5 The Stratified Cox Procedure

D.G. Kleinbaum and M. Klein, *Survival Analysis: A Self-Learning Text, Third Edition,*
Statistics for Biology and Health, DOI 10.1007/978-1-4419-6646-9_5,

Introduction

We begin with an example of the use of the stratified Cox procedure for a single predictor that does not satisfy the PH assumption. We then describe the general approach for fitting a stratified Cox model, including the form of the (partial) likelihood function used to estimate model parameters.

We also describe the assumption of no interaction that is typically incorporated into most computer programs that carry out the stratified Cox procedure. We show how the no-interaction assumption can be tested, and what can be done if interaction is found.

We conclude with a second example of the stratified Cox procedure in which more than one variable is stratified.

Abbreviated Outline

The outline below gives the user a preview of the material to be covered by the presentation. A detailed outline for review purposes follows the presentation.

Objectives

Upon completing the chapter, the learner should be able to:

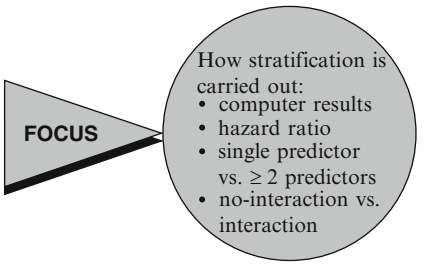
1. Recognize a computer printout for a stratified Cox procedure.
2. State the hazard form of a stratified Cox model for a given survival analysis scenario and/or a given set of computer results for such a model.
3. Evaluate the effect of a predictor of interest based on computer results from a stratified Cox procedure.
4. For a given survival analysis scenario and/or a given set of computer results involving a stratified Cox model.
   - state the no-interaction assumption for the given model;
   - describe and/or carry out a test of the no-interaction assumption;
   - describe and/or carry out an analysis when the no-interaction assumption is not satisfied.

Presentation

I. Preview

Stratified Cox model:

- modification of Cox PH model
- Stratification of predictor not satisfying PH
- includes predictors satisfying PH



The "stratified Cox model" is a modification of the Cox proportional hazards (PH) model that allows for control by "stratification" of a predictor that does not satisfy the PH assumption. Predictors that are assumed to satisfy the PH assumption are included in the model, whereas the predictor being stratified is not included.

In this presentation, we focus on how stratification is carried out by describing the analysis of computer results and the form of the hazard function for a stratified Cox model. We first consider stratifying on a single predictor and then later consider stratifying on two or more predictors. Further, we distinguish between the use of a "no-interaction" version of the stratified Cox model and an alternative approach that allows interaction.

II. An Example

EXAMPLE

Clinical trial: 42 leukemia patients
Response-days in remission

	Coef.	Std. Err.	*P*(*PH*)
log WBC	1.594	0.330	0.828
Rx	1.391	0.457	0.935
Sex	0.263	0.449	0.031

- log WBC and *Rx* satisfy PH
- Sex does not satisfy PH

(Same conclusions using graphical approaches)

Stratified Cox (SC):

- control for sex (stratified);
- simultaneously include log WBC and *Rx* in the model

Consider the computer results shown here for a Cox PH model containing the three variables, log WBC, treatment group (*Rx*), and SEX. These results derive from a clinical trial of 42 leukemia patients, where the response of interest is days in remission.

From the printout, the *P*(*PH*) values for log WBC and treatment group are nonsignificant. However, the *P*(*PH*) value for SEX is significant below the .05 level. These results indicate that log WBC and treatment group satisfy the PH assumption, whereas the SEX variable does not. The same conclusions regarding the PH assumption about these variables would also be made using the graphical procedures described earlier.

Because we have a situation where one of the predictors does not satisfy the PH assumption, we carry out a stratified Cox (SC) procedure for the analysis. Using SC, we can control for the SEX variable – which does not satisfy the PH assumption – by stratification while simultaneously including in the model the log WBC and treatment variables – which do satisfy the PH assumption.

EXAMPLE: (continued)

STATA OUTPUT USING SC:
Stratified Cox regression
Analysis time_t: survt

	Coef.	Std. Err.	p > |z|	Haz. Ratio	[95% Conf.	Interval]
log WBC	1.390	0.338	0.000	4.016	2.072	7.783
Rx	0.931	0.472	0.048	2.537	1.006	6.396

No. of subjects = 42 Log likelihood = −57.560 Stratified by sex

The computer results from a SC procedure are shown here. These results come from the Stata package. (See the Computer Appendix for running a SC procedure in Stata, SAS, SPSS, and R).

Appendix A illustrates SC procedures using Stata, SAS, SPSS, and R

- Log WBC and *Rx* are included in SC model.
- SC model is stratified by SEX.

The computer results show that the log WBC and *Rx* variables are included in the model listing, whereas the SEX variable is not included; rather, the model stratifies on the SEX variable, as indicated at the bottom of the output. Note that the SEX variable is being adjusted by stratification, whereas log WBC is being adjusted by its inclusion in the model along with *Rx*.

Effect of *Rx* adjusted for log WBC and SEX.

- Hazard ratio: $2.537 = e^{0.931}$
- Interpretation: Placebo group ($Rx = 1$) has 2.5 times the hazard as the treatment group (Rx = 0)

In the above output, we have also circled some key information that can be used to assess the effect of the *Rx* variable adjusted for both log WBC and SEX. In particular, we can see that the hazard ratio for the effect of *Rx* adjusted for log WBC and SEX is given by the value 2.537. This value can be obtained by exponentiating the coefficient 0.931 of the *Rx* variable. The hazard ratio value can be interpreted to mean that the placebo group (for which $Rx = 1$) has 2.5 times the hazard for going out of remission as the treatment group (for which $Rx = 0$).

Stratified Cox regression
Analysis time_t: survt

	Coef.	Std. Err.	p > |z|	Haz. Ratio	[95% Conf.	Interval]
log WBC	1.390	0.338	0.000	4.016	2.072	7.783
Rx	0.931	0.472	0.048	2.537	1.006	6.396

No. of subjects = 42 Log likelihood = −57.560 Stratified by sex

95% CI for *Rx* (1.006, 6.396) indicates considerable variability.
CI formula: $\exp(0.931 \pm 1.96 \times 0.472)$

Also, we can see from the output that a 95% confidence interval for the effect of the *Rx* variable is given by the limits 1.006 to 6.396. This is a fairly wide range, thus indicating considerable variability in the 2.537 hazard ratio point estimate. Note that these confidence limits can be obtained by exponentiating the quantity 0.931 plus or minus 1.96 times the standard error 0.472.

Wald test: P = 0.048 (two-tailed), significant at the 0.05 level.

From the above output, a test for the significance of the *Rx* variable adjusted for log WBC and SEX is given by the Wald statistic P-value of 0.048. This is a two-tailed P-value, and the test is (barely) significant at the 0.05 level.

EXAMPLE: (continued)

LR test: Output for reduced model
Stratified Cox regression
Analysis time _t: survt

	Coef.	Std. Err.	p > |z|	Haz. Ratio	[95% Conf.	Interval]
log WBC	1.456	0.320	0.000	4.289	2.291	8.03

No. of subjects = 42 Log likelihood = −59.648 Stratified by sex

$$\text{LR} = (-2 \times -59.648) - (-2 \times -57.560)$$
$$= 119.296 - 115.120 = 4.179\ (\text{P} < 0.05)$$

LR and Wald give same conclusion.

Hazard function for stratified Cox model:

$h_g(t,\mathbf{X}) = h_{0g}(t)\exp[\beta_1 Rx + \beta_2 \log \text{WBC}]$
$g = 1,2;$

g denotes stratum #.

SC model for males and females:
Females ($g = 1$):

$h_1(t,\mathbf{X}) = h_{01}(t)\exp[\beta_1 Rx + \beta_2 \log \text{WBC}]$

Males ($g = 2$):

$h_2(t,\mathbf{X}) = h_{02}(t)\exp[\beta_1 Rx + \beta_2 \log \text{WBC}]$

Rx and log WBC in the model
Sex *not* in the model (stratified)

$\widehat{HR}$ for effect of *Rx* adjusted for log WBC and sex:

$e^{\hat{\beta}_1}$

where β_1 is the coefficient of *Rx*.

An alternative test involves a likelihood ratio (*LR*) statistic that compares the above model (full model) with a reduced model that does not contain the *Rx* variable. The output for the reduced model is shown here. The log-likelihood statistic for the reduced model is −2 times −59.648, which is to be compared with the log-likelihood statistic of −2 times −57.560 for the full model.

The *LR* statistic is therefore 119.296 minus 115.120, which equals 4.179. Under H_0, this statistic has a chi-square distribution with one degree of freedom and is significant at the 0.05 level. Thus, the *LR* and Wald tests lead to the same conclusion.

So far, we have illustrated the results from a stratified Cox procedure without actually describing the model form being used. For the remission data example, we now present the hazard function form for the stratified Cox model, as shown here. This hazard function formula contains a subscript g that indicates the gth stratum.

Thus, in our remission data example, where we have stratified on SEX, g takes on one of two values, so that we have a different baseline hazard function for males and females.

Notice that the hazard function formula contains the variables *Rx* and log WBC, but does not contain the variable SEX. SEX is not included in the model because it doesn't satisfy the PH assumption. So, instead, the SEX variable is controlled by stratification.

Because the variables *Rx* and log WBC are included in the model, we can estimate the effect of each variable adjusted for the other variable and the SEX variable using standard exponential hazard ratio expressions. For example, the estimated hazard ratio for the effect of *Rx*, adjusted for log WBC and SEX, is given by e to the β_1 "hat," where β_1 is the coefficient of the *Rx* variable.

EXAMPLE: (continued)

Cannot estimate *HR* for SEX variable (SEX doesn't satisfy PH).

Nevertheless, because the SEX variable is not included in the model, it is not possible to obtain a hazard ratio value for the effect of SEX adjusted for the other two variables. This is the price to be paid for stratification on the SEX variable. Note that a single value for the hazard ratio for SEX is not appropriate if SEX doesn't satisfy the PH assumption, because the hazard ratio must then vary with time.

Different baseline hazard functions:

$h_{01}(t)$ for females and $h_{02}(t)$ for males.

Same coefficients β_1 and β_2 for both female and male models.

Notice also that the hazard functions for males and females differ only insofar as they have different baseline hazard functions, namely, $h_{01}(t)$ for females and $h_{02}(t)$ for males. However, the coefficients β_1 and β_2 are the same for both female and male models.

Different baselines $\begin{cases} h_{01}(t) \Rightarrow \text{Survival curve for females} \\ h_{02}(t) \Rightarrow \text{Survival curve for males} \end{cases}$

Because there are different baseline hazard functions, the fitted stratified Cox model will yield different estimated survival curves for females and males. These curves will be described shortly.

Females and males:
same β_1 and $\beta_2 \Rightarrow$ same $\widehat{HR}$'s, e.g., $e^{\hat{\beta}_1}$

No interaction assumption (see Section IV)

Note, however, that because the coefficients of *Rx* and log WBC are the same for females and males, estimates of hazard ratios, such as *e* to the β_1 "hat," are the same for both females and males. This feature of the stratified Cox model is called the "no-interaction" assumption. It is possible to evaluate whether this assumption is tenable and to modify the analysis if not tenable. We will discuss this assumption further in Section IV.

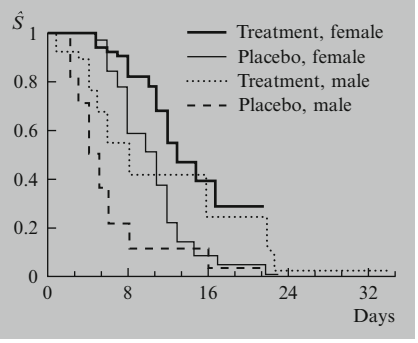
Estimates of β_1 and β_2

Maximize partial likelihood (L), where $L = L_1 \times L_2$
L_1 is the likelihood for females derived from $h_1(t)$,
and L_2 is the likelihood for males derived from $h_2(t)$.

To obtain estimates of β_1 and β_2, a (partial) likelihood function (L) is formed from the model and the data; this function is then maximized using computer iteration. The likelihood function (L) for the stratified Cox (SC) model is different from the nonstratified Cox model. For the SC model, L is obtained by multiplying together likelihood functions for each stratum. Thus, L is equal to the product of L_1 and L_2, where L_1 and L_2 denote the female and male likelihood functions, respectively, which are derived from their respective hazard functions $h_1(t)$ and $h_2(t)$.

