XVII Summer School of the Master's degree in Statistics and Operations Research

Multi-state models: Rates, risks, and pseudo-values

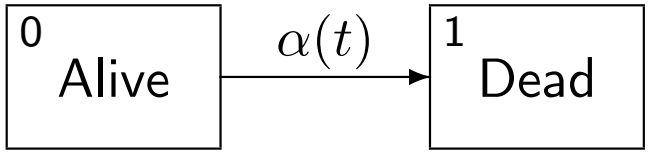
II: **Regression models for intensities (Cox)**

Per Kragh Andersen and Henrik Ravn

II: Regression models for intensities (Cox)

- Two-state model: Cox model for survival data
- Competing risks: Cox model for cause-specific hazard
- Recurrent events

Two-state model (survival data)



Transition intensity: *hazard function*

$$\alpha(t) = \alpha_{01}(t) \approx P(\text{state } 1 \text{ time } t + dt \mid \text{state } 0 \text{ time } t)/dt.$$

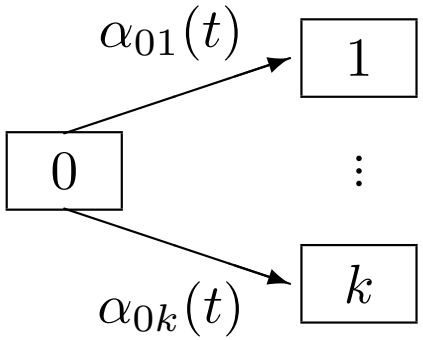
State occupation probabilities: *survival function*

$$Q_0(t) = S(t) = P(\text{state 0 time } t) = P(T > t) = \exp(-\int_0^t \alpha(u)du),$$

and *cumulative probability of death before time* t, Eq. (1.2):

$$Q_1(t) = 1 - Q_0(t) = F(t) = P(\text{state 1 time } t) = 1 - \exp(-\int_0^t \alpha(u)du).$$

Competing risks



Transition intensities: *cause-specific hazards* $h = 1, \ldots, k$:

$$\alpha_h(t) = \alpha_{0h}(t) \approx P(\text{state } h \text{ time } t + dt \mid \text{state } 0 \text{ time } t)/dt.$$

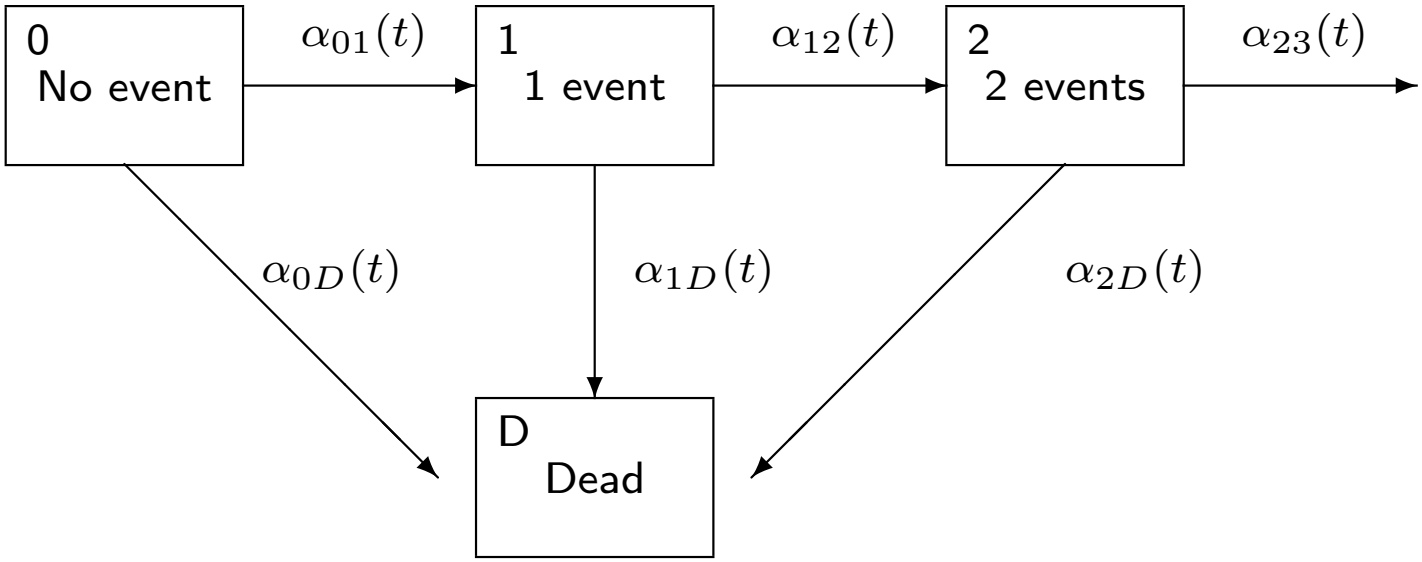
State occupation probabilities: *overall survival function*:

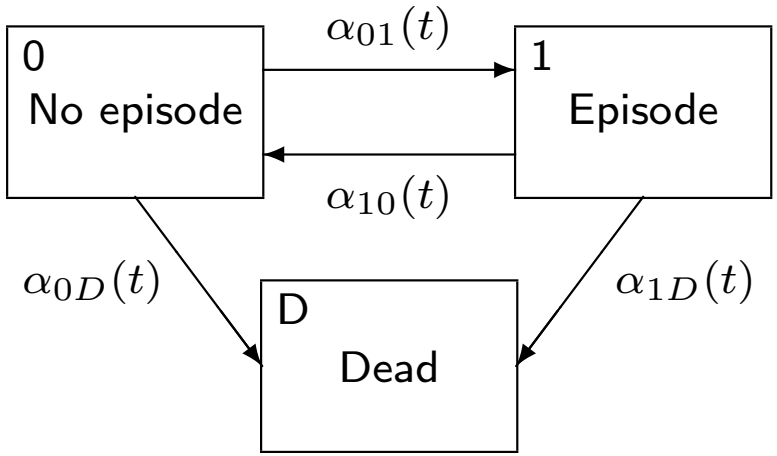
$$Q_0(t) = S(t) = P(\text{alive time } t) = \exp(-\int_0^t (\alpha_{01}(u) + \cdots + \alpha_{0k}(u))du),$$

and *cumulative incidences* $h = 1, \ldots, k$, Eq. (1.3):

