Erratum to

Applied Survival Analysis Using R

Dirk F. Moore

© Springer International Publishing Switzerland 2016
D. Moore, *Applied Survival Analysis Using R*, Use R!,
DOI 10.1007/978-3-319-31245-3

DOI 10.1007/978-3-319-31245-3_13

The original version of this book contain errors which have been corrected. The corrections are given below

Chapter 2: Basic Principles of Survival Analysis

Page 15: Line 1: The equation “ $f(t) = -d/dtS(t)$ ” has been updated.

The correct equation is “ $f(t) = -dS(t)/dt$ ”

Page 17: Line 9: The sentence “For example, we can plot the Weibull survival function with $\alpha = 1$ and $\lambda = 0.03$ by first defining a function “weibSurv” with these parameters and then using the “curve” function to plot the curve as follows (figure not shown):” has been updated.

The correct sentence is “For example, we can plot the Weibull survival function with $\alpha = 1.5$ and $\lambda = 0.03$ by first defining a function “weibSurv” with these parameters and then using the “curve” function to plot the curve as follows (figure not shown):”

Page 21: Line 19: The below exponential distribution likelihood sentence has been updated.

For the exponential distribution the likelihood, we substitute the expressions for the p.d.f. and survival distributions, and simplify as follows:

$$L(\lambda) = \prod_{i=1}^n [\lambda e^{-t_i/\mu}]^{\delta_i} [e^{-\lambda t_i}]^{1-\delta_i} = \lambda^d e^{-\lambda v}$$

The updated online version of the original book can be found at
DOI [10.1007/978-3-319-31245-3](https://doi.org/10.1007/978-3-319-31245-3)

The correct exponential distribution likelihood sentence is given below:

For the exponential distribution the likelihood, we substitute the expressions for the p.d.f. and survival distributions, and simplify as follows:

$$L(\lambda) = \prod_{i=1}^n [\lambda e^{-t_i/\lambda}]^{\delta_i} [e^{-\lambda t_i}]^{1-\delta_i}$$

The updated online version of the original chapter can be found at DOI [10.1007/978-3-319-31245-3_2](https://doi.org/10.1007/978-3-319-31245-3_2)

Chapter 3: Nonparametric Survival Curve Estimation

Page 26: The below Table 3.1 last column has been updated.

Table 3.1 Kaplan-Meier estimate

t_i	n_i	d_i	q_i	$1 - q_i$	$S_i = \prod(1 - q_i)$
2	6	1	0.167	0.833	0.846
4	5	1	0.200	0.800	0.693
6	3	1	0.333	0.667	0.497

The correct table is given below:

Table 3.1 Kaplan-Meier estimate

t_i	n_i	d_i	q_i	$1 - q_i$	$S_i = \prod(1 - q_i)$
2	6	1	0.167	0.833	0.833
4	5	1	0.200	0.800	0.666
6	3	1	0.333	0.667	0.444

Page 29: The “Nelson-Altschuler” estimator is now more commonly known as the “Nelson-Aalen” estimator.

Page 29: The below footnote 2 has been updated

Other names associated with this estimator are Aalen, Fleming, and Harrington.

The correct footnote 2 is:

Other names associated with this estimator are Fleming and Harrington.

Page 35: Line 22: The below code for “Plot” has been updated:

```
plot(result.km, conf.int=T, mark="|", xlab="Time in months",
      xlim=c(0,30), ylab="Survival probability")
```

The correct code for “Plot” is:

```
plot(result.km, conf.int=F, mark="|", xlab="Time in months",
      xlim=c(0,30), ylab="Survival probability")
```

Page 37: Line 20: The below sentence has been updated.

After that patient dies, Patients 2 and 6 enter the risk set, and Patient 4 is censored at time 6.

The correct sentence is:

After that patient dies, Patients 2 and 6 enter the risk set, and Patient 4 is censored at time 8.