EXAMPLE: (continued)

Adjusted Survival Curves for *Rx* from Stratified Cox Model (adjusted for log WBC)



As mentioned above, adjusted survival curves can be obtained for each stratum as shown here. Here we have shown *four* survival curves because we want to compare the survival for two treatment groups over each of two strata.

If we compare treatment and placebo group separately by sex, we can see that the treatment group has consistently better survival prognosis than the placebo group for females and males separately. This supports our findings about the hazard ratio for the treatment effect derived earlier from the computer results for the stratified Cox model.

III. The General Stratified Cox (SC) Model

In the previous example, we illustrated the SC model for one binary predictor not satisfying the PH assumption. We now describe the general form of the SC model that allows for stratification of several predictors over several strata.

Example: one binary predictor
↓
General: several predictors, several strata

We assume that we have k variables not satisfying the PH assumption and p variables satisfying the PH assumption. The variables not satisfying the PH assumption we denote as $Z_1, Z_2, \ldots, Z_k$; the variables satisfying the PH assumption we denote as $X_1, X_2, \ldots, X_p$.

$Z_1, Z_2, \ldots, Z_k$, do not satisfy PH
$X_1, X_2, \ldots, X_p$, satisfy PH

Define a single new variable Z^*:

1. categorize each Z_i
2. form combinations of categories (strata)
3. the strata are the categories of Z^*

To perform the stratified Cox procedure, we define a single new variable, which we call Z^*, from the Z's to be used for stratification. We do this by forming categories of each Z_i, including those Z_i that are interval variables. We then form combinations of categories, and these combinations are our strata. These strata are the categories of the new variable Z^*.

EXAMPLE

		Age		
		Young	Middle	Old
Treatment status	Placebo	1	2	3
	Treatment	4	5	6

Z^* = new variable with six categories stratify on Z^*

For example, suppose k is 2, and the two Z's are age (an interval variable) and treatment status (a binary variable). Then we categorize age into, say, three age groups – young, middle, and old. We then form six age group–by–treatment-status combinations, as shown here. These six combinations represent the different categories of a single new variable that we stratify on in our stratified Cox model. We call this new variable Z^*.

Z^* has k^* categories where k^* = total # of combinations (strata), e.g., $k^* = 6$ in above example.

In general, the stratification variable Z^* will have k^* categories, where k^* is the total number of combinations (or strata) formed after categorizing each of the Z's. In the above example, k^* is equal to 6.

The general SC model:
$h_g(t, \mathbf{X}) = h_{0g}(t)\exp[\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p]$
$g = 1, 2, \ldots, k^*$, strata defined from Z^*

We now present the general hazard function form for the stratified Cox model, as shown here. This formula contains a subscript g which indicates the gth stratum. The strata are defined as the different categories of the stratification variable Z^*, and the number of strata equals k^*.

Z^* not included in the model

$X_1, X_2, \ldots, X_p$ included in the model

Note that the variable Z^* is not explicitly included in the model but that the X's, which are assumed to satisfy the PH assumption, are included in the model.

Different baseline hazard functions: $h_{0g}(t)$, $g = 1, 2, \ldots, k^*$
Same coefficients: $\beta_1, \beta_2, \ldots, \beta_p$

Different baselines:
$\hat{h}_{01}(t) \Rightarrow \hat{S}_1(t)$
$\hat{h}_{02}(t) \Rightarrow \hat{S}_2(t)$
$\vdots$
$\hat{h}_{0k}(t) \Rightarrow \hat{S}_k(t)$
→ Different survival curves

Note also that the baseline hazard function $h_{0g}(t)$ is allowed to be different for each stratum. However, the coefficients $\beta_1, \beta_2, \ldots, \beta_p$ are the same for each stratum.

As previously described by example, the fitted SC model will yield different estimated survival curves for each stratum because the baseline hazard functions are different for each stratum.

$\widehat{HR}$ same for each stratum

(no-interaction assumption, Section IV)

However, because the coefficients of the X's are the same for each stratum, estimates of hazard ratios are the same for each stratum. This latter feature of the SC model is what we previously have called the "no-interaction" assumption to be discussed further in Section IV.

(Partial) likelihood function:

$L = L_1 \times L_2 \times \ldots \times L_{k^*}$

Strata:	1	2	...	k^*
Likelihood:	L_1	L_2	...	L_{k^*}
Hazard:	$h_1(t, \mathbf{X})$	$h_2(t, \mathbf{X})$	...	$h_{k^*}(t, \mathbf{X}^*)$

To obtain estimates of the regression coefficients $\beta_1, \beta_2, \ldots, \beta_p$, we maximize a (partial) likelihood function L that is obtained by multiplying together likelihood functions for each stratum, as shown here. Thus, L is equal to the product of L_1 times L_2, and so on, up until L_{k^*}, where the subscripted L's denote the likelihood functions for different strata, with each of these L's being derived from its corresponding hazard function.

IV. The No-Interaction Assumption and How to Test It

We previously pointed out that the SC model contains regression coefficients, denoted as β's, that do not vary over the strata. We have called this property of the model the "no-interaction assumption." In this section, we explain what this assumption means. We also describe how to evaluate the assumption and what to do if the assumption is violated.

Stratified Cox model

$$h_g(t,\mathbf{X}) = h_{0g}(t)\exp\left[\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p\right]$$

β coefficients do not vary over strata (no-interaction assumption)

- how to evaluate
- what to do if violated

EXAMPLE

No-interaction SC model:
Stratified Cox regression
Analysis time_t: survt

	Coef.	Std. Err.	p>|z|	Haz. Ratio	[95% Conf.	Interval]
log WBC	1.390	0.338	0.000	4.016	2.072	7.783
Rx	0.931	0.472	0.048	2.537	1.006	6.396

No. of subjects = 42 Log likelihood = −57.560 Stratified by sex

Interaction by fitting separate models:
Cox regression (Females)
Analysis time _t: survt

Column name	Coeff.	StErr.	p-value	*HR*	0.95	CI	*p* (*PH*)
log WBC	1.639	0.519	0.002	5.150	1.862	14.242	0.228
Rx	1.859	0.729	0.011	6.418	1.537	26.790	0.603

No. of subjects = 20 Log likelihood = −22.100

Cox regression (Males)
Analysis time _t: survt

Column name	Coeff.	StErr.	p-value	*HR*	0.95	CI	*P* (*PH*)
log WBC	1.170	0.499	0.019	3.222	1.213	8.562	0.674
Rx	0.267	0.566	0.637	1.306	0.431	3.959	0.539

No. of subjects = 22 Log likelihood = −33.736

Which model is more appropriate statistically?

We return to the SC output previously illustrated. Notice that only one set of coefficients, namely, 1.390 for log WBC and 0.931 for *Rx*, are provided, even though there are two strata, one for females and one for males. These results assume no interaction of the sex variable with either log WBC or *Rx*.

If we allow for interaction, then we would expect to obtain different coefficients for each of the (SEX) strata. This would happen if we fit separate hazard models to the female and male data, with each model containing the log WBC and Rx variables. The computer results from fitting separate models are shown here.

Notice that the coefficient of log WBC is 1.639 for females but is 1.170 for males. Also, the coefficient for *Rx* is 1.859 for females but 0.267 for males. These results show different coefficients for females than for males, particularly for the *Rx* variable.

But are corresponding coefficients statistically different? That is, which model is more appropriate statistically, the no-interaction model or the interaction model? To answer this question, we must first look at the hazard function model for the interaction situation.

EXAMPLE: (continued)

Interaction model:
(◆) $h_g(t,\mathbf{X})$
$= h_{0g}(t)\exp[\beta_{1g}\log \text{WBC} + \beta_{2g}Rx]$
where $g = 1$ (females), $g = 2$ (males)

No-interaction model:
$h_g(t,\mathbf{X})$
$= h_{0g}(t)\exp[\beta_1 \log \text{WBC} + \beta_2 Rx]$
where $g = 1$ (females), $g = 2$ (males)

Alternative interaction model:

$$(\star)h_g(t,\mathbf{X}) = h_{0g}(t)\exp\big[\beta_1^* \log \text{WBC} + \beta_2^* Rx + \beta_3^*(\text{SEX}\times\log \text{WBC}) + \beta_4^*\times(\text{SEX}\times Rx)\big]$$

$$\text{where SEX} = \begin{cases} 1 & \text{if female} \\ 0 & \text{if male} \end{cases}$$

$h_{0g}(t)$ are different for $g = 1, 2$
β^* coefficients do not involve g

Equivalence of models (◆) and (⋆):
$g = 1$ (females), so that sex = 1:

$$h_1(t,\boldsymbol{X}) = h_{01}(t)\exp\big[\beta_1^* \log \text{WBC} + \beta_2^* Rx + \beta_3^*(1\times\log \text{WBC}) + \beta_4^*(1\times Rx)\big]$$

$$= h_{01}(t)\exp\Big[\boxed{(\beta_1^*+\beta_3^*)}\log \text{WBC} + \boxed{(\beta_2^*+\beta_4^*)}\,\text{Rx}\Big]$$

$g = 2$ (males), so that sex = 0:

$$h_2(t,\boldsymbol{X}) = h_{02}(t)\exp\big[\beta_1^* \log \text{WBC} + \beta_2^* Rx + \beta_3^*(0\times\log \text{WBC}) + \beta_4^*(0\times Rx)\big]$$

$$= h_{02}(t)\exp\Big[\boxed{\beta_1^*}\log \text{WBC} + \boxed{\beta_2^*}Rx\Big]$$

Interaction models in same format:

Females ($g = 1$): $h_1(t,\mathbf{X})$

$$(\blacklozenge) = h_{01}(t)\exp[\beta_{11}\log \text{WBC} + \beta_{21}Rx]$$
$$(\star) = h_{01}(t)\exp\big[(\beta_1^*+\beta_3^*)\log \text{WBC} + (\beta_2^*+\beta_4^*)Rx\big]$$

Males ($g = 2$): $h_2(t,\mathbf{X})$

$$(\blacklozenge) = h_{02}(t)\exp[\beta_{12}\log \text{WBC} + \beta_{22}Rx]$$
$$(\star) = h_{02}(t)\exp\big[\beta_1^*\log \text{WBC} + \beta_2^*Rx\big]$$

One way to state the hazard model formula when there is **interaction** is shown here (◆). Notice that each variable in this model has a different coefficient for females than for males, as indicated by the subscript g in the coefficients β_{1g} and β_{2g}.

In contrast, in the **no-interaction** model, the coefficient (β_1) of log WBC is the same for females and for males; also, the coefficient (β_2) of Rx is the same for females and for males.

An alternative way to write the interaction model is shown here (⋆). This alternative form contains two product terms, SEX × log WBC and SEX × Rx, as well as the main effects of log WBC and Rx. We have coded the SEX so that 1 denotes female and 0 denotes male.

In this alternative model, note that although the baseline hazards $h_{0g}(t)$ are different for each sex, the β^* coefficients do not involve the subscript g and therefore are the same for each sex.

Nevertheless, this alternative formula (⋆) is equivalent to the interaction formula (◆) above. We show this by specifying the form that the model takes for $g = 1$ (females) and $g = 2$ (males).

Notice that the coefficients of log WBC are different in each formula, namely, $(\beta_1^*+\beta_3^*)$ for females versus β_1^* for males.

Similarly, the coefficients of Rx are different, namely, $(\beta_2^*+\beta_4^*)$ for females versus β_2^* for males.

The preceding formulae indicate that two seemingly different formulae for the interaction model – (◆) versus (⋆), shown earlier – can be written in the same format. We show these formulae here separately for females and males.