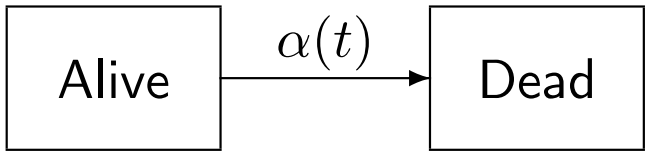
$$Q_h(t) = F_h(t) = P(\text{dead from cause } h \text{ before time } t) = \int_0^t S(u)\alpha_{0h}(u)du.$$

Recurrent events/episodes





Cox model for the two-state model



The *Cox model* for the ith individual ($i = 1, \ldots, n$) is defined as

$$\begin{aligned} \alpha_i(t) &= \alpha_0(t) \exp(\beta_1 Z_{i1} + \beta_2 Z_{i2} + \cdots + \beta_p Z_{ip}) \\ &= \alpha_0(t) \exp(\beta^T Z_i) \\ &= \alpha_0(t) \exp(\mathsf{LP}_i), \end{aligned}$$

where $\beta_1, \beta_2, \ldots, \beta_p$ are *regression parameters*, Z_{i1} is the covariate value for covariate 1 for individual i, *etc*, *baseline hazard* $\alpha_0(t)$, and *linear predictor* $\mathsf{LP}_i = \beta_1 Z_{i1} + \beta_2 Z_{i2} + \cdots + \beta_p Z_{ip}$.

Time t is the time-scale of choice, e.g., time since randomization or age. Only the baseline hazard $\alpha_0(t)$ depends on t.

If all covariates are zero ($\mathsf{LP}_i = 0$) we get $\alpha_i(t) = \alpha_0(t)$: The baseline hazard represents an individual having all covariates equal to zero.

No parametric assumptions assumed for the baseline hazard. The focus is on the regression parameters.

The Cox model

$$\alpha_i(t) \quad = \quad \alpha_0(t) \exp(\mathsf{LP})$$

on the log-scale becomes

$$\begin{aligned} \log(\alpha_i(t)) &= \log((\alpha_0(t) \exp(\mathsf{LP})) \\ &= \log(\alpha_0(t)) + \mathsf{LP}. \end{aligned}$$

This means that the Cox model assumes that efects of covariates are additive and linear on the log-rate scale.

Binary covariate

To make things easier we now study the effect of only one binary covariate, e.g., treatment (yes/no) on the rate of dying

$$Z_i = \begin{cases} 0 & \text{if individual } i \text{ is a untreated} \\ 1 & \text{if individual } i \text{ is a treated} \end{cases}$$

The Cox model is

$$\alpha_i(t) \quad = \quad \alpha_0(t) \exp(\beta Z_i).$$

With Z_i defined as above we get

$$\alpha_i(t) = \begin{cases} \alpha_{\text{untreated}}(t) \quad = \alpha_0(t) & \text{if individual } i \text{ is untreated} \\ \alpha_{\text{treated}}(t) = \alpha_0(t) \exp(\beta) & \text{if individual } i \text{ is treated} \end{cases}$$

Proportional hazards

The *hazard ratio* (HR) between treated and untreated is

$$\text{HR} = \frac{\alpha_{\text{treated}}(t)}{\alpha_{\text{untreated}}(t)} = \frac{\alpha_0(t)\exp(\beta)}{\alpha_0(t)} = \exp(\beta).$$

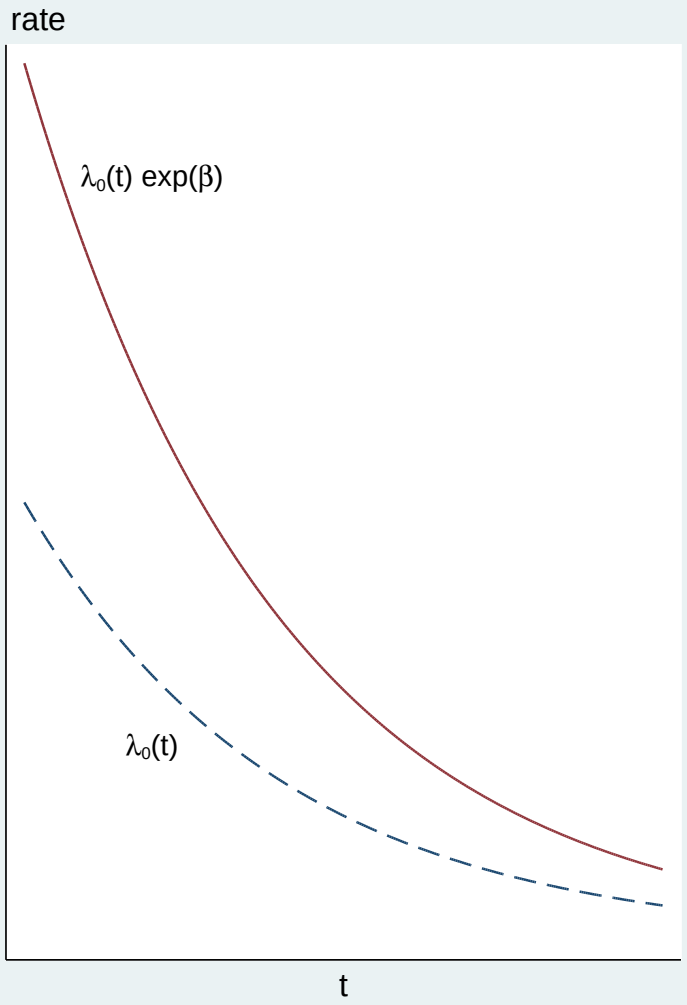
The ratio is independent of time, i.e., *proportional hazards* over time and the untreated is the *reference group*.

