mc5296

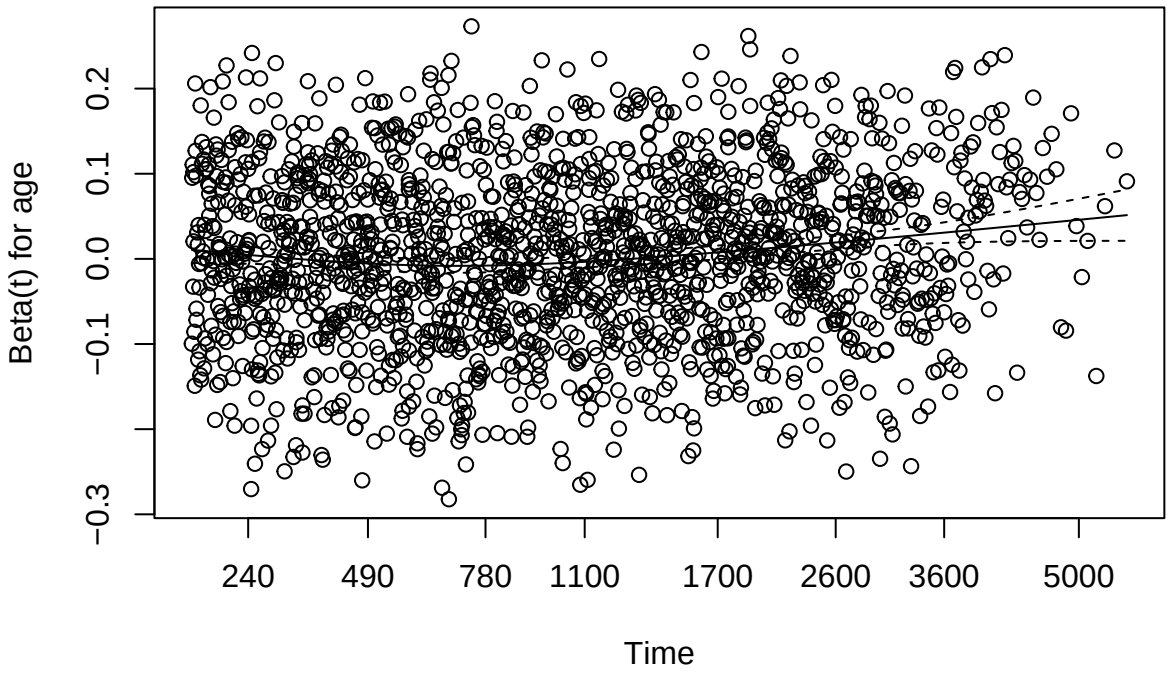
2023-11-28

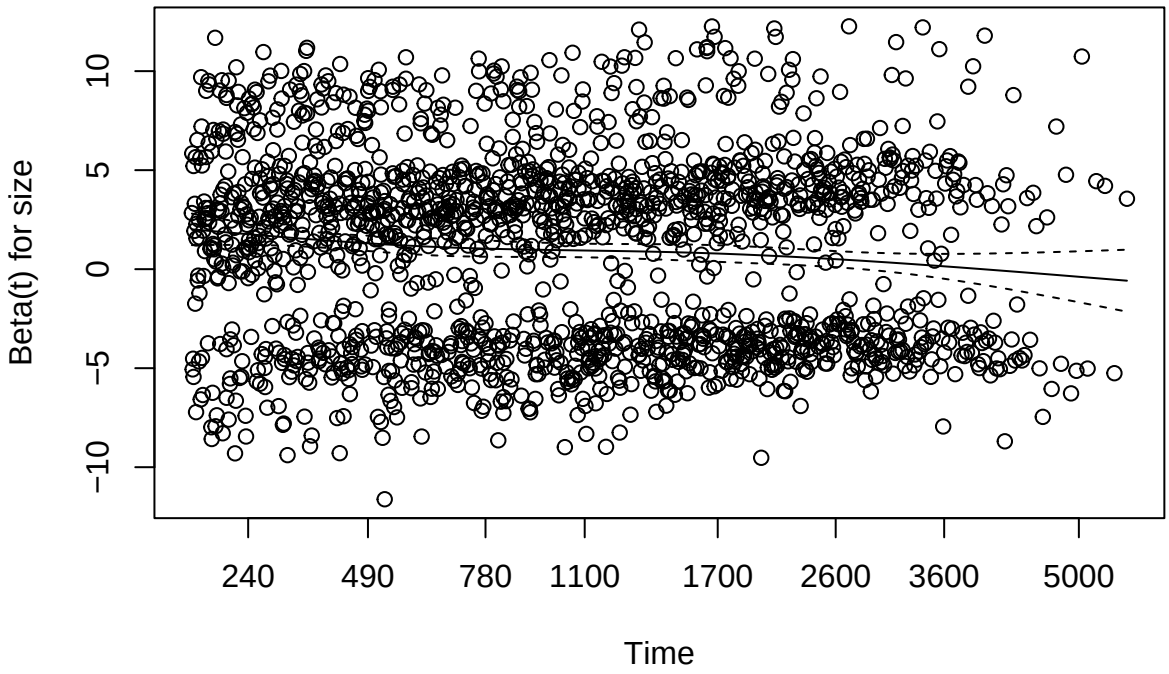
```
library(survival)
library(MASS)
head(rotterdam)
```

