observed number of events in each group to the model-based estimate of the expected number of events. Using the counting process approach, they derive an expression for the covariance matrix of the vector of G sums. They show that their quadratic form test statistic has a chi-square distribution with $G-1$ degrees-of-freedom when the fitted model is the correct model and the sample is large enough that the estimated expected number of events in each group is large. As presented in their paper, the calculations of Grønnesby and Borgan (1996) are not a trivial matter.

May and Hosmer (1998), following the method used by Tsiatis (1980) to derive a goodness-of-fit test in logistic regression, prove that Grønnesby and Borgan's test is the score test for the addition of $G-1$ design variables, based on the G groups, to the fitted proportional hazards model. Thus, the test statistic may be calculated in any package that performs score tests. Using the asymptotic equivalence of score tests and likelihood ratio tests, one may approximate the score test with the partial likelihood ratio test, which may be done in any package.

One may be tempted to define groups based on the subject-specific estimated survival probabilities,

$$\hat{S}\left(t_i, \mathbf{x}_i, \hat{\boldsymbol{\beta}}\right) = \left[\hat{S}_0\left(t_i\right)\right]^{\exp\left(\mathbf{x}_i'\hat{\boldsymbol{\beta}}\right)}.$$

This should not be done as the values of time differ for each subject. If groups are to be based on the survival probability scale, they should be computed using the risk score and a fixed value of time for each subject. For example, in the UIS we could use the estimated one-year survival probability

$$\hat{S}\left(365, \mathbf{x}_i, \hat{\boldsymbol{\beta}}\right) = \left[\hat{S}_0(365)\right]^{\exp\left(\mathbf{x}_i'\hat{\boldsymbol{\beta}}\right)}.$$

Since the choice of a time is arbitrary, one cannot interpret the probability as a prediction of the number of events in each decile of risk. It merely provides another way to express the risk score.

The value of the score test for the inclusion of the nine decile-of-risk design variables to the model in Table 5.11 is 7.86 which, with 9 degrees-of-freedom, has a p-value of 0.549. The partial likelihood ratio test comparing the model in Table 5.11 to the one including the nine design variables is $G = 7.56$ which, with 9 degrees-of-freedom, has a p-value of 0.579. The two test statistics have nearly the same value and

neither is significant, suggesting that there is no evidence that the model does not fit.

May and Hosmer's (1998) result not only greatly simplifies the calculation of the test, but it also suggests that a two by ten table presenting the observed and expected numbers of events in each group is a useful way to summarize the model fit. The individual observed and expected values in the table may be compared by appealing to counting process theory. Under this theory, the counting function is approximately a Poisson variate with mean equal to the cumulative hazard function. Sums of independent count functions will be approximately Poisson distributed, with mean equal to the sum of the cumulative hazard function. This suggests considering the observed counts within each decile of risk to be distributed approximately Poisson, with mean equal to the estimated expected number of counts. Furthermore, the fact that the Poisson distribution may be approximated by the normal for large values of the mean suggests that an easy way to compare the observed and expected counts is to form a z-score by dividing their difference by the square root of the expected. The two-tailed p-value is obtained from the standard normal distribution. There are obvious dependencies in the counts due to the fact that the same estimated parameter vector is used to calculate the individual expected values and some dependency due to grouping subjects into deciles. The effect of these dependencies has not been studied, but it is likely to smooth the counts toward the expected counts. Thus, the proposed cell-wise z-score comparisons should, if anything, be a bit conservative.

Table 6.5 presents the observed and estimated expected numbers of events, the z-score and two-tailed p-value within each decile of risk for the fitted model in Table 5.11. The numbers in Table 6.5 are large enough that we feel comfortable using the normal approximation to the Poisson distribution. With a p-value equal to 0.049, only the sixth decile has a possibly significant difference between the observed and model-based expected count. If we use Bonferroni's method to adjust the 5 percent level of significance for multiple testing to 0.005, then none of the deciles has a significant difference between the observed and expected counts. Thus, we conclude that there is agreement between observed and expected number of events within each of the 10 deciles of risk.

