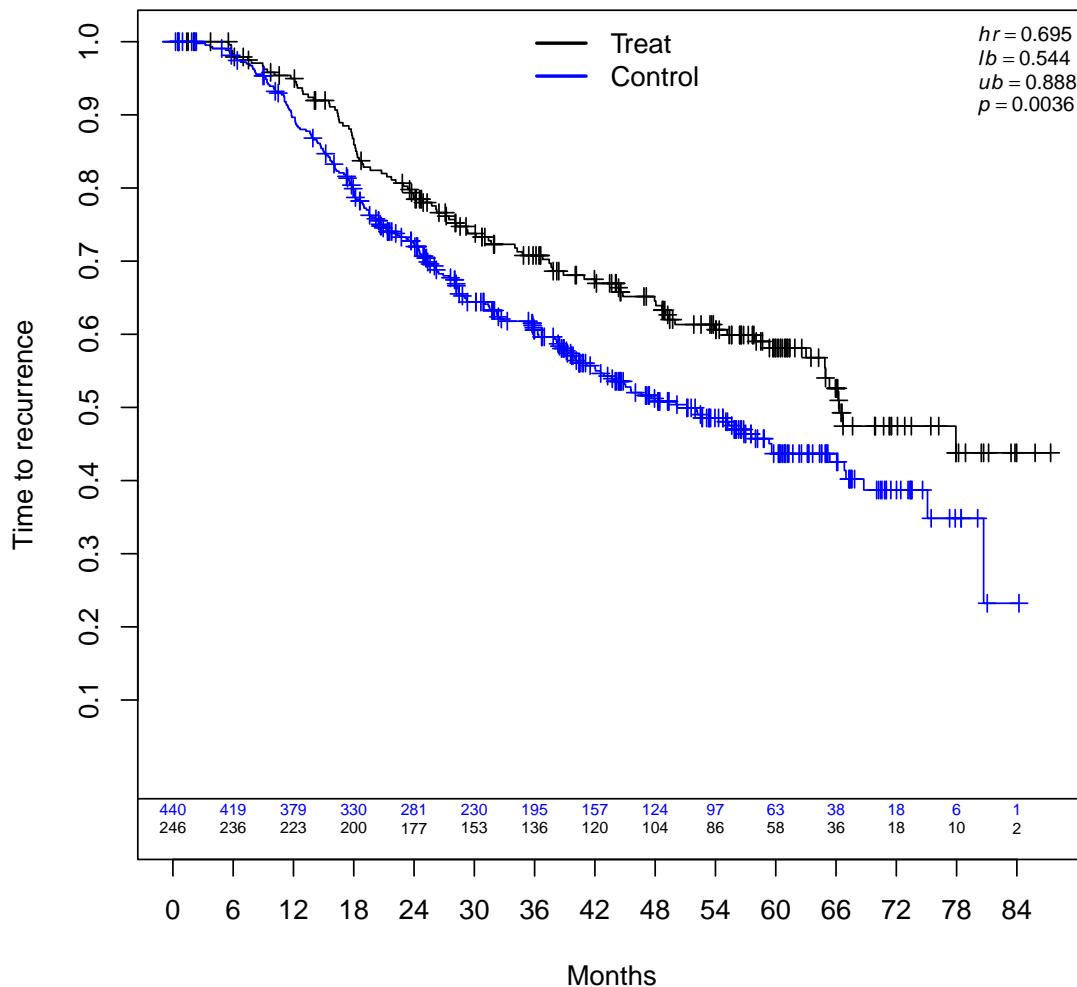
```r
mod.harm <- "null"
this.dgm <- get.dgm4.OC(mod.harm = mod.harm, N = N, k.treat = k.treat, model.index = model.index,
    sol_tol = 10^-8, hrH.target = hrH.target, cens.type = cens.type, out.loc = out.loc,
    file.index = file.index, details = TRUE, parms_torand = FALSE)
```

Figure showing Kaplan-Meier survival curves with legend: Treat (black), Control (blue).

Annotations: $hr = 0.695$, $lb = 0.544$, $ub = 0.888$, $p = 0.0036$

Y-axis: Time to recurrence
X-axis: Months

At-risk table (blue / black):