The updated online version of the original chapter can be found at DOI [10.1007/978-3-319-31245-3_3](https://doi.org/10.1007/978-3-319-31245-3_3)

Chapter 4: Nonparametric Comparison of Survival Distributions

Page 45: Line 3: The below sentence and equation has been updated:

We may sum up over all the tables the differences between the observed and expected values to get a linear test statistic U_0 , and also the sum of the variances V_0 as follows, where N is the number of subjects:

$$U_0 = \sum_{i=1}^N (d_{0i} - e_{0i}) = \sum d_{0i} - \sum e_{0i}$$

The correct sentence and equation is given below:

We may sum up over all the tables the differences between the observed and expected values to get a linear test statistic U_0 , and also the sum of the variances V_0 as follows, where D is the number of failure times:

$$U_0 = \sum_{i=1}^D (d_{0i} - e_{0i}) = \sum d_{0i} - \sum e_{0i}$$

Page 46: Fig 4.1 table: The Fig 4.1 table has been updated.

The correct table is given below:

t_i	n_i	d_i	n_{0i}	d_{0i}	n_{1i}	d_{1i}	e_{0i}	e_{1i}	$v_{0i} = v_{1i}$
6	6	1	3	1	3	0	0.500	0.500	0.2500
10	4	1	1	0	3	1	0.250	0.750	0.1875
15	3	1	1	1	2	0	0.333	0.667	0.2222
25	1	1	0	0	1	1	0.000	1.000	0.0000
				2		2	1.083	2.917	0.6597

Page 47: Lines 14 and 21: The below line has been updated:

“An important generalization of this test makes use of a series of N weights w_i , with which we may define a weighted log-rank test by”

The correct line is given below:

“An important generalization of this test makes use of a series of D weights w_i , with which we may define a weighted log-rank test by”

The below equation has been updated:

$$w_i = N \left\{ \hat{S}(t_i) \right\}^{\rho}$$

The correct equation is given below:

$$w_i = \left\{ \hat{S}(t_i) \right\}^{\rho}$$

Page 47: The below sentence has been updated:

A log-rank test using these weights is called the Fleming-Harrington $G(\rho)$ test [11]. If $\rho = 0$ this test is equivalent to the log-rank test, since then $w_i = n$ for all survival times t_i , and of course the constant n cancels out of the test statistics. If $\rho = 1$, we get what is often known as the Prentice modification (also known as the Peto-Peto modification) of the Gehan-Wilcoxon test. The effect of this test is then to place higher weight on earlier survival differences. The following example illustrates this.

The correct sentence is:

A log-rank test using these weights is called the Fleming-Harrington $G(\rho)$ test [11]. If $\rho = 0$ this test is equivalent to the log-rank test, since then $w_i = 1$ for all survival times t_i . If $\rho = 1$, we get what is often known as the Prentice modification (also known as the Peto-Peto modification) of the Gehan-Wilcoxon test. The effect of this test is then to place higher weight on earlier survival differences. The following example illustrates this.

Page 48: Line 21: The below code has been updated:

```
> pfs <- pmin(progressionOnly , overallSurvival)
```

The correct code is given below:

```
> pfs <- progressionOnly
```

The updated online version of the original chapter can be found at DOI [10.1007/978-3-319-31245-3_4](https://doi.org/10.1007/978-3-319-31245-3_4)

Chapter 5: Regression Analysis Using the Proportional Hazards Model

Page 58: The below sentence has been updated:

The solid curved black line is a plot of the log partial likelihood over a range of values of β . The maximum is indicated by the vertical dashed blue line, and the

value of the l.p.l. at that point is -3.672. Also shown is the value -4.277 of the l.p.l. at the null hypothesis value, $\beta = 0$. The tangent to the $l(\beta)$ curve at $\beta = 0$ is shown by the straight red line. Its slope is the derivative of the log-likelihood (i.e. the score function) evaluated at $\beta = 0$. Interestingly, this is exactly the value of the log-rank statistic U which we obtained in the previous chapter. This simple example illustrates a general principle: *The score function, obtained by taking the derivative of the log partial likelihood, evaluated at $\beta = 0$, is equivalent to the log-rank statistic.* In Sect. 5.4 we will discuss the score test in more detail and also other tests derived from the partial likelihood.