EXAMPLE: (continued)

$$\text{Females } (g=1): \begin{array}{ll} (\blacklozenge) & (\star) \\ \beta_{11} = \beta_1^* + \beta_3^* \\ \beta_{21} = \beta_2^* + \beta_4^* \end{array}$$

$$\text{Males } (g=2): \begin{array}{ll} (\blacklozenge) & (\star) \\ \beta_{12} = \beta_1^* \\ \beta_{22} = \beta_2^* \end{array}$$

Stratified Cox regression
Analysis time _t: survt

	Coef.	Std. Err:	p>|z|	Haz. Ratio	[95% Conf.	Interval]
log WBC	1.170	0.499	0.019	3.222	1.213	8.562
Rx	0.267	0.566	0.637	1.306	0.431	3.959
Sex × log WBC	0.469	0.720	0.515	1.598	0.390	6.549
Sex × *Rx*	1.592	0.923	0.084	4.915	0.805	30.003

No. of subjects = 42 Log likelihood = −55.835 Stratified by sex

Females:

$$\log \text{ WBC} \begin{cases} \beta_{11} = \boxed{1.639} \\ \hat{\beta}_1^* + \hat{\beta}_3^* = 1.170 + 0.469 = \boxed{1.639} \end{cases}$$

$$Rx \begin{cases} \beta_{21} = \boxed{1.859} \\ \hat{\beta}_2^* + \hat{\beta}_4^* = 0.267 + 1.592 = \boxed{1.859} \end{cases}$$

Males:

$$\log WBC \ \hat{\beta}_{12} = \boxed{1.170} = \hat{\beta}_1^*$$

$$Rx \qquad \hat{\beta}_{22} = \boxed{0.267} = \hat{\beta}_2^*$$

Interaction model:

$$h_g(t, \boldsymbol{X}) = h_{0g}(t) \exp\left[\beta_1^* \log \text{WBC} + \beta_2^* Rx + \beta_3^* (\text{SEX} \times \log \text{WBC}) + \beta_4^* \times (\text{SEX} \times Rx)\right]$$

Notice that for females, the coefficient β_{11} in model $(\star)$ must be equivalent to $(\beta_1^* + \beta_3^*)$ in model $(\blacklozenge)$ because both models have the same format, and both β_{11} and $(\beta_1^* + \beta_3^*)$ are coefficients of the same variable, log WBC. Similarly, β_{21} in model $(\star)$ is equivalent to $(\beta_2^* + \beta_4^*)$ in model $(\blacklozenge)$ because both are coefficients of the same variable, *Rx*.

For males, it follows in an analogous way, that the coefficient β_{12} is equivalent to β_1^*, and, similarly, β_{22} equals β_2^*.

Here we provide computer results obtained from fitting the alternative interaction model $(\star)$. The estimated regression coefficients $\hat{\beta}_1^*$, $\hat{\beta}_2^*$, $\hat{\beta}_3^*$, , and $\hat{\beta}_4^*$, respectively, are circled.

We have indicated above that the sums $\hat{\beta}_1^* + \hat{\beta}_3^*$ and $\hat{\beta}_2^* + \hat{\beta}_4^*$ are equal to the coefficients $\hat{\beta}_{11}$ and $\hat{\beta}_{21}$, respectively, in the original interaction model for females.

Also, we have indicated that $\hat{\beta}_1^*$ and $\hat{\beta}_2^*$ are equal to the coefficients $\hat{\beta}_{12}$ and $\hat{\beta}_{22}$, respectively, in the original interaction model for the males. The numerical equivalences are shown here. Note again that the coefficients of log WBC and *Rx* for females are different from males, as is to be expected if sex interacts with each variable.

We have thus seen that the interaction model can be written in a format that contains product terms involving the variable being stratified, SEX, being multiplied by each of the predictors not being stratified. We show this model involving product terms again here. We will use this model to describe a test of the no-interaction assumption.

EXAMPLE: (continued)

Testing the no-interaction assumption:

$LR = -2 \ln L_R - (-2 \ln L_F)$
R = reduced (no-interaction) model
F = full (interaction) model

$LR \dot{\sim} \chi^2_{2\,df}$ under H_0: no interaction
(2 df because two product terms tested in interaction model)

No interaction (reduced model):

Output: –2 log *L*: 115.120

$-2 \ln L_R$

Interaction (full model):

Output: –2 log *L*: 111.670

$-2 \ln L_F$

$LR = 115.120 - 111.670 = 3.45$
($P > 0.05$ not significant).
Thus, the no-interaction model is acceptable.

The test is a likelihood ratio (LR) test which compares log-likelihood statistics for the interaction model and the no-interaction model. That is, the LR test statistic is of the form $-2 \ln L_R$ minus $-2 \ln L_F$, where R denotes the reduced model, which in this case is the no-interaction model, and F denotes the full model, which is the interaction model.

This LR test statistic has approximately a chi-square distribution with 2 degrees of freedom under the null hypothesis that the no-interaction model is correct. The degrees of freedom here is 2 because there are two product terms being tested in the interaction model.

The log-likelihood statistic for the reduced model comes from the computer output for the no-interaction model and is equal to -2 times -57.560, or 115.120.

The log-likelihood statistic for the full model comes from the computer results for the interaction model and is equal to -2 times -55.835, or 111.670.

The LR statistic is therefore 115.120 minus 111.670, which equals 3.45. This value is not significant at the 0.05 level for 2 degrees of freedom. Thus, it appears that despite the numerical difference between corresponding coefficients in the female and male models, there is no statistically significant difference. We can therefore conclude for these data that the no-interaction model is acceptable (at least at the 0.05 level).

Remission data example:

- described no-interaction assumption
- evaluated assumption using LR test
- provided interaction model if needed

Using the remission data example, we have described the no-interaction assumption, have shown how to evaluate this assumption using a likelihood ratio test, and have provided the form of an interaction model that should be used in case the no-interaction assumption does not hold. We now describe this process more generally for any stratified Cox analysis.

Now, we generalize this process.

No-interaction SC model:

$$h_g(t,\mathbf{X}) = h_{0g}(t)\exp[\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p]$$
$g = 1, 2, \ldots, k^*$, strata defined from Z^*

Recall that the general form of the no-interaction model for the stratified Cox procedure is given as shown here. This model allows for several variables being stratified through the use of a newly defined variable called Z^*, whose strata consist of combinations of categories of the variables being stratified.

SC model allowing interaction:

$$h_g(t,\mathbf{X}) = h_{0g}(t)\exp[\beta_{1g} X_1 + \beta_{2g} X_2 + \ldots + \beta_{pg} X_p]$$
$g = 1, 2, \ldots, k^*$, strata defined from Z^*

If, in contrast, we allow for interaction of the Z^* variable with the X's in the model, we can write the model as shown here. Notice that in this interaction model, each regression coefficient has the subscript g, which denotes the gth stratum and indicates that the regression coefficients are different for different strata of Z^*.

Alternative SC interaction model:

- uses product terms involving Z^*
- define $k^* - 1$ dummy variables $Z_1^*, Z_2^*, \ldots, Z_{k^*-1}^*$, from Z^*
- products of the form $Z_i^* \times X_j$, where $i = 1, \ldots, k^* - 1$ and $j = 1, \ldots, p$.

An alternative way to write the interaction model uses product terms involving the variable Z^* with each of the predictors. However, to write this model correctly, we need to use $k^* - 1$ dummy variables to distinguish the k^* categories of Z^*; also, each of these dummy variables, which we denote as $Z_1^*, Z_2^*, \cdots Z_{k^*-1}^*$, needs to be involved in a product term with each of the X's.

$$\begin{aligned} h_g(t,\mathbf{X}) = h_{0g}(t)\exp\big[&\beta_1 X_1 + \cdots + \beta_p X_p \\ &+ \beta_{11}(Z_1^* \times X_1) + \cdots + \beta_{p1}(Z_1^* \times X_p) \\ &+ \beta_{12}(Z_2^* \times X_1) + \cdots + \beta_{p2}(Z_2^* \times X_p) \\ &+ \cdots + \beta_{1,k^*-1}(Z_{k^*-1}^* \times X_1) + \cdots \\ &+ \beta_{p,k^*-1}(Z_{k^*-1}^* \times X_p)\big] \end{aligned}$$
$g = 1, 2, \ldots, k^*$, strata defined from $Z*$

The hazard model formula alternative model is shown here. Notice that the first line of the formula contains the X's by themselves, the next line contains products of each X_j with Z_1^*, the third line contains the products with Z_2^*, and the last line contains products with $Z_{k^*-1}^*$. Note also that the subscript g occurs only with the baseline hazard function $h_{0g}(t)$, and is not explicitly used in the β coefficients.

EXAMPLE: (Remission Data)

$Z^* = \text{sex}, k^* = 2,$

$Z_1^* = \text{sex}(0, 1),$

$X_1 = \log \text{WBC}, X_2 = Rx \ (p = 2)$

$$\begin{aligned} h_g(t, \boldsymbol{X}) &= h_{0g}(t) \exp[\beta_1 X_1 + \beta_2 X_2 \\ &\quad + \beta_{11}(Z_1^* \times X_1) \\ &\quad + \beta_{21}(Z_1^* \times X_2)] \\ &= h_{0g}(t) \exp[\beta_1^* \log \text{WBC} \\ &\quad + \beta_2^* \text{RX} + \beta_3^* (\text{sex} \times \log \text{WBC}) \\ &\quad + \beta_4^* (\text{sex} \times Rx)] \end{aligned}$$

$g = 1, 2.$

$\beta_1 = \beta_1^*, \ \beta_2 = \beta_2^*, \ \beta_{11} = \beta_3^*, \text{ and } \beta_{21} = \beta_4^*$

In our previous example involving the remission data, the stratification variable (Z^*) was the variable SEX, and k^* was equal to 2; thus, we have only one dummy variable Z_1^*, which uses a (0, 1) coding to indicate sex, and we have only (p equal to) two predictors: X_1 equal to log WBC and X_2 equal to Rx. The interaction model is then written in either of the forms shown here.

The latter version of the interaction model is what we previously presented for the remission data example. Because the two versions presented here are equivalent, it follows that $\beta_1^* = \beta_1, \ \beta_2 = \beta_2^*, \ \beta_{11} = \beta_3^*$, and $\beta_{21} = \beta_4^*$.

We have thus seen that the interaction model can be written in a format that contains product terms involving dummy variables (i.e., Z_i^*) for the variable being stratified being multiplied by each of the predictors (i.e., X_i) not being stratified. We will use this model to describe a test of the no-interaction assumption.

Testing the no-interaction assumption:

$LR = -2 \ln L_\text{R} - (-2 \ln L_\text{F})$

$R =$ reduced (no-interaction) model

$F =$ full (interaction) model contains product terms

$$H_0 : \begin{cases} \beta_{11} = \cdots = \beta_{p1} = 0 \\ \beta_{12} = \cdots = \beta_{p2} = 0 \\ \vdots \\ \beta_{1,k^*-1} = \cdots = \beta_{p,k^*-1} = 0 \end{cases}$$

The test is a likelihood ratio (LR) test which compares log likelihood statistics for the interaction model and the no-interaction model. That is, the LR test statistic is of the form $-2 \ln L_\text{R}$ minus $-2 \ln L_\text{F}$, where R denotes the reduced model, which in this case is the no-interaction model, and F denotes the full model, which is the interaction model.

The no-interaction model differs from the interaction model in that the latter contains additional product terms. Thus, one way to state the null hypothesis of no interaction is that the coefficients of each of these product terms are all zero.

$$LR \dot{\sim} \chi^2_{p(k^*-1) \text{ df}} \quad \text{under } H_0 : \text{no interaction}$$

The LR test statistic has approximately a chi-square distribution with $p(k^* - 1)$ degrees of freedom under the null hypothesis. The degrees of freedom here is $p(k^* - 1)$ because this value gives the number of product terms that are being tested in the interaction model.

$p(k^* - 1)$ gives number of product terms being tested in interaction model

EXAMPLE: (Remission Data)

$Z^* = \text{sex}, k^* = 2,$

$Z_1^* = \text{sex}(0, 1)$

$X_1 = \log \text{WBC}, X_2 = \text{Rx}\ (p = 2)$

$p(k^* - 1) = 2$, so

$LR \dot{\sim} \chi^2_{2\,\text{df}}$ under H_0: no interaction

Returning to the remission data example, for which $p = 2$ and $k^* = 2$, the value of $p(k^* - 1)$ is equal to two times $(2 - 1)$, which equals two. Thus, to test whether the SEX variable interacts with the log WBC and *Rx* predictors, the degrees of freedom for the *LR* statistic is two, as previously described.

