A Handbook of Statistical Analyses Using R

Brian S. Everitt and Torsten Hothorn

CHAPTER 9

Survival Analysis: Glioma Treatment and Breast Cancer Survival

9.1 Introduction

9.2 Survival Analysis

9.3 Analysis Using R

9.3.1 Glioma Radioimmunotherapy

Figure 9.1 leads to the impression that patients treated with the novel radioimmunotherapy survive longer, regardless of the tumor type. In order to assess if this informal finding is reliable, we may perform a log-rank test via

```
R> survdiff(Surv(time, event) ~ group, data = g3)
```

```
Call:
survdiff(formula = Surv(time, event) ~ group, data = g3)

                N Observed Expected (O-E)^2/E (O-E)^2/V
group=Control  6        4     1.49      4.23      6.06
group=RIT     11        2     4.51      1.40      6.06

 Chisq= 6.1  on 1 degrees of freedom, p= 0.0138
```

which indicates that the survival times are indeed different in both groups. However, the number of patients is rather limited and so it might be dangerous to rely on asymptotic tests. As shown in Chapter **??**, conditioning on the data and computing the distribution of the test statistics without additional assumptions is one alternative. The function `surv_test` from package *coin* (Hothorn et al., 2005) can be used to compute an exact conditional test answering the question whether the survival times differ for grade III patients:

```
R> library("coin")
R> surv_test(Surv(time, event) ~ group, data = g3,
+     distribution = "exact")
```

```
        Exact Logrank Test

data:  Surv(time, event) by groups Control, RIT
Z = 2.1711, p-value = 0.02877
alternative hypothesis: two.sided
```

```
R> data("glioma", package = "coin")
R> library("survival")
R> layout(matrix(1:2, ncol = 2))
R> g3 <- subset(glioma, histology == "Grade3")
R> plot(survfit(Surv(time, event) ~ group, data = g3),
+     main = "Grade III Glioma", lty = c(2, 1), ylab = "Probability",
+     xlab = "Survival Time in Month", legend.bty = "n",
+     legend.text = c("Control", "Treated"))
R> g4 <- subset(glioma, histology == "GBM")
R> plot(survfit(Surv(time, event) ~ group, data = g4),
+     main = "Grade IV Glioma", ylab = "Probability",
+     lty = c(2, 1), xlab = "Survival Time in Month",
+     xlim = c(0, max(glioma$time) * 1.05))
```

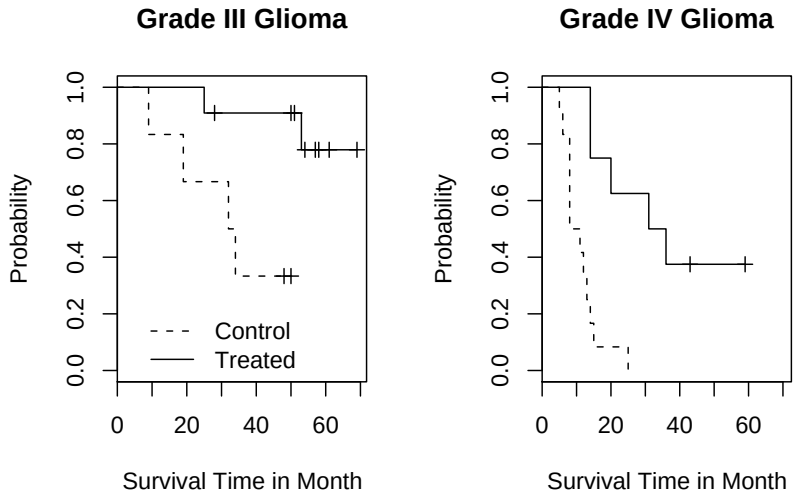

Figure 9.1 Survival times comparing treated and control patients.

which, in this case, confirms the above results. The same exercise can be performed for patients with grade IV glioma

```
R> surv_test(Surv(time, event) ~ group, data = g4,
+     distribution = "exact")

        Exact Logrank Test

data:  Surv(time, event) by groups Control, RIT
Z = 3.2215, p-value = 0.0001588
alternative hypothesis: two.sided
```

which shows a difference as well. However, it might be more appropriate to

answer the question whether the novel therapy is superior for both groups of tumors simultaneously. This can be implemented by *stratifying*, or *blocking*, with respect tumor grading:

```
R> surv_test(Surv(time, event) ~ group | histology,
+     data = glioma, distribution = approximate(B = 10000))

        Approximative Logrank Test

data:  Surv(time, event) by
         groups Control, RIT
         stratified by histology
Z = 3.6704, p-value = 1e-04
alternative hypothesis: two.sided
```

Here, we need to approximate the exact conditional distribution since the exact distribution is hard to compute. The result supports the initial impression implied by Figure 9.1

