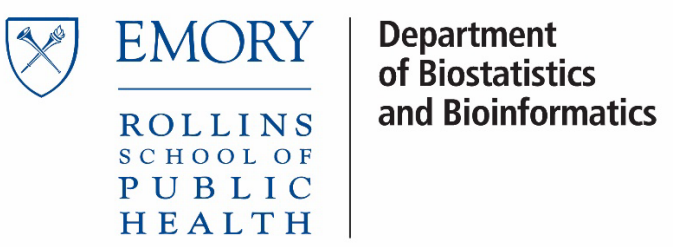


BIOS 522: Survival Analysis Methods

Reading 8:

Cox model extensions

This week, we will study several extensions to the model. We will learn about stratified Cox models which allow the baseline hazard function to vary across strata. We will study time-dependent covariates and time-varying effects to increase the flexibility of the model.

Part 1. The stratified Cox model

Motivation for a stratified model

Consider data from the ACTG 320 trial. The primary purpose of the trial was to evaluate the effect of indinavir on survival in an HIV-infected population. The trial included two distinct subgroups: patients with CD4 cell count ≤50 (sickest patients), and patients with CD4 cell count >50 (less sick patients). One approach to model indinavir and simultaneously adjust for CD4 cell count group is to include two terms in our Cox model:

$$h_i(t) = h_0(t)\exp\left(\beta_{trt}X_{i,trt} + \beta_{CD4}I\left[X_{i,CD4} \leq 50\right]\right)$$

This model assumes: (i) the effect of treatment is the same across CD4 groups, and (ii) participants with lower CD4 can have different hazard functions, but the hazard functions are the same shape and differ by a constant ratio.

Where we are unwilling to assume that the treatment effect is common across CD4 groups, we can add an interaction term, but we still assume a common shape for the baseline hazard function across CD4 groups. This is because we assume a single baseline hazard for all four groups (treated & low CD4, untreated & low CD4, treated & higher CD4, untreated & higher CD4).

Stratified Cox model formulation

Where we are unwilling to assume that hazards are proportional across CD4 groups, a reasonable approach is to fit a separate baseline hazard function for each CD4 group. This is known as a **stratified Cox model**.

$$h_{si}(t) = h_{s0}(t)\exp\left(\beta_{trt}X_{si,trt}\right)$$

where $h_{s0}(t)$ is the stratum-specific baseline hazard function for stratum s. $h_{si}(t)$ is the hazard function for individual i in stratum s, who has covariate $X_{si,trt}$. Note that we still model the effect of treatment with a single coefficient β_{trt}. Thus, we assume a common effect of treatment across strata.

The form of the partial likelihood for the sth stratum is identical to the partial likelihood used in the unstratified model, but it includes an additional subscript s indicating the stratum. The full stratified partial likelihood is then the product of these partial likelihoods across strata:

$$L(\beta) = \prod_{s=1}^{S}\prod_{i=1}^{m_s}\frac{\exp(\beta X_{si})}{\sum_{j\in R(T_{si})}\exp(\beta X_{si})}$$

In the above, there are $s = 1, \ldots, S$ strata, each of which has m_s unique failure times. For each stratum, we take the product over these unique failure times $T_{s1}, \ldots, T_{sm_s}$. Note that the risk set $R(T_{si})$ in the denominator includes only individuals from the same stratum s.

A stratified Cox model allows us to fit common log hazard ratios $\hat{\beta}$ across all strata (e.g. a common effect of treatment for both CD4 strata) but allows each stratum to have its own baseline hazard function.

Notes on the stratified Cox model

Stratification is an extremely useful extension of the standard Cox model. It is one way to relax the proportional hazards assumption. By stratifying on CD4 level, it is no longer necessary to assume that the two CD4 groups have proportional hazards.

But the extension comes at a price. When there is actually a single baseline hazard function, the stratified model is less efficient than an unstratified model.

One also cannot compare the model easily to an unstratified model using an information criterion (e.g. AIC, BIC). One cannot readily test the effect of the stratification variable(s). There is no summary statistic for how the stratum-specific baseline hazard functions compare to each other (nothing similar to a hazard ratio for low versus high CD4). The best use of stratification is when the effects of the stratification variables are of no direct interest.