Arjas (1988) suggests plotting the cumulative observed versus the cumulative estimated expected number of events for subjects with observed, not censored, survival times within partitions of the data to assess model fit. If the model is the correct one, the points should follow a 45

Table 6.5 Observed Number of Events, Estimated Number of Events, z-Scores and Two-Tailed p-Values within Each Decile of Risk Based on the Model in Table 5.11

Decile of Risk	Observed Number of Events	Estimated Number of Events	z	p-Value
1	34	33.96	0.007	0.994
2	43	36.35	1.103	0.270
3	37	44.88	–1.176	0.240
4	44	45.77	–0.262	0.744
5	46	52.82	–0.939	0.348
6	51	38.73	1.972	0.049
7	49	49.28	–0.041	0.968
8	53	52.73	0.037	0.971
9	52	53.19	–0.164	0.870
10	55	56.3	–0.173	0.863
Total	464	464		

degree line beginning at the origin. Arjas suggests forming groups based on covariate values, such as the treatment variable for the model in Table 5.11. Rather than using groups based on only a few covariates, we feel that a partition based on the risk score is a convenient way to incorporate all study covariates into the grouping strategy. The Arjas plot for each of the deciles of risk in Table 6.5 provides graphical support for the conclusion that the model fits within each decile of risk. These plots do, in fact, support model fit. For illustrative purposes we demonstrate the Arjas plots in Figure 6.8 using quartiles of risk instead of deciles of risk.

The plots in Figure 6.8 are obtained as follows: first, create groups based on quartiles of risk and sort on risk score within each group; second, compute the cumulative sum of the zero-one censoring variable and the cumulative sum of the estimated cumulative hazard function within each group; third, plot the pairs of cumulative sums within each group only for subjects with an observed survival time.

In Figures 6.8a–6.8d the polygons connecting the points are each close to the 45 degree line. They display small departures that do not necessarily indicate a poorly fitting model. Thus, these Arjas plots do not contradict earlier conclusions regarding the model's fit to the data.

As in all regression analyses, some measure analogous to R^2 may be of interest as a measure of model performance. As shown in a detailed study by Schemper and Stare (1996), there is not a single, simple, easy

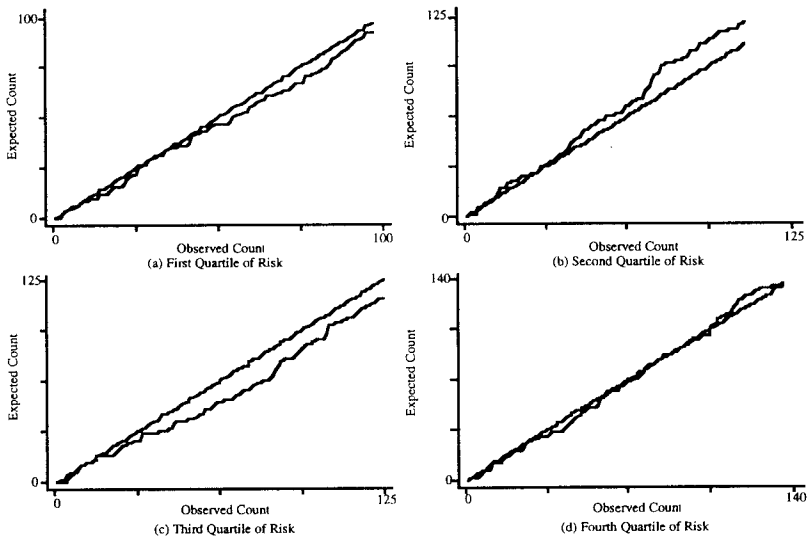

Figure 6.8 Plots of the cumulative estimated expected count versus the cumulative observed count within each quartile of risk based on the fitted model in Table 5.11, for subjects with an observed survival time.

to calculate, useful, easy to interpret measure for a proportional hazards regression model. In particular, all measures depend on the proportion of values that are censored. A perfectly adequate model may have what, at face value, seems like a terribly low R^2 due to a high percent of censored data. In our opinion, further work needs to be done before we can recommend one measure over another. However, if one must compute such a measure, then

$$R_p^2 = 1 - \left\{ \exp\left[\frac{2}{n}\left(L_0 - L_p\right) \right] \right\}$$

is perhaps the easiest and best one to use, where L_p is the log partial likelihood for the fitted model with p covariates, and L_0 is the log partial likelihood for model zero, the model with no covariates. For the fitted model in Table 5.11, the value is

$$R_p^2 = 1 - \left(\exp\left\{ \left[\frac{2}{575} \right] \times [(-2663.985) - (-2630.418)] \right\} \right) = 0.11.$$

The model displayed in Table 5.11 has passed all the tests for a good fitting model. We are now in a position to discuss the interpretation of this model and how to best present the results to the audience of interest.