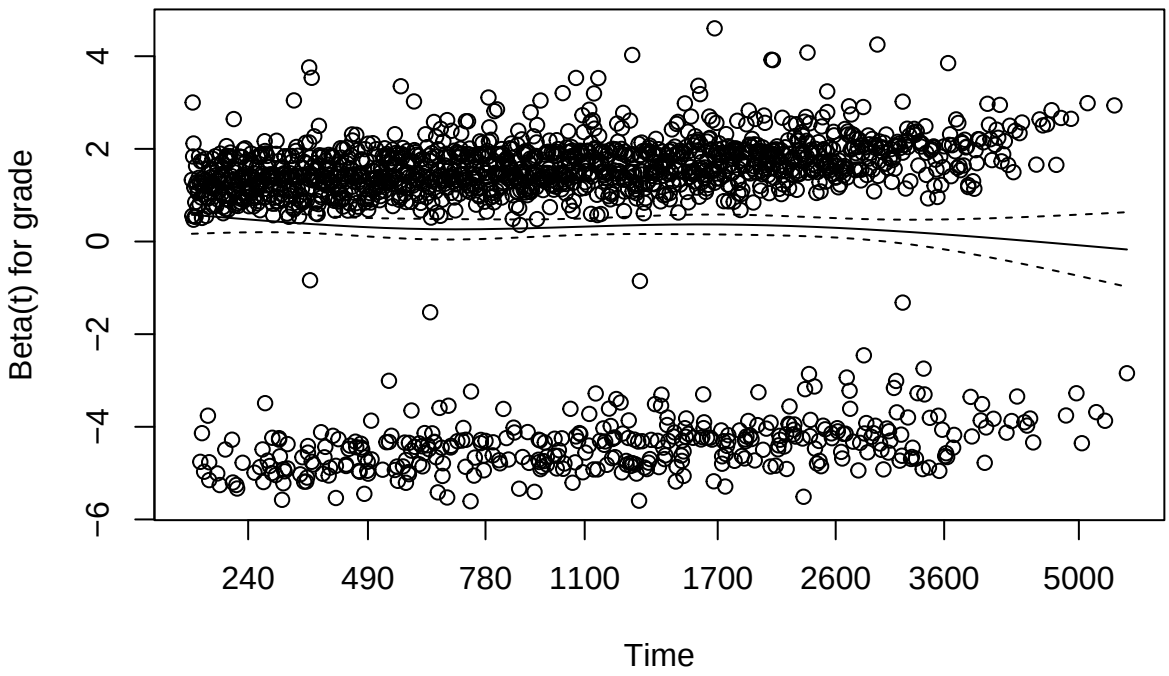
```
##      pid year age meno  size grade nodes pgr  er hormon chemo rtime recur dtime
## 1393   1 1992  74    1  <=20     3     0  35 291      0     0  1799     0  1799
## 1416   2 1984  79    1 20-50     3     0  36 611      0     0  2828     0  2828
## 2962   3 1983  44    0  <=20     2     0 138   0      0     0  6012     0  6012
## 1455   4 1985  70    1 20-50     3     0   0  12      0     0  2624     0  2624
## 977    5 1983  75    1  <=20     3     0 260 409      0     0  4915     0  4915
## 617    6 1983  52    0  <=20     3     0 139 303      0     0  5888     0  5888
##      death
## 1393     0
## 1416     0
## 2962     0
## 1455     0
## 977      0
## 617      0
```