	0	6	12	18	24	30	36	42	48	54	60	66	72	78	84
	440	419	379	330	281	230	195	157	124	97	63	38	18	6	1
	246	236	223	200	177	153	136	120	104	86	58	36	18	10	2

```
## 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         3.76        62.00       3.58     1
## 2           3        -4.53       638.00       1.13     1
## 11          4        -5.52       312.00       1.65     2
## 4           7        -4.82       329.00       1.51     2
##    leaf.node control.mean control.size control.se depth
## 1          2         3.76        62.00       3.58     1
## GRF subgroup NOT found
## NO GRF cuts meeting delta(RMST): dmin.grf= 12
## # of continuous/categorical characteristics 4 6
## Continuous characteristics: size noise1 noise2 noise3
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## Default cuts included (1st 20)
## Categorical: z1 z2 z3 z4 z5 grade3
## # of candidate subgroup factors= 19
##  [1] "size <= 29.4"   "size <= 25"     "size <= 20"     "size <= 35"
##  [5] "noise1 <= 0"    "noise1 <= -0.6" "noise1 <= 0.7"  "noise2 <= 0"
##  [9] "noise2 <= -0.8" "noise2 <= 0.7"  "noise3 <= 0"    "noise3 <= -0.7"
## [13] "noise3 <= 0.6"  "z1"             "z2"             "z3"
## [17] "z4"             "z5"             "grade3"
## Number of factors evaluated= 19
## Confounders per grf screening q15 q17 q9 q4 q11 q18 q16 q6 q7 q14 q8 q5 q10 q19 q13 q1 q12 q3 q2
##           Factors Labels VI(grf)
## 15             z2    q15  0.1353
## 17             z4    q17  0.0950
## 9   noise2 <= -0.8   q9  0.0796
## 4       size <= 35   q4  0.0664
## 11     noise3 <= 0  q11  0.0537
## 18             z5   q18  0.0518
## 16             z3   q16  0.0513
## 6   noise1 <= -0.6   q6  0.0508
## 7    noise1 <= 0.7   q7  0.0491
## 14             z1   q14  0.0427
## 8      noise2 <= 0   q8  0.0416
## 5      noise1 <= 0   q5  0.0391
## 10   noise2 <= 0.7  q10  0.0388
## 19         grade3   q19  0.0384
## 13   noise3 <= 0.6  q13  0.0371
## 1     size <= 29.4   q1  0.0361
## 12  noise3 <= -0.7  q12  0.0344
## 3       size <= 20   q3  0.0311
## 2       size <= 25   q2  0.0278
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 741
## Approximately 5% of max_count met: minutes 0.001033333
## Approximately 10% of max_count met: minutes 0.001716667
## Approximately 20% of max_count met: minutes 0.003183333
## Approximately 33% of max_count met: minutes 0.0049
## Approximately 50% of max_count met: minutes 0.007033333
## Approximately 75% of max_count met: minutes 0.01045
## Approximately 90% of max_count met: minutes 0.01266667
## # of subgroups evaluated based on (up to) maxk-factor combinations 741
## % 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 52
```

```
## # of subgroups with sample size less than criteria 80
## # of subgroups meeting all criteria = 645
## # of subgroups fitted (Cox model estimable) = 645
## *Subgroup Searching Minutes=* 0.0138
## 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) q15.0 q15.1
## 1 82 31 19 1.49  0.72     0     0
## 2 86 41 24 1.31  0.70     0     0
## Consistency 0.7075
## Consistency 0.6075
## Subgroup Consistency Minutes= 0.02393333
## NO subgroup found (FS)
## Minutes forestsearch overall= 0.0541
## 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         3.76        62.00       3.58     1
## 2          3        -4.53       638.00       1.13     1
## 11         4        -5.52       312.00       1.65     2
## 4          7        -4.82       329.00       1.51     2
##   leaf.node control.mean control.size control.se depth
## 1         2         3.76        62.00       3.58     1
## GRF subgroup NOT found
## NO GRF cuts meeting delta(RMST): dmin.grf= 12
## # of continuous/categorical characteristics 4 6
## Continuous characteristics: size noise1 noise2 noise3
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## CV lambda = 0.02580757
## 10 x 1 sparse Matrix of class "dgCMatrix"
##                s0
## z1      0.07626064
## z2      .
## z3      .
## z4      0.48897013
## z5     -0.75595164
## size    .
## grade3  .
## noise1 -0.04100400
## noise2  .
## noise3  .
## Cox-LASSO selected: z1 z4 z5 noise1
## Cox-LASSO not selected: z2 z3 size grade3 noise2 noise3
## Default cuts included from Lasso:
## Categorical after Lasso: z1 z4 z5
## # of candidate subgroup factors= 16
##  [1] "size <= 29.4"   "size <= 25"     "size <= 20"     "size <= 35"
##  [5] "noise1 <= 0"    "noise1 <= -0.6" "noise1 <= 0.7"  "noise2 <= 0"
##  [9] "noise2 <= -0.8" "noise2 <= 0.7"  "noise3 <= 0"    "noise3 <= -0.7"
## [13] "noise3 <= 0.6"  "z1"             "z4"             "z5"
## Number of factors evaluated= 16
## Confounders per grf screening q15 q9 q4 q16 q11 q7 q6 q8 q13 q14 q5 q10 q12 q3 q1 q2
##           Factors Labels VI(grf)
```

```
## 15                  z4    q15  0.1167
## 9   noise2 <= -0.8    q9  0.1028
## 4      size <= 35     q4  0.0829
## 16                 z5  q16  0.0753
## 11    noise3 <= 0    q11  0.0734
## 7   noise1 <= 0.7     q7  0.0689
## 6   noise1 <= -0.6    q6  0.0649
## 8      noise2 <= 0    q8  0.0515
## 13  noise3 <= 0.6    q13  0.0512
## 14                 z1  q14  0.0504
## 5      noise1 <= 0    q5  0.0491
## 10  noise2 <= 0.7    q10  0.0469
## 12 noise3 <= -0.7    q12  0.0438
## 3      size <= 20     q3  0.0435
## 1    size <= 29.4     q1  0.0408
## 2      size <= 25     q2  0.0378
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 528
## Approximately 5% of max_count met: minutes 0.0006833333
## Approximately 10% of max_count met: minutes 0.001366667
## Approximately 20% of max_count met: minutes 0.0044
## Approximately 33% of max_count met: minutes 0.005416667
## Approximately 50% of max_count met: minutes 0.00725
## Approximately 75% of max_count met: minutes 0.009616667
## Approximately 90% of max_count met: minutes 0.01168333
## # of subgroups evaluated based on (up to) maxk-factor combinations 528
## % 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 37 44
## # of subgroups with sample size less than criteria 66
## # of subgroups meeting all criteria = 447
## # of subgroups fitted (Cox model estimable) = 447
## *Subgroup Searching Minutes=* 0.01265
## 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) q15.0 q15.1
## 1 82 31 19 1.49  0.72     1     0
## 2 86 41 24 1.31  0.70     0     0
## Consistency 0.7075
## Consistency 0.6075
## Subgroup Consistency Minutes= 0.0272
## NO subgroup found (FS)
## Minutes forestsearch overall= 0.05358333
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 48.41962 2
##    leaf.node control.mean control.size control.se depth
## 1          2        -5.31       675.00       1.08     1
## 11         4        -6.58       373.00       1.49     2
## 2          5         5.93        68.00       2.97     2
## 4          7        -6.95       227.00       1.85     2
##   leaf.node control.mean control.size control.se depth
## 2         5         5.93        68.00       2.97     2
## GRF subgroup NOT found
```

```
## NO GRF cuts meeting delta(RMST): dmin.grf= 12
## # of continuous/categorical characteristics 4 6
## Continuous characteristics: size noise1 noise2 noise3
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## Default cuts included (1st 20)
## Categorical: z1 z2 z3 z4 z5 grade3
## # of candidate subgroup factors= 19
##  [1] "size <= 28.5"   "size <= 25"    "size <= 20"    "size <= 35"
##  [5] "noise1 <= 0"    "noise1 <= -0.7" "noise1 <= 0.7"  "noise2 <= 0"
##  [9] "noise2 <= -0.7" "noise2 <= 0.7"  "noise3 <= 0.1"  "noise3 <= -0.5"
## [13] "noise3 <= 0.6"  "z1"             "z2"             "z3"
## [17] "z4"             "z5"             "grade3"
## Number of factors evaluated= 19
## Confounders per grf screening q18 q14 q7 q11 q5 q6 q16 q8 q17 q9 q12 q3 q2 q15 q13 q10 q19 q4 q1
##           Factors Labels VI(grf)
## 18             z5    q18  0.1482
## 14             z1    q14  0.1062
## 7   noise1 <= 0.7     q7  0.0844
## 11  noise3 <= 0.1    q11  0.0716
## 5     noise1 <= 0     q5  0.0699
## 6  noise1 <= -0.7     q6  0.0678
## 16             z3    q16  0.0475
## 8     noise2 <= 0     q8  0.0456
## 17             z4    q17  0.0419
## 9  noise2 <= -0.7     q9  0.0414
## 12 noise3 <= -0.5    q12  0.0375
## 3      size <= 20     q3  0.0357
## 2      size <= 25     q2  0.0346
## 15             z2    q15  0.0317
## 13  noise3 <= 0.6    q13  0.0305
## 10  noise2 <= 0.7    q10  0.0296
## 19         grade3    q19  0.0262
## 4      size <= 35     q4  0.0250
## 1    size <= 28.5     q1  0.0248
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 741
## Approximately 5% of max_count met: minutes 0.001383333
## Approximately 10% of max_count met: minutes 0.002333333
## Approximately 20% of max_count met: minutes 0.004916667
## Approximately 33% of max_count met: minutes 0.007883333
## Approximately 50% of max_count met: minutes 0.01211667
## Approximately 75% of max_count met: minutes 0.01816667
## Approximately 90% of max_count met: minutes 0.02141667
## # of subgroups evaluated based on (up to) maxk-factor combinations 741
## % 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 39 48
## # of subgroups with sample size less than criteria 81
## # of subgroups meeting all criteria = 644
## # of subgroups fitted (Cox model estimable) = 644
## *Subgroup Searching Minutes=* 0.02381667
## 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) q18.0 q18.1
## 1 159 62 35 1.40  0.84     0     1
## 2  69 41 23 1.36  0.73     0     0
## Consistency 0.805
## Consistency 0.65
## Subgroup Consistency Minutes= 0.0505
## NO subgroup found (FS)
## Minutes forestsearch overall= 0.1270167
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 48.41962 2
##    leaf.node control.mean control.size control.se depth
## 1          2        -5.31       675.00       1.08     1
## 11         4        -6.58       373.00       1.49     2
## 2          5         5.93        68.00       2.97     2
## 4          7        -6.95       227.00       1.85     2
##   leaf.node control.mean control.size control.se depth
## 2         5         5.93        68.00       2.97     2
## GRF subgroup NOT found
## NO GRF cuts meeting delta(RMST): dmin.grf= 12
## # of continuous/categorical characteristics 4 6
## Continuous characteristics: size noise1 noise2 noise3
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## CV lambda = 0.03552225
## 10 x 1 sparse Matrix of class "dgCMatrix"
##                 s0
## z1      0.06182051
## z2       .
## z3       .
## z4      0.36809804
## z5     -0.82786782
## size     .
## grade3   .
## noise1   .
## noise2   .
## noise3   .
## Cox-LASSO selected: z1 z4 z5
## Cox-LASSO not selected: z2 z3 size grade3 noise1 noise2 noise3
## Default cuts included from Lasso:
## Categorical after Lasso: z1 z4 z5
## # of candidate subgroup factors= 16
##  [1] "size <= 28.5"  "size <= 25"     "size <= 20"     "size <= 35"
##  [5] "noise1 <= 0"    "noise1 <= -0.7" "noise1 <= 0.7"  "noise2 <= 0"
##  [9] "noise2 <= -0.7" "noise2 <= 0.7"  "noise3 <= 0.1"  "noise3 <= -0.5"
## [13] "noise3 <= 0.6"  "z1"             "z4"             "z5"
## Number of factors evaluated= 16
## Confounders per grf screening q16 q14 q7 q6 q5 q11 q8 q15 q2 q9 q3 q12 q10 q13 q4 q1
##           Factors Labels VI(grf)
## 16             z5    q16  0.1550
## 14             z1    q14  0.1169
## 7   noise1 <= 0.7     q7  0.0911
## 6  noise1 <= -0.7     q6  0.0803
## 5     noise1 <= 0     q5  0.0796
## 11  noise3 <= 0.1    q11  0.0753
## 8     noise2 <= 0     q8  0.0572
## 15             z4    q15  0.0478
```

```
## 2       size <= 25      q2  0.0459
## 9  noise2 <= -0.7       q9  0.0451
## 3       size <= 20      q3  0.0402
## 12 noise3 <= -0.5      q12  0.0376
## 10  noise2 <= 0.7      q10  0.0357
## 13  noise3 <= 0.6      q13  0.0333
## 4       size <= 35      q4  0.0313
## 1     size <= 28.5      q1  0.0277
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 528
## Approximately 5% of max_count met: minutes 0.00075
## Approximately 10% of max_count met: minutes 0.00185
## Approximately 20% of max_count met: minutes 0.004083333
## Approximately 33% of max_count met: minutes 0.006433333
## Approximately 50% of max_count met: minutes 0.009116667
## Approximately 75% of max_count met: minutes 0.01358333
## Approximately 90% of max_count met: minutes 0.01658333
## # of subgroups evaluated based on (up to) maxk-factor combinations 528
## % 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 35 41
## # of subgroups with sample size less than criteria 67
## # of subgroups meeting all criteria = 446
## # of subgroups fitted (Cox model estimable) = 446
## *Subgroup Searching Minutes=* 0.01781667
## 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) q16.0
## 1 69 41 23 1.36  0.73      0
## Consistency 0.65
## Subgroup Consistency Minutes= 0.02436667
## NO subgroup found (FS)
## Minutes forestsearch overall= 0.08005
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 48.09494 2
##    leaf.node control.mean control.size control.se depth
## 2         3        -3.64       651.00       1.09     1
## 21        5        -4.92       354.00       1.48     2
## 3         6         4.19       157.00       2.14     2
## 4         7        -7.27       160.00       2.22     2
##   leaf.node control.mean control.size control.se depth
## 3         6         4.19       157.00       2.14     2
## GRF subgroup NOT found
## NO GRF cuts meeting delta(RMST): dmin.grf= 12
## # of continuous/categorical characteristics 4 6
## Continuous characteristics: size noise1 noise2 noise3
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## Default cuts included (1st 20)
## Categorical: z1 z2 z3 z4 z5 grade3
## # of candidate subgroup factors= 20
##  [1] "size <= 29.8"   "size <= 26.5"   "size <= 21"     "size <= 35"
##  [5] "noise1 <= -0.1" "noise1 <= -0.7" "noise1 <= 0.6"  "noise2 <= 0.1"
##  [9] "noise2 <= 0"    "noise2 <= -0.7" "noise2 <= 0.8"  "noise3 <= 0"
```

```
## [13] "noise3 <= -0.6" "noise3 <= 0.7"  "z1"            "z2"
## [17] "z3"             "z4"            "z5"            "grade3"
## Number of factors evaluated= 20
## Confounders per grf screening q11 q6 q10 q18 q14 q5 q17 q4 q3 q19 q20 q16 q12 q1 q7 q2 q15 q8 q9 q13
##           Factors Labels VI(grf)
## 11  noise2 <= 0.8    q11  0.0994
## 6  noise1 <= -0.7     q6  0.0763
## 10 noise2 <= -0.7    q10  0.0654
## 18             z4    q18  0.0612
## 14  noise3 <= 0.7    q14  0.0592
## 5  noise1 <= -0.1     q5  0.0538
## 17             z3    q17  0.0502
## 4      size <= 35     q4  0.0500
## 3      size <= 21     q3  0.0469
## 19             z5    q19  0.0467
## 20         grade3    q20  0.0450
## 16             z2    q16  0.0444
## 12    noise3 <= 0    q12  0.0439
## 1     size <= 29.8    q1  0.0422
## 7    noise1 <= 0.6     q7  0.0405
## 2    size <= 26.5     q2  0.0398
## 15             z1    q15  0.0381
## 8   noise2 <= 0.1     q8  0.0374
## 9     noise2 <= 0     q9  0.0338
## 13 noise3 <= -0.6    q13  0.0257
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 820
## Approximately 5% of max_count met: minutes 0.001883333
## Approximately 10% of max_count met: minutes 0.00295
## Approximately 20% of max_count met: minutes 0.0062
## Approximately 33% of max_count met: minutes 0.0105
## Approximately 50% of max_count met: minutes 0.01481667
## Approximately 75% of max_count met: minutes 0.02273333
## Approximately 90% of max_count met: minutes 0.02653333
## # of subgroups evaluated based on (up to) maxk-factor combinations 820
## % 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 45 65
## # of subgroups with sample size less than criteria 85
## # of subgroups meeting all criteria = 712
## # of subgroups fitted (Cox model estimable) = 712
## *Subgroup Searching Minutes=* 0.03126667
## Number of subgroups meeting HR threshold 3
## Subgroup candidate(s) found (FS)
## # of candidate subgroups (meeting HR criteria) =  3
## SGs (1st 10) meeting screening thresholds sorted by sg_focus= hr
##    n  E d1   HR L(HR) q11.0 q11.1 q6.0
## 1 82 46 22 1.43  0.80     0     0    0
## 2 73 40 20 1.37  0.74     0     0    0
## 3 87 51 28 1.26  0.73     0     0    0
## Consistency 0.75
## Consistency 0.66
## Subgroup Consistency Minutes= 0.0475
## NO subgroup found (FS)
## Minutes forestsearch overall= 0.1649333
```

```
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 48.09494 2
##    leaf.node control.mean control.size control.se depth
## 2          3       -3.64       651.00        1.09     1
## 21         5       -4.92       354.00        1.48     2
## 3          6        4.19       157.00        2.14     2
## 4          7       -7.27       160.00        2.22     2
##   leaf.node control.mean control.size control.se depth
## 3         6        4.19       157.00        2.14     2
## GRF subgroup NOT found
## NO GRF cuts meeting delta(RMST): dmin.grf= 12
## # of continuous/categorical characteristics 4 6
## Continuous characteristics: size noise1 noise2 noise3
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## CV lambda = 0.00340862
## 10 x 1 sparse Matrix of class "dgCMatrix"
##                  s0
## z1      0.211265280
## z2     -0.548294391
## z3      0.415741499
## z4      0.721068682
## z5     -0.926005067
## size   -0.002816545
## grade3 -0.071092243
## noise1  0.042627920
## noise2  0.070435098
## noise3 -0.006669633
## Cox-LASSO selected: z1 z2 z3 z4 z5 size grade3 noise1 noise2 noise3
## Cox-LASSO not selected:
## Default cuts included from Lasso:
## Categorical after Lasso: z1 z2 z3 z4 z5 grade3
## # of candidate subgroup factors= 20
##  [1] "size <= 29.8"  "size <= 26.5"  "size <= 21"    "size <= 35"
##  [5] "noise1 <= -0.1" "noise1 <= -0.7" "noise1 <= 0.6" "noise2 <= 0.1"
##  [9] "noise2 <= 0"    "noise2 <= -0.7" "noise2 <= 0.8" "noise3 <= 0"
## [13] "noise3 <= -0.6" "noise3 <= 0.7"  "z1"            "z2"
## [17] "z3"            "z4"            "z5"            "grade3"
## Number of factors evaluated= 20
## Confounders per grf screening q11 q6 q10 q18 q14 q5 q17 q4 q3 q19 q20 q16 q12 q1 q7 q2 q15 q8 q9 q13
##            Factors Labels VI(grf)
## 11  noise2 <= 0.8    q11  0.0994
## 6   noise1 <= -0.7    q6  0.0763
## 10 noise2 <= -0.7    q10  0.0654
## 18             z4    q18  0.0612
## 14  noise3 <= 0.7    q14  0.0592
## 5   noise1 <= -0.1    q5  0.0538
## 17             z3    q17  0.0502
## 4      size <= 35     q4  0.0500
## 3      size <= 21     q3  0.0469
## 19             z5    q19  0.0467
## 20         grade3    q20  0.0450
## 16             z2    q16  0.0444
## 12    noise3 <= 0    q12  0.0439
## 1     size <= 29.8    q1  0.0422
## 7    noise1 <= 0.6    q7  0.0405
```

```
## 2     size <= 26.5      q2  0.0398
## 15               z1     q15  0.0381
## 8    noise2 <= 0.1      q8  0.0374
## 9      noise2 <= 0      q9  0.0338
## 13 noise3 <= -0.6      q13  0.0257
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 820
## Approximately 5% of max_count met: minutes 0.001666667
## Approximately 10% of max_count met: minutes 0.002783333
## Approximately 20% of max_count met: minutes 0.007133333
## Approximately 33% of max_count met: minutes 0.01223333
## Approximately 50% of max_count met: minutes 0.02028333
## Approximately 75% of max_count met: minutes 0.02891667
## Approximately 90% of max_count met: minutes 0.03356667
## # of subgroups evaluated based on (up to) maxk-factor combinations 820
## % 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 45 65
## # of subgroups with sample size less than criteria 85
## # of subgroups meeting all criteria = 712
## # of subgroups fitted (Cox model estimable) = 712
## *Subgroup Searching Minutes=* 0.03708333
## Number of subgroups meeting HR threshold 3
## Subgroup candidate(s) found (FS)
## # of candidate subgroups (meeting HR criteria) =  3
## SGs (1st 10) meeting screening thresholds sorted by sg_focus= hr
##    n  E d1   HR L(HR) q11.0 q11.1 q6.0
## 1 82 46 22 1.43  0.80     0     0    0
## 2 73 40 20 1.37  0.74     0     0    0
## 3 87 51 28 1.26  0.73     0     0    0
## Consistency 0.75
## Consistency 0.66
## Subgroup Consistency Minutes= 0.04646667
## NO subgroup found (FS)
## Minutes forestsearch overall= 0.12795

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)
##   20000    20000   20000    20000    20000    20000   20000   20000

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=20000_m4a-Noise=3_N=700_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         700           1     0      0     700
```

```
## 2:       1          0          0         700          1     0     0     700
## 3:       1          0          0         700          1     1    69     631
## 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
##     <lgcl>    <num>      <num>      <num>   <int>   <int>      <num>
## 1:     NA 1.0000000          1          NA       0       0          0
## 2:     NA 1.0000000          1          NA       0       0          0
## 3:     NA 0.9014286          1           0      NA      NA         NA
## 4:     NA 1.0000000          1          NA      NA      NA         NA
## 5:     NA 1.0000000          1          NA       0       0          0
## 6:     NA 1.0000000          1          NA       0       0          0
##    found.al3 hr.H.true hr.Hc.true hr.H.hat hr.Hc.hat   b1.H  b2.H       b1.Hc
##       <num>    <lgcl>     <num>    <num>     <num> <lgcl> <lgcl>      <num>
## 1:        0        NA  0.6155054       NA 0.6155054     NA     NA  0.00000000
## 2:        0        NA  0.6155054       NA 0.6155054     NA     NA  0.00000000
## 3:       NA        NA  0.6155054 2.406676 0.5325424     NA     NA -0.08296298
## 4:       NA        NA  0.6155054       NA 0.6155054     NA     NA  0.00000000
## 5:        0        NA  0.6155054       NA 0.6155054     NA     NA  0.00000000
## 6:        0        NA  0.6155054       NA 0.6155054     NA     NA  0.00000000
##         b2.Hc    p.cens analysis   taumax    hr.itt    l.itt    u.itt
##         <num>    <num>  <char>    <num>    <num>    <num>    <num>
## 1: -0.08552157 0.4642857      FSl 81.69660 0.6155054 0.5014207 0.755547
## 2: -0.08552157 0.4642857     FSlg 49.01796 0.6155054 0.5014207 0.755547
## 3: -0.16848455 0.4642857      GRF 81.69660 0.6155054 0.5014207 0.755547
## 4: -0.08552157 0.4642857   GRF.60 49.01796 0.6155054 0.5014207 0.755547
## 5: -0.08552157 0.4642857   VT(24) 81.69660 0.6155054 0.5014207 0.755547
## 6: -0.08552157 0.4642857  VT#(24) 81.69660 0.6155054 0.5014207 0.755547
##    hr.adj.itt l.adj.itt u.adj.itt l.H.true u.H.true l.Hc.true u.Hc.true
##         <num>    <num>    <num>   <lgcl>   <lgcl>    <num>     <num>
## 1:  0.5906047  0.480604 0.7257825       NA       NA 0.5014207  0.755547
## 2:  0.5906047  0.480604 0.7257825       NA       NA 0.5014207  0.755547
## 3:  0.5906047  0.480604 0.7257825       NA       NA 0.5014207  0.755547
## 4:  0.5906047  0.480604 0.7257825       NA       NA 0.5014207  0.755547
## 5:  0.5906047  0.480604 0.7257825       NA       NA 0.5014207  0.755547
## 6:  0.5906047  0.480604 0.7257825       NA       NA 0.5014207  0.755547
##    l.H.hat  u.H.hat  l.Hc.hat  u.Hc.hat
##      <num>    <num>     <num>     <num>
## 1:      NA       NA 0.5014207 0.7555470
## 2:      NA       NA 0.5014207 0.7555470
## 3: 1.043941 5.548289 0.4297038 0.6599928
## 4:      NA       NA 0.5014207 0.7555470
## 5:      NA       NA 0.5014207 0.7555470
## 6:      NA       NA 0.5014207 0.7555470
## Subgroup HRs: H, H^c, Causal= NA 0.701027 0.701027
## Simulations= 20000
## Avg censoring= 0.4617969
## Min,Max,Avg tau.max= 71.81152 83.99787 81.70761
## P(H) approximation at causal(Hrc), n=60, approx= 0.701027 0.03396618
## P(H) approximation at plim(Hrc), n=60, approx= 0.7066595 0.03564707
## Minutes,hours 402.7611 6.712684
##                   FSl    FSlg     GRF   GRF.60   VT(24)  VT#(24)   VT(36)  VT#(36)
## any.H           0.250   0.230   0.610    0.280    0.040    0.020    0.060    0.030
## sensH             NaN     NaN     NaN      NaN      NaN      NaN      NaN      NaN
```

```
## sensHc           0.970    0.970    0.920    0.970    1.000    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)         89.000   89.000   93.000   81.000   79.000   76.000   80.000   78.000
## minH            61.000   61.000   60.000   60.000   60.000   60.000   60.000   60.000
## maxH           426.000  307.000  328.000  202.000  222.000  164.000  174.000  172.000
## Avg(#Hc)       678.000  680.000  643.000  678.000  697.000  698.000  695.000  698.000
## minHc          274.000  393.000  372.000  498.000  478.000  536.000  526.000  528.000
## 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])      1.887    1.886    1.694    1.575    1.294    1.429    1.365    1.600
## hat(Hc*)         0.759    0.761    0.727    0.722    0.736    0.742    0.737    0.745
## hat(hat[Hc])     0.672    0.673    0.643    0.655    0.690    0.688    0.686    0.681
## 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   0.664    0.664    0.664    0.664    0.664    0.664    0.664    0.664

missC <- tab_tests(res = res)

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

Table 1: Average classification rates: $avg|\hat{H}|$, $min|\hat{H}|$, and $max|\hat{H}|$, denote the average, minimum, and maximum of the number of subjects in the estimated subgroup $\hat{H}$ (analogously for $\hat{H}^c$). Note that under the null $sens(\hat{H})$ is undefined and $ppv(\hat{H}) = 0$.
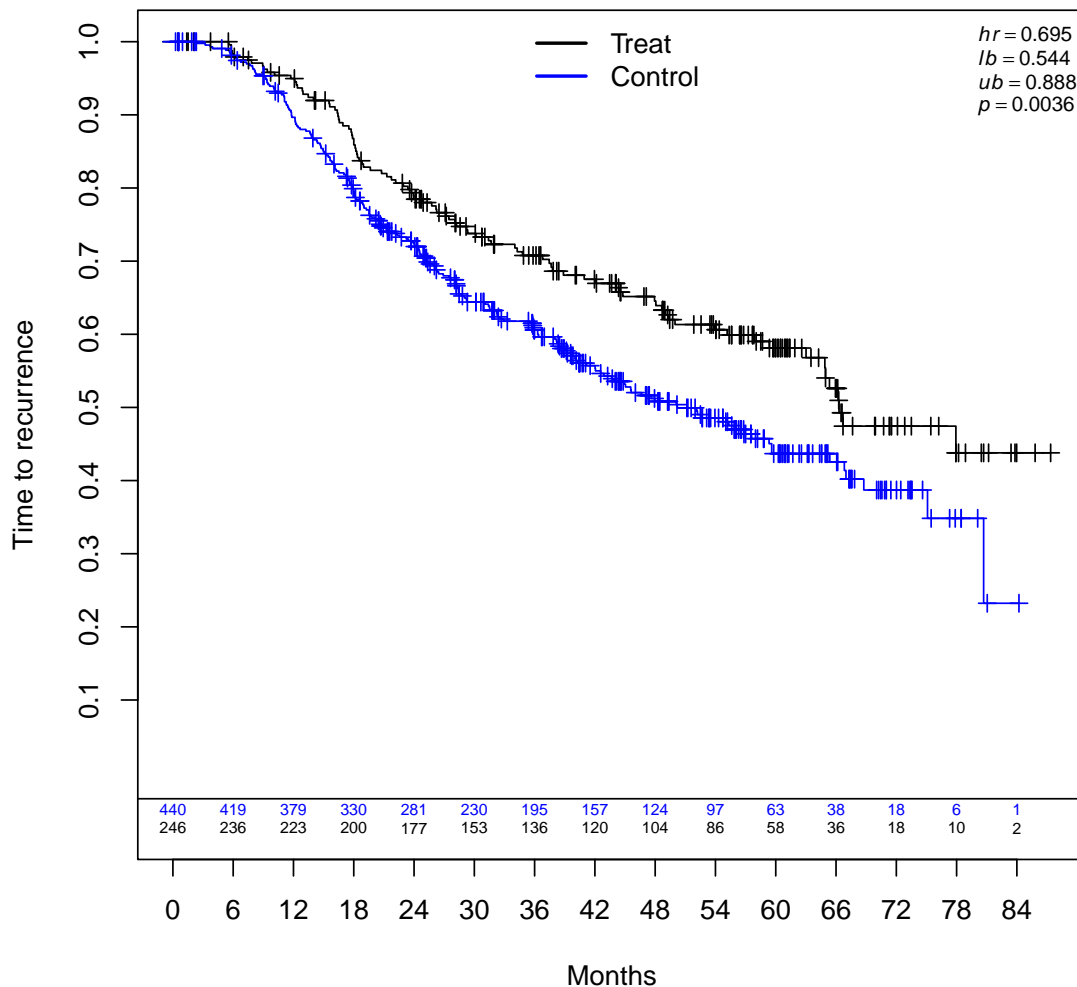
	FS_g	FS_{lg}	GRF	GRF_{60}	$VT(24)$	$VT^{\#}(24)$	$VT(36)$	$VT^{\#}(36)$
Finding H								
any(H)	0.25	0.23	0.61	0.28	0.04	0.02	0.06	0.03
$sens(\hat{H})$	.	.	.	.	.	.	.	.
$sens(\hat{H}^C)$	0.97	0.97	0.92	0.97	1	1	0.99	1
$ppv(\hat{H})$	0	0	0	0	0	0	0	0
$ppv(\hat{H}^C)$	1	1	1	1	1	1	1	1
Size of H and H-complement								
$avg|\hat{H}|$	89	89	93	81	79	76	80	78
$min|\hat{H}|$	61	61	60	60	60	60	60	60
$max|\hat{H}|$	426	307	328	202	222	164	174	172
$avg|\hat{H}^C|$	678	680	643	678	697	698	695	698
$min|\hat{H}^C|$	274	393	372	498	478	536	526	528
$max|\hat{H}^C|$	700	700	700	700	700	700	700	700

Note: Number of simulations= 20000 .

Note: Probability approximation= 0.0356 .

