EXAMPLE 3 (continued)

Want HR for effect of Rx adjusted for log WBC.

Placebo subject:
$X^* = (X_1^* = 1, X_2^* = \log \text{WBC}, X_3^* = 1 \times \log \text{WBC})$

Treated subject:
$X = (X_1 = 0, X_2 = \log \text{WBC}, X_3 = 0 \times \log \text{WBC})$

To obtain the hazard ratio for the effect of Rx adjusted for log WBC using model 3, we consider X^* and X vectors which have three components, one for each variable in the model. The $\boldsymbol{X}^*$ vector, which denotes a treated subject, has components $X_1^* = 1$, $X_2^* = \log$ WBC and $X_3^* = 1$ times log WBC. The $\boldsymbol{X}$ vector, which denotes a placebo subject, has components $X_1 = 0$, $X_2 = \log$ WBC and $X_3 = 0$ times log WBC. Note again that, as with the previous example, the value for log WBC is treated as fixed, though unspecified.

$$\widehat{HR} = \exp\left[\sum_{i=1}^{3} \hat{\beta}_i \left(X_i^* - X_i\right)\right]$$

$$\begin{aligned}\widehat{HR} = \exp[&2.355(1-0)\\ &+ 1.803\,(\log \text{WBC} - \log \text{WBC})\\ &+ (-0.342)(1 \times \log \text{WBC} - 0 \times \log \text{WBC})]\\ = \exp[&2.355 - 0.342 \log \text{WBC}]\end{aligned}$$

Using the general formula for the hazard ratio, we must now compute the exponential of the sum of three quantities, corresponding to the three variables in the model. Substituting the values from the printout and the values of the vectors $\boldsymbol{X}^*$ and $\boldsymbol{X}$ into this formula, we obtain the exponential expression shown here. Using algebra, this expression simplifies to the exponential of 2.355 minus 0.342 times log WBC.

log WBC = 2:
$$\widehat{HR} = \exp[2.355 - 0.342\,(2)] = e^{1.671} = 5.32$$

log WBC = 4:
$$\widehat{HR} = \exp[2.355 - 0.342\,(4)] = e^{0.987} = 2.68$$

In order to get a numerical value for the hazard ratio, we must specify a value for log WBC. For instance, if log WBC = 2, the estimated hazard ratio becomes 5.32, whereas if log WBC = 4, the estimated hazard ratio becomes 2.68. Thus, we get different hazard ratio values for different values of log WBC, which should make sense since log WBC is an effect modifier in model 3.

General rule for (0,1) exposure variables when there are product terms:

$$\widehat{HR} = \exp\left[\hat{\beta} + \sum \hat{\delta}_j W_j\right]$$

where
$\hat{\beta}$ = coefficient of E
$\hat{\delta}_j$ = coefficient of $E \times W_j$

($\widehat{HR}$ does not contain coefficients of nonproduct terms)

The example we have just described using model 3 illustrates a general rule which states that the hazard ratio for the effect of a (0,1) exposure variable in a model which contains product terms involving this exposure with other X's can be written as shown here. Note that β "hat" denotes the coefficient of the exposure variable and the δ "hats" are coefficients of product terms in the model of the form ExW_j. Also note that this formula does not contain coefficients of nonproduct terms other than E.

EXAMPLE

Model 3:

$E \rightarrow$ $\leftarrow W_1$

$\hat{\beta}$ = coefficient of Rx

$\hat{\delta}_1$ = coefficient of $Rx \times$ log WBC

$\widehat{HR}$ (model 3) $= \exp[\hat{\beta} + \hat{\delta}_1 \log \text{WBC}]$
$= \exp[2.355 - 0.342 \log \text{WBC}]$