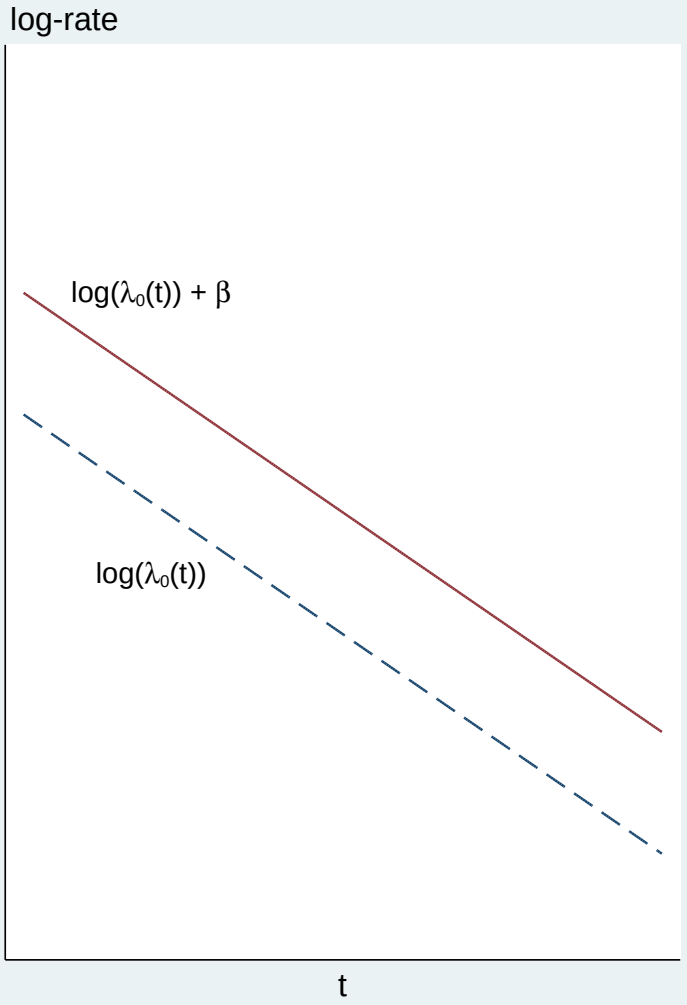
In the log-scale

$$\begin{aligned} \log(\alpha_{\text{treated}}(t)) - \log(\alpha_{\text{untreated}}(t)) &= \log(\alpha_0(t)\exp(\beta)) - \log(\alpha_0(t)) \\ &= \beta. \end{aligned}$$

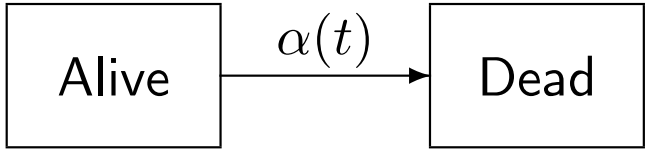
The proportionality assumption is the same as a constant difference between the log-rates at any time t.

Proportional hazards





Interpretation



$$\text{HR} = \frac{\alpha_0(t)\exp(\beta)}{\alpha_0(t)} = \exp(\beta)$$

- HR < 1 ($\beta < 0$), treated have lower mortality *rate* than untreated
- HR $= 1$ ($\beta = 0$), no difference in mortality *rate*
- HR > 1 ($\beta > 0$), treated have higher mortality *rate* than untreated

As we for the two-state model, we have the **one-to-one relationship**

$$S(t \mid Z) = \exp(-\int_0^t \alpha_0(u)\exp(\beta Z)du),$$

we also have

- HR < 1 ($\beta < 0$), treated have lower *probability* to die than ...
- HR $= 1$ ($\beta = 0$), no difference in *probability* to die
- HR > 1 ($\beta > 0$), treated have higher *probability* to die than ...

Cox's partial likelihood function

Let $t_1, \ldots, t_n$ be observed event times for n individuals and $D_i = dN_i(t_i)$ the event indicators (1=event, 0=censoring). Cox's partial likelihood function is

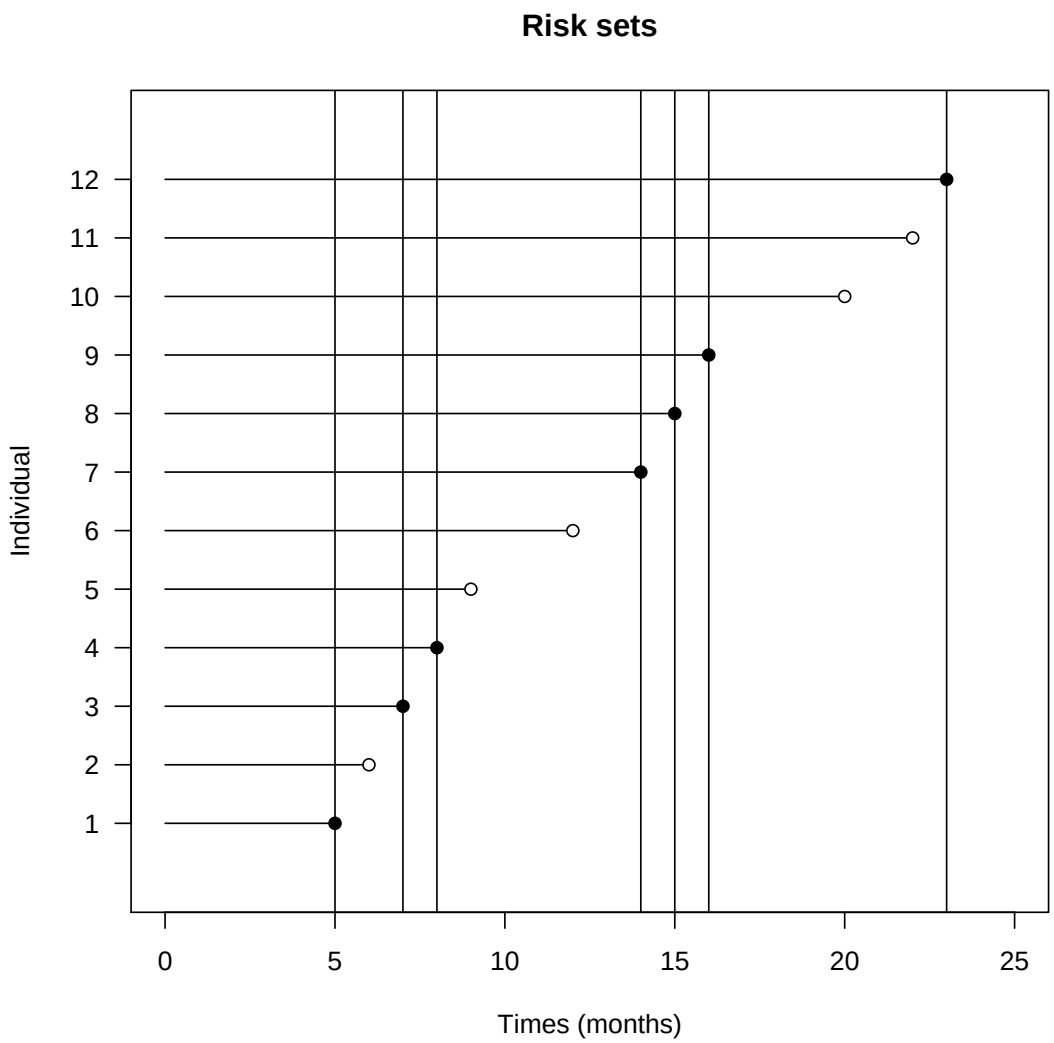
$$L(\beta) = \prod_{i=1}^{n} \Big(\frac{\exp(\beta Z_i)}{\sum_{j \in R(t_i)} \exp(\beta Z_j)} \Big)^{D_i}.$$