```
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        -4.35       509.00       1.23     1
## 2         3         2.53       191.00       2.28     1
## 3         4         2.24        95.00       2.92     2
## 4         5        -5.87       414.00       1.34     2
## 5         6         5.88       146.00       2.61     2
##   leaf.node control.mean control.size control.se depth
## 5         6         5.88       146.00       2.61     2
## GRF subgroup NOT found
## NO GRF cuts meeting delta(RMST): dmin.grf= 12
## # of continuous/categorical characteristics 4 6
## Continuous characteristics: size noise1 noise2 noise3
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## Default cuts included (1st 20)
## Categorical: z1 z2 z3 z4 z5 grade3
## # of candidate subgroup factors= 19
##  [1] "size <= 29.4"  "size <= 25"    "size <= 20"    "size <= 35"
##  [5] "noise1 <= 0"   "noise1 <= -0.6" "noise1 <= 0.7" "noise2 <= 0"
##  [9] "noise2 <= -0.8" "noise2 <= 0.7" "noise3 <= 0"   "noise3 <= -0.7"
## [13] "noise3 <= 0.6" "z1"            "z2"            "z3"
## [17] "z4"            "z5"            "grade3"
## Number of factors evaluated= 19
## Confounders per grf screening q15 q16 q14 q9 q17 q11 q18 q10 q4 q6 q5 q7 q8 q19 q12 q13 q1 q2 q3
##          Factors Labels VI(grf)
## 15            z2    q15  0.2299
## 16            z3    q16  0.1589
## 14            z1    q14  0.1469
## 9  noise2 <= -0.8   q9  0.0496
## 17            z4    q17  0.0361
## 11    noise3 <= 0   q11  0.0345
## 18            z5    q18  0.0344
## 10  noise2 <= 0.7   q10  0.0331
## 4       size <= 35   q4  0.0325
## 6  noise1 <= -0.6    q6  0.0304
## 5     noise1 <= 0    q5  0.0273
## 7   noise1 <= 0.7    q7  0.0261
## 8     noise2 <= 0    q8  0.0248
## 19         grade3   q19  0.0246
## 12 noise3 <= -0.7   q12  0.0244
## 13  noise3 <= 0.6   q13  0.0226
## 1     size <= 29.4   q1  0.0222
## 2       size <= 25   q2  0.0216
## 3       size <= 20   q3  0.0201
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 741
## Approximately 5% of max_count met: minutes 0.001483333
## Approximately 10% of max_count met: minutes 0.0026
## Approximately 20% of max_count met: minutes 0.004583333
## Approximately 33% of max_count met: minutes 0.00675
## Approximately 50% of max_count met: minutes 0.009666667
## Approximately 75% of max_count met: minutes 0.01353333
## Approximately 90% of max_count met: minutes 0.0157
## # of subgroups evaluated based on (up to) maxk-factor combinations 741
## % 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 51
## # of subgroups with sample size less than criteria 80
## # of subgroups meeting all criteria = 645
## # of subgroups fitted (Cox model estimable) = 645
## *Subgroup Searching Minutes=* 0.01708333
## 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) q15.0 q15.1 q16.0 q16.1 q14.0 q14.1 q9.0 q9.1 q17.0
## 1   95 78 44 1.74  1.11     0     0     0     1     0     1    0    0     0
## 2   79 47 26 1.54  0.87     0     0     0     1     0     0    0    1     0
## 3   82 31 19 1.50  0.73     0     0     0     0     0     0    0    1     1
## 4   74 39 18 1.49  0.79     0     0     0     1     0     0    0    0     0
## 5  122 95 52 1.48  0.98     0     1     0     0     0     1    0    0     0
## 6  100 80 41 1.47  0.94     0     0     0     0     0     1    0    0     0
## 7   77 55 30 1.41  0.83     0     0     0     0     0     1    0    0     0
## 8   86 41 24 1.37  0.73     0     0     0     0     0     0    0    1     0
## 9   93 53 30 1.34  0.78     0     1     0     0     0     0    0    1     0
## 10  90 64 32 1.31  0.80     0     0     0     0     0     1    0    0     0
##    q17.1
## 1      0
## 2      0
## 3      0
## 4      0
## 5      0
## 6      0
## 7      0
## 8      0
## 9      0
## 10     0
## Consistency 0.9675
## # of splits= 400
## Model, % Consistency Met= {z3} {z1} 0.9675
## Number of subgroups meeting consistency criteria=
##     Pcons     N      g     m     K    M.1    M.2
##      <num> <num> <char> <char> <num> <char> <char>
## 1: 0.9675    95     3     1     2   {z3}   {z1}
## [1] "{z3}" "{z1}"
## % consistency criteria met= 0.9675
## SG focus= hr
## Subgroup Consistency Minutes= 0.0157
## Subgroup found (FS)
## Minutes forestsearch overall= 0.04253333
## 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        -4.35       509.00       1.23     1
## 2         3         2.53       191.00       2.28     1
## 3         4         2.24        95.00       2.92     2
## 4         5        -5.87       414.00       1.34     2
## 5         6         5.88       146.00       2.61     2
##   leaf.node control.mean control.size control.se depth
## 5         6         5.88       146.00       2.61     2
## GRF subgroup NOT found
```