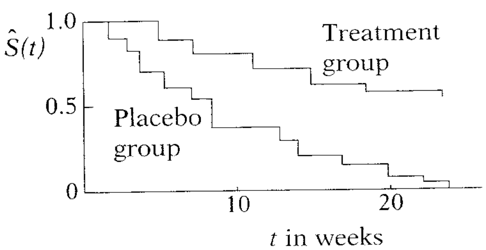
For model 3, β "hat" is the coefficient of the Rx variable, and there is only one δ "hat" in the sum, which is the coefficient of the product term $Rx \times$ log WBC. Thus, there is only one W, namely W_1 = log WBC. The hazard ratio formula for the effect of exposure is then given by exponentiating β "hat" plus δ_1 "hat" times log WBC. Substituting the estimates from the printout into this formula yields the expression obtained previously, namely the exponential of 2.355 minus 0.342 times log WBC.

VI. Adjusted Survival Curves Using the Cox PH Model

Two primary quantities:
1. estimated hazard ratios
2. estimated survival curves

The two primary quantities desired from a survival analysis point of view are estimated hazard ratios and estimated survival curves. Having just described how to compute hazard ratios, we now turn to estimation of survival curves using the Cox model.

No model: use KM curves



Recall that if no model is used to fit survival data, a survival curve can be estimated using a Kaplan–Meier method. Such KM curves are plotted as step functions as shown here for the remission data example.

Cox model: adjusted survival curves (also step functions).

When a Cox model is used to fit survival data, survival curves can be obtained that adjust for the explanatory variables used as predictors. These are called **adjusted survival curves,** and, like KM curves, these are also plotted as step functions.

Cox model hazard function:

$$h(t,\mathbf{X}) = \hat{h}_0(t)\, e^{\sum_{i=1}^{p} \beta_i X_i}$$

Cox model survival function:

$$S(t,\mathbf{X}) = \left[S_0(t)\right]^{e^{\sum \beta_i X_i}}$$

The hazard function formula for the Cox PH model, shown here again, can be converted to a corresponding survival function formula as shown below. This survival function formula is the basis for determining adjusted survival curves. Note that this formula says that the survival function at time t for a subject with vector $\mathbf{X}$ as predictors is given by a baseline survival function $S_0(t)$ raised to a power equal to the exponential of the sum of β_i times X_i.

Estimated survival function:

$$\hat{S}(t,\mathbf{X}) = \left[\hat{S}_0(t)\right]^{e^{\sum \hat{\beta}_i X_i}}$$

The expression for the estimated survival function can then be written with the usual "hat" notation as shown here.

$\hat{S}_0(t)$ and $\hat{\beta}_i$ are provided by the computer program. The X_i must be specified by the investigator.

The estimates of $\hat{S}_0(t)$ and $\hat{\beta}_i$ are provided by the computer program that fits the Cox model. The X's, however, must first be specified by the investigator before the computer program can compute the estimated survival curve.

EXAMPLE: Model 2 Remission Data

$$\hat{h}(t,\mathbf{X}) = \hat{h}_0(t)e^{1.294\,Rx + 1.604 \log \text{WBC}}$$

$$\hat{S}(t,\mathbf{X}) = [\hat{S}_0(t)]^{e^{1.294\,Rx + 1.604 \log \text{WBC}}}$$

Specify values for $\mathbf{X} = (Rx, \log \text{WBC})$:

$Rx = 1$, $\log \text{WBC} = 2.93$:

$$\hat{S}(t,\mathbf{X}) = [\hat{S}_0(t)]^{e^{1.294(1) + 1.604(2.93)}} = [\hat{S}_0(t)]^{e^{5.99}} = [\hat{S}_0(t)]^{400.9}$$

$Rx = 0$, $\log \text{WBC} = 2.93$:

$$\hat{S}(t,\mathbf{X}) = [\hat{S}_0(t)]^{e^{1.294(0) + 1.604(2.93)}} = [\hat{S}_0(t)]^{e^{4.70}} = [\hat{S}_0(t)]^{109.9}$$

For example, if we consider model 2 for the remission data, the fitted model written in terms of both the hazard function and corresponding survival function is given here.