V. A Second Example Involving Several Stratification Variables

EXAMPLE

vets.dat: survival time in days, $n = 137$

Veteran's Administration Lung Cancer Trial
Column 1: Treatment (standard = 1, test = 2)
Column 2: Cell type 1 (large = 1, other = 0)
Column 3: Cell type 2 (adeno = 1, other = 0)
Column 4: Cell type 3 (small = 1, other = 0)
Column 5: Cell type 4 (squamous = 1, other = 0)
Column 6: Survival time (days)
Column 7: Performance status (0 = worst,. . ., 100 = best)
Column 8: Disease duration (months)
Column 9: Age
Column 10: Prior therapy (none = 0, some = 10)
Column 11: Status (0 = censored, 1 = died)

Cox regression
Analysis time_t: survt

	Coef.	Std. Err.	p > |z|	Haz. Ratio	[95% Conf. Interval]		P (PH)
Treatment	0.290	0.207	0.162	1.336	0.890	2.006	0.628
Large cell	0.400	0.283	0.157	1.491	0.857	2.594	0.033
Adeno cell	1.188	0.301	0.000	3.281	1.820	5.915	0.081
Small cell	0.856	0.275	0.002	2.355	1.374	4.037	0.078
Perf. Stat	−0.033	0.006	0.000	0.968	0.958	0.978	0.000
Dis. Durat.	0.000	0.009	0.992	1.000	0.982	1.018	0.919
Age	−0.009	0.009	0.358	0.991	0.974	1.010	0.198
Pr. Therapy	0.007	0.023	0.755	1.007	0.962	1.054	0.145

No. of subjects = 137 Log likelihood = −475.180

Variables not satisfying PH:

- cell type (3 dummy variables)
- performance status
- prior therapy (possibly)

SC model: stratifies on cell type and performance status

The dataset "vets.dat" considers survival times in days for 137 patients from the Veteran's Administration Lung Cancer Trial cited by Kalbfleisch and Prentice in their text (*The Statistical Analysis of Survival Time Data*, Wiley, pp. 223–224, 1980). The exposure variable of interest is treatment status. Other variables of interest as control variables are cell type (four types, defined in terms of dummy variables), performance status, disease duration, age, and prior therapy status. Failure status is defined by the status variable. A complete list of the variables is shown here.

Here we provide computer output obtained from fitting a Cox PH model to these data. Using the *P*(*PH*) information in the last column, we can see that at least four of the variables listed have *P*(*PH*) values below the 0.100 level. These four variables are labeled in the output as large cell (0.033), adeno cell (0.081), small cell (0.078), and Perf. Stat (0.000). Notice that the three variables, large cell, adeno cell, and small cell, are dummy variables that distinguish the four categories of cell type.

Thus, it appears from the *P*(*PH*) results that the variables cell type (defined using dummy variables) and performance status do not satisfy the PH assumption.

Based on the conclusions just made about the PH assumption, we now describe a stratified Cox analysis that stratifies on the variables, cell type and performance status.

EXAMPLE: (continued)

Z^* given by combinations of categories:

- cell type (four categories)
- performance status (interval) change to
- PSbin (two categories)

Z^* has $k^* = 4 \times 2 = 8$ categories

Four other variables considered as X's:

- treatment status
- disease duration
- age
- prior therapy

Here, we use treatment status and age as X's

Stratified Cox regression
Analysis time_t: survt

	Coef.	Std. Err.	p>|z|	Haz. Ratio	[95% Conf.	Interval]
Treatment	0.125	0.208	0.548	1.134	0.753	1.706
Age	−0.001	0.010	0.897	0.999	0.979	1.019

No. of subjects = 137 Log likelihood = −262.020 Stratified by Z^*

No-interaction model

$\widehat{HR} = 1.134 (P = 0.548)$

Treatment effect (adjusted for age and Z^*) is nonsignificant

No-interaction model:
$h_g(t,\mathbf{X})$
$= h_{0g}(t)\exp[\beta_1 \text{ Treatment} + \beta_2 \text{ Age}]$
$g = 1, 2, \ldots, 8$ (=# of strata defined from Z^*)

Interaction model:
$h_g(t,\mathbf{X})$
$= h_{0g}(t)\exp[\beta_{1g} \text{ Treatment} + \beta_{2g} \text{ Age}]$
$g = 1, 2, \ldots, 8$

Because we are stratifying on two variables, we need to form a single new categorical variable Z^* whose categories represent combinations of categories of the two variables. The cell type variable has four categories by definition. The performance status variable, however, is an interval variable ranging between 0 for worst to 100 for best, so it needs to be categorized. We categorize this variable into two groups using a cutpoint of 60, and we denote this binary variable as PSbin. Thus, the number of categories for our Z^* variable is 4×2, or 8; that is, $k^* = 8$.

In addition to the two stratification variables, cell type and performance status, there are four other variables to be considered as predictors in the stratified Cox model. These are treatment status, disease duration, age, and prior therapy.

For illustrative purposes here, we use only treatment status and age as predictors. The other two variables, disease duration and prior therapy, are considered in exercises following this presentation.

Here we show computer output from fitting a stratified Cox model that stratifies on cell type and performance status using the eight-category stratification variable Z^*. This model also includes treatment and age as predictors. These results consider a no-interaction model, because only one regression coefficient is provided for the treatment and age predictors. Notice that the estimated hazard ratio is 1.134 for the effect of the treatment variable adjusted for age and Z^*, the latter being adjusted by stratification. The p-value for this adjusted treatment effect is 0.548, which is highly nonsignificant.

The no-interaction model we have just described has the hazard function formula shown here.

To evaluate whether the no-interaction model is appropriate, we need to define an interaction model that allows different regression coefficients for different strata. One way to write this interaction model is shown here.

EXAMPLE: (continued)

Alternative interaction model:

$$h_g(t,X) = h_{0g}(t)\exp[\beta_1 \text{ Treatment} + \beta_2 \text{ Age} + \beta_{11}(Z_1^* \times \text{Treatment}) + \cdots + \beta_{17}(Z_7^* \times \text{Treatment}) + \beta_{21}(Z_1^* \times \text{Age}) + \cdots + \beta_{27}(Z_7^* \times \text{Age})]$$

$$g = 1, 2, \ldots, 8$$

An alternative version of this interaction model that involves product terms is shown here. This version uses seven dummy variables denoted as Z_1^*, Z_2^* up through Z_7^* to distinguish the eight categories of the stratification variable Z^*. The model contains the main effects of treatment and age plus interaction terms involving products of each of the seven dummy variables with each of the two predictors.

Another version of interaction model:
Replace $Z_1^*, \ldots, Z_7^*$ by

$Z_1^* =$ large cell (binary)
$Z_2^* =$ adeno cell (binary)
$Z_3^* =$ small cell (binary)
$Z_4^* =$ PSbin (binary)
$Z_5^* = Z_1^* \times Z_4^*$
$Z_6^* = Z_2^* \times Z_4^*$
$Z_7^* = Z_3^* \times Z_4^*$

Yet another version of the interaction model is to replace the seven dummy variables Z_1^* to Z_7^* by the seven variables listed here. These variables are three of the binary variables making up the cell type variable, the binary variable for performance status, plus three product terms involving each of the cell type dummy variables multiplied by the PSbin dummy variable (Z_4^*).

$$h_g(t,\boldsymbol{X}) = h_{0g}(t)\exp[\beta_1 \text{ Treatment} + \beta_2 \text{ Age} + \beta_{11}(\text{tr } Z_1^*) + \beta_{12}(\text{tr } Z_2^*) + \beta_{13}(\text{tr } Z_3^*) + \beta_{14}(\text{tr } Z_4^*) + \beta_{15}(\text{tr } Z_1^*Z_4^*) + \beta_{16}(\text{tr } Z_2^*Z_4^*) + \beta_{17}(\text{tr } Z_3^*Z_4^*) + \beta_{21}(\text{AGE } Z_1^*) + \beta_{22}(\text{AGE } Z_2^*) + \beta_{23}(\text{AGE } Z_3^*) + \beta_{24}(\text{AGE } Z_4^*) + \beta_{25}(\text{AGE } Z_1^*Z_4^*) + \beta_{26}(\text{AGE } Z_2^*Z_4^*) + \beta_{27}(\text{AGE } Z_3^*Z_4^*)]$$

The latter interaction model is shown here. In this model, the variable tr Z_1^* denotes the product of treatment status with the large cell dummy Z_1^*, the variable tr Z_2^* denotes the product of treatment status with the adeno cell variable Z_2^*, and so on. Also, the variable tr $Z_1^*Z_4^*$ denotes the triple product of treatment status times the large cell variable Z_1^* times the PSbin variable Z_4^*, and so on, for the other triple product terms involving treatment. Similarly, for the terms involving age, the variable Age Z_1^* denotes the product of age with Z_1^*, and the variable Age $Z_1^*Z_4^*$ denotes the triple product of age times Z_1^* times Z_4^*.

Note that we are only considering the interaction between the stratified variables and the predictors. We could also (but do not) consider the interaction between the two predictors, treatment, and age.

EXAMPLE: (continued)

Stratified Cox Regression Analysis on Variable: Z^*
Response: Surv. Time

	Coef.	Std. Err.	p > |z|	Haz. Ratio	[95% Conf. Interval]	
Treatment	0.286	0.664	0.667	1.331	0.362	4.893
Age	0.000	0.030	0.978	0.999	0.942	1.060
tr Z_1^*	2.351	1.772	0.184	10.495	0.326	337.989
tr Z_2^*	−1.158	0.957	0.226	0.314	0.048	2.047
tr Z_3^*	0.582	0.855	0.496	1.790	0.335	9.562
tr Z_4^*	−1.033	0.868	0.234	0.356	0.065	1.950
tr $Z_1^*Z_4^*$	−0.794	1.980	0.688	0.452	0.009	21.882
tr $Z_2^*Z_4^*$	2.785	1.316	0.034	16.204	1.229	213.589
tr $Z_3^*Z_4^*$	0.462	1.130	0.683	1.587	0.173	14.534
Age Z_1^*	0.078	0.064	0.223	1.081	0.954	1.225
Age Z_2^*	−0.047	0.045	0.295	0.954	0.873	1.042
Age Z_3^*	−0.059	0.042	0.162	0.943	0.868	1.024
Age Z_4^*	0.051	0.048	0.287	1.053	0.958	1.157
Age $Z_1^*Z_4^*$	−0.167	0.082	0.042	0.847	0.721	0.994
Age $Z_2^*Z_4^*$	−0.045	0.068	0.511	0.956	0.838	1.092
Age $Z_3^*Z_4^*$	0.041	0.061	0.499	1.042	0.924	1.175

No. of subjects = 137 Log likelihood = −249.972 Stratified by Z^*

Eight possible combinations of Z_1^* to Z_4^*:

$g = 1: \ Z_1^* = Z_2^* = Z_3^* = Z_4^* = 0$
$g = 2: \ Z_1^* = 1, \ Z_2^* = Z_3^* = Z_4^* = 0$
$g = 3: \ Z_2^* = 1, \ Z_1^* = Z_3^* = Z_4^* = 0$
$g = 4: \ Z_3^* = 1, \ Z_1^* = Z_2^* = Z_4^* = 0$
$g = 5: \ Z_1^* = Z_2^* = Z_3^* = 0, \ Z_4^* = 1$
$g = 6: \ Z_1^* = 1, \ Z_2^* = Z_3^* = 0, \ Z_4^* = 1$
$g = 7: \ Z_2^* = 1, \ Z_1^* = Z_3^* = 0, \ Z_4^* = 1$
$g = 8: \ Z_3^* = 1, \ Z_1^* = Z_2^* = 0, \ Z_4^* = 1$

$g = 1: \ Z_1^* = Z_2^* = Z_3^* = Z_4^* = 0$
(Squamous cell type and PSbin = 0)

All product terms are zero:
$h_1(t,\mathbf{X})$
$= h_{01}(t)\exp[\beta_1 \text{ Treatment} + \beta_2 \text{ Age}]$,
where $\hat{\beta}_1 = 0.286$,
$\hat{\beta}_2 = 0.000$, so that
$\hat{h}_1(t,\mathbf{X}) = \hat{h}_{01}(t)\exp[(0.283)\,\text{Treatment}]$

$g = 2: \ Z_1^* = 1, \ Z_2^* = Z_3^* = Z_4^* = 0$
(Large cell type and PSbin = 0)

Nonzero product terms	Coefficients
Age Z_1^* = Age	β_{21}
tr Z_1^* = Treatment	β_{11}

Here we provide the computer results from fitting the interaction model just described. Notice that the first two variables listed are the main effects of treatment status and age. The next seven variables are product terms involving the interaction of treatment status with the seven categories of Z^*. The final seven variables are product terms involving the interaction of age with the seven categories of Z^*. As defined on the previous page, the seven variables used to define Z^* consist of three dummy variables Z_1^*, Z_2^* and Z_3^* for cell type, a binary variable Z_4^* for performance status and products of Z_4^* with each of Z_1^*, Z_2^*, and Z_3^*. Note that once the variables Z_1^*, Z_2^* Z_3^*, and Z_4^* are specified, the values of the three product terms are automatically determined.

We can use these results to show that the interaction model being fit yields different regression coefficients for each of the eight categories defined by the subscript g for the stratification variable Z^*. These eight categories represent the possible combinations of the four variables Z_1^* to Z_4^*, as shown here.

Consider the hazard function when the variables Z_1^* through Z_4^* are all equal to zero. This stratum is defined by the combination of squamous cell type and a binary performance status value of 0. In this case, all product terms are equal to zero and the hazard model contains only the main effect terms treatment and age. The estimated hazard function for this stratum uses the coefficients 0.286 for treatment and 0.000 for age, yielding the expression shown here. Note that age drops out of the expression because its coefficient is zero to three decimal places.