9.3.2 Breast Cancer Survival

Before fitting a Cox model to the `GBSG2` data, we again derive a Kaplan-Meier estimate of the survival function of the data, here stratified with respect to whether a patient received a hormonal therapy or not (see Figure 9.2). Fitting a Cox model follows roughly the same rules are shown for linear models in Chapters **??**, **??** or **??** with the exception that the response variable is again coded as a *Surv* object. For the `GBSG2` data, the model is fitted via

```
R> GBSG2_coxph <- coxph(Surv(time, cens) ~ ., data = GBSG2)
```

and the results as given by the `summary` method are given in Figure 9.3. Since we are especially interested in the relative risk for patients who underwent a hormonal therapy, we can compute an estimate of the relative risk and a corresponding confidence interval via

```
R> ci <- confint(GBSG2_coxph)
R> exp(cbind(coef(GBSG2_coxph), ci))["horThyes", ]
              2.5 %    97.5 %
0.7073155 0.5492178 0.9109233
```

This result implies that patients treated with a hormonal therapy had a lower risk and thus survived longer compared to women who were not treated this way. Model checking and model selection for proportional hazards models are complicated by the fact that easy to use residuals, such as those discussed in Chapter **??** for linear regression model are not available, but several possibilities do exist. A check of the proportional hazards assumption can be done by looking at the parameter estimates $\beta_1, \ldots, \beta_q$ over time. We can safely assume proportional hazards when the estimates don't vary much over time. The null hypothesis of constant regression coefficients can be tested, both globally as well as for each covariate, by using the `cox.zph` function

```
R> data("GBSG2", package = "ipred")
R> plot(survfit(Surv(time, cens) ~ horTh, data = GBSG2),
+     lty = 1:2, mark.time = FALSE, ylab = "Probability",
+     xlab = "Survival Time in Days")
R> legend(250, 0.2, legend = c("yes", "no"), lty = c(2,
+     1), title = "Hormonal Therapy", bty = "n")
```