We can obtain a specific survival curve by specifying values for the vector $\mathbf{X}$, whose component variables are Rx and log WBC.

For instance, if $Rx = 1$ and log WBC = 2.93, the estimated survival curve is obtained by substituting these values in the formula as shown here, and carrying out the algebra to obtain the expression circled. Note that the value 2.93 is the overall mean log WBC for the entire dataset of 42 subjects.

Also, if $Rx = 0$ and log WBC = 2.93, the estimated survival curve is obtained as shown here.

EXAMPLE (continued)

Adjusted Survival Curves

$Rx = 1$, log WBC = 2.93:

$\hat{S}(t,\mathbf{X}) = [\hat{S}_0(t)]^{400.9}$

$Rx = 0$, log WBC = 2.93:

$\hat{S}(t,\mathbf{X}) = [\hat{S}_0(t)]^{109.9}$

Each of the circled expressions gives **adjusted** survival curves, where the adjustment is for the values specified for the X's. Note that for each expression, a survival probability can be obtained for any value of t.

The two formulae just obtained, again shown here, allow us to compare survival curves for different treatment groups adjusted for the covariate log WBC. Both curves describe estimated survival probabilities over time assuming the same value of log WBC, in this case, the value 2.93.

Typically, use $X = \overline{X}$ or X_{median}.

Computer uses $\overline{X}$.

Typically, when computing adjusted survival curves, the value chosen for a covariate being adjusted is an average value like an arithmetic mean or a median. In fact, most computer programs for the Cox model automatically use the mean value over all subjects for each covariate being adjusted.

EXAMPLE (continued)

Remission data ($n = 42$):

$\overline{\log \text{WBC}} = 2.93$

In our example, the mean log WBC for all 42 subjects in the remission data set is 2.93. That is why we chose this value for log WBC in the formulae for the adjusted survival curve.

General formulae for adjusted survival curves comparing two groups:

Exposed subjects:

$$\hat{S}(t,\mathbf{X}_1) = \left[\hat{S}_0(t)\right]^{e^{\hat{\beta}_1(1) + \sum_{i \neq 1} \hat{\beta}_i \overline{X}_i}}$$

Unexposed subjects:

$$\hat{S}(t,\mathbf{X}_0) = \left[\hat{S}_0(t)\right]^{e^{\hat{\beta}_1(0) + \sum_{i \neq 1} \hat{\beta}_i \overline{X}_i}}$$

More generally, if we want to compare survival curves for two levels of an exposure variable, and we want to adjust for several covariates, we can write the formula for each curve as shown here. Note that we are assuming that the exposure variable is variable X_1, whose estimated coefficient is β_1 "hat," and the value of X_1 is 1 for exposed and 0 for unexposed subjects.

General formula for adjusted survival curve for all covariates in the model:

$$\hat{S}(t,\overline{\mathbf{X}}) = \left[\hat{S}_0(t)\right]^{e^{\sum \hat{\beta}_i \overline{X}_i}}$$

Also, if we want to obtain an adjusted survival curve which adjusts for all covariates in the model, the general formula which uses the mean value for each covariate is given as shown here. This formula will give a single adjusted survival curve rather than different curves for each exposure group.

EXAMPLE (continued)

Single survival curve for Cox model containing Rx and log WBC:

$\overline{Rx} = 0.50$

$\overline{\log \text{WBC}} = 2.93$

$$\hat{S}(t,\bar{\mathbf{X}}) = [\hat{S}_0(t)]^{e^{\hat{\beta}_1 \overline{Rx} + \hat{\beta}_2 \overline{\log \text{WBC}}}}$$

$$= [\hat{S}_0(t)]^{e^{1.294(0.5) + 1.604(2.93)}}$$

$$= [\hat{S}_0(t)]^{e^{5.35}} = [\hat{S}_0(t)]^{210.6}$$

To illustrate this formula, suppose we again consider the remission data, and we wish to obtain a single survival curve that adjusts for both Rx and log WBC in the fitted Cox model containing these two variables. Using the mean value of each covariate, we find that the mean value for Rx is 0.5 and the mean value for log WBC is 2.93, as before.