Now consider the hazard function when the variable Z_1^* equals 1 and Z_2^* through Z_4^* are equal to zero. This stratum is defined by the combination of large cell type and a PSbin value of 0. In this case, the only nonzero product terms are Age Z_1^* and tr Z_1^*, whose coefficients are β_{21} and β_{11}, respectively.

EXAMPLE: (continued)

$h_2(t,\mathbf{X}) = h_{02}(t)\exp[(\beta_1 + \beta_{11})\,\text{Treatment} + (\beta_2 + \beta_{21})\,\text{Age}]$

$\hat{\beta}_1 = 0.286,\ \hat{\beta}_2 = 0.000$

$\hat{\beta}_{11} = 2.351,\ \hat{\beta}_{21} = 0.078$

The hazard function for this second stratum is shown here. Notice that the coefficients of the treatment and age variables are $(\beta_1 + \beta_{11})$ and $(\beta_2 + \beta_{21})$, respectively. The estimated values of each of these coefficients are given here.

Hazard functions for interaction model:

$g = 1: (Z_1^* = Z_2^* = Z_3^* = Z_4^* = 0):$
$\hat{h}_1(t,\mathbf{X}) = \hat{h}_{01}(t)\exp[(0.286)\,\text{Treatment}]$

$g = 2: (Z_1^* = 1, Z_2^* = Z_3^* = Z_4^* = 0):$
$\hat{h}_2(t,\mathbf{X}) = \hat{h}_{02}(t)\exp[(2.637)\,\text{Treatment} + (0.078)\text{Age}]$

The corresponding *estimated* hazard function for the second stratum (i.e., $g = 2$) is shown here. For comparison, we repeat the estimated hazard function for the first stratum.

$g = 3: (Z_2^* = 1, Z_1^* = Z_3^* = Z_4^* = 0):$
$\hat{h}_3(t,\mathbf{X}) = \hat{h}_{03}(t)\exp[(-0.872)\,\text{Treatment} + (-0.047)\text{Age}]$

$g = 4: (Z_3^* = 1, Z_1^* = Z_2^* = Z_4^* = 0):$
$\hat{h}_4(t,\mathbf{X}) = \hat{h}_{04}(t)\exp[(0.868)\,\text{Treatment} + (-0.059)\text{Age}]$

$g = 5: (Z_1^* = Z_2^* = Z_3^* = 0, Z_4^* = 1):$
$\hat{h}_5(t,\mathbf{X}) = \hat{h}_{05}(t)\exp[(0.747)\,\text{Treatment} + (-0.051)\text{Age}]$

$g = 6: (Z_1^* = 1, Z_2^* = Z_3^* = 0, Z_4^* = 1):$
$\hat{h}_6(t,\mathbf{X}) = \hat{h}_{06}(t)\exp[(0.810)\,\text{Treatment} + (-0.038)\text{Age}]$

$g = 7: (Z_2^* = 1, Z_1^* = Z_3^* = 0, Z_4^* = 1):$
$\hat{h}_7(t,\mathbf{X}) = \hat{h}_{07}(t)\exp[(0.880)\,\text{Treatment} + (-0.041)\text{Age}]$

$g = 8: (Z_3^* = 1, Z_1^* = Z_2^* = 0, Z_4^* = 1):$
$\hat{h}_8(t,\mathbf{X}) = \hat{h}_{08}(t)\exp[(0.297)\,\text{Treatment} + (0.033)\text{Age}]$

The estimated hazard functions for the remaining strata are provided here. We leave it up to the reader to verify these formulae. Notice that the coefficients of treatment are all different in the eight strata, and the coefficients of age also are all different in the eight strata.

LR test to compare no-interaction model with interaction model:

We have presented computer results for both the no-interaction and the interaction models. To evaluate whether the no-interaction assumption is satisfied, we need to carry out a likelihood ratio test to compare these two models.

H_0: no-interaction model acceptable, i.e., Treatment: $\beta_{11} = \beta_{12} = \cdots = \beta_{17} = 0$ and Age: $\beta_{21} = \beta_{22} = \cdots = \beta_{27} = 0$

The null hypothesis being tested is that the no-interaction model is acceptable. Equivalently, this null hypothesis can be stated by setting the coefficients of all product terms in the interaction model to zero. That is, the seven coefficients of product terms involving treatment and the seven coefficients of the product terms involving age are set equal to zero as shown here.

14 coefficients $\Rightarrow$ df = 14

$LR = -2\ln L_R - (2\ln L_F)$

R = reduced (no-interaction) model

F = full (interaction) model

Because the null hypothesis involves 14 coefficients, the degrees of freedom of the LR chi-square statistic is 14. The test statistic takes the usual form involving the difference, between log-likelihood statistics for the reduced and full models, where the reduced model is the no-interaction model and the full model is the interaction model.

EXAMPLE (continued)

$LR \dot{\sim} \chi^2_{14\,df}$ under H_0: no interaction

$LR = (-2 \times -262.020) - (-2 \times -249.972)$
$= 524.040 - 499.944 = 24.096$

$P = 0.045$ (significant at 0.05)

Conclusion:
Reject H_0: interaction model is preferred.

Might use further testing to simplify interaction model, e.g., test for seven products involving treatment or test for seven products involving age.

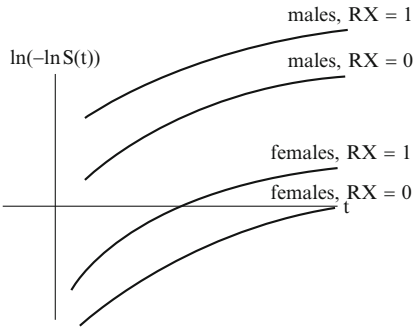
Thus, under the null hypothesis, the *LR* statistic is approximately chi-square with 14 degrees of freedom.

The computer results for the no-interaction and interaction models give log-likelihood values of 524.040 and 499.944, respectively. The difference is 24.096. A chi-square value of 24.096 with 14 degrees of freedom yields a p-value of 0.045, so that the test gives a significant result at the 0.05 level. This indicates that the no-interaction model is not acceptable and the interaction model is preferred.

Note, however, that it may be possible from further statistical testing to simplify the interaction model to have fewer than 14 product terms. For example, one might test for only the seven product terms involving treatment or only the seven product terms involving age.

VI. A Graphical View of the Stratified Cox Approach

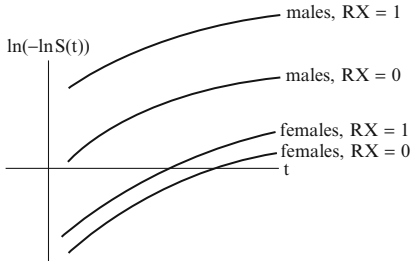
a. $h(t) = h_0(t)\exp(\beta_1 RX + \beta_2 SEX)$
$\ln(-\ln S(t)) = \ln(-\ln S_0(t)) + \beta_1 RX + \beta_2 SEX$



In this section we examine four log–log survival plots illustrating the assumptions underlying a stratified Cox model with or without interaction. Each of the four models considers two dichotomous predictors: treatment (coded RX = 1 for placebo and RX = 0 for new treatment) and SEX (coded 0 for females and 1 for males). The four models are as follows (see left).

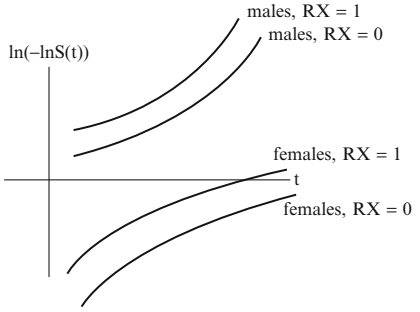
a. This model assumes the PH assumption for both RX and SEX and also assumes no interaction between RX and SEX. Notice all four log–log curves are parallel (PH assumption) and the effect of treatment is the same for females and males (no interaction). The effect of treatment (controlling for SEX) can be interpreted as the distance between the log–log curves from RX = 1 to RX = 0, for males and for females, separately.

b. $h(t) = h_0(t) \exp(\beta_1 RX + \beta_2 SEX + \beta_3 RX \times SEX)$
$\ln(-\ln S(t)) = \ln(-\ln S_0(t)) + \beta_1 RX + \beta_2 SEX + \beta_3 RX \times SEX$



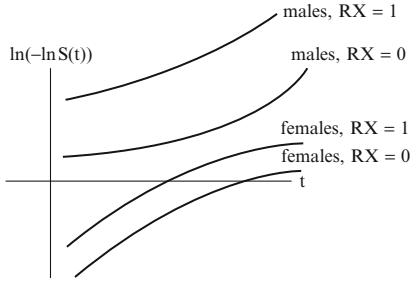
b. This model assumes the PH assumption for both RX and SEX and allows for interaction between these two variables. All four log–log curves are parallel (PH assumption) but the effect of treatment is larger for males than females as the distance from RX = 1 to RX = 0 is greater for males.

c. $h(t) = h_{0g}(t) \exp(\beta_1 RX)$
(g = 1 for males, g = 0 for females)
$\ln(-\ln S(t)) = \ln(-\ln S_{0g}(t)) + \beta_1 RX$



c. This is a stratified Cox model in which the PH assumption is not assumed for SEX. Notice the curves for males and females are not parallel. However, the curves for RX are parallel within each stratum of SEX indicating that the PH assumption is satisfied for RX. The distance between the log–log curves from RX = 1 to RX = 0 is the same for males and females indicating no interaction between RX and SEX.

d. $h(t) = h_{0g}(t) \exp(\beta_1 RX + \beta_2 RX \times SEX)$
(g = 1 for males, g = 0 for females)
$\ln(-\ln S(t)) = \ln(-\ln S_{0g}(t)) + \beta_1 RX + \beta_2 RX \times SEX$



d. This is a stratified Cox model allowing for interaction of RX and SEX. The curves for males and females are not parallel although the PH assumption is satisfied for RX within each stratum of SEX.The distance between the log–log curves from RX = 1 to RX = 0 is greater for males than females indicating interaction between RX and SEX.

VII. The Stratified Cox Likelihood

- Cox PH Likelihood (L) described in Chapter 3, Section VIII
- L now extended for SC model

At the end of the presentation from Chapter 3 (Section VIII), we illustrated the Cox likelihood using the dataset shown on the left. In this section, we extend that discussion to illustrate the likelihood for a stratified Cox model.

ID	TIME	STATUS	SMOKE
Barry	2	1	1
Gary	3	1	0
Harry	5	0	0
Larry	8	1	1

TIME = Survival time (in years)
STATUS = 1 for event, 0 for censorship
SMOKE = 1 for a smoker, 0 for a non-smoker

To review: The data indicate that Barry got the event at TIME = 2 years. Gary got the event at 3 years, Harry was censored at 5 years and Larry got the event at 8 years. Furthermore, Barry and Larry were smokers while Gary and Harry were non-smokers.

Cox PH model: $h(t) = h_0(t)e^{\beta_1 SMOKE}$

Cox PH Likelihood:

$$L = \left[\frac{h_0(t)e^{\beta_1}}{h_0(t)e^{\beta_1} + h_0(t)e^0 + h_0(t)e^0 + h_0(t)e^{\beta_1}}\right] \times \left[\frac{h_0(t)e^0}{h_0(t)e^0 + h_0(t)e^0 + h_0(t)e^{\beta_1}}\right] \times \left[\frac{h_0(t)e^{\beta_1}}{h_0(t)e^{\beta_1}}\right]$$

In Chapter 3, we constructed the Cox likelihood with one predictor, SMOKE, in the model. The model and the likelihood are shown on the left.

The likelihood is a product of three terms, one term for each event time t_f (TIME = 2, 3, and 8). The denominator of each term is the sum of the hazards from the subjects still in the risk set at time t_f, including the censored subject, Harry. The numerator of each term is the hazard of the subject who got the event at t_f. The reader may wish to reread Section VIII of Chapter 3.

Stratified Cox model:

$h_g(t) = h_{0g}(t)e^{\beta_1 SMOKE}$

g=1 history of hypertension
g=2 no history of hypertension

Now consider a stratified Cox model, in which the stratified variable is a dichotomous indicator of whether the subject has or does not have a history of hypertension. The predictor SMOKE remains in the model (shown at left).

PH assumption:

- Violated overall
- Assumed to hold within categories of stratified variable

The model allows for a violation of the PH assumption. However, within each category of hypertension, the PH assumption is assumed to hold.