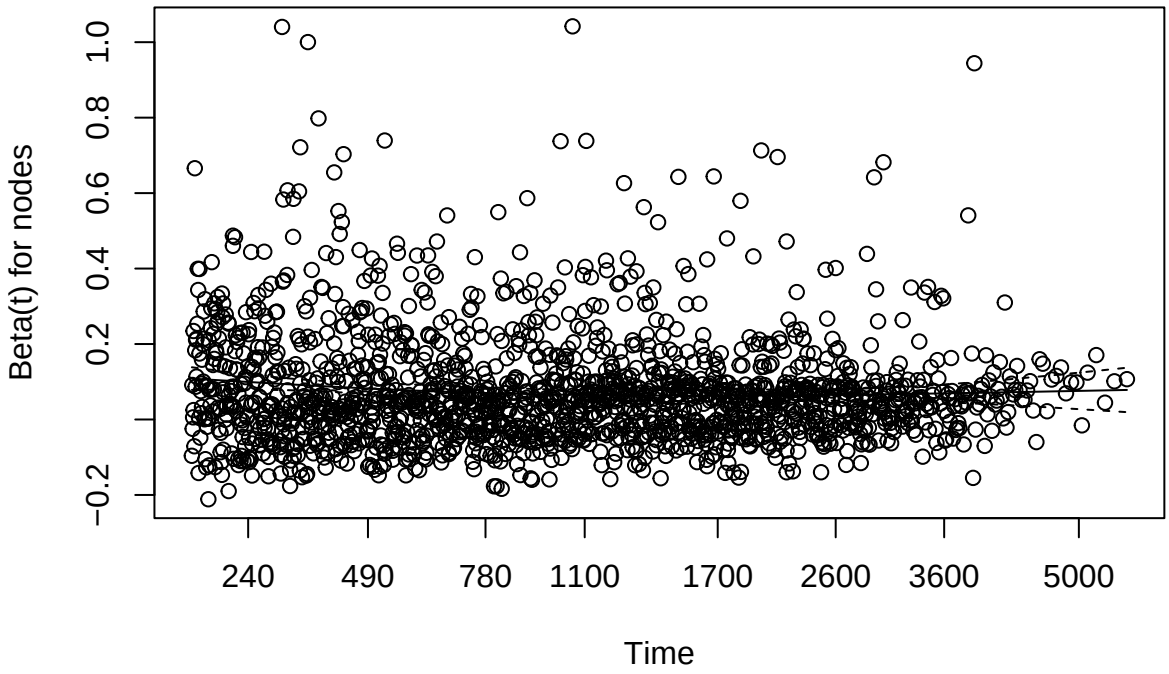
```
rfs  <- pmax(rotterdam$recur, rotterdam$death)
rfstime <- with(rotterdam, ifelse(recur==1, rtime, dtime))
rotterdam$tstart = c(rep(0,2982))
surv_object <- with(rotterdam, Surv(tstart, rfstime, rfs)) #with consideration of time-varying covariat
```