6.6 INTERPRETATION AND PRESENTATION OF THE FINAL MODEL

The model fit to the UIS data, shown in Table 5.11, is reported again in Table 6.6. It is an excellent model for teaching purposes, as it contains an example of just about every possible covariate one is likely to encounter in practice. The model contains two simple dichotomous covariates (treatment and recent IV drug use), a continuous linear covariate (Beck score), a continuous non-linear covariate (number of prior drug treatments), an interaction between a continuous and a dichotomous covariate (age and site) and and interaction between two dichotomous covariates (race and site). In this section, when we refer to "the model" we are referring to the one in Table 6.6.

Table 6.6 Estimated Coefficients, Standard Errors, *z*-Scores, Two-Tailed *p*-Values and 95% Confidence Intervals for the Final Proportional Hazards Model for the UIS (n = 575)

Variable	Coeff.	Std. Err.	z	$P>\|z\|$	95% CIE
AGE	–0.041	0.010	–4.18	<0.001	–0.061, –0.022
BECKTOTA	0.009	0.005	1.76	0.078	–0.001, 0.018
NDRUGFP1	–0.574	0.125	–4.59	<0.001	–0.820, –0.329
NDRUGFP2	–0.215	0.049	–4.42	<0.001	–0.310, –0.119
IVHX_3	0.228	0.109	2.10	0.036	0.015, 0.441
RACE	–0.467	0.135	–3.47	0.001	–0.731, –0.203
TREAT	–0.247	0.094	–2.62	0.009	–0.432, –0.062
SITE	–1.317	0.531	–2.48	0.013	–2.359, –0.275
AGE×SITE	0.032	0.016	2.02	0.044	0.001, 0.064
RACE×SITE	0.850	0.248	3.43	0.001	0.365, 1.336

Log-likelihood = –2630.418

We begin by discussing how to prepare point and interval estimates of hazard ratios for the covariates. We wish to call attention to the fact that we have assiduously avoided including any exponentiated coefficients in tables of estimated coefficients in Chapters 5 and 6. While most software packages automatically provide these quantities, they probably will be useful summary statistics for only a few model covariates. We feel it is best not to even attempt estimating any hazard ratios until one has completed all steps in both model development and model checking: i.e., the model fits, satisfies the proportional hazards assumption, and any and all highly influential subjects have been dealt with in a scientifically appropriate manner.

Only the covariates for Beck score, recent IV drug use and treatment have hazard ratios that may be estimated by exponentiating their estimated coefficients. This is because the other covariates are either involved in interactions or are nonlinear in the model. It is convenient to display these estimated hazard ratios and their confidence intervals in a table similar to Table 6.7.

The estimated hazard ratio for a 10-point increase in the Beck score is $1.09 = \exp(10 \times 0.009)$, which shows a slight increase in the rate of return to drug use. The interpretation is that subjects with the 10-point higher score are returning to drug use at a rate that is 9 percent higher than for subjects at the lower score. The 95 percent confidence interval suggests that an increased rate of return to drug use as high as 20 percent or even a decreased rate of 1 percent is consistent with the data. Since the model is linear in the Beck score, this interpretation holds over the observed range of Beck scores.

The estimated hazard ratio for recent IV drug use is 1.26. The interpretation of this is that subjects who have a recent history of IV drug use are returning to drug use at a rate that is 26 percent higher than for subjects who are not recent IV drug users. The confidence interval indicates that the rate could actually be as much as 55 percent higher or as little as 2 percent higher.

The hazard ratio for treatment of 0.78 means that subjects in the longer or extended treatment program are returning to drug use at a rate that is 22 percent lower than for subjects with the shorter treatment. The 95 percent confidence interval suggests that the rate could be as much as 35 percent lower to only 6 percent lower. The estimated hazard ratio points to a significant benefit for the longer of the two treatments, controlling for all other model covariates. In studies in which there is a single covariate of primary interest, such as a treatment covariate, one may encounter tables of results in which a summary statistic for

Table 6.7 Estimated Hazard Ratios and 95% Confidence Intervals for Beck Score, Recent IV Drug Use and Treatment for the UIS (n = 575)

Variable	Hazard Ratio	95% CIE
BECKTOTA*	1.09	0.99, 1.20
IVHX_3	1.26	1.02, 1.55
TREAT	0.78	0.65, 0.94

* Hazard ratio for a 10-point increase.

this covariate only is presented, with the other covariates in the model relegated to footnote status. We feel that this is not good statistical or scientific practice. With such an oversimplified summary, the reader has no way of evaluating whether an appropriate model building and model checking paradigm has been followed or what the actual fitted model contains. We feel that the full model should be presented in a table similar to Table 6.6 at some point in the results section.