The correct sentence is given below:

The solid curved black line is a plot of the log partial likelihood over a range of values of β . The maximum is indicated by the vertical dashed blue line, and the value of the l.p.l. at that point is -3.672. Also shown is the value -4.277 of the l.p.l. at the null hypothesis value, $\beta = 0$. The tangent to the $l(\beta)$ curve at $\beta = 0$ is shown by the straight red line. Its slope is the derivative of the log-likelihood (i.e. the score function) evaluated at $\beta = 0$. Interestingly, this is exactly the (absolute value of) the log-rank statistic U_0 which we obtained in the previous chapter. This simple example illustrates a general principle: *The score function, obtained by taking the derivative of the log partial likelihood, evaluated at $\beta = 0$, is equivalent to the log-rank statistic.* In Sect. 5.4 we will discuss the score test in more detail and also other tests derived from the partial likelihood.

Page 63: Section 5.4. Line 10:

The below sentence has been updated:

For each factor, (i.e., for each i), the denominator is the sum of all risks in the risk set R_i (denoted by “ $k \in R_j$ ”) of $h_k = h_0\psi_k$ and the numerator is the hazard $h_i = h_0\psi_i$ for the individual in the risk set R_i who experienced the failure. As we saw previously, the baseline hazards $h_0(t_j)$ cancel out of all of the terms, as follows:

The correct sentence is given below:

For each factor, (i.e., for each j), the denominator is the sum of all risks in the risk set R_i (denoted by “ $k \in R_j$ ”) of $h_k = h_0\psi_k$ and the numerator is the hazard $h_j = h_0\psi_j$ for the individual in the risk set R_j who experienced the failure. As we saw previously, the baseline hazards $h_0(t_j)$ cancel out of all of the terms, as follows:

The below line is added after Equation (5.4.1)

where the product is taken over all of the deaths.

Page 64: Sec 5.5: The below equation has been updated:

$$h_0(t_i) = \frac{d_i}{\sum_{j \in R_j} \exp(z_j \hat{\beta})}.$$

The correct equation is given below:

$$h_0(t_i) = \frac{d_i}{\sum_{j \in R_i} \exp(z_j \hat{\beta})}.$$

The updated online version of the original chapter can be found at DOI [10.1007/978-3-319-31245-3_5](https://doi.org/10.1007/978-3-319-31245-3_5)

Chapter 6: Model Selection and Interpretation

Page 79: Line 24: The below sentence has been updated:

We also see that those with part-time employment have a higher hazard when compared to the baseline (which we noted above is the “full-time” group), with a log-hazard ratio of 0.526 and a p-value of 0.056, which may be seen as not quite statistically significant at the 0.05 level.

The correct sentence is given below:

We also see that those with other employment have a higher hazard when compared to the baseline (which we noted above is the “full-time” group), with a log-hazard ratio of 0.526 and a p-value of 0.056, which may be seen as not quite statistically significant at the 0.05 level.

Page 79: The below sentence and equation have been updated:

Let us begin by determining if “ageGroup4” belongs in the model by comparing Models A and C. The null hypothesis is that the three coefficients for “ageGroup4” are zero, and the alternative is that they are not all zero. The likelihood ratio statistic is

$$2 \left(l(\hat{\beta}_{\text{full}}) - l(\hat{\beta}_{\text{reduced}}) \right) = 2(-377.7597 + 380.043) = 4.567.$$

The correct sentence and equation are:

Let us begin by determining if “employment” belongs in the model by comparing Models A and C. The null hypothesis is that the three coefficients for “employment” are zero, and the alternative is that they are not all zero. The likelihood ratio statistic is

$$2 \left(l(\hat{\beta}_{\text{full}}) - l(\hat{\beta}_{\text{reduced}}) \right) = 2(-377.7597 + 380.043) = 4.567.$$

Page 80: Lines 10–22 have been updated:

and we would conclude that the effect of “ageGroup4” is not statistically significant when “employment” is included in the model.