To obtain the single survival curve that adjusts for Rx and log WBC, we then substitute the mean values in the formula for the adjusted survival curve for the model fitted. The formula and the resulting expression for the adjusted survival curve are shown here. (Note that for the remission data, where it is of interest to compare two exposure groups, the use of a single survival curve is not appropriate.)

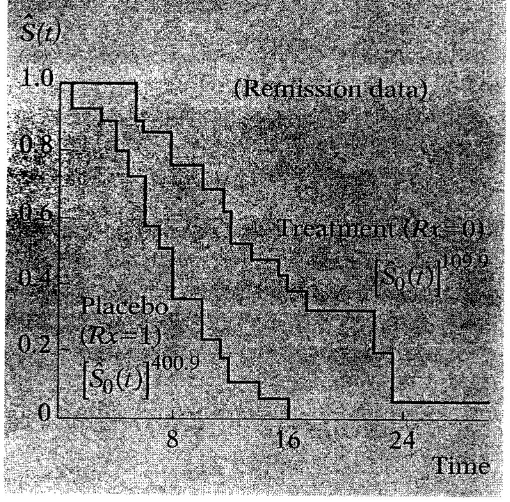
Compute survival probability by specifying value for t in $\hat{S}(t,\bar{\mathbf{X}}) = [\hat{S}_0(t)]^{210.6}$

Computer uses t's which are failure times.

From this expression for the survival curve, a survival probability can be computed for any value of t that is specified. When graphing this survival curve using a computer package, the values of t that are chosen are the failure times of all persons in the study who got the event. This process is automatically carried out by the computer without having the user specify each failure time.

EXAMPLE

Adjusted Survival Curves for Treatment and Placebo Groups



The graph of adjusted survival curves obtained from fitting a Cox model is usually plotted as a step function. For example, we show here the step functions for the two adjusted survival curves obtained by specifying either 1 or 0 for treatment status and letting log WBC be the mean value 2.93.

Next section: PH assumption
- explain meaning
- when PH **not** satisfied

We now turn to the concept of the proportional hazard (PH) assumption. In the next section, we explain the meaning of this assumption and we give an example of when this assumption is not satisfied.

Later presentations:
- how to evaluate PH
- analysis when PH not met

In later presentations, we expand on this subject, describing how to evaluate statistically whether the assumption is met and how to carry out the analysis when the assumption is not met.

VII. The Meaning of the PH Assumption

PH: *HR* is constant over time, i.e.,

$$\hat{h}(t,\mathbf{X}^*) = \text{constant} \times \hat{h}(t,\mathbf{X})$$

The PH assumption requires that the *HR* is constant over time, or equivalently, that the hazard for one individual is proportional to the hazard for any other individual, where the proportionality constant is independent of time.

$$\widehat{HR} = \frac{\hat{h}(t,\mathbf{X}^*)}{\hat{h}(t,\mathbf{X})} = \frac{\hat{h}_0(t)\, e^{\sum_{i=1}^{p} \hat{\beta}_i X_i^*}}{\hat{h}_0(t)\, e^{\sum_{i=1}^{p} \hat{\beta}_i X_i}} = \exp\left[\sum_{i=1}^{p} \hat{\beta}_i \left(X_i^* - X_i\right)\right]$$

where $\mathbf{X}^* = (X_1^*, X_2^*, \ldots, X_p^*)$
and $\mathbf{X} = (X_1, X_2, \ldots, X_p)$
denote the set of *X*'s for two individuals.