The number of previous drug treatments is modeled with two nonlinear terms, so any hazard ratio will depend on the values of the number of previous drug treatments being compared. The graph in Figure 5.1d of the log hazard using the two non-linear terms shows an initial decrease in risk followed by a progressive nonlinear increase. One possible strategy would be to compare the hazard ratio for an increase of one in the number of previous drug treatments (i.e., hazard ratios for 0 vs. 1, 1 vs. 2, 2 vs. 3, etc.). These hazard ratios could either be tabulated or presented graphically, along with their confidence limits. We will do the latter.

One must proceed carefully when calculating hazard ratios for nonlinear functions of a covariate. The first step is to write down the expression for the log-hazard function, keeping all the other covariates constant. For ease of presentation, let the log-hazard function computed at a particular value of the number of previous drug treatments, NDRUG, holding all other covariates fixed and denoted as $\mathbf{z}$, be

$$g(\text{NDRUGTX},\mathbf{z}) = \beta_1 \text{NDRUGFP1} + \beta_2 \text{NDRUGFP2} + \boldsymbol{\beta}'\mathbf{z},$$

where

$$\text{NDRUGFP1} = [10/(\text{NDRUGTX}+1)]$$

and

$$\text{NDRUGFP2} = \text{NDRUGFP1} \times \ln[(\text{NDRUGTX}+1)/10].$$

The next step is to write down the equation for the difference of interest in the log hazard function. In this case it is, for a one-unit increase,

$$g(\text{NDRUGTX}+1,\mathbf{z}) - g(\text{NDRUGTX},\mathbf{z}).$$

Note that the first term in the difference in the log-hazard functions is the log-hazard evaluated at a one-unit increase in the previous number of drug treatments, not a one-unit increase in the nonlinear transformation. If we denote the values of the transformed variables at the increased value as

$$\text{NDRUGFP11} = [10/(\text{NDRUGTX}+2)]$$

and

$$\text{NDRUGFP21} = \text{NDRUGFP11} \times \ln[(\text{NDRUGTX}+2)/10],$$

then the difference in the log-hazard functions is

$$g(\text{NDRUGTX}+1,\mathbf{z}) - g(\text{NDRUGTX},\mathbf{z}) = a\beta_1 + b\beta_2, \qquad (6.27)$$

where

$$\begin{aligned} a &= \text{NDRUGFP11} - \text{NDRUGFP1} \\ &= [10/(\text{NDRUGTX}+2)] - [10/(\text{NDRUGTX}+1)] \end{aligned}$$

and

$$b = \text{NDRUGFP21} - \text{NDRUGFP2}.$$

The estimated difference in the log-hazard function is obtained by evaluating (6.27) using the values of the coefficients from Table 6.6, $\hat{\beta}_1 = -0.574$ and $\hat{\beta}_2 = -0.215$, and NDRUGTX = 0, 1, 2, 3, etc. The estimated hazard ratios are then obtained by exponentiating the estimated differences in the log-hazard functions,

$$\hat{\text{HR}}(\text{NDRUGTX}+1,\text{NDRUGTX},\mathbf{z}) = \exp\left(a\hat{\beta}_1 + b\hat{\beta}_2\right). \qquad (6.28)$$

The estimator of the endpoints of the $100(1-\alpha)$ percent confidence interval for the difference in the log-hazard functions is

$$\left(a\hat{\beta}_1 + b\hat{\beta}_2\right) \pm z_{1-\alpha/2}\hat{\text{SE}}\left(a\hat{\beta}_1 + b\hat{\beta}_2\right), \qquad (6.29)$$

where

$$\hat{\text{SE}}\left(a\hat{\beta}_1 + b\hat{\beta}_2\right) = \left[a^2\hat{\text{Var}}\left(\hat{\beta}_1\right) + b^2\hat{\text{Var}}\left(\hat{\beta}_2\right) + 2ab\hat{\text{Cov}}\left(\hat{\beta}_1,\hat{\beta}_2\right)\right]^{0.5}. \qquad (6.30)$$