$R(t_i)$ is the *risk set* at time t_i, i.e., the set of individuals at risk of just before time t_i.

$$\begin{aligned} L(\beta) &= \prod_{i=1}^{d} \frac{\exp(\beta Z_i)}{\sum_{j \in R(t_i)} \exp(\beta Z_j)} \\ &= \frac{\exp(\beta Z_1)}{\sum_{j \in R(t_1)} \exp(\beta Z_j)} \cdot \frac{\exp(\beta Z_2)}{\sum_{j \in R(t_2)} \exp(\beta Z_j)} \cdots \frac{\exp(\beta Z_d)}{\sum_{j \in R(t_d)} \exp(\beta Z_j)}, \end{aligned}$$

where d is number of total events and i now only those with an event.

Risk sets


Risk sets
Individual
12
11
10
9
8
7
6
5
4
3
2
1
0
5
10
15
20
25
Times (months)

Time

The time-variable t is adjusted for by comparing individuals at the same time t (think about the risk sets). If you, e.g., have chosen age as the time-variable you have automatically adjusted for age. However, we don't get an estimate of the effect of the time-variable on the event, but may model interactions with time and covariates.

$$L(\beta) = \prod_{i=1}^{n} \left(\frac{\exp(\beta Z_i)}{\sum_{j \in R(t_i)} \exp(\beta Z_j)} \right)^{D_i}.$$

Cox score

The log-likelihood is

$$\ell(\beta) = \log(L(\beta)) = \sum_{i=1}^{n} D_i \big(\beta Z_i - \log \sum_{j \in R(t_i)} \exp(\beta Z_j)\big)$$

and the *Cox score*, Eq. (3.17), is

$$U(\beta) = \frac{d}{d\beta}\ell(\beta) = \sum_{i=1}^{n} \int_0^{\infty} \big(Z_i - \frac{\sum_j Z_j Y_j(t) \exp(\beta Z_j)}{\sum_j Y_j(t) \exp(\beta Z_j)}\big) dN_i(u),$$

with *at risk process*

$$Y_i(t) = \text{ indictor for } i \text{ being alive at time } t - .$$

Solving the Cox score equation, $U(\beta) = 0$, an estimate of β is obtained.

Inference and baseline hazard (Breslow estimator)

Let $\widehat{\beta}$ be the Cox maximum (partial) likelihood estimator from solving the Cox score equation ($U(\beta) = 0$).

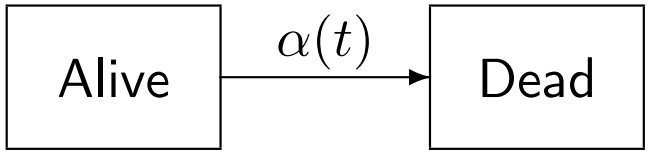
Large-sample inference for $\widehat{\beta}$ may be based on standard likelihood results and model-based standard deviations (SD) of $\widehat{\beta}$ may be obtained from the second derivative of the log-likelihood. *Wald tests, score tests, and likelihood-ratio tests* are thus for our disposition.

The score test from a Cox model with only a binary covariate corresponds to the *logrank test*.

With solution $\widehat{\beta}$, the *cumulative baseline hazard*, $A_0(t) = \int_0^t \alpha_0(u)du$, can be estimated by the Breslow estimator, Eq. (3.18):

$$\widehat{A}_0(t) = \int_0^t \frac{\sum_i dN_i(u)}{\sum_i Y_i(u)\exp(\widehat{\beta}Z_i)}.$$

Survival probability and hazard



From the one-to-one relationship between survival probability and hazards function, Eq. (1.2), we have

$$S(t \mid Z) = \exp(-\int_0^t \alpha_0(u)\exp(\mathsf{LP})du) = \exp(-A_0(t)\exp(\mathsf{LP})).$$

Using the complementary log-log transform, *cloglog*, of a distribution function $F(t)$

$$\mathsf{cloglog}(F(t)) = \log(-\log(1-F(t))) = \log(-\log(S(t))),$$

we get the regression model in the cloglog scale, Eq. (4.6),

$$\log(-\log(S(t \mid Z))) = \log(A_0(t)) + \mathsf{LP}.$$

The cloglog function is the *link function* which takes us from the marginal parameter, $S(t)$, to the linear predictor.

PBC3 data

Variable	Description
`id`	patient id
`unit`	hospital
`days`	follow-up time in days
`status`	0 = censoring, 1 = transplantation, 2 = death without transplantation
`tment`	0 = placebo, 1 = CyA
`sex`	0 = female, 1 = male
`age`	age (years)
`bili`	bilirubin (micromoles/L)
`alb`	albumin (g/L)
`stage`	disease stage: 2 = I-II, 3 = III, 4 = IV

Cox in R

```
library(survival)
summary(coxph(Surv(days, status != 0) ~ tment, data = pbc3))

n= 349, number of events= 90

          coef exp(coef) se(coef)      z Pr(>|z|)
tment -0.05874   0.94295  0.21092 -0.278    0.781

      exp(coef) exp(-coef) lower .95 upper .95
tment     0.943       1.06    0.6237     1.426

Likelihood ratio test= 0.08  on 1 df,   p=0.8
Wald test            = 0.08  on 1 df,   p=0.8
Score (logrank) test = 0.08  on 1 df,   p=0.8
```

Categorical covariate

PBC3 trial:

$$\text{stage} = \begin{cases} 2 & \text{disease stage I or II} \\ 3 & \text{disease stage III} \\ 4 & \text{disease stage IV} \end{cases}$$

Aim is to estimate the HR between the three groups. The Cox model will need two indicator functions, e.g.,

$$\text{stage2} = \begin{cases} 1 & \text{disease stage I-II} \\ 0 & \text{otherwise} \end{cases}$$

and

$$\text{stage3} = \begin{cases} 1 & \text{disease stage III} \\ 0 & \text{otherwise} \end{cases}$$

We only need two indicators, because if both are zero the patient will be in disease stage IV, which is then the reference. The Cox model becomes

$$\alpha(t) \quad = \quad \alpha_0(t) \exp(\beta_1 \cdot \text{stage2} + \beta_2 \cdot \text{stage3}),$$

where index i is removed for readiness. The Cox model assumes

$$\alpha(t) = \begin{cases} \alpha_0(t) \exp(\beta_1) & \text{if disease stage I-II} \\ \alpha_0(t) \exp(\beta_2) & \text{if disease stage III} \\ \alpha_0(t) & \text{if disease stage IV} \end{cases}$$

An overall test for the categorical variable ($\beta_1 = \beta_2 = 0$) with 2 degrees of freedom is done using standard tests.