To understand the PH assumption, we need to reconsider the formula for the *HR* that compares two different specifications $\mathbf{X}^*$ and $\mathbf{X}$ for the explanatory variables used in the Cox model. We derived this formula previously in Section V, page 99, and we show this derivation again here. Notice that the baseline hazard function $\hat{h}_0(t)$ appears in both the numerator and denominator of the hazard ratio and cancels out of the formula.

$$\frac{\hat{h}(t,\mathbf{X}^*)}{\hat{h}(t,\mathbf{X})} = \exp\left[\sum_{i=1}^{p} \hat{\beta}_i \left(X_i^* - X_i\right)\right]$$

does not involve *t*.

The final expression for the hazard ratio therefore involves the estimated coefficients β_1 "hat" and the values of $\mathbf{X}^*$ and $\mathbf{X}$ for each variable. However, because the baseline hazard has canceled out, the final expression does not involve time *t*.

Let (Constant)

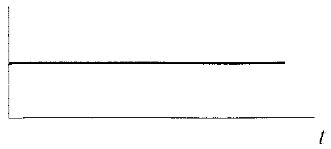
$$\hat{\theta} = \exp\left[\sum_{i=1}^{p} \hat{\beta}_i \left(X_i^* - X_i\right)\right]$$

then

$$\frac{\hat{h}(t,\mathbf{X}^*)}{\hat{h}(t,\mathbf{X})} = \hat{\theta}$$

Thus, once the model is fitted and the values for $\mathbf{X}^*$ and $\mathbf{X}$ are specified, the value of the exponential expression for the estimated hazard ratio is a constant, which does not depend on time. If we denote this constant by θ "hat," then we can write the hazard ratio as shown here. This is a mathematical expression which states the proportional hazards assumption.

$\widehat{HR}$ ($\mathbf{X}^*$ versus $\mathbf{X}$)



Graphically, this expression says that the estimated hazard ratio comparing any two individuals, plots as a constant over time.

$$\hat{h}(t, X^*) = \hat{\theta}\hat{h}(t, X)$$

Proportionality constant (not dependent on time)

Another way to write the proportional hazards assumption mathematically expresses the hazard function for individual $\mathbf{X}^*$ as θ "hat" times the hazard function for individual $\mathbf{X}$, as shown here. This expression says that the hazard function for one individual is proportional to the hazard function for another individual, where the proportionality constant is θ "hat," which does not depend on time.

EXAMPLE: Remission Data

$$\hat{h}(t,\mathbf{X}) = \hat{h}_0(t)e^{1.294\,Rx + 1.604\,\log \text{WBC}}$$

$$\widehat{HR} = \frac{\hat{h}(t, Rx = 1, \log \text{WBC} = 2.93)}{\hat{h}(t, Rx = 0, \log \text{WBC} = 2.93)} = \exp[1.294] = 3.65 \text{ Constant}$$

Placebo

$$\hat{h}(t, Rx = 1, \log \text{WBC} = 2.93) = 3.65\hat{h}(t, Rx = 0, \log \text{WBC} = 2.93)$$

Treatment

3.65 = proportionality constant

To illustrate the proportional hazard assumption, we again consider the Cox model for the remission data involving the two variables Rx and log WBC. For this model, the estimated hazard ratio that compares placebo ($Rx = 1$) with treated ($Rx = 0$) subjects controlling for log WBC is given by e to the 1.294, which is 3.65, a constant.

Thus, the hazard for placebo group ($Rx = 1$) is 3.65 times the hazard for the treatment group ($Rx = 0$), and the value, 3.65, is the same regardless of time. In other words, using the above model, the hazard for the placebo group is proportional to the hazard for the treatment group, and the proportionality constant is 3.65.

EXAMPLE: PH Not Satisfied

Cancer patients → Surgery; Radiation with no surgery

$$E = \begin{cases} 0 & \text{if surgery} \\ 1 & \text{if no surgery} \end{cases}$$

$$h(t,\mathbf{X}) = h_0(t)e^{\beta E}$$

To further illustrate the concept of proportional hazards, we now provide an example of a situation for which the proportional hazards assumption is *not* satisfied.