ID	TIME	STATUS	SMOKE	HT
Barry	2	1	1	1
Gary	3	1	0	2
Harry	5	0	0	1
Larry	8	1	1	2

The data are shown on the left. The additional variable is HT which is the variable we wish to stratify on. Barry and Harry have a history of hypertension (coded HT=1) while Gary and Larry do not (coded HT=2).

HT = History of hypertension
(1=yes, 2=no)

Formulate likelihood in pieces:

The stratified Cox likelihood is formulated in pieces. Each piece represents a stratified category. Within each piece, the likelihood is formulated similarly as the likelihood formulated for a Cox PH model.

Data for HT=1

ID	TIME	STATUS	SMOKE	HT
Barry	2	1	1	1
Harry	5	0	0	1

Data for HT=2

The first piece (stratified category) is defined among observations in which HT=1. The second piece is defined among observations in which HT=2.

ID	TIME	STATUS	SMOKE	HT
Gary	3	1	0	2
Larry	8	1	1	2

Stratified Cox model:

$h_g(t) = h_{0g}(t)e^{\beta_1 SMOKE}$

g=1 history of hypertension
g=2 no history of hypertension

Here, for convenience, we again show you the SC model for these data on the left.

Among HT=1
Barry has only event
Barry and Harry at risk at TIME=2

$$L_1 = \frac{h_{01}e^{\beta_1}}{h_{01}e^{\beta_1} + h_{01}e^{0}}$$

For the first stratum (HT=1), Barry gets an event at TIME=2. Barry and Harry are in the risk set when Barry gets his event. Barry is the only event from this stratum as Harry is censored at TIME=5. Therefore the likelihood for this piece, L_1 (shown at left), contains one term. That one term has Barry's hazard in the numerator. Barry's and Harry's hazard are summed in the denominator.

Among HT=2
Gary and Larry both get events
Gary and Larry at risk at TIME=3
Larry at risk at TIME=8

$$L_2 = \frac{h_{02}e^{0}}{h_{02}e^{0} + h_{02}e^{\beta_1}} \times \frac{h_{02}e^{\beta_1}}{h_{02}e^{\beta_1}}$$

For the second stratum (HT=2), Gary gets an event at TIME=3. Gary and Larry are in the risk set when Gary gets his event. Larry gets an event at TIME=8 and is the only one in the risk set at the time of his event. Therefore, the likelihood for this piece, L_2 (shown at left), is a product of two factors. Each of the two factors corresponds to an event.

$$L = L_1 \times L_2$$
$$= \left[\frac{h_{01}e^{\beta_1}}{h_{01}e^{\beta_1} + h_{01}e^{0}}\right]\left[\frac{h_{02}e^{0}}{h_{02}e^{0} + h_{02}e^{\beta_1}} \times \frac{h_{02}e^{\beta_1}}{h_{02}e^{\beta_1}}\right]$$

The stratified Cox likelihood can be formulated by taking the product of each piece (L_1 and L_2). Each piece was formulated within a stratum.

The baseline hazard cancels in L

$$L = \left[\frac{e^{\beta_1}}{e^{\beta_2} + e^{0}}\right] \times \left[\frac{e^{0}}{e^{0} + e^{\beta_1}} \times \frac{e^{\beta_1}}{e^{\beta_1}}\right]$$

Notice that the baseline hazard cancels from the likelihood. As with the Cox PH model, the stratified likelihood is just determined by the order of events (not the baseline hazard).

$$L = \left[\frac{e^{\beta_1}}{e^{\beta_2} + e^{0}}\right] \times \left[\frac{e^{0}}{e^{0} + e^{\beta_1}} \times \frac{e^{\beta_1}}{e^{\beta_1}}\right]$$

Same β_1 in each piece

One additional point about this model is that the β_1 in the first piece of the likelihood (L_1) is the same β_1 that is in the second piece of the likelihood (L_2). In other words, this is a no interaction model. The effect of smoking (expressed as β_1) does not depend on hypertension status.

No interaction model:

Effect of smoking is same for those with or without hypertension

VIII. Summary

We now summarize the most important features of the stratified Cox (SC) model described in this presentation.

Stratified Cox (SC) model:

- stratification of predictors not satisfying PH assumption
- includes predictors satisfying PH
- does not include stratified variables

The SC model is a modification of the Cox PH model to allow for control by "stratification" of predictors not satisfying the PH assumption. Variables that are assumed to satisfy the assumption are included in the model as predictors; the stratified variables are not included in the model.

Computer Results

The computer results for a SC model provides essentially the same type of output as provided for a Cox PH model without stratification. An example of SC output using the remission data is shown here. The variables included as predictors in the model are listed in the first column followed by their estimated coefficients, standard errors, p-values, hazard ratio values, and 95% confidence limits. Such information cannot be provided for the variables being stratified, because these latter variables are not explicitly included in the model.

Stratified Cox regression
Analysis time _t: survt

	Coef.	Std. Err.	p > |z|	Haz. Ratio	[95% Conf.	Interval]
log WBC	1.390	0.338	0.000	4.016	2.072	7.783
RX	0.931	0.472	0.048	2.537	1.006	6.396

No. of subjects = 42 | Log likelihood = −57.560 | Stratified by sex

Hazard function for no-interaction stratified Cox model:

$$h_g(t,\mathbf{X}) = h_{0g}(t)\exp[\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p]$$

$g = 1, 2, \ldots, k^*$, strata defined from Z^*

Z^* has k^* categories

$X_1, X_2, \ldots, X_p$ satisfy PH

The general hazard function form for the no-interaction stratified Cox model is shown here. This formula contains a subscript g that indicates the gth stratum, where the strata are different categories of the stratification variable Z^* and the number of strata equals k^*. Notice that the baseline hazard functions are different in each stratum.

Stratification variable Z^*:

- identify $Z_1, Z_2, \ldots, Z_k$ not satisfying PH
- categorize each Z
- form combinations of categories (strata)
- each combination is a stratum of Z^*

The variable Z^* is defined by first identifying the Z_i variables not satisfying the PH assumption. We then categorize each Z and form combinations of categories of each of the Z's. Each combination represents a different stratum making up the variable Z^*.

No-interaction model:
Same coefficients $\beta_1, \beta_2, \ldots, \beta_p$ for each g, i.e., Z^* does not interact with the X's.

The above model is designated as a "no-interaction" model because the β's in the model are the same for each subscript g. The no-interaction assumption means that the variables being stratified are assumed *not* to interact with the X's in the model.

$$\text{Different baselines} \left\{ \begin{array}{lcl} h_{01}(t) & \Rightarrow & \hat{S}_1(t) \\ h_{02}(t) & \Rightarrow & \hat{S}_2(t) \\ & \vdots & \\ h_{0k}(t) & \Rightarrow & \hat{S}_{k^*}(t) \end{array} \right\} \text{Different survival curves}$$

For the no-interaction model, the fitted SC model will yield different estimated survival curves for each stratum because the baseline hazard functions are different for each stratum.

$\widehat{HR}$ same for each stratum

However, because the coefficients of the X's are the same for each stratum, estimates of hazard ratios are the same for each stratum.

(Partial) likelihood function:

$$L = L_1 \times L_2 \times \ldots \times L_{k^*}$$

Regression coefficients in the SC model are estimated by maximizing a partial likelihood function that is obtained by multiplying likelihood functions for each stratum.

Stratified Cox model allowing interaction:

$$h_g(t,\mathbf{X}) = h_{0g}(t)\exp[\beta_{1g} X_1 + \beta_{2g} X_2 + \ldots + \beta_{pg} X_p]$$

$g = 1, 2, \ldots, k^*$, strata defined from Z^*.

In order to evaluate the no-interaction assumption, we must define an interaction model for comparison. One version of the interaction model is shown here. This version shows regression coefficients with different subscripts in different strata; that is, each β coefficient has a subscript g.

Alternative stratified Cox interaction model:

- uses product terms involving Z^*
- define $k^* - 1$ dummy variables from Z^*
- products of the form $Z_i^* \times X_j$

An alternative way to write the interaction model uses product terms involving the Z^* variable with each predictor. This model uses $k^* - 1$ dummy variables to distinguish the k^* categories of Z^*. Each of these dummy variables is included as a product term with each of the X's.

Testing the no-interaction assumption:

$LR = -2 \ln L_R - (-2 \ln L_F)$
R = reduced (no-interaction) model
F = full (interaction) model contains product terms
$LR \dot{\sim} \chi^2_{p(k^*-1)\text{df}}$ under H_0: no interaction

To evaluate the no-interaction assumption, we can perform a likelihood ratio test that compares the (reduced) no-interaction model to the (full) interaction model. The null hypothesis is that the no-interaction assumption is satisfied. The test statistic is given by the difference between the log-likelihood statistics for the no-interaction and interaction models. This statistic is approximately chi-square under the null hypothesis. The degrees of freedom is $p(k^* - 1)$ where p denotes the number of X's and k^* is the number of categories making up Z^*.

Presentation Complete!

Chapters

1. Introduction to Survival Analysis
2. Kaplan–Meier Survival Curves and the Log–Rank Test
3. The Cox Proportional Hazards Model and Its Characteristics
4. Evaluating the Proportional Hazards Assumption
✓5. The Stratified Cox Procedure

Next:

6. Extension of the Cox Proportional Hazards Model for Time-Dependent Variables

This presentation is now complete. We suggest that the reader review this presentation using the detailed outline that follows. Then answer the practice exercises and the test that follow.

The next Chapter (6) is entitled "Extension of the Cox PH Model for Time-Dependent Variables." There we show how an "extended" Cox model can be used as an alternative to the stratified Cox model when one or more predictors do not satisfy the PH assumption. We also discuss more generally what is a time-dependent variable, and show how such a variable can be evaluated using an extended Cox model.

Detailed Outline

I. Preview (page 204)

A. Focus on how stratified Cox (SC) procedure is carried out:
- analysis of computer results from SC procedure;
- hazard function for SC model;
- stratifying on a single predictor versus two or more predictors;
- no-interaction versus interaction models.

II. An Example (pages 204–208)

A. Cox PH results for remission data yield $P(PH) = 0.031$ for SEX.

B. SC model used: control for SEX (stratified); include log WBC and *Rx* in model.

C. Analysis of *Rx* effect from stratified Cox results: $\widehat{HR} = 2.537$; 95% CI: (1.006,6.396); LR and Wald tests: $P < 0.05$.

D. Hazard model: $h_g(t,\mathbf{X}) = h_{0g}(t)\exp[\beta_1 \log \text{WBC} + \beta_2 Rx]$, $g = 1,2$
- different baseline hazard functions and survival curves for females and males;
- same coefficients β_1 and β_2 for both females and males (no-interaction assumption);
- obtain estimates by maximizing partial likelihood $L = L_1 \times L_2$.

E. Graph of four adjusted survival curves for *Rx* (adjusted for log WBC).

III. The General Stratified Cox (SC) Model (pages 208–209)

A. $\boxed{h_g(t,\mathbf{X}) = h_{0g}(t)\exp[\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p],\ g = 1, 2,\ldots,k^*}$

where the strata are defined from the stratification variable Z^*.

B. Z^* defined from $Z_1,Z_2,\ldots,Z_k$ variables that do not satisfy PH:
- categorize each Z_i
- form combinations of categories
- each combination is a stratum of Z^*

C. Different baseline hazard functions and survival curves for each stratum.

D. Assumes no interaction: same coefficients $\beta_1,\beta_2,\ldots,\beta_p$ for each g; i.e., Z^* does not interact with the X's; i.e., estimated HR is same for each stratum.

E. Obtain estimates by maximizing partial likelihood $L = L_1 \times L_2 \times \ldots \times L_{k^*}$, where L_i is likelihood for ith stratum.

IV. The No-Interaction Assumption and How to Test It (pages 210–216)

A. Assumes same coefficients $\beta_1,\beta_2,\ldots,\beta_p$ for each g.

B. Interaction model:

$$h_g(t,\mathbf{X}) = h_{0g}(t)\mathbf{exp}[\beta_{1g}X_1 + \beta_{2g}X_2 + \ldots + \beta_{pg}X_p],$$

$g = 1, 2,\ldots,k^*$ strata defined from Z^*.