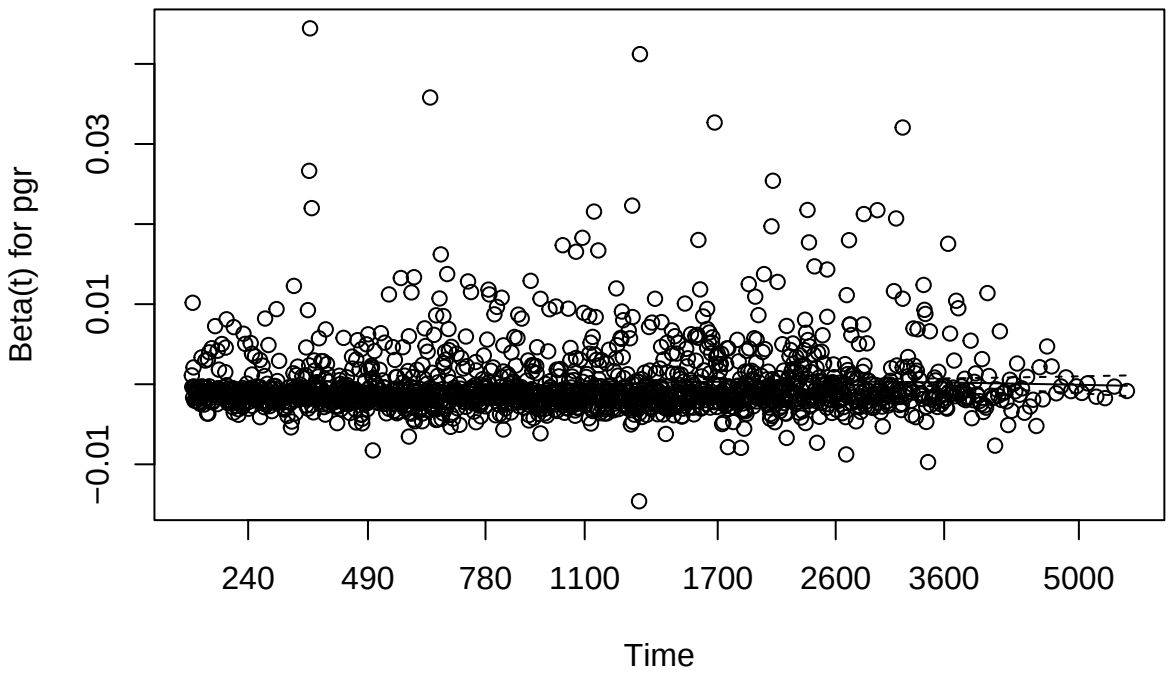
```
# Fit a basic Cox proportional hazards model
cox_model <- coxph(Surv(rfstime, rfs) ~ age + size + grade + nodes + pgr + er + hormon + chemo, data = r
# Checking the Proportional Hazards assumption
cox.zph_model <- cox.zph(cox_model)
plot(cox.zph_model)
```