Similarly we can compare Models B and C to test for the importance of “employment” in the presence of “age”:

$$2 \left(l(\hat{\beta}_{\text{full}}) - l(\hat{\beta}_{\text{reduced}}) \right) = 2(-377.7597 + 385.1232) = 14.727.$$

We compare this to a chi-square distribution with $5 - 2 = 3$ degrees for freedom:

```
> pchisq(14.727, df=3, lower.tail=F)
[1] 0.002065452
```

We thus conclude that “employment” belongs in the model if “ageGroup4” is also included, since the p-value for the former is extremely small.

Should “ageGroup4” even be in the model? To carry out a likelihood ratio test for this factor we need to refer to what we shall call the “null” model, one with no covariates. We may evaluate this as follows:

The corrected texts are given below:

and we would conclude that the effect of “employment” is not statistically significant when “ageGroup4” is included in the model.

Similarly we can compare Models B and C to test for the importance of “ageGroup4” in the presence of “employment”:

$$2 \left(l(\hat{\beta}_{\text{full}}) - l(\hat{\beta}_{\text{reduced}}) \right) = 2(-377.7597 + 385.1232) = 14.727.$$

We compare this to a chi-square distribution with $5 - 2 = 3$ degrees for freedom:

```
> pchisq(14.727, df=3, lower.tail=F)
[1] 0.002065452
```

We thus conclude that “ageGroup4” belongs in the model if “employment” is also included, since the p-value for the former is extremely small.

Should “ageGroup4” alone be in the model? To carry out a likelihood ratio test for this factor we need to refer to what we shall call the “null” model, one with no covariates. We may evaluate this as follows:

Page 80: Line 11 from bottom: The equation has been updated:

$$2 \left(l(\hat{\beta}_{\text{full}}) - l(\hat{\beta}_{\text{reduced}}) \right) = 2(-380.043 + 386.1533) = 12.2206$$

The correct equation is:

$$2 \left(l(\hat{\beta}_{\text{full}}) - l(\hat{\beta}_{\text{reduced}}) \right) = 2(-380.043 + 386.1533) = 12.2206$$

Page 80: Line 4 from bottom: The below sentence has been updated:

Since the reference model is always the null model, another way to carry out the likelihood ratio test for, for example, Model B to Model C, is to take the differences

of the printed log-likelihood statistics, e.g., $16.8 - 12.2 = 4.6$, which (to one decimal place) is identical to the value we computed using the “logLik” function.

The correct sentence is:

Since the reference model is always the null model, another way to carry out the likelihood ratio test for, for example, Model A to Model C, is to take the differences of the printed log-likelihood statistics, e.g., $16.8 - 12.2 = 4.6$, which (to one decimal place) is identical to the value we computed using the “logLik” function.

The updated online version of the original chapter can be found at DOI [10.1007/978-3-319-31245-3_6](https://doi.org/10.1007/978-3-319-31245-3_6)

Chapter 7: Model Diagnostics

Page 87: Line 18: The below sentence has been updated:

These residuals, which originate from the counting process theory underlying the Cox proportional hazards model, sum to 1, range in value from $-\infty$ to a maximum of 1, and each has an expected value of zero.

The correct sentence is given below:

These residuals, which originate from the counting process theory underlying the Cox proportional hazards model, sum to 0, range in value from $-\infty$ to a maximum of 1, and each has an expected value of zero.