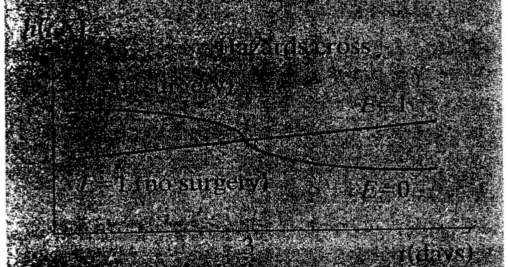
For our example, we consider a study in which cancer patients are randomized to either surgery or radiation therapy without surgery. Thus, we have a (0,1) exposure variable denoting surgery status, with 0 if a patient receives surgery and 1 if not. Suppose further that this exposure variable is the only variable of interest, so that a Cox PH model for the analysis of this data, as shown here, will contain only the one variable E, denoting exposure.

EXAMPLE (continued)

Is the above Cox PH model appropriate?

Note:

Serious surgery ⇒ High risk for death early



$$\frac{h(t=2, E=1)}{h(t=2, E=0)} < 1$$

but

$$\frac{h(t=5, E=1)}{h(t=5, E=0)} > 1$$

Given the above description, ***HR* is not constant over time.**

Cox PH model inappropriate because PH model assumes constant *HR*

[illegible]

Now the question we consider here is whether the above Cox model containing the variable E is an appropriate model to use for this situation. To answer this question we note that when a patient undergoes serious surgery, as when removing a cancerous tumor, there is usually a high risk for complications from surgery or perhaps even death early in the recovery process, and once the patient gets past this early critical period, the benefits of surgery, if any, can then be observed.

Thus, in a study that compares surgery to no surgery, we might expect to see hazard functions for each group that appear as shown here. Notice that these two functions cross at about three days, and that prior to three days the hazard for the surgery group is higher than the hazard for the no surgery group, whereas after three days, the hazard for the surgery group is lower than the hazard for the no surgery group.

Looking at the above graph more closely, we can see that at 2 days, when $t = 2$, the hazard ratio of nonsurgery ($E = 1$) to surgery ($E = 0$) patients yields a value less than 1. In contrast, at $t = 5$ days, the hazard ratio of nonsurgery to surgery yields a value greater than 1.

Thus, if the above description of the hazard functions for each group is accurate, the hazard ratios are not constant over time. That is, the hazard ratio is some number less than 1 before three days and greater than 1 after three days.

It is therefore inappropriate to use a Cox PH model for this situation, because the PH model assumes a constant hazard ratio across time, whereas our situation yields a hazard ratio that varies with time.

In fact, if we use a Cox PH model, shown here again, the estimated hazard ratio comparing exposed to unexposed patients at any time is given by the constant value e to the β "hat," which does not vary over time.

General rule:
If the hazards cross, then a Cox PH model is not appropriate.

This example illustrates the general rule that if the hazards cross, then the PH assumption cannot be met, so that a Cox PH model is inappropriate.

Analysis when Cox PH model not appropriate? See Chapters 5 and 6.

It is natural to ask at this point, if the Cox PH model is inappropriate, how should we carry out the analysis? The answer to this question is discussed in Chapters 5 and 6. However, we will give a brief reply with regard to the surgery study example just described.

EXAMPLE (continued)

Surgery study analysis options:
- stratify by exposure (use KM curves)
- start analysis at three days; use Cox PH model
- fit PH model for < 3 days and for > 3 days; get $\widehat{HR}$ (< 3 days) and $\widehat{HR}$ (≥ 3 days)
- include time-dependent variable (e.g., $E \times t$); use extended Cox model

Actually, for the surgery study there are several options available for the analysis. These include:

- analyze by stratifying on the exposure variable; that is, do not fit any model, and, instead obtain Kaplan–Meier curves for each exposure group separately;
- start the analysis at three days, and use a Cox PH model on three-day survivors;
- fit Cox model for less than three days and a different Cox model for greater than three days to get two different hazard ratio estimates, one for each of these two time periods;
- fit a modified Cox model that includes a time-dependent variable which measures the interaction of exposure with time. This model is called an **extended Cox model.**

Different options may lead to different conclusions.