Cox in R

```
sumamry(coxph(Surv(days, status != 0) ~ factor(stage), data = pbc3))

  n= 291, number of events= 77
   (58 observations deleted due to missingness)

                    coef exp(coef) se(coef)      z Pr(>|z|)
factor(stage)2 -2.1477    0.1167   0.3265 -6.577 4.79e-11 ***
factor(stage)3 -0.9827    0.3743   0.2770 -3.548 0.000388 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

               exp(coef) exp(-coef) lower .95 upper .95
factor(stage)2    0.1167      8.565   0.06156    0.2214
factor(stage)3    0.3743      2.672   0.21748    0.6441

Likelihood ratio test= 56.65  on 2 df,   p=5e-13
Wald test            = 46.94  on 2 df,   p=6e-11
Score (logrank) test = 61.99  on 2 df,   p=3e-14
```

Quantitative covariate

Using bilirubin measured in micromoles/L (variable `bili`) from the PBC3 trial

$$\alpha(t) = \alpha_0(t) \exp(\beta \cdot \text{bilirubin}),$$

or on the log-rate scale

$$\log(\alpha(t)) = \log(\alpha_0(t)) + \beta \cdot \text{bilirubin}.$$

For all t this is a straight line with intercept $\log(\alpha_0(t))$ and slope β. The log-rate increases (or decreases) with β for each unit increase in bilirubin.

What does $\alpha_0(t)$ mean here?

Cox in R

```
summary(coxph(Surv(days, status != 0) ~ bili, data = pbc3))

  n= 349, number of events= 90

          coef exp(coef)  se(coef)     z Pr(>|z|)
bili 0.0093362 1.0093799 0.0008913 10.47   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

     exp(coef) exp(-coef) lower .95 upper .95
bili     1.009     0.9907     1.008     1.011

Likelihood ratio test= 70.71  on 1 df,   p=<2e-16
Wald test            = 109.7  on 1 df,   p=<2e-16
Score (logrank) test = 160.6  on 1 df,   p=<2e-16
```

Multiple Cox regression in R

```
summary(coxph(Surv(days, status != 0) ~ tment+alb+log2(bili)+sex+age,
        data = pbc3))

  n= 343, number of events= 88
   (6 observations deleted due to missingness)

                coef exp(coef) se(coef)      z Pr(>|z|)
tment       -0.51005   0.60047  0.22338 -2.283  0.02242
alb         -0.07132   0.93117  0.02293 -3.110  0.00187
log2(bili)   0.73783   2.09140  0.07768  9.499  < 2e-16
sex          0.58615   1.79706  0.26740  2.192  0.02838
age          0.03073   1.03121  0.01199  2.563  0.01038

            exp(coef) exp(-coef) lower .95 upper .95
tment          0.6005     1.6654    0.3876    0.9303
alb            0.9312     1.0739    0.8902    0.9740
log2(bili)     2.0914     0.4781    1.7960    2.4353
sex            1.7971     0.5565    1.0640    3.0351
age            1.0312     0.9697    1.0073    1.0557
```

Stratified Cox model

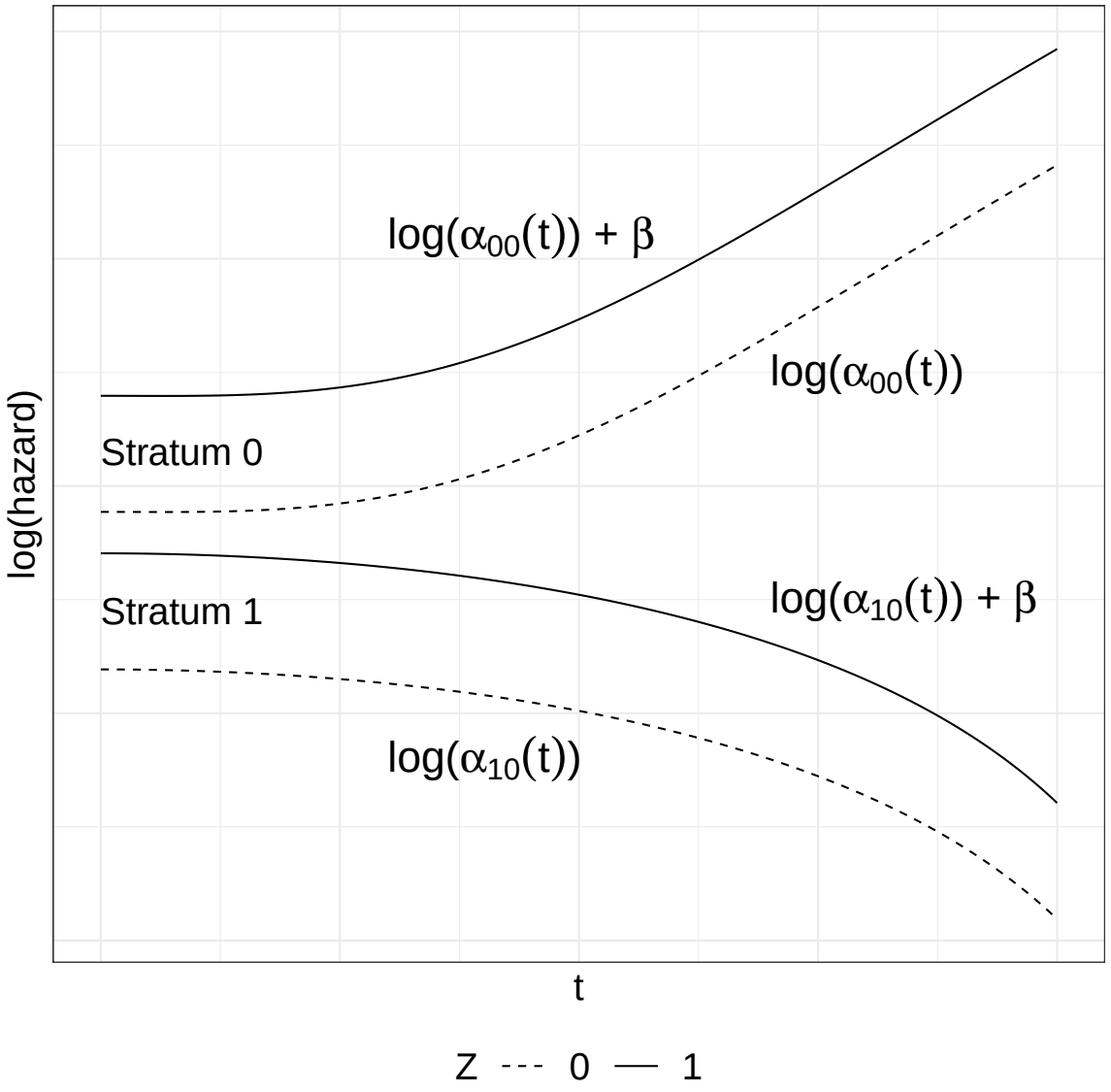
If the assumption of proportional hazards is not needed for a categorical covariate with m categories it is possible to expand the Cox model to include different baseline hazards for each category:

$$\alpha(t) = \alpha_{k0}(t) \exp(\beta Z),$$

where $\alpha_{k0}(t)$ for $k = 0, \ldots, m$, is the baseline hazard in each of the strata. These baseline hazards are allowed to depend on time, but no other assumptions are made.

The effect of the stratified covariate is not estimated, but other covariates are adjusted for the stratified covariate. The effect of the (other) covariates are assumed equal over strata, i.e., no interaction between the stratifying variable and the covariates (but this can be relaxed).

Stratified Cox model

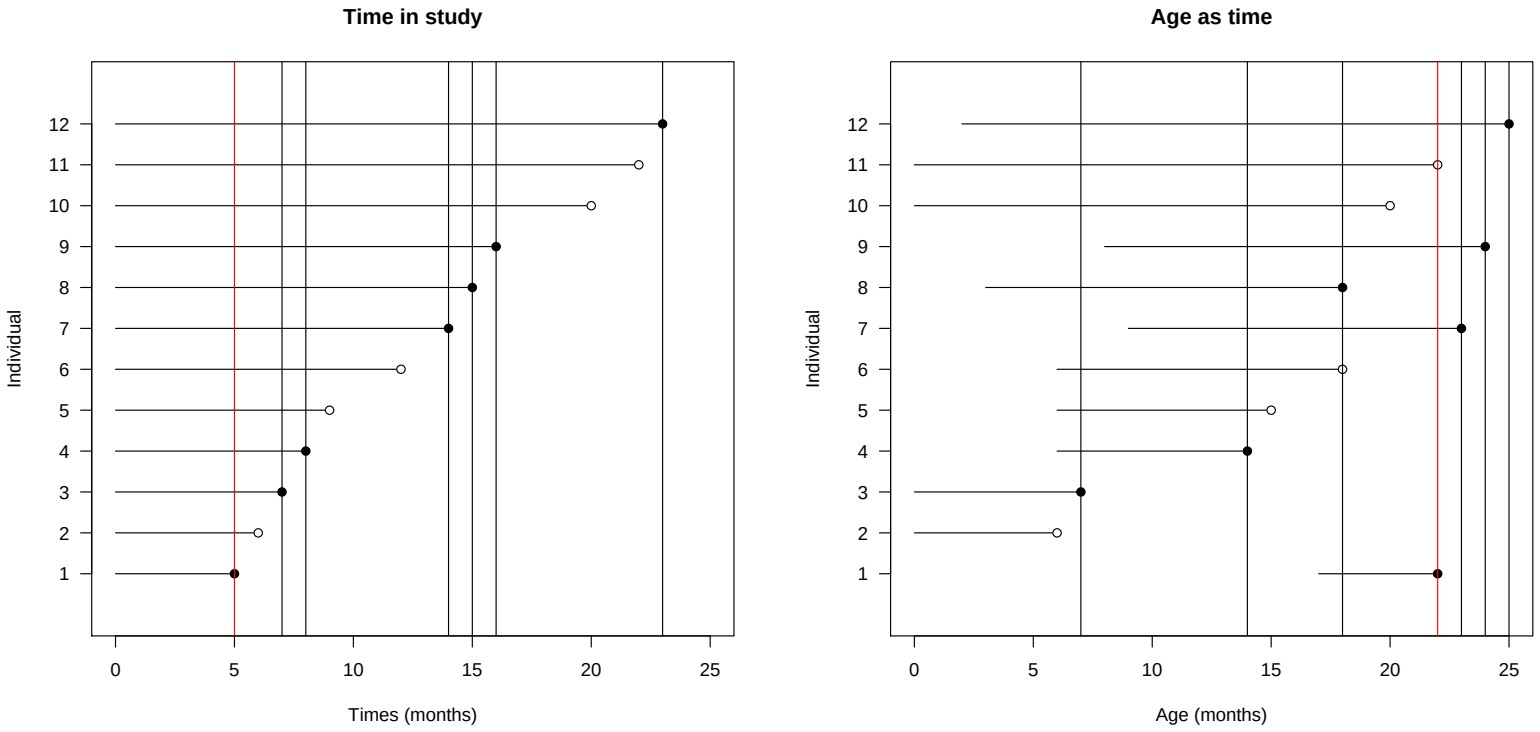


Stratified Cox in R

```
 table(pbc3$unit)
  1   2   3   4   5   6
 23 150  46  79  23  28

 coxph(Surv(days, status != 0) ~ tment + strata(unit), data = pbc3)

  n= 349, number of events= 90

         coef exp(coef) se(coef)      z Pr(>|z|)
tment -0.1117    0.8944   0.2140 -0.522    0.602

      exp(coef) exp(-coef) lower .95 upper .95
tment    0.8944      1.118     0.588      1.36

Likelihood ratio test= 0.27  on 1 df,   p=0.6
Wald test            = 0.27  on 1 df,   p=0.6
Score (logrank) test = 0.27  on 1 df,   p=0.6
```

Delayed entry aka left-truncation

Often in epidemiological studies subjects are only becoming at risk at a certain age or time. To be included in the sample, a subject must survive until the date that the sample is identified. This type of incomplete observation is denoted *left-truncation* or *delayed entry*.