Page 96: Line 23: The below sentence has been updated:

The “expected” value of the covariate at that time is the weighted sum (weighted by p_i). To clarify how we compute these, let us return to our simple example from Chap. 3 comparing two sets of three observations. We first need $\hat{\beta}$, the log hazard ratio estimate, which we may compute as follows:

The correct sentence is given below:

The “expected” value of the covariate at that time is the weighted sum (weighted by p_i). To clarify how we compute these, let us return to our simple example from Chap. 4 comparing two sets of three observations. We first need $\hat{\beta}$, the log hazard ratio estimate, which we may compute as follows:

Page 97: The below equation has been updated:

t_i	n_{0i}	n_{1i}	$p(\beta, z_k = 0)$	$p(\beta, z_k = 1)$
6	3	3	$\frac{1}{3 + 3e^{-1.326}}$	$\frac{e^{-1.326}}{3 + 3e^{-1.326}}$
10	1	3	$\frac{1}{1 + 3e^{-1.326}}$	$\frac{e^{-1.326}}{1 + 3e^{-1.326}}$
15	1	2	$\frac{1}{1 + 2e^{-1.326}}$	$\frac{e^{-1.326}}{1 + 2e^{-1.326}}$
25	1	1	$\frac{1}{e^{-1.326}}$	$\frac{e^{-1.326}}{e^{-1.326}} = 1$

The correct equation is given below:

t_i	n_{0i}	n_{1i}	$p(\beta, z_k = 0)$	$p(\beta, z_k = 1)$
6	3	3	$\frac{1}{3+3e^{-1.326}}$	$\frac{e^{-1.326}}{3+3e^{-1.326}}$
10	1	3	$\frac{1}{1+3e^{-1.326}}$	$\frac{e^{-1.326}}{1+3e^{-1.326}}$
15	1	2	$\frac{1}{1+2e^{-1.326}}$	$\frac{e^{-1.326}}{1+2e^{-1.326}}$
25	0	1	$\frac{1}{e^{-1.326}}$	$\frac{e^{-1.326}}{e^{-1.326}} = 1$

The updated online version of the original chapter can be found at DOI [10.1007/978-3-319-31245-3_7](https://doi.org/10.1007/978-3-319-31245-3_7)

Chapter 8: Time Dependent Covariates

Page 103: The below Fig. 8.1 caption has been updated:

Fig. 8.1 Sample of six patients from the Stanford heart transplant data set. In this plot, death is denoted by an “X”, and the time of transplant (for Patients 1, 3, and 6) by a *solid dot*. In the plot on the right, the timelines of patients who received a transplant are split into pre- and post-transplant components

The correct Fig. 8.1 caption is given below:

Fig. 8.1 Sample of six patients from the Stanford heart transplant data set. In this plot, death is denoted by an “X”, and the time of transplant (for Patients 10, 28, and 95) by a *solid dot*. In the plot on the right, the timelines of patients who received a transplant are split into pre- and post-transplant components

Page 104: Line 14: The below sentence has been updated:

For example, from Table 8.1 and Fig. 8.1, we see that Patient #2 is the first to fail, at $t = 5$. At this time, all six patients are at risk, but only one, Patient #2, has had a transplant at this time.

The correct sentence is given below:

For example, from Table 8.1 and Fig. 8.1, we see that Patient #2 is the first to fail, at $t = 5$. At this time, all six patients are at risk, but only one, Patient #95, has had a transplant at this time.

Page: 106: The following code has been updated:

```
> tdata <- jasa[, -c(1:4, 11:14)]
> tdata$futime <- pmax(.5, tdata$futime)
> indx <- {{tdata$wait.time == tdata$futime} &
+         !is.na(tdata$wait.time)}
> tdata$wait.time[indx] <- tdata$wait.time[indx] - .5
> sdata <- tmerge(tdata, tdata, id=1:nrow(tdata),
+               death = event(futime, fustat),
+               trans = tdc(wait.time))
> jasa.counting <- sdata[,c(7:11, 2:3)]
> head(jasa.counting)
```