Further discussion of these options is beyond the scope of this presentation. We point out, however, that different options may lead to different conclusions, so that the investigator may have to weigh the relative merits of each option in light of the data actually obtained before deciding on any particular option as best.

Hazards cross $\Rightarrow$ PH not met

but

? $\Rightarrow$ PH met

See Chapter 4: Evaluating PH Assumption

One final comment before concluding this section: although we have shown that when the hazards cross, the PH assumption is not met, we have not shown how to decide when the PH assumption is met. This is the subject of Chapter 4 entitled, "Evaluating the PH Assumption."

VIII. Summary

In this section we briefly summarize the content covered in this presentation.

1. Review: $S(t)$, $h(t)$, data layout, etc.
2. Computer example of Cox model:
   - estimate HR
   - test hypothesis about HR
   - obtain confidence intervals

- We began with a computer example that uses the Cox PH model. We showed how to use the output to estimate the HR, and how to test hypotheses and obtain confidence intervals about the hazard ratio.

3. Cox model formula:

$$h(t,\mathbf{X}) = h_0(t)\, e^{\sum_{i=1}^{p} \beta_i X_i}$$

- We then provided the formula for the hazard function for the Cox PH model and described basic features of this model. The most important feature is that the model contains two components, namely, a baseline hazard function of time and an exponential function involving X's but not time.

4. Why popular: Cox PH model is "robust"

- We discussed reasons why the Cox model is popular, the primary reason being that the model is "robust" for many different survival analysis situations.

5. ML estimation: maximize a partial likelihood
   L = joint probability of observed data = $L(\beta)$

- We then discussed ML estimation of the parameters in the Cox model, and pointed out that the ML procedure maximizes a "partial" likelihood that focuses on probabilities at failure times only.

6. Hazard ratio formula:

$$\widehat{HR} = \exp\left[\sum_{i=1}^{p} \hat{\beta}_i \left(X_i^* - X_i\right)\right]$$

- Next, we gave a general formula for estimating a hazard ratio that compared two specifications of the X's, defined as $\mathbf{X}^*$ and $\mathbf{X}$. We illustrated the use of this formula when comparing two exposure groups adjusted for other variables.

7. Adjusted survival curves:
   Comparing E groups:

$$\hat{S}(t,\mathbf{X}) = \left[\hat{S}_0(t)\right]^{e^{\hat{\beta}_1 (E) + \sum_{i \neq 1} \hat{\beta}_i \bar{X}_i}}$$

(E = 0 or 1)

   Single curve:

$$\hat{S}(t,\bar{\mathbf{X}}) = \left[\hat{S}_0(t)\right]^{e^{\sum \hat{\beta}_i \bar{X}_i}}$$

- We then defined an adjusted survival curve and presented formulas for adjusted curves comparing two groups adjusted for other variables in the model and a formula for a single adjusted curve that adjusts for all X's in the model. Computer packages for these formulae use the mean value of each X being adjusted in the computation of the adjusted curve.

8. PH assumption:

$$\frac{\hat{h}(t,\mathbf{X}^*)}{\hat{h}(t,\mathbf{X})} = \hat{\theta} \quad \text{(a constant over } t\text{)}$$

   i.e., $\hat{h}(t,\mathbf{X}^*) = \hat{\theta}\hat{h}(t,\mathbf{X})$

Hazards cross $\Rightarrow$ PH not met

- Finally, we described the PH assumption as meaning that the hazard ratio is constant over time, or equivalently that the hazard for one individual is proportional to the hazard for any other individual, where the proportionality constant is independent of time. We also showed that for study situations in which the hazards cross, the PH assumption is not met.