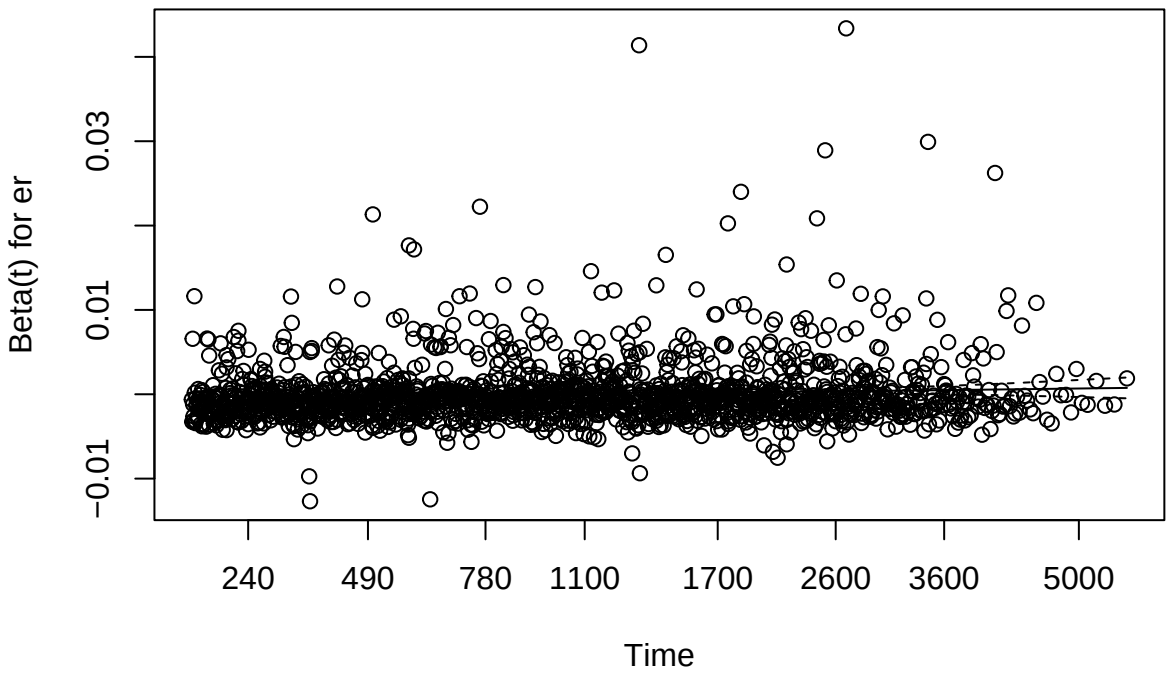

Beta(t) for age
−0.3
−0.1
0.0
0.1
0.2
240
490
780
1100
1700
2600
3600
5000
Time

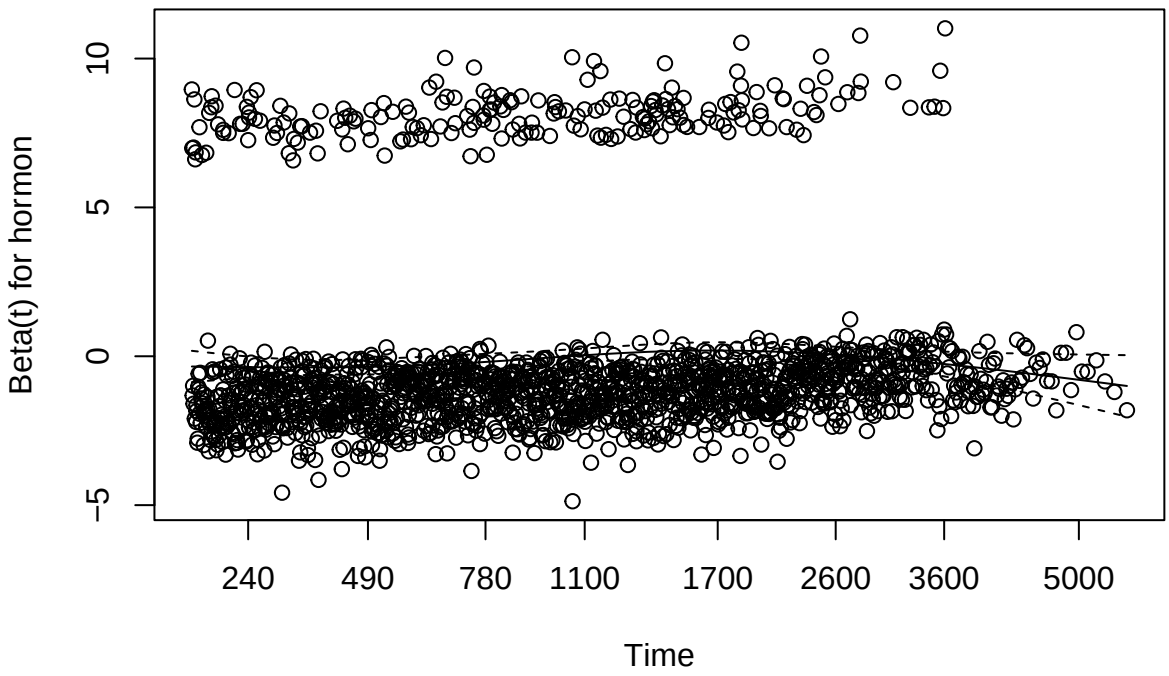

Beta(t) for size
10
5
0
−5
−10
240
490
780
1100
1700
2600