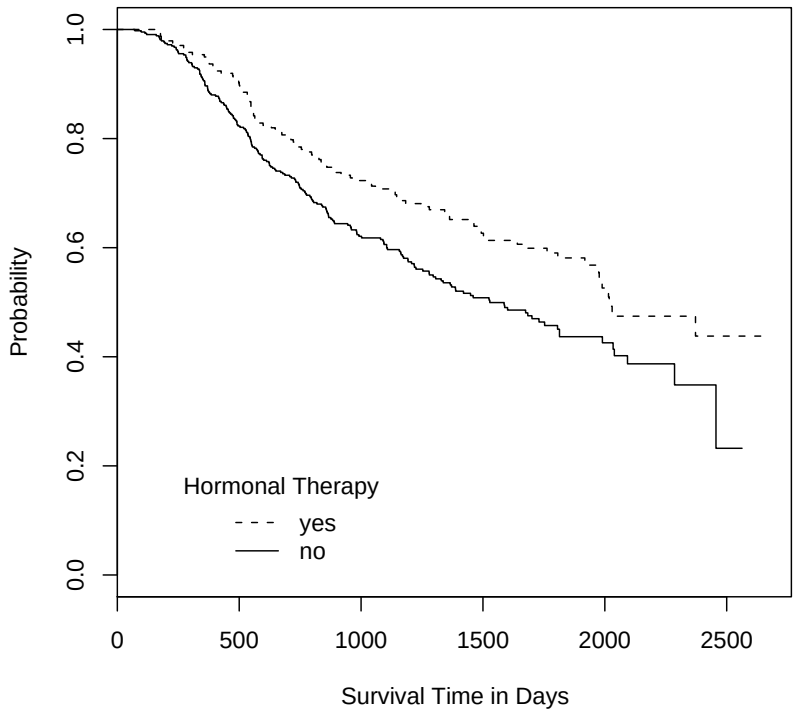

Figure 9.2 Kaplan-Meier estimates for breast cancer patients who either received a hormonal therapy or not.

```
R> GBSG2_zph <- cox.zph(GBSG2_coxph)
R> GBSG2_zph
```

```
                  rho    chisq       p
horThyes     -2.54e-02 1.96e-01 0.65778
age           9.40e-02 2.96e+00 0.08552
menostatPost -1.19e-05 3.75e-08 0.99985
tsize        -2.50e-02 1.88e-01 0.66436
tgrade.L     -1.30e-01 4.85e+00 0.02772
```

```
R> summary(GBSG2_coxph)

Call:
coxph(formula = Surv(time, cens) ~ ., data = GBSG2)

  n= 686
                 coef exp(coef) se(coef)      z       p
horThyes     -0.346278     0.707 0.129075 -2.683 7.3e-03
age          -0.009459     0.991 0.009301 -1.017 3.1e-01
menostatPost  0.258445     1.295 0.183476  1.409 1.6e-01
tsize         0.007796     1.008 0.003939  1.979 4.8e-02
tgrade.L      0.551299     1.736 0.189844  2.904 3.7e-03
tgrade.Q     -0.201091     0.818 0.121965 -1.649 9.9e-02
pnodes        0.048789     1.050 0.007447  6.551 5.7e-11
progrec      -0.002217     0.998 0.000574 -3.866 1.1e-04
estrec        0.000197     1.000 0.000450  0.438 6.6e-01

             exp(coef) exp(-coef) lower .95 upper .95
horThyes         0.707      1.414     0.549     0.911
age              0.991      1.010     0.973     1.009
menostatPost     1.295      0.772     0.904     1.855
tsize            1.008      0.992     1.000     1.016
tgrade.L         1.736      0.576     1.196     2.518
tgrade.Q         0.818      1.223     0.644     1.039
pnodes           1.050      0.952     1.035     1.065
progrec          0.998      1.002     0.997     0.999
estrec           1.000      1.000     0.999     1.001

Rsquare= 0.142   (max possible= 0.995 )
Likelihood ratio test= 105  on 9 df,   p=0
Wald test            = 115  on 9 df,   p=0
Score (logrank) test = 121  on 9 df,   p=0
```

Figure 9.3 R output of the `summary` method for `GBSG2_coxph`.

```
tgrade.Q      3.22e-03 3.14e-03 0.95530
pnodes        5.84e-02 5.98e-01 0.43941
progrec       5.65e-02 1.20e+00 0.27351
estrec        5.46e-02 1.03e+00 0.30967
GLOBAL              NA 2.27e+01 0.00695
```

There seems to be some evidence of time-varying effects, especially for age and tumor grading. A graphical representation of the estimated regression coefficient over time is shown in Figure 9.4. We refer to Therneau and Grambsch (2000) for a detailed theoretical description of these topics. The tree-structured regression models applied to continuous and binary responses in Chapter **??** are applicable to censored responses in survival analysis as well. Such a simple prognostic model with only a few terminal nodes might be

```
R> plot(GBSG2_zph, var = "age")
```

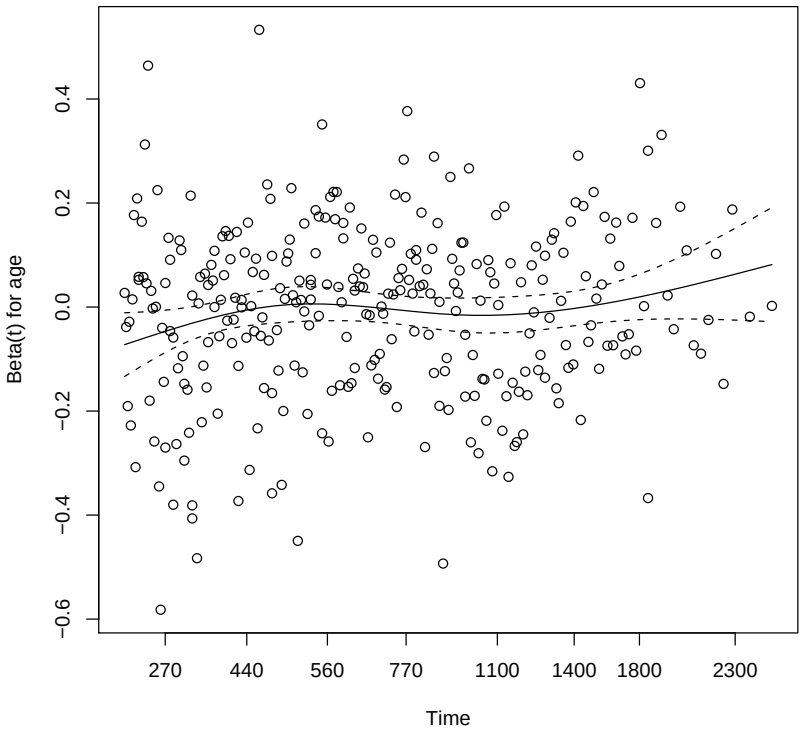

Figure 9.4 Estimated regression coefficient for `age` depending on time for the `GBSG2` data.