Part 2. Time-dependent covariates

Time-dependent covariate formulation

So far, we have considered the following Cox proportional hazards model:

$$h_i(t) = h_0(t)\exp\left(\beta_1 X_{i1} + \cdots + \beta_k X_{ik}\right)$$

In this model, two central features are:

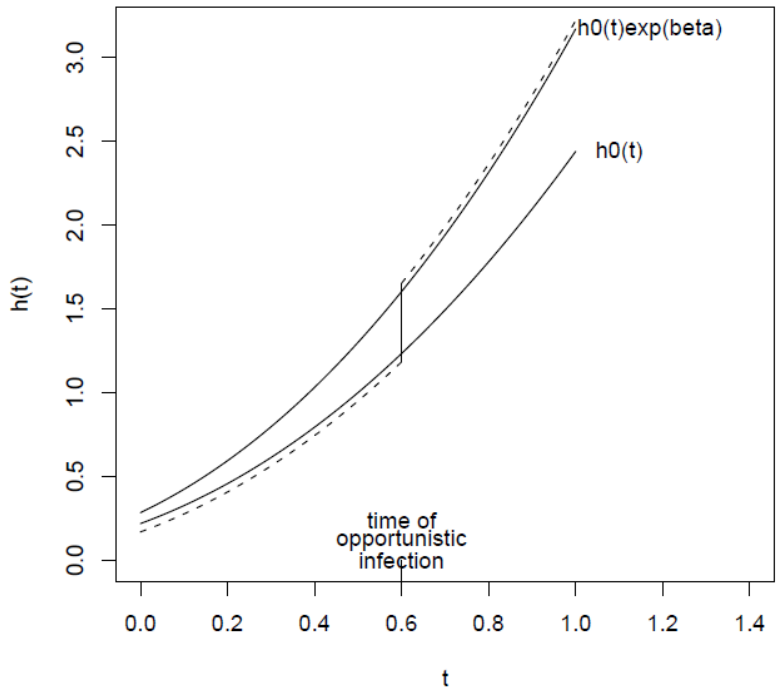
(1) The baseline hazard $h_0(t)$ depends on t, but not on the covariates $X_{i1}, \ldots, X_{ik}$.
(2) The hazard ratio $\exp(\beta_1 X_{i1} + \cdots + \beta_k X_{ik})$ depends on the covariates $X_{i1}, \ldots, X_{ik}$, but not on time t.

Frequently, we include the baseline values of the covariates in our model. These are the values measured at the start of the study. These covariates do not change over time.

An important advantage of the Cox model is that it can be extended to allow covariates to change during the course of the study. These are known as **time-dependent covariates**. Instead of a value fixed at the start of the study X_{ij}, the covariates are represented as functions of time:

$$h_i(t) = h_0(t)\exp\left(\beta_1 X_{i1}(t) + \cdots + \beta_k X_{ik}(t)\right)$$

The following figure nicely demonstrates the concept of time-dependent covariates:



Imagine we are studying survival time for a cohort of AIDS patients. We predict survival using a model with a single time-dependent covariate for whether a patient currently has an opportunistic infection. Observe in the plot how the patient acquires an opportunistic infection at time 0.6, at which time their covariate for having an infection "turns on" and the patient's hazard function switches to the higher hazard function for an individual with an opportunistic infection.

Note that the log hazard ratios $\beta_1, \ldots, \beta_k$ are not functions of time. Thus, the hazard ratio at time t depends (only) on the value of the covariates at time t, and it does not depend on time itself. Using the example above, the relative increase in hazard associated with an opportunistic infection is the same whether the opportunistic infection occurs at time 0.2, time 0.6, time 1.0 or any time. What changes is only when the patient switches from the $h_0(t)$ to the $h_0(t)\exp(\beta)$ curve.