3600
5000
Time

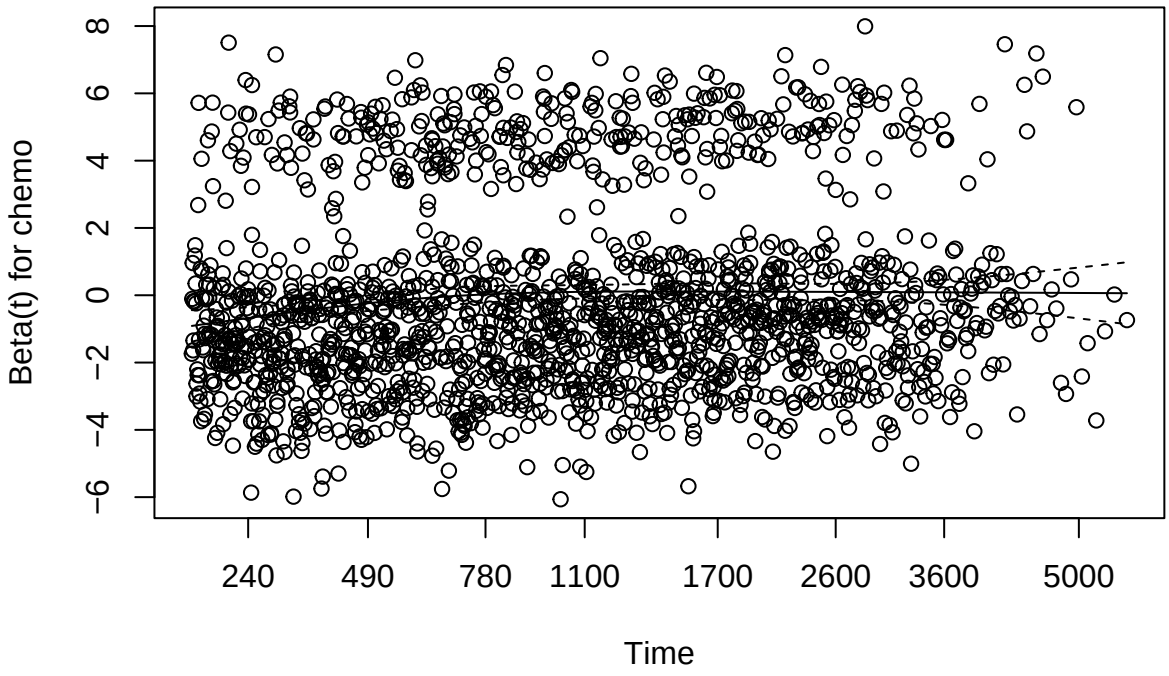

Beta(t) for grade
4
2
0
−2
−4
−6
240
490
780
1100
1700
2600
3600
5000
Time


Beta(t) for nodes
−0.2
0.2
0.4
0.6
0.8
1.0
240
490
780
1100
1700
2600
3600
5000
Time