helpful for relating the risk to certain subgroups of patients. Both `rpart` and the `ctree` function from package *party* can be applied to the GBSG2 data, where the conditional trees of the latter selects cutpoints based on log-rank statistics;

```
R> GBSG2_ctree <- ctree(Surv(time, cens) ~ ., data = GBSG2)
```

and the `plot` method applied to this tree produces the graphical representation in Figure 9.6. The number of positive lymph nodes (`pnodes`) is the most important variable in the tree, this corresponds to the p-value associated with this variable in Cox's regression, see Figure 9.3. Women with not more than three positive lymph nodes who have undergone a hormonal therapy seem to have the best prognosis whereas a large number of positive lymph nodes and a small value of the progesterone receptor indicates a bad prognosis.

```
R> layout(matrix(1:3, ncol = 3))
R> res <- residuals(GBSG2_coxph)
R> plot(res ~ age, data = GBSG2, ylim = c(-2.5, 1.5),
+     pch = ".", ylab = "Martingale Residuals")
R> abline(h = 0, lty = 3)
R> plot(res ~ pnodes, data = GBSG2, ylim = c(-2.5,
+     1.5), pch = ".", ylab = "")
R> abline(h = 0, lty = 3)
R> plot(res ~ log(progrec), data = GBSG2, ylim = c(-2.5,
+     1.5), pch = ".", ylab = "")
R> abline(h = 0, lty = 3)
```

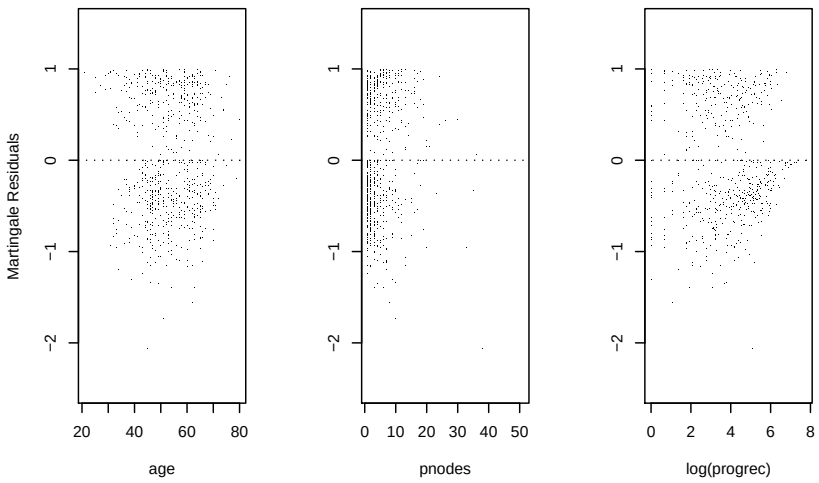

Figure 9.5 Martingale residuals for the `GBSG2` data.

```
R> plot(GBSG2_ctree)
```

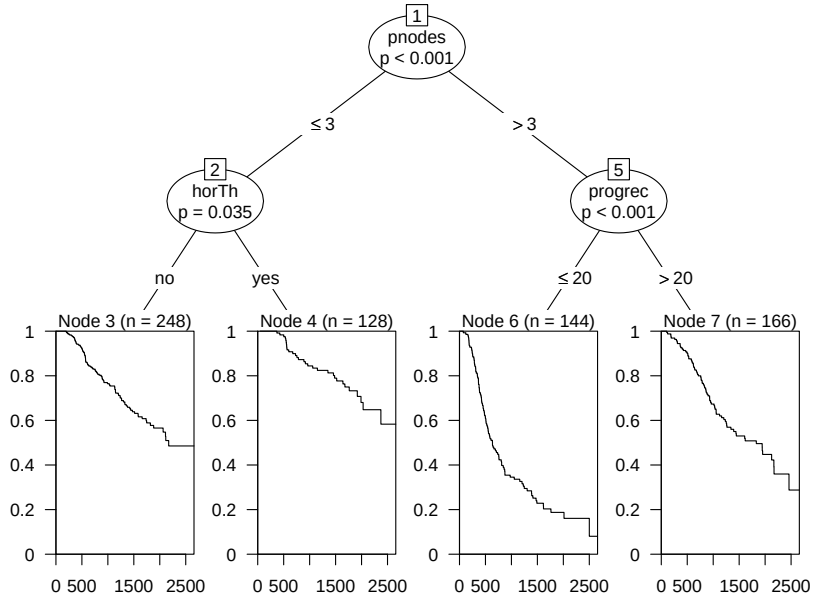

Figure 9.6 `GBSG2` data: Conditonal inference tree with the survival function, estimated by Kaplan-Meier, shown for every subgroup of patients identified by the tree.

Bibliography

Hothorn, T., Hornik, K., van de Wiel, M., and Zeileis, A. (2005), *coin: Conditional Inference Procedures in a Permutation Test Framework*, URL `http://CRAN.R-project.org`, R package version 0.4-7.

Therneau, T. M. and Grambsch, P. M. (2000), *Modeling Survival Data: Extending the Cox Model*, New York, USA: Springer.