The estimators of the variances and covariance in (6.30) may be obtained from output of the covariance matrix of the estimated coefficients from software packages. In the current example these values are $\hat{\text{Var}}(\hat{\beta}_1) = 0.015672$, $\hat{\text{Var}}(\hat{\beta}_2) = 0.002361$ and $\hat{\text{Cov}}(\hat{\beta}_1, \hat{\beta}_2) = 0.006022$. The endpoints of the confidence interval estimator for the hazard ratio are obtained by exponentiating the estimators in (6.29).

We calculated the value of the hazard ratio in (6.28) for the entire range of values for number of previous drug treatments, 0–40, and observed that, after about 10 previous treatments, there was not much change in the hazard ratio for a one-unit increase. Thus, we present in Figure 6.9 the graph of the hazard ratio and its confidence interval for up to 10 previous drug treatments.

The point estimate of the hazard ratio at 0 previous drug treatments in Figure 6.9 is 0.70. The interpretation is that subjects who have had one previous drug treatment are returning to drug use at a rate that is 30 percent lower than subjects who have had no previous drug treatments.

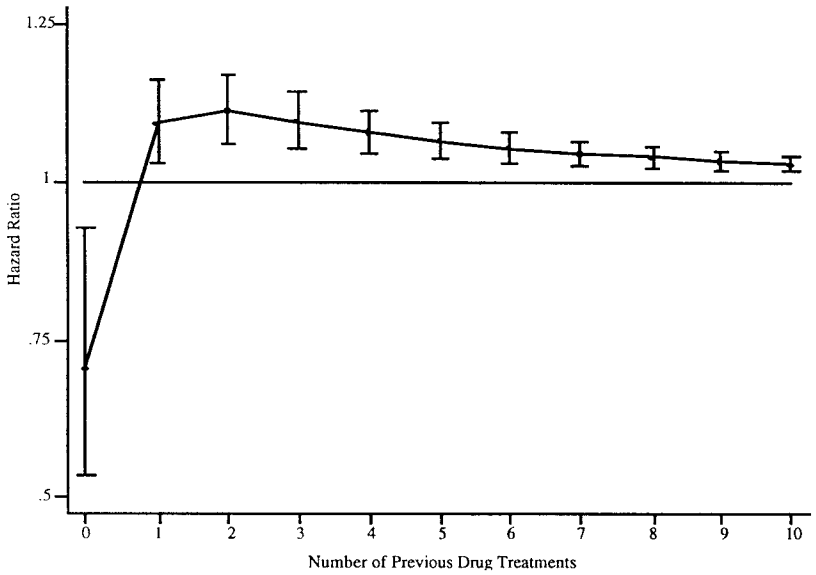

Figure 6.9 Graph of the estimated hazard ratio and associated 95 percent confidence interval for a one unit increase in the number of drug treatments from the labeled value.

The point estimate of the hazard ratio at 1 in Figure 6.9 is 1.09. This means that subjects with two previous treatments are returning to drug use at a rate that is 9 percent higher than subjects with one previous drug treatment. The estimate of the hazard ratio at 2 is 1.11, and then it slowly falls to 1.03 at 10. Since the estimated hazard ratios in Figure 6.9 exceed 1 at one or more previous drug treatments, the graph indicates a continuing increase in risk as the number of treatments increase; but the increase is progressively less, that is, the change in rate from 1 to 2 is much greater than the change from 9 to 10. We note that, since none of the confidence intervals include 1.0, the increase in rate of return to drugs is significant at all values.

An alternative presentation is obtained if we define one previous treatment as a common reference value. The resulting estimated hazard ratios could be either graphed or tabulated. We present, in Table 6.8, the estimated hazard ratios and corresponding confidence intervals comparing NDRUGTX = 0,2,5,10 to NDRUGTX = 1. These results are obtained by using (6.27)–(6.30), with a change in the values being compared. The results in Table 6.8, in the column labeled "0," are the reciprocals of the values shown in Figure 6.9 for 0 versus 1 previous drug treatments, while the values for column "2" are the same as those in Figure 6.9 for 1 versus 2 previous drug treatments. The values in the columns for 5 and 10 previous treatments cannot be obtained from Figure 6.9. The results in Table 6.8 demonstrate in a more direct manner the increase in hazard rate relative to the modeled minimum at 1. We note that the rates at 0 and 5 previous treatments are about 40 percent higher than the rate at 1. Although not shown, the rates increase progressively, so that the point estimate for 40 versus 1 is about 2.5 (1.70, 3.74).