C. Alternative stratified Cox interaction model:

- uses product terms involving Z^*
- define k^*-1 dummy variables $Z_1^*,\ Z_2^*,\ldots,Z_{k^*-1}^*$ from Z^*
- products of the form $Z_i^* \times X_j$, where $i = 1,\ldots, k^* - 1; j = 1,\ldots, p$
- hazard function: $g = 1, 2,\ldots, k^*$ strata defined from Z^*

$$\begin{aligned} h_g(t,\boldsymbol{X}) = h_{0g}(t)\exp\big[&\beta_1 X_1 + \cdots + \beta_p X_p + \beta_{11}(Z_1^* \times X_1) \\ &+ \cdots + \beta_{p1}(Z_1^* \times X_p) + \beta_{12}(Z_2^* \times X_1) + \cdots + \beta_{p2}(Z_2^* \times X_p) \\ &+ \cdots + \beta_{1,k^*-1}(Z_{k^*-1}^* \times X_1) + \cdots + \beta_{p,k^*-1}(Z_{k^*-1}^* \times X_p)\big] \end{aligned}$$

D. Testing the no-interaction assumption: use LR statistic given by $LR = -2\ln L_{\mathrm{R}} - (-2\ln L_{\mathrm{F}})$ where R = reduced (no interaction) model and F = full (interaction) model
$LR \dot{\sim} \chi^2_{p(k^*-1)\mathrm{df}}$ under H_0: no interaction, i.e., $\beta_{11} = \beta_{12} = \ldots = \beta_{p,k^*-1} = 0$

V. A Second Example Involving Several Stratification Variables (pages 216–221)

A. Dataset "vets.dat" from Veteran's Administration Lung Cancer Trial; $n = 137$; survival time in days.

B. Variables are: treatment status, cell type (four types), performance status, disease duration, age, and prior therapy status.

C. Cox PH results indicate [using $P(PH)$] that cell type and performance status do not satisfy PH assumption.

D. Example stratifies on cell type and performance status using four categories of cell type and two categories of performance status, so that Z^* has $k^* = 8$ strata.

E. X's considered in model are treatment status and age.

F. Computer results for no-interaction model: estimated HR for effect of treatment adjusted for age and Z^* is 1.134 ($P = 0.548$); not significant.

G. Hazard function for no-interaction model:

$h_g(t,\mathbf{X}) = h_{0g}(t)\exp[\beta_1 \text{ Treatment} + \beta_2 \text{ Age}]$,
$g = 1, 2, \ldots, 8$

H. Hazard function for interaction model:

$h_g(t,\mathbf{X}) = h_{0g}(t)\exp[\beta_{1g} \text{ Treatment} + \beta_{2g} \text{ Age}]$,
$g = 1, 2, \ldots, 8$

I. Alternative version of interaction model:

$$h_g(t,\mathbf{X}) = h_{0g}(t)\mathbf{exp}[\beta_1 \text{ Treatment} + \beta_2 \text{ Age} + \beta_{11}(Z_1^* \times \text{Treatment}) + \cdots + \beta_{17}(Z_7^* \times \text{Treatment}) + \beta_{21}(Z_1^* \times \text{Age}) + \cdots + \beta_{27}(Z_7^* \times \text{Age})],$$

$g = 1, 2, \ldots 8$
where Z_1^* = large cell (binary), Z_2^* = adeno cell (binary), Z_3^* = small cell (binary), Z_4^* = PSbin (binary), $Z_5^* = Z_1^* \times Z_4^*$, $Z_6^* = Z_2^* \times Z_4^*$, $Z_7^* = Z_3^* \times Z_4^*$

J. Demonstration that alternative interaction version (in item I) is equivalent to original interaction formulation (in item H) using computer results for the alternative version.

K. Test of no-interaction assumption:
- null hypothesis: $\beta_{11} = \beta_{12} = \ldots = \beta_{17} = 0$ and $\beta_{21} = \beta_{22} = \ldots = \beta_{27} = 0$
- $LR \dot{\sim} \chi^2_{14\text{df}}$ under H_0: no interaction
- $LR = 524.040 - 499.944 = 24.096$ $(P = 0.045)$

  Conclusion: Reject null hypothesis; interaction model is preferred.

VI. A Graphical View of the Stratified Cox Approach (pages 221–222)
Comparison of log–log survival curves
1. Describe interaction of Rx and Sex.
2. Describe violation of PH assumption for Sex.

VII. The Stratified Cox Likelihood (pages 223–225)

VIII. Summary (pages 225–227)

Practice Exercises

The following questions derive from the dataset **vets.dat** concerning the Veteran's Administration Lung Cancer Trial that we previously considered in this chapter. Recall that survival times are in days and that the study size contains 137 patients. The exposure variable of interest is treatment status (standard = 1, test = 2). Other variables of interest as control variables are cell type (four types, defined in terms of dummy variables), performance status, disease duration, age, and prior therapy status. Failure status is defined by the status variable (0 = censored, 1 = died).

1. Consider the following two edited printouts obtained from fitting a Cox PH model to these data.

Cox regression
Analysis time _t:

survt	Coef.	Std. Err.	p > |z|	Haz. Ratio	[95% Conf.	Interval]	*P*(*PH*)
Treatment	0.290	0.207	0.162	1.336	0.890	2.006	0.628
Large cell	0.400	0.283	0.157	1.491	0.857	2.594	0.033
Adeno cell	1.188	0.301	0.000	3.281	1.820	5.915	0.081
Small cell	0.856	0.275	0.002	2.355	1.374	4.037	0.078
Perf.Stat	−0.033	0.006	0.000	0.968	0.958	0.978	0.000
Dis.Durat.	0.000	0.009	0.992	1.000	0.982	1.018	0.919
Age	−0.009	0.009	0.358	0.991	0.974	1.010	0.198
Pr.Therapy	0.007	0.023	0.755	1.007	0.962	1.054	0.145

No. of subjects = 137 Log likelihood = −475.180

Cox regression
Analysis time _t:

survt	Coef.	Std. Err.	p > |z|	Haz. Ratio	[95% Conf.	Interval]	*P*(*PH*)
Treatment	0.298	0.197	0.130	1.347	0.916	1.981	0.739
Small cell	0.392	0.210	0.062	1.481	0.981	2.235	0.382
Perf.Stat	−0.033	0.005	0.000	0.968	0.958	0.978	0.000
Dis.Durat.	−0.001	0.009	0.887	0.999	0.981	1.017	0.926
Age	−0.006	0.009	0.511	0.994	0.976	1.012	0.211
Pr.Therapy	−0.003	0.023	0.884	0.997	0.954	1.042	0.146

No. of subjects = 137 Log likelihood = −487.770

How do the printouts differ in terms of what the *P*(*PH*) information says about which variables do not satisfy the PH assumption?

2. Based on the above information, if you were going to stratify on the cell type variable, how would you define the strata? Explain.

3. Consider a stratified analysis that stratifies on the variables Z_1 = "small cell" and Z_2 = "performance status." The small cell variable is one of the dummy variables for cell type defined above. The performance status variable is dichotomized into high (60 or above) and low (below 60) and is denoted as PSbin. The stratification variable which combines categories from Z_1 and Z_2 is denoted as SZ^* and consists of four categories. The predictors included (but not stratified) in the analysis are treatment status, disease duration, age, and prior therapy. The computer results are as follows:

Stratified Cox regression
Analysis time _t: survt

	Coef.	Std. Err.	p > |z|	Haz. Ratio	[95% Conf.	Interval]
Treatment	0.090	0.197	0.647	1.095	0.744	1.611
Dis.Durat.	0.000	0.010	0.964	1.000	0.982	1.019
Age	0.002	0.010	0.873	1.002	0.983	1.021
Pr.Therapy	−0.010	0.023	0.656	0.990	0.947	1.035

No. of subjects = 137 Log likelihood = −344.848 Stratified by SZ^*

Based on these results, describe the point and interval estimates for the hazard ratio for the treatment effect adjusted for the other variables, including SZ^*. Is this hazard ratio meaningfully and/or statistically significant? Explain.

4. State the form of the hazard function for the model being fit in question 3. Why does this model assume no interaction between the stratified variables and the predictors in the model?
5. State two alternative ways to write the hazard function for an "interaction model" that allows for the interaction of the stratified variables with the treatment status variable, but assumes no other type of interaction.
6. State two alternative versions of the hazard function for an interaction model that allows for the interaction of the stratified variables (small cell and performance status) with each of the predictors treatment status, disease duration, age, and prior therapy.
7. For the interaction model described in question 6, what is the formula for the hazard ratio for the effect of treatment adjusted for the other variables? Does this formula give a different hazard ratio for different strata? Explain.

8. State two alternative versions of the null hypothesis for testing whether the no-interaction assumption is satisfied for the stratified Cox model. Note that one of these versions should involve a set of regression coefficients being set equal to zero.
9. State the form of the likelihood ratio statistic for evaluating the no-interaction assumption. How is this statistic distributed under the null hypothesis, and with what degrees of freedom?
10. Provided below are computer results for fitting the interaction model described in question 6. In this printout the variable Z_1^* denotes the small cell variable and the variable Z_2^* denotes the PSbin variable. The variable DDZ_1^* denotes the product of Z_1^* with disease duration, and other product terms are defined similarly.

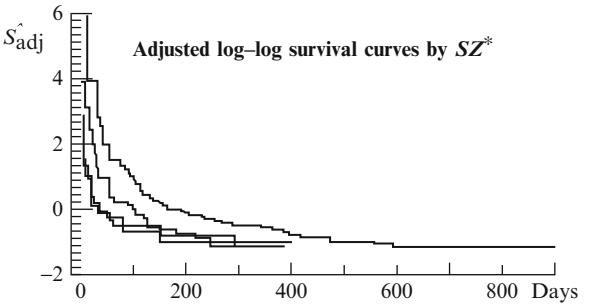
Stratified Cox regression
Analysis time _t:

survt	Coef.	Std. Err.	p > |z|	Haz. Ratio	[95% Conf.	Interval]
Treatment	0.381	0.428	0.374	1.464	0.632	3.389
Dis.Durat.	0.015	0.021	0.469	1.015	0.975	1.057
Age	0.000	0.017	0.994	1.000	0.968	1.033
Pr.Therapy	0.023	0.041	0.571	1.023	0.944	1.109
DDZ_1^*	−0.029	0.024	0.234	0.971	0.926	1.019
$AgeZ_1^*$	−0.055	0.037	0.135	0.946	0.880	1.018
PTZ_1^*	0.043	0.075	0.564	1.044	0.901	1.211
DDZ_2^*	0.025	0.032	0.425	1.026	0.964	1.092
$AgeZ_2^*$	0.001	0.024	0.956	1.001	0.956	1.049
PTZ_2^*	−0.078	0.054	0.152	0.925	0.831	1.029
$DDZ_1\ Z_2^*$	−0.071	0.059	0.225	0.931	0.830	1.045
$AgeZ_1Z_2^*$	0.084	0.049	0.084	1.088	0.989	1.196
$PTZ_1Z_2^*$	−0.005	0.117	0.963	0.995	0.791	1.250
trZ_1^*	0.560	0.732	0.444	1.751	0.417	7.351
trZ_2^*	−0.591	0.523	0.258	0.554	0.199	1.543
$trZ_1Z_2^*$	−0.324	0.942	0.731	0.723	0.114	4.583

No. of subjects = 137 Log likelihood = −335.591 Stratified by SZ^*

Use the above computer results to state the form of the **estimated** hazard model for each of the four strata of the stratification variable SZ^*. Also, for each strata, compute the hazard ratio for the treatment effect adjusted for disease duration, age, and prior therapy.

11. Carry out the likelihood ratio test to evaluate the no-interaction model described in question 4. In carrying out this test, make sure to state the null hypothesis in terms of regression coefficients being set equal to zero in the interaction model fitted in question 10. Also, determine the p-value for this test and state your conclusions about significance as well as which model you prefer, the no-interaction model or the interaction model.
12. The adjusted log–log survival curves for each of the four strata defined by the stratification variable SZ^* (adjusted for treatment status, disease duration, age, and prior therapy) are presented below.



Using this graph, what can you conclude about whether the PH assumption is satisfied for the variables, small cell type and PSbin?
13. Comment on what you think can be learned by graphing adjusted survival curves that compare the two treatment groups for each of the four strata of SZ^*.