```
## NO GRF cuts meeting delta(RMST): dmin.grf= 12
## # of continuous/categorical characteristics 4 6
## Continuous characteristics: size noise1 noise2 noise3
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## CV lambda = 0.02852468
## 10 x 1 sparse Matrix of class "dgCMatrix"
##                 s0
## z1      0.29461175
## z2          .
## z3      0.02987702
## z4      0.42856317
## z5     -0.72819804
## size        .
## grade3      .
## noise1 -0.03160709
## noise2      .
## noise3      .
## Cox-LASSO selected: z1 z3 z4 z5 noise1
## Cox-LASSO not selected: z2 size grade3 noise2 noise3
## Default cuts included from Lasso:
## Categorical after Lasso: z1 z3 z4 z5
## # of candidate subgroup factors= 17
##  [1] "size <= 29.4"  "size <= 25"     "size <= 20"     "size <= 35"
##  [5] "noise1 <= 0"    "noise1 <= -0.6" "noise1 <= 0.7"  "noise2 <= 0"
##  [9] "noise2 <= -0.8" "noise2 <= 0.7"  "noise3 <= 0"    "noise3 <= -0.7"
## [13] "noise3 <= 0.6"  "z1"             "z3"             "z4"
## [17] "z5"
## Number of factors evaluated= 17
## Confounders per grf screening q15 q14 q9 q17 q10 q16 q4 q6 q11 q12 q5 q8 q3 q7 q13 q2 q1
##           Factors Labels VI(grf)
## 15              z3    q15  0.2991
## 14              z1    q14  0.1735
## 9   noise2 <= -0.8    q9  0.0614
## 17              z5    q17  0.0460
## 10   noise2 <= 0.7   q10  0.0426
## 16              z4    q16  0.0418
## 4       size <= 35    q4  0.0411
## 6   noise1 <= -0.6    q6  0.0362
## 11     noise3 <= 0   q11  0.0353
## 12 noise3 <= -0.7   q12  0.0331
## 5      noise1 <= 0    q5  0.0319
## 8      noise2 <= 0    q8  0.0298
## 3       size <= 20    q3  0.0293
## 7    noise1 <= 0.7    q7  0.0289
## 13   noise3 <= 0.6   q13  0.0251
## 2       size <= 25    q2  0.0226
## 1     size <= 29.4    q1  0.0223
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 595
## Approximately 5% of max_count met: minutes 0.0006666667
## Approximately 10% of max_count met: minutes 0.001466667
## Approximately 20% of max_count met: minutes 0.00395
## Approximately 33% of max_count met: minutes 0.006083333
## Approximately 50% of max_count met: minutes 0.0085
## Approximately 75% of max_count met: minutes 0.01193333
## Approximately 90% of max_count met: minutes 0.01418333
```