If the truncation mechanism satisfies a condition of *independent truncation* similar to that of independent censoring, then handling of left-truncation is done by careful control of the risk sets $R(t_i)$ in the likelihood function:

$$L(\beta) = \prod_{i=1}^{d} \frac{\exp(\beta Z_i)}{\sum_{j \in R(t_i)} \exp(\beta Z_j)}$$

Only individuals at risk and under observation is included in the risk set $R(t_i)$ at time t_i.

Remember, that the time-variable is automatically adjusted for, which means that if the time-variable in the Cox model is changed, then is the adjustment as well.

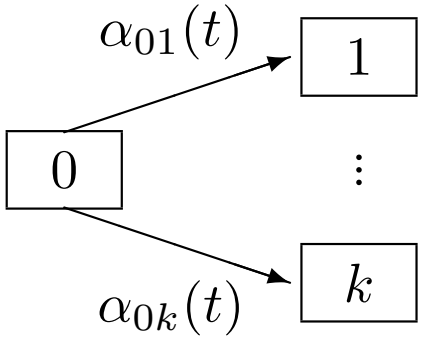
A *time of entry* (e.g., age at entry) is needed in the data.

Checking proportional hazards

This has developed into a whole 'industry' within survival analysis and many approaches have been put forward.

- Graphical methods based on the stratified model, e.g., by ploting $\log(\widehat{\alpha})_{0j}(t)$ against t (or $\log(t)$) for each stratum j and see if curves have constant vertical distance
- Modeling time-dependent effects
- Methods based on residuals
  - Martingale residuals
  - Score (Schoenfeld) residuals
  - Pseudo-residuals (more on this later)

Competing risks



Cause-specific hazards:

$$\alpha_h(t) = \alpha_{0h}(t) \approx P(\text{state } h \text{ time } t + dt \mid \text{state } 0 \text{ time } t)/dt.$$

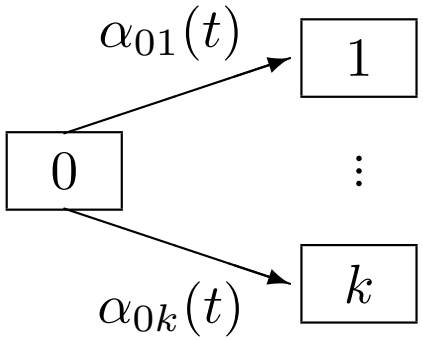
Assume that only one of the competing events can happen at time t.

Good news! We may proceed and analyse each event type using the cause-specific hazards for event of interest by **censoring for competing event types**. We don't need other assumptions about this 'censoring'.

However, if some parameters of interest are *shared* between event types, then other approaches are needed.

The reason for the good news comes from a factorization of the likelihood function over the event types (Sec. 3.1).

Cause-specific hazards



All standard hazard-based models for survival data apply when analyzing cause-specific hazards by **censoring for competing events**:

- Nelson-Aalen estimator
- Logrank test
- Cox model (Poisson and Aalen models)

Cause-specific hazards – PBC3

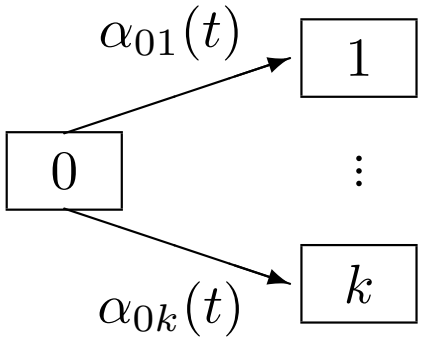
Table 2.13 *PBC3 trial in liver cirrhosis: Estimated coefficients (and SD) from Cox models for death without transplantation, transplantation, and failure of medical treatment, respectively.*

Event type	Covariate		$\widehat{\beta}$	SD
Death without transplantation	Treatment	CyA vs. placebo	-0.420	0.268
	Albumin	per 1 g/L	-0.070	0.029
	$\log_2$(Bilirubin)	per doubling	0.692	0.093
	Sex	male vs. female	-0.486	0.319
	Age	per year	0.073	0.016
Transplantation	Treatment	CyA vs. placebo	-0.673	0.413
	Albumin	per 1 g/L	-0.094	0.039
	$\log_2$(Bilirubin)	per doubling	0.832	0.147
	Sex	male vs. female	-0.204	0.563
	Age	per year	-0.048	0.021
Failure of medical treatment	Treatment	CyA vs. placebo	-0.510	0.223
	Albumin	per 1 g/L	-0.071	0.023
	$\log_2$(Bilirubin)	per doubling	0.738	0.078
	Sex	male vs. female	-0.585	0.267
	Age	per year	0.031	0.012

Cause-specific hazards in R

```
library(survival)

# Death without transplantation
coxph(Surv(days, status==2)~ tment+alb+log2(bili)+age+sex, data = pbc3)

# Transplantation
coxph(Surv(days, status==1)~ tment+alb+log2(bili)+age+sex, data = pbc3)
```

Risks and cause-specific hazards



Remember the relationship between cumulative incidences (risks) and cause-specific hazards

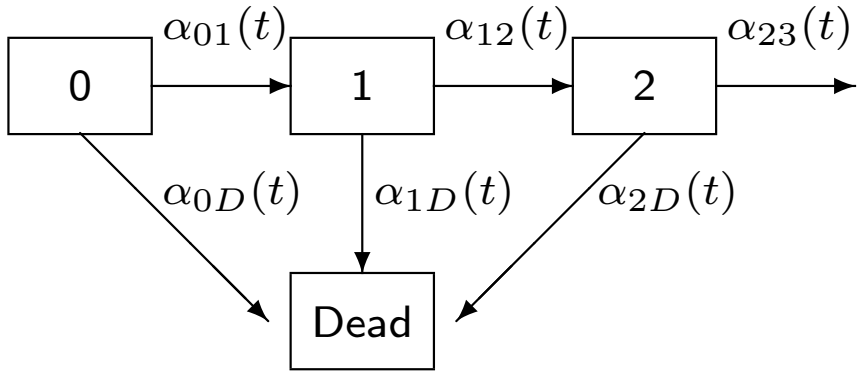
$$Q_h(t) = F_h(t) = P(\text{dead from cause } h \text{ before time } t) = \int_0^t S(u)\alpha_{0h}(u)du,$$