The correct code is given below:

```
> tdata <- jasa[, -c(1:4, 11:14)]
> tdata$futime <- pmax(.5, tdata$futime)
> indx <- {{tdata$wait.time == tdata$futime} &
+         !is.na(tdata$wait.time)}
> tdata$wait.time[indx] <- tdata$wait.time[indx] - .5
> id <- 1:nrow(tdata)
> tdata$id <- id
> sdata <- tmerge(tdata, tdata, id=id,
+                 death = event(futime, fustat),
+                 trans = tdc(wait.time))
> jasa.counting <- sdata[,c(7:11, 2:3)]
> head(jasa.counting)
```

Page 107: Line 20: The below sentence has been updated:

The p-value (0.14) for the likelihood ratio test, which is similar to that from the log-rank test in Sect. 4.1, shows little evidence of a group difference, as we saw there.

The correct sentence is given below:

The p-value (0.119) for the likelihood ratio test, which is similar to that from the log-rank test in Sect. 4.1, shows little evidence of a group difference, as we saw there.

Page 109: Line 21: The below sentence has been updated:

We see that the log hazard increases with increasing age, with a p-value of 0.042.

The correct sentence is given below:

We see that the log hazard increases with increasing age, with a p-value of 0.040.

The updated online version of the original chapter can be found at
DOI [10.1007/978-3-319-31245-3_8](https://doi.org/10.1007/978-3-319-31245-3_8)

Chapter 9: Multiple Survival Outcomes and Competing Risks

Page 114: Line 26: The below sentence in Example 9.2 has been updated:

For each patient on eye was randomly assigned to receive the laser treatment and the other was untreated.

The correct sentence in Example 9.2 is given below:

For each patient one eye was randomly assigned to receive the laser treatment and the other was untreated.

Page 125: The below table has been updated:

Time	n.risk	n.event.1	n.event.2	n.event.any	Survival	h.1	h.2	CI.1	CI.2
2	6	1	0	1	0.833	1/6	0	0.167	0.000
3	5	0	1	1	0.667	0	1/5	0.167	0.167
5	3	1	0	1	0.444	1/3	0	0.389	0.167
7	1	0	1	1	0.1000	0	1	0.389	0.611

The correct table is given below:

Time	n.risk	n.event.1	n.event.2	n.event.any	Survival	h.1	h.2	CI.1	CI.2
2	6	1	0	1	0.833	1/6	0	0.167	0.000
3	5	0	1	1	0.667	0	1/5	0.167	0.167
5	3	1	0	1	0.444	1/3	0	0.389	0.167
7	1	0	1	1	0.000	0	1	0.389	0.611

Page 126: Section 9.2.3 : The following text and code have been updated:

Returning to the prostate cancer example of Fig. 9.1, we may now estimate the competing risks cumulative incidence functions using the “Cuminc” function in the R package “mstate” as follows:

```
> library(mstate)
> ci.prostate <- Cuminc(time=prostateSurvival.highrisk$survTime,
+   status=prostateSurvival.highrisk$status)
```

The first few lines of the resulting file “ci.prostate” are as follows:

```
> head(ci.prostate)
  time Surv    CI.1    CI.2 seSurv seCI.1 seCI.2
1    0 1.000 0.00000 0.00000 0.00000 0.00000 0.00000
2    1 0.994 0.00000 0.00592 0.00264 0.00000 0.00264
3    2 0.988 0.00602 0.00592 0.00376 0.00269 0.00264
4    3 0.984 0.00848 0.00715 0.00430 0.00319 0.00291
5    4 0.983 0.00973 0.00715 0.00447 0.00342 0.00291
6    5 0.978 0.01477 0.00715 0.00511 0.00423 0.00291
```

The first column, “time” is the time in months. The column “Surv” is the Kaplan-Meier survival estimate for time to death from any cause (prostate or something else). The next two columns are the cumulative incidence function estimates for causes 1 (prostate) and 2 (other). The remaining columns are standard errors of the respective estimates. We may plot the cause-specific cumulative incidence functions as follows:

```
cil <- ci.prostate$CI.1 # CI.1 is for prostate cancer
ci2 <- ci.prostate$CI.2 # CI.2 is for other causes
times <- ci.prostate$time/12 # convert months to years
Rci2 <- 1 - ci2
```