```
## # of subgroups evaluated based on (up to) maxk-factor combinations 595
## % 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 44
## # of subgroups with sample size less than criteria 67
## # of subgroups meeting all criteria = 513
## # of subgroups fitted (Cox model estimable) = 513
## *Subgroup Searching Minutes=* 0.01565
## Number of subgroups meeting HR threshold 10
## Subgroup candidate(s) found (FS)
## # of candidate subgroups (meeting HR criteria) =  10
## SGs (1st 10) meeting screening thresholds sorted by sg_focus= hr
##       n  E d1   HR L(HR) q15.0 q15.1 q14.0 q14.1 q9.0 q9.1 q17.0 q17.1 q10.0
## 1    95 78 44 1.74  1.11     0     1     0     1    0    0     0     0     0
## 2    79 47 26 1.54  0.87     0     1     0     0    0    1     0     0     0
## 3    82 31 19 1.50  0.73     0     0     0     0    0    1     0     0     0
## 4    74 39 18 1.49  0.79     0     1     0     0    0    0     0     0     0
## 5   100 80 41 1.47  0.94     0     0     0     1    0    0     1     0     0
## 6    77 55 30 1.41  0.83     0     0     0     1    0    0     0     0     0
## 7    86 41 24 1.37  0.73     0     0     0     0    0    1     0     0     0
## 8    90 64 32 1.31  0.80     0     0     0     1    0    0     0     0     0
## 9    67 39 21 1.30  0.69     0     1     0     0    0    0     0     0     0
## 10  146 84 48 1.27  0.83     0     1     0     0    0    0     0     0     0
##    q10.1
## 1      0
## 2      0
## 3      0
## 4      0
## 5      0
## 6      0
## 7      0
## 8      0
## 9      0
## 10     0
## Consistency 0.9675
## # of splits= 400
## Model, % Consistency Met= {z3} {z1} 0.9675
## Number of subgroups meeting consistency criteria=
##      Pcons     N     g     m     K   M.1   M.2
##      <num> <num> <char> <char> <num> <char> <char>
## 1: 0.9675    95     1     1     2   {z3}   {z1}
## [1] "{z3}" "{z1}"
## % consistency criteria met= 0.9675
## SG focus= hr
## Subgroup Consistency Minutes= 0.01843333
## Subgroup found (FS)
## Minutes forestsearch overall= 0.0538
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 49.18697 2
##    leaf.node control.mean control.size control.se depth
## 1          2        -3.93       675.00       1.13     1
## 11         4        -5.16       372.00       1.55     2
## 3          6        -8.64       164.00       2.03     2
## 4          7         3.14       118.00       2.95     2
```