Age and site are present in the model, with both main effects and their interaction. Since site is at two levels and is fixed by design of the study, we present hazard ratios for age at each site rather than for site at each age. The process is essentially the same as the one used to compute hazard ratios for the number of previous drug treatments, but it is a

Table 6.8 Estimated Hazard Ratios and 95% Confidence Intervals for the Stated Number of Previous Drug Treatments versus One Treatment.

	0	2	5	10
HR	1.42	1.09	1.44	1.82
95% CIE	(1.08, 1.87)	(1.03, 1.16)	(1.22, 1.71)	(1.40, 2.36)

bit simpler as we don't have to deal with nonlinear scaling. Again, the first step is to write down the equation for the hazard ratio as a function of the variables of interest, holding all the others fixed,

$$g(\text{AGE},\text{SITE},\mathbf{z}) = \beta_1\text{AGE} + \beta_2\text{SITE} + \beta_3\text{AGE}\times\text{SITE} + \boldsymbol{\beta}'\mathbf{z}.$$

The second step is to write down the expression for the difference of interest, in this case an increase of c years of age holding SITE fixed:

$$\begin{aligned} &g(\text{AGE}+c,\text{SITE},\mathbf{z}) - g(\text{AGE},\text{SITE},\mathbf{z}) \\ &\quad = \{\beta_1(\text{AGE}+c) + \beta_2\text{SITE} + \beta_3(\text{AGE}+c)\times\text{SITE} + \boldsymbol{\beta}'\mathbf{z}\} \\ &\qquad - \{\beta_1(\text{AGE}) + \beta_2\text{SITE} + \beta_3(\text{AGE})\times\text{SITE} + \boldsymbol{\beta}'\mathbf{z}\} \\ &\quad = \beta_1 c + \beta_3 c\times\text{SITE}. \end{aligned} \tag{6.31}$$

The next step is to choose a value for c, say 5 years, and to estimate the value of (6.31) using the estimated coefficient of AGE from Table 6.6, $\hat{\beta}_1 = -0.041$, and the estimated coefficient of the interaction of AGE and SITE, $\hat{\beta}_3 = 0.032$. The estimated hazard ratio for an increase of 5 years of AGE at SITE = 0 is

$$\hat{\text{HR}}(\text{AGE}+5,\text{AGE},\text{SITE}=0) = e^{5\times(-0.041)} = 0.815,$$

and at SITE = 1 it is

$$\hat{\text{HR}}(\text{AGE}+5,\text{AGE},\text{SITE}=1) = e^{5\times(-0.041)+5\times0.032} = 0.956.$$

The endpoints of the $100(1-\alpha)$ percent confidence interval estimator of the hazard ratio are computed by exponentiating the endpoints of the confidence interval of the estimator of (6.32), which are

$$\hat{\beta}_1 c + \hat{\beta}_3 c\times\text{SITE} \pm z_{1-\alpha/2}\hat{\text{SE}}\left(\hat{\beta}_1 c + \hat{\beta}_3 c\times\text{SITE}\right), \tag{6.32}$$

where

$$\hat{\text{SE}}\left(\hat{\beta}_1 c + \hat{\beta}_3 c\times\text{SITE}\right) = \left[\begin{array}{c} c^2\times\hat{\text{Var}}\left(\hat{\beta}_1\right) + c^2\times\text{SITE}^2\times\hat{\text{Var}}\left(\hat{\beta}_3\right) \\ +2c^2\times\text{SITE}\times\hat{\text{Cov}}\left(\hat{\beta}_1,\hat{\beta}_3\right) \end{array}\right]^{0.5}, \tag{6.33}$$

and since SITE is coded zero or one, $\text{SITE}^2 = \text{SITE}$. Again the values of the variances and covariance needed to compute the standard error are available from software packages. In this example, these are

$$\hat{\text{Var}}\left(\hat{\beta}_1\right) = 0.000098,$$

$$\hat{\text{Var}}\left(\hat{\beta}_3\right) = 0.000259$$

and

$$\hat{\text{Cov}}\left(\hat{\beta}_1,\hat{\beta}_3\right) = -0.00009.$$

Using these values, the 95 percent confidence interval at SITE = 0 is (0.739, 0.898) and at SITE = 1 is (0.839, 1.089). The fact that age is linear within site means that these results hold at all ages.