Fitting a model with a time-dependent covariate

To estimate the coefficient $\hat{\beta}$ for our time-dependent Cox model, we will maximize the partial likelihood as before. The difference is that the value of the covariate can now change over time. At each distinct failure time T_i, the model examines who is at risk and the current value of their covariate at T_i.

$$L(\beta) = \prod_{i=1}^{m} \frac{\exp\left(\beta X_i(T_i)\right)}{\sum_{j \in R(T_i)} \exp\left(\beta X_i(T_i)\right)}$$

Estimation is achieved in R (and other software) by creating a new **pseudo-observation** each time an individual's time-dependent covariate changes its value. There may be more than one pseudo-observation per individual. For those interested in learning more about the topic, a vignette is linked below.[1]

Part 3. Time-varying effects

Distinguishing time-dependent covariates and time-varying effects

A related but distinct concept is that of **time-varying effects**. Recall that a central assumption of the proportional hazards model is that the hazard ratio capturing the effect of a covariate is constant over time. When the hazard ratio is not constant over time, the covariate is said to have a **time-varying effect**. For example, the effect of treatment may fade over time, or a treatment may only become active after some time lag/induction period. This change in the magnitude of the effect over time violates the proportional hazards assumption.

[1] https://cran.r-project.org/web/packages/survival/vignettes/timedep.pdf

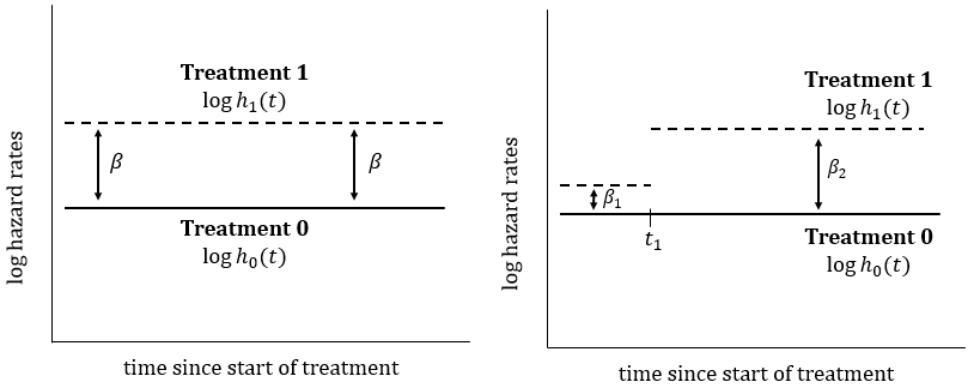
Though the concepts are related, a time-varying effect is slightly different from a time-dependent covariate. For a time-dependent covariate, the value of the *covariate* changes over time, but the relationship between the covariate and the hazard is constant. Consider our example of a model for time until hospital-acquired infection for hospitalized patients. A time-dependent covariate is whether the patient is on antibiotics. As a time-dependent covariate, the effect of the antibiotic on reducing the hazard of hospital acquired-infection is the same at the beginning or at the end of a hospital stay. With a *time-varying effect*, a patient's antibiotic status might not change over time, but one could model that, the *longer* a patient is on antibiotics, the *less effective* they might become.

We can model the effect of covariate X_i with a time-varying effect $\beta(t)$ as shown in the following model:

$$h_i(t) = h_0(t)\exp(\beta(t)X_i)$$

This model allows the effect $\beta(t)$ of X_i to change with time.

For example, consider two treatments (treatment 1 and treatment 0, e.g. placebo). The log hazard rates are plotted below. The x-axis is time since start of treatment.



In the left-hand example, the effect of treatment 1 is β. This measures the log hazard ratio. Note that this effect is constant at all times, satisfying the proportional hazards assumption.

In the right-hand example, the effect of treatment 1 varies over time. It is low before time t_1 (β_1 positive but near 0) and high after time t_1 (β_2 much greater than 0). The effect is modeled as **piecewise constant**, with time separated into periods, and a constant effect within each of the two time periods.

Note that, in our model, treatment status X_i does not change over time. Individuals are in one of two groups, and they do not cross-over between groups. While the covariates do not change, we can use time-dependent covariates to express a time-varying effect.