```
##   leaf.node control.mean control.size control.se depth
## 4        7         3.14       118.00         2.95     2
## GRF subgroup NOT found
## NO GRF cuts meeting delta(RMST): dmin.grf= 12
## # of continuous/categorical characteristics 4 6
## Continuous characteristics: size noise1 noise2 noise3
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## Default cuts included (1st 20)
## Categorical: z1 z2 z3 z4 z5 grade3
## # of candidate subgroup factors= 19
##  [1] "size <= 28.5"  "size <= 25"     "size <= 20"     "size <= 35"
##  [5] "noise1 <= 0"    "noise1 <= -0.7" "noise1 <= 0.7"  "noise2 <= 0"
##  [9] "noise2 <= -0.7" "noise2 <= 0.7"  "noise3 <= 0.1"  "noise3 <= -0.5"
## [13] "noise3 <= 0.6"  "z1"             "z2"             "z3"
## [17] "z4"             "z5"             "grade3"
## Number of factors evaluated= 19
## Confounders per grf screening q16 q18 q7 q17 q15 q14 q8 q5 q6 q11 q9 q19 q12 q3 q4 q2 q13 q10 q1
##          Factors Labels VI(grf)
## 16            z3    q16  0.1798
## 18            z5    q18  0.1156
## 7   noise1 <= 0.7    q7  0.0784
## 17            z4    q17  0.0705
## 15            z2    q15  0.0684
## 14            z1    q14  0.0542
## 8    noise2 <= 0     q8  0.0503
## 5    noise1 <= 0     q5  0.0501
## 6  noise1 <= -0.7    q6  0.0481
## 11  noise3 <= 0.1   q11  0.0452
## 9  noise2 <= -0.7    q9  0.0331
## 19        grade3   q19  0.0302
## 12 noise3 <= -0.5   q12  0.0286
## 3      size <= 20    q3  0.0281
## 4      size <= 35    q4  0.0267
## 2      size <= 25    q2  0.0257
## 13  noise3 <= 0.6   q13  0.0246
## 10  noise2 <= 0.7   q10  0.0245
## 1    size <= 28.5    q1  0.0177
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 741
## Approximately 5% of max_count met: minutes 0.0012
## Approximately 10% of max_count met: minutes 0.0038
## Approximately 20% of max_count met: minutes 0.0068
## Approximately 33% of max_count met: minutes 0.01068333
## Approximately 50% of max_count met: minutes 0.01581667
## Approximately 75% of max_count met: minutes 0.02376667
## Approximately 90% of max_count met: minutes 0.02825
## # of subgroups evaluated based on (up to) maxk-factor combinations 741
## % 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 39 47
## # of subgroups with sample size less than criteria 81
## # of subgroups meeting all criteria = 644
## # of subgroups fitted (Cox model estimable) = 644
## *Subgroup Searching Minutes=* 0.03165
## Number of subgroups meeting HR threshold 15
```

```
## Subgroup candidate(s) found (FS)
## # of candidate subgroups (meeting HR criteria) =  15
## SGs (1st 10) meeting screening thresholds sorted by sg_focus= hr
##      n   E d1   HR L(HR) q16.0 q16.1 q18.0 q18.1 q7.0 q7.1 q17.0 q17.1 q15.0
## 1   159  66 40 1.75  1.07     0     1     0     1    0    0     0     0     0
## 2   104  83 47 1.65  1.06     0     1     0     0    0    0     0     0     0
## 3   148  87 51 1.58  1.03     0     1     0     0    0    0     0     0     0
## 4    77  46 30 1.47  0.80     0     1     0     0    0    0     0     0     0
## 5   139  80 50 1.45  0.92     0     1     0     0    0    0     0     0     0
## 6    79  43 23 1.43  0.79     0     1     0     0    0    0     0     0     0
## 7    76  42 26 1.41  0.75     0     1     0     0    0    0     0     0     0
## 8    69  40 22 1.38  0.74     0     0     0     0    0    0     0     0     0
## 9   231 132 78 1.38  0.97     0     1     0     0    0    0     0     0     0
## 10 166  97 59 1.35  0.90     0     1     0     0    0    0     0     0     0
##    q15.1
## 1      0
## 2      0
## 3      0
## 4      0
## 5      0
## 6      0
## 7      0
## 8      0
## 9      0
## 10     0
## Consistency 0.98
## # of splits= 400
## Model, % Consistency Met= {z3} {z5} 0.98
## Number of subgroups meeting consistency criteria=
##    Pcons     N     g     m     K    M.1     M.2
##    <num> <num> <char> <char> <num> <char> <char>
## 1:  0.98   159     1     1     2   {z3}    {z5}
## [1] "{z3}" "{z5}"
## % consistency criteria met= 0.98
## SG focus= hr
## Subgroup Consistency Minutes= 0.03065
## Subgroup found (FS)
## Minutes forestsearch overall= 0.1484333
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 49.18697 2
##    leaf.node control.mean control.size control.se depth
## 1          2        -3.93       675.00       1.13     1
## 11         4        -5.16       372.00       1.55     2
## 3          6        -8.64       164.00       2.03     2
## 4          7         3.14       118.00       2.95     2
##   leaf.node control.mean control.size control.se depth
## 4         7         3.14       118.00       2.95     2
## GRF subgroup NOT found
## NO GRF cuts meeting delta(RMST): dmin.grf= 12
## # of continuous/categorical characteristics 4 6
## Continuous characteristics: size noise1 noise2 noise3
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## CV lambda = 0.03646668
## 10 x 1 sparse Matrix of class "dgCMatrix"
##                 s0
```