The interpretation is that being older by 5 years at SITE = 0 reduces the rate of return to drug use by about 18 percent, and the fact that 1.0 is not contained in the confidence interval points to a significant age effect at this site. At the other site, SITE = 1, the rate is only 5 percent lower and is not significant.

The remaining hazard ratio involves race and site. We present the hazard ratio and its confidence interval for race at each site. These may be obtained by using (6.27)–(6.33) with $c = 1$, reflecting the fact that race is dichotomous and has been recoded zero and one. The estimated hazard ratio and 95 percent confidence interval for RACE = other versus RACE = white at SITE = 0 is

$$\hat{\text{HR}}(\text{other},\text{white},\text{SITE}=0) = e^{-0.467} = 0.627$$

and (0.481, 0.816) while at SITE = 1 they are

$$\hat{\text{HR}}(\text{other},\text{white},\text{SITE}=1) = e^{-0.467+0.850} = 1.467$$

and (0.972, 2.214). The interpretation is that non-whites are returning to drug use at a rate that is about 37 percent lower than whites at SITE = 0. The confidence interval for the ratio suggests that the rate could be as much as 52 percent lower to only about 18 percent lower. The reverse seems to be the case at SITE = 1, where the non-whites are returning to drug use at much higher rate than whites (about 46 percent

higher), and the confidence interval suggests that this could be as high as 121 percent or even slightly negative. In any event, the results point not only to important racial differences, but to differences between the two sites. This was discussed with the study team and deemed to be an appropriate interpretation.

In this section we have had to emphasize both the calculation and interpretation of the estimated hazard ratios. In practice, the estimated hazard ratios and their confidence intervals would likely be tabulated with no computational details presented and thus lend themselves to a discussion with more continuity than was possible here. However, for a complicated nonlinear variable like the number of previous drug treatments, inclusion of an appendix providing an outline of how the graphed (or tabulated) hazard ratios and their confidence intervals have been computed can be a helpful addition to a paper.

We conclude our presentation of the fitted model in Table 6.6 with graphs of the covariate-adjusted survivorship functions for the two levels of treatment. Since the model is complicated, it is not clear what we could use for a mean or median subject, so we use the modified risk score method discussed in Section 4.3 and illustrated in (4.29) and (4.30). The modified risk score is calculated for each subject as $r\hat{m}_i = \hat{r}_i - (-0.247)\text{TREAT}_i$ and the median is $r\hat{m}_{50} = -2.088$. The plotted points for the covariate-adjusted survivorship function for the shorter treatment are

$$\hat{S}(t_i, r\hat{m}_{50}) = \left[\hat{S}_0(t_i)\right]^{\exp(-2.088)}$$

and for the longer treatment are

$$\hat{S}(t_i, r\hat{m}_{50}) = \left[\hat{S}_0(t_i)\right]^{\exp(-2.088-0.2468)}.$$

Graphs of these two functions, at all observed values of time, are shown in Figure 6.10.

The figure shows that, at all times, the covariate-adjusted proportion of subjects who have not returned to drug use is higher for the longer of the two treatments. The covariate-adjusted estimated median times to return to drug may be obtained from the graphs in Figure 6.10 or from a time-sorted list of the functions. The median times are 157 days for the shorter treatment and 190 days for the longer treatment. As noted in Section 4.5, there is no easily computed confidence interval for the estimator of the median time from a modified risk-score-adjusted survivorship function.