Test

The following questions consider a dataset from a study by Caplehorn et al. ("Methadone Dosage and Retention of Patients in Maintenance Treatment," *Med. J. Aust.*, 1991). These data comprise the times in days spent by heroin addicts from entry to departure from one of two methadone clinics. Two other covariates, namely, prison record and maximum methadone dose, are believed to affect the survival times. The dataset name is **addicts.dat**. A listing of the variables is given below:

Column 1: Subject ID
Column 2: Clinic (1 or 2)

Column 3: Survival status (0 = censored, 1 = departed from clinic)
Column 4: Survival time in days
Column 5: Prison record (0 = none, 1 = any)
Column 6: Maximum methadone dose (mg/day)

1. The following edited printout was obtained from fitting a Cox PH model to these data:

Cox regression
Analysis time _t:

survt	Coef.	Std. Err.	p > |z|	Haz. Ratio	[95% Conf.	Interval]	*P*(*PH*)
Clinic	−1.009	0.215	0.000	0.365	0.239	0.556	0.001
Prison	0.327	0.167	0.051	1.386	0.999	1.924	0.332
Dose	−0.035	0.006	0.000	0.965	0.953	0.977	0.341

No. of subjects = 238 Log likelihood = −673.403

Based on the *P*(*PH*) information in the above printout, it appears that clinic does not satisfy the PH assumption; this conclusion is also supported by comparing log–log curves for the two clinics and noticing strong nonparallelism. What might we learn from fitting a stratified Cox (SC) model stratifying on the clinic variable? What is a drawback to using a SC procedure that stratifies on the clinic variable?

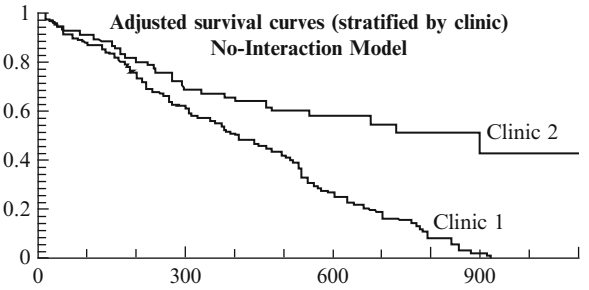
2. The following printout was obtained from fitting a SC PH model to these data, where the variable being stratified is clinic:

Stratified Cox regression
Analysis time _t:

survt	Coef.	Std. Err.	p > |z|	Haz. Ratio	[95% Conf.	Interval]
Prison	0.389	0.169	0.021	1.475	1.059	2.054
Dose	−0.035	0.006	0.000	0.965	0.953	0.978

No. of subjects = 238 Log likelihood = −597.714 Stratified by clinic

Using the above fitted model, we can obtain the adjusted curves below that compare the adjusted survival probabilities for each clinic (i.e., stratified by clinic) adjusted for the variables, prison and maximum methadone dose.



Based on these adjusted survival curves, what conclusions can you draw about whether the survival experience is different between the two clinics? Explain.

3. State the hazard function model being estimated in the above computer results. Why is this model a no-interaction model?
4. Using the above computer results, provide point and interval estimates for the effect of prison adjusted for clinic and dose. Is this adjusted prison effect significant? Explain.
5. The following computer results consider a SC model that allows for interaction of the stratified variable clinic with each of the predictors, prison and dose. Product terms in the model are denoted as clinpr = clinic × prison and clindos = clinic × dose.

Stratified Cox regression
Analysis time _t: survt

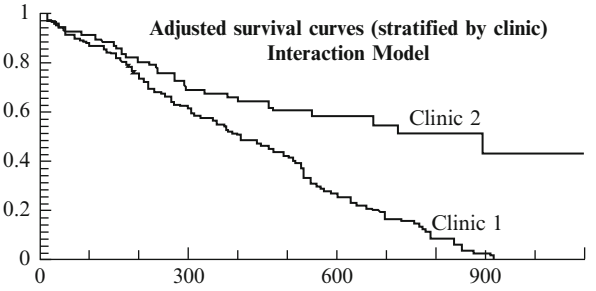
	Coef.	Std. Err.	P > |z|	Haz. Ratio	[95% Conf.	Interval]
prison	1.087	0.539	0.044	2.966	1.032	8.523
dose	−0.035	0.020	0.079	0.966	0.929	1.004
clinpr	−0.585	0.428	0.172	0.557	0.241	1.290
clindos	−0.001	0.015	0.942	0.999	0.971	1.028

No. of subjects = 238 Log likelihood = −596.779 Stratified by clinic

State two alternative versions of the interaction model being estimated by the above printout, where one of these versions should involve the product terms used in the above printout.

6. Using the computer results above, determine the estimated hazard models for each clinic. (Note that the clinics are coded as 1 or 2.)

7. below are the adjusted survival curves for each clinic based on the interaction model results above. These curves are adjusted for the prison and dose variables.



   Compare the survival curves by clinic obtained for the interaction model with the corresponding curves previously shown for the no-interaction model. Do both graphs indicate the similar conclusions about the clinic effect? Explain.
8. Carry out a likelihood ratio test to determine whether the no-interaction model is appropriate. In doing so, make use of the computer information described above, state the null hypothesis, state the form of the likelihood statistic and its distribution under the null hypothesis, and compute the value of the likelihood statistic and evaluate its significance. What are your conclusions?

Answers to Practice Exercises

1. The first printout indicates that the variables large cell, adeno cell, small cell, and performance status do not satisfy the PH assumption at the 0.10 level. The second printout considers a different model that does not contain the large cell and adeno cell variables. This latter printout indicates that small cell satisfies the PH assumption, in contrast to the first printout. The performance status variable, however, does not satisfy the PH assumption as in the first printout.
2. The cell type variable is defined to have four categories, as represented by the three dummy variables in the first printout. The "small cell" variable dichotomizes the cell type variable into the categories small cell type versus the rest. From the second printout, the small cell variable does not appear by itself to violate the PH assumption. This result conflicts with the results of the first printout, for which the cell type variable considered in four categories does not satisfy the PH assumption at the 0.10 level of significance.

We therefore think it is more appropriate to use a SC procedure only if four strata are to be used. A drawback to using four strata, however, is that the number of survival curves to be plotted is larger than for two strata; consequently, a large number of curves is more difficult to interpret graphically than when there are only two curves. Thus, for convenience of interpretation, we may choose to dichotomize the cell type variable instead of considering four strata. We may also consider dichotomies other than those defined by the small cell variable. For instance, we might consider dichotomizing on either the adeno or large cell variables instead of the small cell variable. Alternatively, we may combine categories so as to compare, say, large and adeno cell types with small and squamous types. However, a decision to combine categories should not be just a statistical decision, but should also be based on biologic considerations.

3. $\widehat{HR}_{\text{adj}} = 1.095$, 95% CI: (0.744,1.611), two-tailed p-value is 0.647, not significant. The estimated hazard ratio for treatment is neither meaningfully nor statistically significant. The point estimate is essentially 1, which says that there is no meaningful effect of treatment adjusted for the predictors in the model and for the stratified predictor SZ^*.

4. $h_g(t,\mathbf{X}) = h_{0g}(t)\exp[\beta_1 \text{ Treatment} + \beta_2 DD + \beta_3 \text{ Age} + \beta_4 PT]$, $g = 1,\ldots, 4$, where the strata are defined from the stratification variable SZ^*, DD = disease duration, and PT = prior therapy. This model assumes no interaction because the coefficient of each predictor in the model is not subscripted by g, i.e., the regression coefficients are the same for each stratum.

5. Version 1: $h_g(t,\mathbf{X}) = h_{0g}(t)\exp[\beta_{1g} \text{ Treatment} + \beta_2 DD + \beta_3 \text{ Age} + \beta_4 PT]$, $g = 1,\ldots, 4$.

   Version 2: $h_g(t,\mathbf{X}) = h_{0g}(t)\exp[\beta_1 \text{ Treatment} + \beta_2 DD + \beta_3 \text{Age} + \beta_4 PT + \beta_5(Z_1^* \times \text{Treatment}) + \beta_6(Z_2^* \times \text{Treatment}) + \beta_7(Z_1^* \times Z_2^* \times \text{Treatment})]$, where $Z_1^* =$ small cell type (0, 1), $Z_2^* =$ Psbin (0, 1), and $g = 1,\ldots, 4$.

6. Version 1: $h_g(t,\mathbf{X}) = h_{0g}(t)\exp[\beta_{1g}\,\text{Treatment} + \beta_{2g}\,DD + \beta_{3g}\,\text{Age} + \beta_{4g}\,PT]$, $g = 1,\ldots, 4$.

   Version 2: $h_g(t,\mathbf{X}) = h_{0g}(t)\exp[\beta_1\,\text{Treatment} + \beta_2 DD + \beta_3\,\text{Age} + \beta_4\,PT + \beta_5(Z_1^* \times \text{Treatment}) + \beta_6(Z_1^* \times DD) + \beta_7(Z_1^* \times \text{Age}) + \beta_8(Z_1^* \times PT) + \beta_9(Z_2^* \times \text{Treatment}) + \beta_{10}(Z_2^* \times DD) + \beta_{11}(Z_2^* \times \text{Age}) + \beta_{12}(Z_2^* \times PT) + \beta_{13}(Z_1^* \times Z_2^* \times \text{Treatment}) + \beta_{14}(Z_1^* \times Z_2^* \times DD) + \beta_{15}(Z_1^* \times Z_2^* \times \text{Age}) + \beta_{16}(Z_1^* \times Z_2^* \times PT)]$, $g = 1,\ldots,4$.

7. $HR_g = \exp(\beta_{1g})$, using version 1 model form. Yes, this formula gives different hazard ratios for different strata because the value of the hazard ratio changes with the subscript g.

8. H_0: No interaction assumption is satisfied.

   H_0: $\beta_{11} = \beta_{12} = \beta_{13} = \beta_{14}$, $\beta_{21} = \beta_{22} = \beta_{23} = \beta_{24}$, $\beta_{31} = \beta_{32} = \beta_{33} = \beta_{34}$, $\beta_{41} = \beta_{42} = \beta_{43} = \beta_{44}$ from version 1.
   H_0: $\beta_5 = \beta_6 = \beta_7 = \beta_8 = \beta_9 = \beta_{10} = \beta_{11} = \beta_{12} = \beta_{13} = \beta_{14} = \beta_{15} = \beta_{16} = 0$ from version 2.

9. $LR = -2\ln L_R - (-2\ln L_F)$, where R denotes the reduced (no-interaction) model and F denotes the full (interaction) model. Under the null hypothesis, LR is approximately a chi-square with 12 degrees of freedom.

10. Estimated hazard models for each stratum:

    $g = 1;\ Z_1^* = Z_2^* = 0$:
    $\hat{h}_1,\mathbf{X}) = \hat{h}_{01}(t)\exp[(0.381)\text{Treatment} + (0.015)DD + (0.000)\text{Age} + (0.023)PT]$

    $g = 2;\ Z_1^* = 1,\ Z_2^* = 0$:
    $\hat{h}_2(t,\mathbf{X}) = \hat{h}_{02}(t)\exp[(0.941)\text{Treatment} + (-0.014)DD + (-0.055)\text{Age} + (0.066)PT]$

    $g = 3;\ Z_1^* = 0,\ Z_2^* = 1$:
    $\hat{h}_3(t,\mathbf{X}) = \hat{h}_{03}(t)\exp[(-0.210)\text{Treatment} + (0.040)DD + (0.001)\text{Age} + (-0.055)PT]$

    $g = 4;\ Z_1^* = 1,\ Z_2^* = 1$:
    $\hat{h}_4(t,\mathbf{X}) = \hat{h}_{04}(t)\exp[(0.026)\text{Treatment} + (-0.060)DD + (0.030)\text{Age} + (-0.017)PT]$

Estimated hazard ratios for treatment effect adjusted for DD, Age, and PT:

$$g = 1 : \widehat{HR}_1 = \exp(0.381) = 1.464$$
$$g = 2 : \widehat{HR}_2 = \exp(0.941) = 2.563$$
$$g = 3 : \widehat{HR}_3 = \exp(-0.210) = 0.811$$
$$g = 4 : \widehat{HR}_4 = \exp(0.026) = 1.026$$

11. H_0: $\beta_5 = \beta_6 = \beta_7 = \beta_8 = \beta_9 = \beta_{10} = \beta_{11} = \beta_{12} = \beta_{13} = \beta_{14} = \beta_{15} = \beta_{16} = 0$

    $LR = 689.696 - 671.182 = 18.514$, which is approximately chi-square with 12 df.

    $P = 0.101$, which is not significant below the .05 level. *Conclusion*: Accept the null hypothesis and conclude that the no-interaction model is preferable to the interaction model.

12. The three curves at the bottom of the graph appear to be quite non-parallel. Thus, the PH assumption is not satisfied for one or both of the variables, small cell type and PSbin. Note, however, that because both these variables have been stratified together, it is not clear from the graph whether only one of these variables fails to satisfy the PH assumption.

13. If we graph adjusted survival curves that compare the two treatment groups for each of the four strata, we will be able to see graphically how the treatment effect, if any, varies over time within each strata. The difficulty with this approach, however, is that eight adjusted survival curves will be produced, so that if all eight curves are put on the same graph, it may be difficult to see what is going on.