The correct code is given below

Returning to the prostate cancer example of Fig. 9.1, we may now estimate the competing risks cumulative incidence functions as follows:

```
> sf <- survfit(Surv(survTime, status, type="mstate") ~ 1,
+   data=prostateSurvival.highrisk)
```

We may plot the cause-specific cumulative incidence functions as follows:

```
> tt <- sf$time
> CIs <- sf$pstate
> ci1 <- CIs[,1]
> ci2 <- CIs[,2]
> times <- tt/12
> Rci2 <- 1 - ci2
```

The updated online version of the original chapter can be found at DOI [10.1007/978-3-319-31245-3_9](https://doi.org/10.1007/978-3-319-31245-3_9)

Chapter 10: Parametric Models

Page 138: The below sentence has been updated:

The Weibull survival distribution, as expressed in Sect. 2.4, has hazard and survival functions $h(t) = \alpha\lambda^\alpha t^{\alpha-1}$ and $S(t) = e^{-(\lambda t)^\alpha}$. Later, when we use the Weibull distribution to assess the effects of covariates on survival, we shall find it convenient to use the scale parameter $\sigma = 1/\alpha$, and the mean parameter $\mu = -\log \lambda$. Then

The correct sentence is given below:

The Weibull survival distribution, as expressed in Sect. 2.4, has hazard and survival functions $h(t) = \alpha\lambda^\alpha t^{\alpha-1}$ and $S(t) = e^{-(\lambda t)^\alpha}$. Later, when we use the Weibull distribution to assess the effects of covariates on survival, we shall find it convenient to use the scale parameter $\sigma = 1/\alpha$, with parameterization $\mu = -\log \lambda$. Then

The updated online version of the original chapter can be found at DOI [10.1007/978-3-319-31245-3_10](https://doi.org/10.1007/978-3-319-31245-3_10)

Chapter 12: Additional Topics

Page 182: The following table has been updated:

diagGrp	trt	
	0	1
1	1 / 21	0 / 24
2	1 / 7	1 / 18
3	1 / 12	- / -

The correct table is given below:

diagGrp	trt	
	0	1
1	1 / 21	0 / 24
2	1 / 7	1 / 18
3	- / -	1 / 12

Page 184: The following code has been updated:

```
> prost.80plus.poor <- prostateSurvival[{{grade == "poor"}} &
+   {ageGroup == "80+"}},]
> prost.80plus.poor$status.all <-
+   as.numeric(prost.80plus.poor$status >= 1)
> prost.80plus.poor$T2 <-
+   as.numeric(prost.80plus.poor$stage == "T2")
> prost.80plus.poor$id <- 1:nrow(prost.80plus.poor)
> head(prost.80plus.poor)
```

The correct code is given below:

```
> attach(prostateSurvival)
> prost.80plus.poor <- prostateSurvival[{{grade == "poor"}} &
+   {ageGroup == "80+"}},]
> prost.80plus.poor$status.all <-
+   as.numeric(prost.80plus.poor$status >= 1)
> prost.80plus.poor$T2 <-
+   as.numeric(prost.80plus.poor$stage == "T2")
> prost.80plus.poor$id <- 1:nrow(prost.80plus.poor)
> head(prost.80plus.poor)
```

Page 185: The following code has been updated:

```
> prost.split.s <- survSplit(data=prost.80plus.poor,
+   cut=tau.s, end="survTime", start="t0",
+   event="status.all", episode="survGrp")
> prost.split.s$expo <- prost.split.s$survTime - prost.split.s$t0
```

The correct code is given below:

```
> zeroInd <- {prost.80plus.poor$survTime == 0}
> prost.80plus.poor$survTime[zeroInd] <- 0.1
```

The updated online version of the original chapter can be found at
DOI [10.1007/978-3-319-31245-3_12](https://doi.org/10.1007/978-3-319-31245-3_12)