```
## z1        0.3056654
## z2         .
## z3         .
## z4        0.3544789
## z5       -0.8085053
## size      .
## grade3    .
## noise1    .
## noise2    .
## noise3    .
## Cox-LASSO selected: z1 z4 z5
## Cox-LASSO not selected: z2 z3 size grade3 noise1 noise2 noise3
## Default cuts included from Lasso:
## Categorical after Lasso: z1 z4 z5
## # of candidate subgroup factors= 16
##  [1] "size <= 28.5"   "size <= 25"     "size <= 20"     "size <= 35"
##  [5] "noise1 <= 0"    "noise1 <= -0.7" "noise1 <= 0.7"  "noise2 <= 0"
##  [9] "noise2 <= -0.7" "noise2 <= 0.7"  "noise3 <= 0.1"  "noise3 <= -0.5"
## [13] "noise3 <= 0.6"  "z1"             "z4"             "z5"
## Number of factors evaluated= 16
## Confounders per grf screening q16 q15 q7 q14 q11 q8 q5 q6 q3 q2 q9 q4 q10 q12 q13 q1
##            Factors Labels VI(grf)
## 16              z5    q16  0.1480
## 15              z4    q15  0.0972
## 7    noise1 <= 0.7    q7  0.0959
## 14              z1    q14  0.0839
## 11   noise3 <= 0.1   q11  0.0698
## 8      noise2 <= 0    q8  0.0697
## 5      noise1 <= 0    q5  0.0695
## 6   noise1 <= -0.7    q6  0.0617
## 3       size <= 20    q3  0.0456
## 2       size <= 25    q2  0.0418
## 9   noise2 <= -0.7    q9  0.0398
## 4       size <= 35    q4  0.0383
## 10   noise2 <= 0.7   q10  0.0381
## 12  noise3 <= -0.5   q12  0.0363
## 13   noise3 <= 0.6   q13  0.0339
## 1       size <= 28.5   q1  0.0306
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 528
## Approximately 5% of max_count met: minutes 0.00115
## Approximately 10% of max_count met: minutes 0.001883333
## Approximately 20% of max_count met: minutes 0.004066667
## Approximately 33% of max_count met: minutes 0.006066667
## Approximately 50% of max_count met: minutes 0.009183333
## Approximately 75% of max_count met: minutes 0.01428333
## Approximately 90% of max_count met: minutes 0.01635
## # of subgroups evaluated based on (up to) maxk-factor combinations 528
## % 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 35 41
## # of subgroups with sample size less than criteria 67
## # of subgroups meeting all criteria = 446
## # of subgroups fitted (Cox model estimable) = 446
## *Subgroup Searching Minutes=* 0.01888333
```

```
## 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) q16.0
## 1 69 40 22 1.38  0.74     0
## Consistency 0.675
## Subgroup Consistency Minutes= 0.0257
## NO subgroup found (FS)
## Minutes forestsearch overall= 0.1060833
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 49.86585 2
##   leaf.node control.mean control.size control.se depth
## 1         2        -3.61       523.00       1.23     1
## 2         3         2.60       177.00       2.47     1
## 3         4        -4.53       461.00       1.30     2
## 4         5         3.17        62.00       3.64     2
## 5         6         8.00       115.00       3.06     2
## 6         7        -7.43        62.00       3.88     2
##   leaf.node control.mean control.size control.se depth
## 5         6         8.00       115.00       3.06     2
## GRF subgroup found
## All splits
## [1] "z1 <= 0"        "noise3 <= 1.23" "noise1 <= 0.34"
## Terminating node at max.diff (sg.harm.id)
## [1] "noise1 <= 0.340726504862128"
## # of continuous/categorical characteristics 4 6
## Continuous characteristics: size noise1 noise2 noise3
## 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 noise3 <= 1.23 noise1 <= 0.34
## Initial GRF cuts included z1 <= 0 noise3 <= 1.23 noise1 <= 0.34
## # of candidate subgroup factors= 22
##  [1] "noise3 <= 1.23" "noise1 <= 0.34" "size <= 29.8"   "size <= 26.5"
##  [5] "size <= 21"     "size <= 35"     "noise1 <= -0.1" "noise1 <= -0.7"
##  [9] "noise1 <= 0.6"  "noise2 <= 0.1"  "noise2 <= 0"    "noise2 <= -0.7"
## [13] "noise2 <= 0.8"  "noise3 <= 0"    "noise3 <= -0.6" "noise3 <= 0.7"
## [17] "z1"             "z2"             "z3"             "z4"
## [21] "z5"             "grade3"
## Number of factors evaluated= 22
## Confounders per grf screening q17 q20 q8 q22 q13 q19 q1 q16 q5 q7 q21 q3 q6 q9 q18 q14 q15 q10 q4 q1
##           Factors Labels VI(grf)
## 17             z1    q17  0.1455
## 20             z4    q20  0.0814
## 8  noise1 <= -0.7    q8  0.0725
## 22         grade3   q22  0.0557
## 13  noise2 <= 0.8   q13  0.0554
## 19             z3    q19  0.0545
## 1  noise3 <= 1.23    q1  0.0543
## 16  noise3 <= 0.7   q16  0.0443
## 5      size <= 21    q5  0.0413
## 7  noise1 <= -0.1    q7  0.0391
## 21             z5    q21  0.0362
## 3    size <= 29.8    q3  0.0340
```

```
## 6        size <= 35      q6   0.0324
## 9    noise1 <= 0.6      q9   0.0317
## 18                z2    q18  0.0304
## 14     noise3 <= 0     q14  0.0297
## 15 noise3 <= -0.6      q15  0.0285
## 10  noise2 <= 0.1      q10  0.0281
## 4     size <= 26.5      q4   0.0276
## 11     noise2 <= 0     q11  0.0267
## 12 noise2 <= -0.7      q12  0.0256
## 2   noise1 <= 0.34      q2   0.0251
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 990
## Approximately 5% of max_count met: minutes 0.003766667
## Approximately 10% of max_count met: minutes 0.00565
## Approximately 20% of max_count met: minutes 0.009833333
## Approximately 33% of max_count met: minutes 0.01368333
## Approximately 50% of max_count met: minutes 0.01875
## Approximately 75% of max_count met: minutes 0.02531667
## Approximately 90% of max_count met: minutes 0.02966667
## # of subgroups evaluated based on (up to) maxk-factor combinations 990
## % 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 60 79
## # of subgroups with sample size less than criteria 127
## # of subgroups meeting all criteria = 834
## # of subgroups fitted (Cox model estimable) = 834
## *Subgroup Searching Minutes=* 0.03358333
## Number of subgroups meeting HR threshold 22
## Subgroup candidate(s) found (FS)
## # of candidate subgroups (meeting HR criteria) =  22
## SGs (1st 10) meeting screening thresholds sorted by sg_focus= hr
##       n  E d1   HR L(HR) q17.0 q17.1 q20.0 q20.1 q8.0 q8.1 q22.0 q22.1 q13.0
## 1    82 55 29 1.99  1.16     0     1     0     0    0    0     0     0     0
## 2    82 49 25 1.79  1.02     0     0     0     0    0    0     0     1     0
## 3    86 70 37 1.70  1.06     0     1     0     0    0    0     0     0     0
## 4   113 74 39 1.70  1.07     0     1     0     0    0    0     0     0     0
## 5    75 49 23 1.59  0.90     0     1     0     0    0    0     0     0     0
## 6    82 47 24 1.55  0.88     0     0     0     0    0    1     0     0     0
## 7    73 41 21 1.53  0.83     0     0     0     0    0    1     0     0     0
## 8    69 44 23 1.50  0.82     0     0     0     0    0    0     0     1     0
## 9   107 64 30 1.49  0.91     0     0     0     0    0    0     0     1     0
## 10   86 55 30 1.47  0.86     0     0     0     1    0    1     0     0     0
##    q13.1
## 1      0
## 2      0
## 3      0
## 4      0
## 5      0
## 6      0
## 7      0
## 8      0
## 9      0
## 10     0
## Consistency 0.995
## # of splits= 400
```

```
## Model, % Consistency Met= {z1} {noise1 <= -0.1} 0.995
## Number of subgroups meeting consistency criteria=
##     Pcons     N     g     m     K     M.1                 M.2
##     <num> <num> <char> <char> <num> <char>            <char>
## 1: 0.995    82     3     1     2   {z1} {noise1 <= -0.1}
## [1] "{z1}"            "{noise1 <= -0.1}"
## % consistency criteria met= 0.995
## SG focus= hr
## Subgroup Consistency Minutes= 0.02741667
## Subgroup found (FS)
## Minutes forestsearch overall= 0.1466
## FS: GRF stage for cut selection with dmin,tau= 12 0.6
## tau, maxdepth= 49.86585 2
##   leaf.node control.mean control.size control.se depth
## 1         2        -3.61       523.00       1.23     1
## 2         3         2.60       177.00       2.47     1
## 3         4        -4.53       461.00       1.30     2
## 4         5         3.17        62.00       3.64     2
## 5         6         8.00       115.00       3.06     2
## 6         7        -7.43        62.00       3.88     2
##   leaf.node control.mean control.size control.se depth
## 5         6         8.00       115.00       3.06     2
## GRF subgroup found
## All splits
## [1] "z1 <= 0"       "noise3 <= 1.23" "noise1 <= 0.34"
## Terminating node at max.diff (sg.harm.id)
## [1] "noise1 <= 0.340726504862128"
## # of continuous/categorical characteristics 4 6
## Continuous characteristics: size noise1 noise2 noise3
## Categorical characteristics: z1 z2 z3 z4 z5 grade3
## CV lambda = 0.005854487
## 10 x 1 sparse Matrix of class "dgCMatrix"
##                 s0
## z1      0.431463308
## z2     -0.503920433
## z3      0.525553842
## z4      0.701493120
## z5     -0.881297175
## size   -0.003727524
## grade3 -0.010964732
## noise1  0.045387860
## noise2  0.044998344
## noise3   .
## Cox-LASSO selected: z1 z2 z3 z4 z5 size grade3 noise1 noise2
## Cox-LASSO not selected: noise3
## Default cuts included from Lasso:
## Categorical after Lasso: z1 z2 z3 z4 z5 grade3
## Factors per GRF: z1 <= 0 noise3 <= 1.23 noise1 <= 0.34
## Initial GRF cuts included z1 <= 0 noise3 <= 1.23 noise1 <= 0.34
## Factors included per GRF (not in lasso) noise3 <= 1.23 noise1 <= 0.34
## # of candidate subgroup factors= 22
##  [1] "noise3 <= 1.23" "noise1 <= 0.34" "size <= 29.8"   "size <= 26.5"
##  [5] "size <= 21"     "size <= 35"     "noise1 <= -0.1" "noise1 <= -0.7"
##  [9] "noise1 <= 0.6"  "noise2 <= 0.1"  "noise2 <= 0"    "noise2 <= -0.7"
## [13] "noise2 <= 0.8"  "noise3 <= 0"    "noise3 <= -0.6" "noise3 <= 0.7"
```