Continuing with the piecewise constant example, the effect of treatment is modeled by distinct log hazard ratios before and after the cut point t_1:

$$\beta(t) = \begin{cases} \beta_1 & \text{for } t \leq t_1 \\ \beta_2 & \text{for } t > t_1 \end{cases}$$

β_1 is the log hazard ratio for the effect of treatment before time t_1, and β_2 is the log hazard ratio for the effect of treatment after time t_1. To incorporate this into a Cox model, we can add indicator variables that are a *function of time t*. For example:

$$h_i(t) = h_0(t) \exp(\beta_1 I[t \leq t_1] X_i + \beta_2 I[t > t_1] X_i)$$

At times before t_1, the effect of X_i is modeled by β_1. After t_1, the effect of X_i is modeled by β_2. Depending on how we divide time into intervals, the piecewise constant model can approximate any shape of $\beta(t)$. The intervals can be of any size and in any position.

EXAMPLE FROM THE LITERATURE.

Waning Tdap Effectiveness in Adolescents

Nicola P. Klein, MD, PhD, Joan Bartlett, MPH, MPP, Bruce Fireman, MA, Roger Baxter, MD

Goal: Researchers investigated the effectiveness of tetanus toxoid, reduced diphtheria toxoid, and acellular pertussis (Tdap) for protection against pertussis. The researchers were interested in waning of vaccine effectiveness over time.

Population: The researchers used routinely collected data from adolescents within the Kaiser Permanente Northern California health care network. The study population included network members starting at age 10 years.

Outcome variable: The primary outcome was time until first occurrence of testing positive by polymerase chain reaction (PCR) for pertussis. The time origin was the child's 10th birthday. Children were censored at time of receipt of a second Tdap vaccine (booster), disenrollment from the health care network, or end of follow-up.

Predictor variables: The predictor variable of greatest interest was Tdap vaccination status. Tdap vaccination status was coded using a set of indicator variables to allow the effect of vaccine to vary over time. Individuals in the population were:

- Unvaccinated
- Too-recently-vaccinated-to-benefit (within 1-7 days of vaccination)
- Vaccinated in the previous 8 days to <1 year ("year 1")
- Vaccinated in the previous 1 to <2 years ("year 2")
- Vaccinated in the previous 2 to <3 years ("year 3")
- Vaccinated ≥ 3 years ago ("year 4+")

Vaccine efficacy was assessed for each of the 4 ranges of vaccinated person-time beginning 8 days after receipt of Tdap. Adolescents were considered unvaccinated until they received Tdap and then moved through the year-since-vaccination indicator variables with each additional year of follow-up.

Note that here vaccination is modeled as both a *time-dependent covariate* (each person has a different time from their 10th birthday until vaccination) and a *time-varying effect* (the effect of vaccine is allowed to change as a function of time since vaccination).

Other key covariates include gender, age group, birth year, and race/ethnicity, and facility.

Statistical analysis: Researchers fit a Cox regression model to estimate the hazard ratio of pertussis for each Tdap time interval compared with the unvaccinated reference period. For example, the hazard ratio for the 1 to <2 years since vaccination variable estimates the hazard of pertussis in an adolescent who received Tdap 1 to <2 years ago divided by the hazard in an otherwise similar unvaccinated adolescent. The regression was adjusted for birth year, gender, race, and facility.

Results: Incidence of pertussis varied by year, peaking sharply during outbreaks in 2010 and 2014. The study population included 1207 pertussis cases among 279,493 persons contributing 792,418 person-years. Table 3 summarizes the hazard ratio for each Tdap time interval. Tdap vaccine efficacy (VE = 1-HR) steadily decreased each additional year after vaccination.

TABLE 3 Tdap VE by Year After Tdap Vaccination

Year After Tdap (Time Since Tdap)	HR (95% CI)	Tdap VE(95% CI)
Year 1 (8 d to <1 y)	0.31 (0.24 to 0.40)	68.8 (59.7 to 75.9)
Year 2 (1 to <2 y)	0.43 (0.32 to 0.59)	56.9 (41.3 to 68.4)
Year 3 (2 to <3 y)	0.75 (0.54 to 1.04)	25.2 (−4.3 to 46.4)
Year 4+ (≥3 y)	0.91 (0.64 to 1.31)	8.9 (−30.6 to 36.4)

■

Part 4. Looking ahead

Next week we will continue our discussion of regression methods, but move on to parametric approaches, like accelerated failure time models and parametric proportional hazards models.