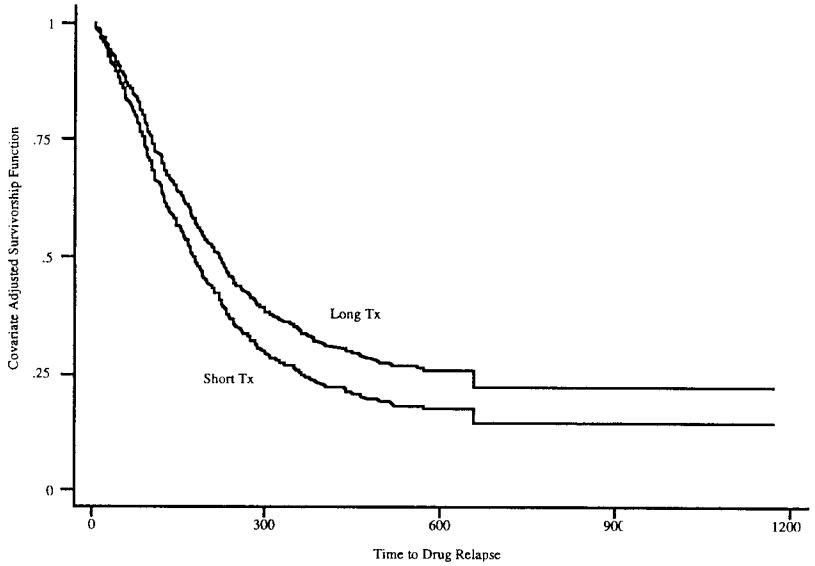

Figure 6.10 Graphs of the covariate-adjusted survivorship functions for the longer and shorter treatments computed using the model in Table 6.6.

In conclusion, the fitted model, shown in Table 6.6, has allowed description of a number of interesting relationships between time to return to drug use and study factors. The notable results include, besides a significant treatment effect, a differential effect of race within treatment site, the nonlinear effect of the number of previous drug treatments and significant effects due to Beck score and recent IV drug use.

In the next chapter we consider alternative methods for modeling study covariates. These methods are of interest in themselves, but they often also provide alternatives to models that are not adequate due to poor fit or violations of the proportional hazards assumption.

EXERCISES

1. Using data from the HMO-HIV+ study, assess the fit the proportional hazards model containing AGE and DRUG. This assessment of fit should include the following steps: evaluation of the proportional haz-

ards assumption for each of the two covariates, examination of diagnostic statistics, and an overall test of fit. If the model does not fit or adhere to the proportional hazards assumption what would you do next? Note: the goal is to obtain a model to estimate the effect of AGE and DRUG on the survivorship experience.

2. Using the model obtained at the conclusion of problem 1, present a table of estimated hazard ratios, with confidence intervals. Present graphs of the age-adjusted, at the mean age, estimated survivorship functions for the two drug use groups. Use the estimated survivorship functions to estimate the age-adjusted median survival time for each of the two drug use groups.

3. In Section 6.4 diagnostic statistics were plotted and a few subjects were identified as being possibly influential. Fit the model shown in Table 6.6 deleting these subjects one at a time and then, collectively, calculate the percent change in all coefficients with each deletion. Do you agree or disagree with the conclusion in Section 6.4 to keep all subjects in the analysis? Explain the rationale for your decision.

4. A considerable amount of the material presented in this chapter dealt with the evaluation of fit, and the presentation and interpretation of the fitted model shown in Table 5.11 (and repeated in Table 6.6). Repeat the entire process for the fitted model shown in Table 5.13. This model contains an interaction b 'ween AGE and NDRUGFP1 and, as a result, estimation and presentation of hazard ratios for age, controlling for the number of previous drug treatments and for the number of previous drug treatments controlling for age, is a major challenge.

5. Repeat the full model evaluation and presentation process using the fitted model developed for the WHAS in problems 3 of Chapter 5.

CHAPTER 7

Extensions of the Proportional Hazards Model

7.1 INTRODUCTION

Up to this point we have made several simplifying assumptions in developing and interpreting proportional hazards models. We have used a proportional hazards model with a common unspecified baseline hazard function where all the study covariates had values that remained fixed over the follow-up period. Additionally, we have assumed that the observations of the time variable were continuous and subject only to right censoring. In some settings one or more of these assumptions may not be appropriate.

We may have data from a study in which subjects were randomized within study sites. If we account for site by including it as a covariate, the model forces the baseline hazards to be proportional across study sites. This may not be justified and, if it isn't, a careful analysis of the proportional hazards assumption (as discussed in Chapter 6) for site should reveal the problem. One possible solution is to use site as a stratification variable, whereby each site would have a separate baseline hazard function.

When study subjects are observed on a regular basis during the follow-up period, the course of some covariates over time may be more predictive of survival experience than the original baseline values. For example, continued survival of intensive care unit patients may depend more on changes in their physiologic condition since admission than on their absolute state at admission. Covariates whose values change over time are commonly called *time-varying* or *time-dependent* covariates. These may include measurements on individual subjects or measure-