```
## [17] "z1"            "z2"              "z3"             "z4"
## [21] "z5"            "grade3"
## Number of factors evaluated= 22
## Confounders per grf screening q17 q20 q8 q22 q13 q19 q1 q16 q5 q7 q21 q3 q6 q9 q18 q14 q15 q10 q4 q1
##         Factors Labels VI(grf)
## 17             z1    q17  0.1455
## 20             z4    q20  0.0814
## 8  noise1 <= -0.7    q8  0.0725
## 22         grade3   q22  0.0557
## 13  noise2 <= 0.8   q13  0.0554
## 19             z3    q19  0.0545
## 1  noise3 <= 1.23    q1  0.0543
## 16  noise3 <= 0.7   q16  0.0443
## 5      size <= 21    q5  0.0413
## 7  noise1 <= -0.1    q7  0.0391
## 21             z5    q21  0.0362
## 3    size <= 29.8    q3  0.0340
## 6      size <= 35    q6  0.0324
## 9   noise1 <= 0.6    q9  0.0317
## 18             z2    q18  0.0304
## 14    noise3 <= 0   q14  0.0297
## 15 noise3 <= -0.6   q15  0.0285
## 10  noise2 <= 0.1   q10  0.0281
## 4    size <= 26.5    q4  0.0276
## 11    noise2 <= 0   q11  0.0267
## 12 noise2 <= -0.7   q12  0.0256
## 2  noise1 <= 0.34    q2  0.0251
## Number of possible configurations (<= maxk): maxk, # <= maxk 2 990
## Approximately 5% of max_count met: minutes 0.001533333
## Approximately 10% of max_count met: minutes 0.003283333
## Approximately 20% of max_count met: minutes 0.006566667
## Approximately 33% of max_count met: minutes 0.01165
## Approximately 50% of max_count met: minutes 0.0172
## Approximately 75% of max_count met: minutes 0.02381667
## Approximately 90% of max_count met: minutes 0.029
## # of subgroups evaluated based on (up to) maxk-factor combinations 990
## % 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 60 79
## # of subgroups with sample size less than criteria 127
## # of subgroups meeting all criteria = 834
## # of subgroups fitted (Cox model estimable) = 834
## *Subgroup Searching Minutes=* 0.0319
## Number of subgroups meeting HR threshold 22
## Subgroup candidate(s) found (FS)
## # of candidate subgroups (meeting HR criteria) =  22
## SGs (1st 10) meeting screening thresholds sorted by sg_focus= hr
##      n  E d1    HR L(HR) q17.0 q17.1 q20.0 q20.1 q8.0 q8.1 q22.0 q22.1 q13.0
## 1    82 55 29 1.99  1.16     0     1     0     0    0    0     0     0     0
## 2    82 49 25 1.79  1.02     0     0     0     0    0    0     0     1     0
## 3    86 70 37 1.70  1.06     0     1     0     0    0    0     0     0     0
## 4   113 74 39 1.70  1.07     0     1     0     0    0    0     0     0     0
## 5    75 49 23 1.59  0.90     0     1     0     0    0    0     0     0     0
## 6    82 47 24 1.55  0.88     0     0     0     0    0    1     0     0     0
```

```
## 7    73 41 21 1.53  0.83    0     0     0     0     0    1      0      0      0
## 8    69 44 23 1.50  0.82    0     0     0     0     0    0      0      1      0
## 9   107 64 30 1.49  0.91    0     0     0     0     0    0      0      1      0
## 10   86 55 30 1.47  0.86    0     0     0     1     0    1      0      0      0
##     q13.1
## 1      0
## 2      0
## 3      0
## 4      0
## 5      0
## 6      0
## 7      0
## 8      0
## 9      0
## 10     0
## Consistency 0.995
## # of splits= 400
## Model, % Consistency Met= {z1} {noise1 <= -0.1} 0.995
## Number of subgroups meeting consistency criteria=
##     Pcons     N     g     m     K   M.1                M.2
##     <num> <num> <char> <char> <num> <char>          <char>
## 1: 0.995    82     3     1     2   {z1} {noise1 <= -0.1}
## [1] "{z1}"          "{noise1 <= -0.1}"
## % consistency criteria met= 0.995
## SG focus= hr
## Subgroup Consistency Minutes= 0.02255
## Subgroup found (FS)
## Minutes forestsearch overall= 0.09725

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


print(table(res$analysis))

##
##      FSl     FSlg      GRF   GRF.60   VT(24)   VT(36)  VT#(24)  VT#(36)
##    20000    20000    20000    20000    20000    20000    20000    20000

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 = "FSl")

## [1] "results/oc_sims=20000_m4a-Noise=3_N=700_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        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= 20000
## Avg censoring= 0.4502911
## Min,Max,Avg tau.max= 69.89395 83.99237 81.61975
## P(H) approximation at causal(H), n(sg)=60, approx= 2.000007 60 0.8281929
## P(H) approximation at causal(H), Avg(n(sg)), approx= 2.000007 89 0.8997886
## P(H) approximation at plim(H), Avg(n(sg)), approx= 2.114003 89 0.9262042
## Minutes,hours 392.5442 6.542403
##                  FSl    FSlg     GRF  GRF.60  VT(24) VT#(24)  VT(36) VT#(36)
## any.H          0.930   0.870   0.940   0.700   0.440   0.490   0.420   0.560
## sensH          0.790   0.700   0.660   0.520   0.370   0.440   0.340   0.510
## sensHc         0.980   0.970   0.930   0.960   0.990   0.990   0.990   0.990
## ppH            0.790   0.680   0.590   0.470   0.360   0.430   0.330   0.500
## ppHc           0.970   0.960   0.950   0.940   0.920   0.930   0.920   0.940
## Avg(#H)       88.000  91.000 106.000 101.000  92.000  92.000  92.000  92.000
```

```
## minH              61.000  61.000  60.000  60.000  60.000  60.000  60.000  60.000
## maxH             273.000 290.000 420.000 327.000 251.000 190.000 210.000 238.000
## Avg(#Hc)         618.000 621.000 601.000 629.000 660.000 655.000 661.000 649.000
## minHc            427.000 410.000 280.000 373.000 449.000 510.000 490.000 462.000
## maxHc            700.000 700.000 700.000 700.000 700.000 700.000 700.000 700.000
## hat(H*)            2.172   2.195   2.146   2.228   2.378   2.373   2.409   2.369
## hat(hat[H])        2.248   2.192   1.989   1.967   2.181   2.249   2.222   2.294
## hat(Hc*)           0.654   0.655   0.653   0.648   0.653   0.652   0.653   0.651
## hat(hat[Hc])       0.654   0.655   0.643   0.648   0.661   0.656   0.661   0.652
## hat(H*)all         2.114   2.114   2.114   2.114   2.114   2.114   2.114   2.114
## hat(Hc*)all        0.652   0.652   0.652   0.652   0.652   0.652   0.652   0.652
## hat(ITT)all        0.755   0.755   0.755   0.755   0.755   0.755   0.755   0.755
## 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)	0.93	0.87	0.94	0.7	0.44	0.49	0.42	0.56
$sens(\hat{H})$	0.79	0.7	0.66	0.52	0.37	0.44	0.34	0.51
$sens(\hat{H}^C)$	0.98	0.97	0.93	0.96	0.99	0.99	0.99	0.99
$ppv(\hat{H})$	0.79	0.68	0.59	0.47	0.36	0.43	0.33	0.5
$ppv(\hat{H}^C)$	0.97	0.96	0.95	0.94	0.92	0.93	0.92	0.94
Size of H and H-complement								
$avg|\hat{H}|$	88	91	106	101	92	92	92	92
$min|\hat{H}|$	61	61	60	60	60	60	60	60
$max|\hat{H}|$	273	290	420	327	251	190	210	238
$avg|\hat{H}^C|$	618	621	601	629	660	655	661	649
$min|\hat{H}^C|$	427	410	280	373	449	510	490	462
$max|\hat{H}^C|$	700	700	700	700	700	700	700	700

Note: Number of simulations= 20000 .
Note: Probability approximation= 0.8998 .

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

## Minutes and hours to finish 795.5621 13.25937

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 397.781 6.629684

# 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: Apple M1 Ultra

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= 20 69
```