Beta(t) for pgr
0.03
0.01
−0.01
240
490
780
1100
1700
2600
3600
5000
Time


Beta(t) for er
0.03
0.01
−0.01
240
490
780
1100
1700
2600
3600
5000
Time


Beta(t) for hormon
10
5
0
−5
240
490
780
1100
1700
2600
3600
5000
Time




```
fit1 <- coxph(surv_object ~ pspline(age) + size + grade + nodes + pgr + er + hormon + chemo, data = rot
summary(fit1)
```

```
## Call:
## coxph(formula = surv_object ~ pspline(age) + size + grade + nodes +
##     pgr + er + hormon + chemo, data = rotterdam)
##
##   n= 2982, number of events= 1713
##
##                       coef       se(coef)  se2       Chisq  DF   p
## pspline(age), linear  4.409e-03 2.118e-03 2.118e-03   4.33 1.00 3.7e-02
## pspline(age), nonlin                                  35.87 3.05 8.6e-08
## size20-50             3.337e-01 5.479e-02 5.478e-02  37.09 1.00 1.1e-09
## size>50               5.829e-01 8.278e-02 8.272e-02  49.59 1.00 1.9e-12
## grade                 3.216e-01 6.020e-02 6.018e-02  28.53 1.00 9.2e-08
## nodes                 7.620e-02 4.420e-03 4.417e-03 297.23 1.00 1.3e-66
## pgr                  -1.090e-04 9.704e-05 9.688e-05   1.26 1.00 2.6e-01
## er                   -2.711e-05 9.424e-05 9.400e-05   0.08 1.00 7.7e-01
## hormon               -1.077e-01 7.784e-02 7.775e-02   1.92 1.00 1.7e-01
## chemo                -9.963e-02 7.012e-02 6.988e-02   2.02 1.00 1.6e-01
##
##           exp(coef) exp(-coef) lower .95 upper .95
## ps(age)3     0.9295     1.0759    0.5901     1.464
## ps(age)4     0.8642     1.1571    0.4200     1.778
## ps(age)5     0.7667     1.3043    0.3412     1.723
```

```
## ps(age)6     0.6359     1.5725    0.2851     1.419
## ps(age)7     0.5794     1.7258    0.2636     1.274
## ps(age)8     0.6057     1.6511    0.2758     1.330
## ps(age)9     0.6622     1.5102    0.3008     1.458
## ps(age)10    0.7532     1.3277    0.3415     1.661
## ps(age)11    0.8970     1.1148    0.4055     1.985
## ps(age)12    1.1518     0.8682    0.5089     2.607
## ps(age)13    1.5048     0.6645    0.5966     3.795
## ps(age)14    1.9730     0.5068    0.6055     6.429
## size20-50    1.3961     0.7163    1.2540     1.554
## size>50      1.7913     0.5583    1.5230     2.107
## grade        1.3793     0.7250    1.2258     1.552
## nodes        1.0792     0.9266    1.0699     1.089
## pgr          0.9999     1.0001    0.9997     1.000
## er           1.0000     1.0000    0.9998     1.000
## hormon       0.8979     1.1138    0.7708     1.046
## chemo        0.9052     1.1048    0.7889     1.039
##
## Iterations: 7 outer, 18 Newton-Raphson
##      Theta= 0.9748169
## Degrees of freedom for terms= 4 2 1 1 1 1 1 1
## Concordance= 0.675  (se = 0.007 )
## Likelihood ratio test= 526.2  on 12.03 df,   p=<2e-16
```

```
stepwise_model <- stepAIC(fit1, direction = "both", trace = FALSE)
summary(stepwise_model)
```

```
## Call:
## coxph(formula = surv_object ~ pspline(age) + size + grade + nodes,
##     data = rotterdam)
##
##   n= 2982, number of events= 1713
##
##                       coef     se(coef) se2      Chisq  DF   p
## pspline(age), linear 0.004963 0.001790 0.001790   7.69 1.00 5.6e-03
## pspline(age), nonlin                              36.30 3.05 7.0e-08
## size20-50            0.330228 0.054578 0.054567  36.61 1.00 1.4e-09
## size>50              0.572162 0.082475 0.082403  48.13 1.00 4.0e-12
## grade                0.330550 0.059609 0.059584  30.75 1.00 2.9e-08
## nodes                0.074478 0.004362 0.004359 291.59 1.00 2.2e-65
##
##           exp(coef) exp(-coef) lower .95 upper .95
## ps(age)3     0.9243     1.0819    0.5891     1.450
## ps(age)4     0.8545     1.1703    0.4178     1.748
## ps(age)5     0.7523     1.3293    0.3369     1.680
## ps(age)6     0.6214     1.6093    0.2803     1.378
## ps(age)7     0.5708     1.7521    0.2612     1.247
## ps(age)8     0.6038     1.6563    0.2769     1.317
## ps(age)9     0.6624     1.5097    0.3035     1.446
## ps(age)10    0.7549     1.3247    0.3454     1.650
## ps(age)11    0.9001     1.1110    0.4106     1.973
## ps(age)12    1.1508     0.8690    0.5136     2.579
## ps(age)13    1.4946     0.6691    0.5993     3.728
## ps(age)14    1.9482     0.5133    0.6056     6.268
```

```
## size20-50    1.3913     0.7188    1.2501     1.548
## size>50      1.7721     0.5643    1.5076     2.083
## grade        1.3917     0.7185    1.2383     1.564
## nodes        1.0773     0.9282    1.0682     1.087
##
## Iterations: 7 outer, 19 Newton-Raphson
##      Theta= 0.9752975
## Degrees of freedom for terms= 4.1 2.0 1.0 1.0
## Concordance= 0.673  (se = 0.007 )
## Likelihood ratio test= 520.7  on 8.05 df,   p=<2e-16
```