where

$$S(t) = \exp(-\int_0^t (\alpha_{01}(u) + \cdots + \alpha_{0k}(u))du).$$

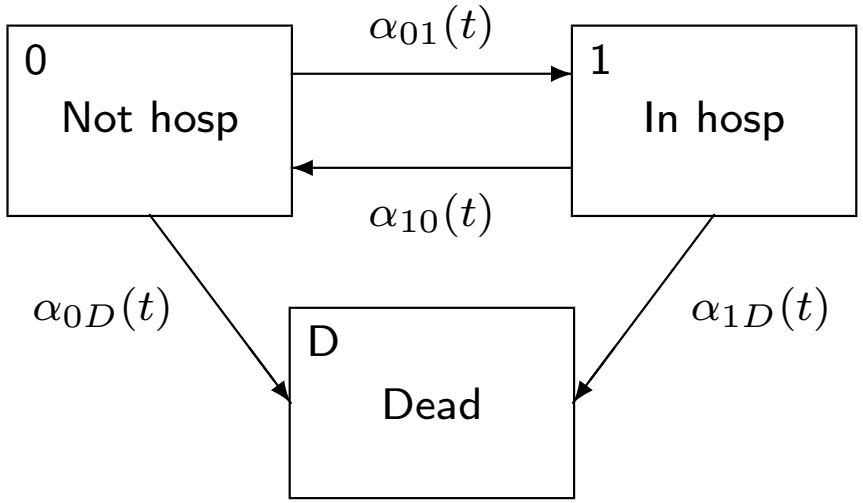
Cox models (regression coefficients and Breslow estimates) for each of the cause-specific hazards may be plugged-in to obtain estimates of $F_h(t \mid Z)$ but will not give a simple relationship between a covariate and the risk as in the two-state model.

Recurrent events



- Cox model per event transition, censor for death, using time since $t = 0$ (time since randomisation or diagnosis) and *delayed entry* is needed.
- Cox model per event transition, censor for death, using time since entry into state, known as *gap time models*.
- Andersen-Gill model: $\alpha(t \mid Z) = \alpha_0(t) \exp(\beta Z)$, a common baseline hazard.
- PWP model: $\alpha(t \mid Z) = \alpha_{0,k}(t) \exp(\beta Z)$, stratified version with separate baseline hazards for event number k (PWP: Prentice, Williams, and Peterson).

Recurrent episodes in affective disorders

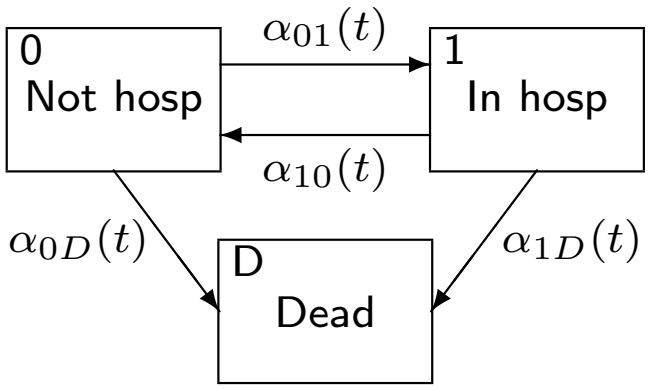


The *illness-death model with recovery*, applicable for recurrent episodes (hospitalisations) in affective disorders.

Recurrent episodes in affective disorders

Variable name	Description
`id`	patient id
`episode`	number of affective episodes
`state`	Status at time `start`:
	0 = no current affective episode, 1 = current affective episode
`start`	start time in state (months)
`stop`	last time seen in state (months)
`status`	status at time `stop`:
	0 = transition to state 0
	1 = transition to state 1
	2 = transition to death
	3 = censoring
`prev`	'start' of time to next transition to state 1, even if in state 1
`bip`	0 = unipolar, 1 = bipolar
`sex`	0 = female, 1 = male
`age`	age (years)
`year`	year of initial episode

Recurrent episodes in affective disorders



	id	episode	state	start	stop	status	bip
54	8	1	1	0.00	11.00	0	0
55	8	1	0	11.00	61.00	1	0
56	8	2	1	61.00	62.00	0	0
57	8	2	0	62.00	130.00	1	0
58	8	3	1	130.00	144.00	0	0
59	8	3	0	144.00	266.00	1	0
60	8	4	1	266.00	267.75	0	0
61	8	4	0	267.75	268.75	3	0
62	9	1	1	0.00	6.00	0	0
63	9	1	0	6.00	295.00	2	0

Recurrent episodes in affective disorders

Table 2.14 *Recurrent episodes of affective disorder: Estimated coefficients (and SD) from Cox models per episode, AG model, and PWP model for bipolar vs. unipolar disease.*

		Time since diagnosis		Gap time model	
Model	Episode	$\widehat{\beta}$	SD	$\widehat{\beta}$	SD
Cox model	1	0.356	0.250	0.399	0.249
	2	0.189	0.260	0.217	0.258
	3	-0.117	0.301	-0.111	0.287
	4	1.150	0.354	0.596	0.318
AG model		0.366	0.094	0.126	0.094
PWP model		0.242	0.112	0.028	0.100

Recurrent events in R

```
# Cox model for 1st episode, time since diagnosis
coxph(Surv(start, stop, status == 1) ~ bip,
      method = "breslow", data = subset(affective, episode == 1 & state == 0))
# AG model, time since diagnosis
coxph(Surv(start, stop, status == 1) ~ bip,
      method = "breslow", data = subset(affective, state == 0))
# PWP model, time since diagnosis
coxph(Surv(start, stop, status == 1) ~ strata(episode) + bip,
      method = "breslow", data = subset(affective, state == 0))
# Cox model for 1st episode, gap time
affective$wait <- with(affective, stop - start)
coxph(Surv(wait, status == 1) ~ bip,
      method = "breslow", data = subset(affective, episode == 1 & state == 0))
# AG model, gap time
coxph(Surv(wait, status == 1) ~ bip,
      method = "breslow", data = subset(affective, state == 0))
# PWP model, gap time
coxph(Surv(wait, status == 1) ~ strata(episode) + bip,
      method = "breslow", data = subset(affective, state == 0))
```

Exercises

1. Consider the data from the Copenhagen Holter study and the composite end-point stroke-free survival. Fit a Cox model and estimate the hazard ratio between subjects with or without ESVEA.
2. Fit a Cox model as before, now also adjusting for sex, age, and systolic blood pressure.
3. Consider the data from the Copenhagen Holter study and fit Cox models for the **cause-specific hazards** for the two outcomes (1) stroke and (2) death without stroke including ESVEA, sex, age, and systolic blood pressure. Compare to previous exercise.
4. Reproduce the results in Table 2.14 (as many as time allows).