Marginal analysis of competing risks

(Sections 4.2.2, 5.5.3)

Per Kragh Andersen

Section of Biostatistics, University of Copenhagen

DSBS Course
Survival Analysis in Clinical Trials
2025

Overview

- The EBMT data
- Recap of intensity models (cause-specific hazards)
- Direct marginal analysis: The Fine-Gray model

PK Andersen, RB Geskus, T de Witte, H Putter (2012). Competing risks in epidemiology: Possibilities and pitfalls. *Int. J. Epidemiol.* **41**, 861-870.

The EBMT data

Example

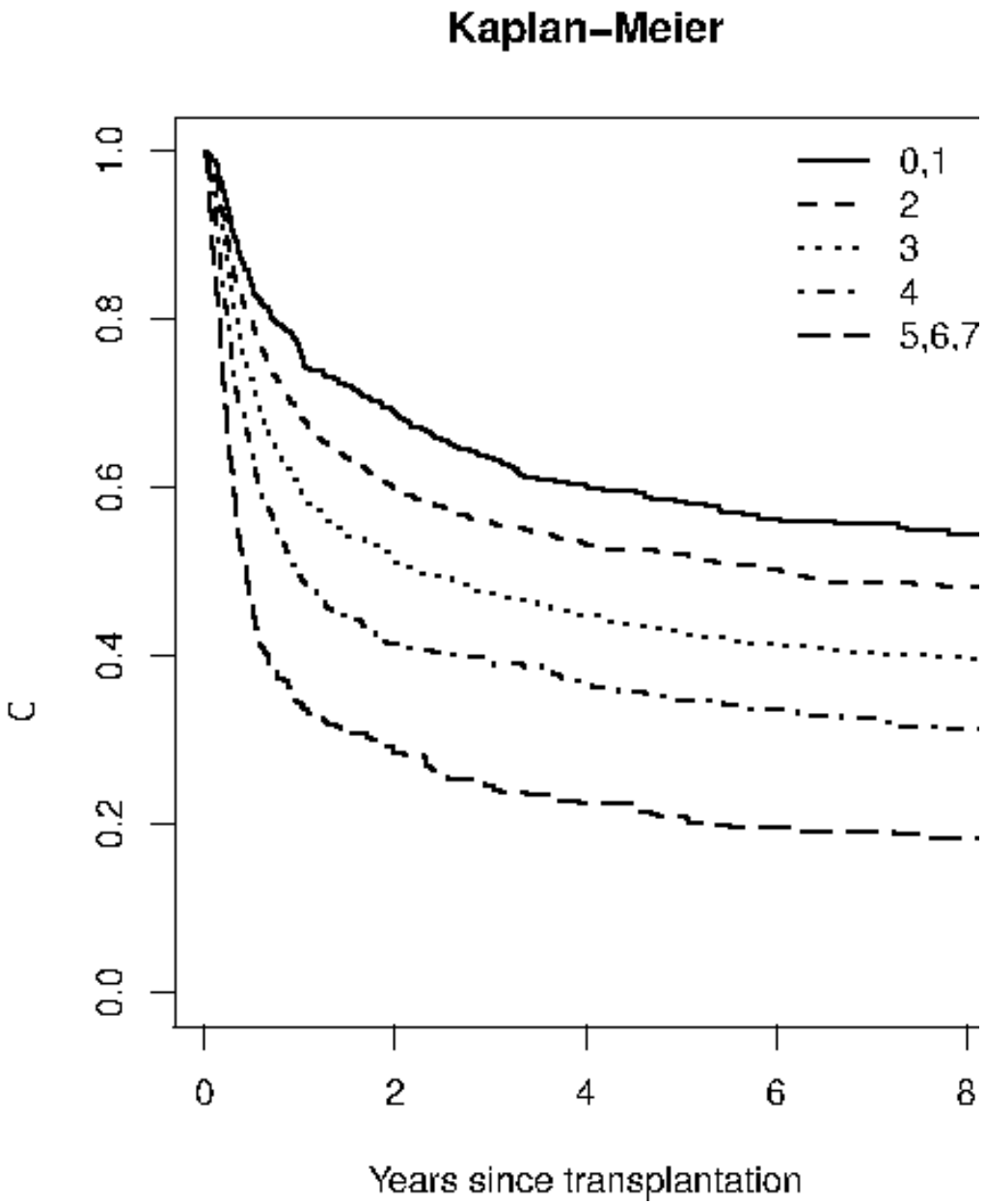
- Data from the European group for Blood and Marrow Transplantation (EBMT)
- All (3982) chronic myeloid leukemia (CML) patients with an allogeneic stem cell transplantation from an HLA-identical sibling or a matched unrelated donor during the years 1997–2000.
- Study effect of EBMT risk score with values 0–7, here grouped into five groups: 0, 1 ($n = 506$), 2 ($n = 1159$), 3 ($n = 1218$), 4 ($n = 745$), and 5, 6, 7 ($n = 354$).
- Points obtained from: donor type (2), stage (3), age (3: 20,40), female-to-male (2), time from diagnosis (2: 12 mo.)
- Failure from transplantation may either be due to relapse or to non-relapse mortality (NRM). Often these two endpoints are taken together to relapse-free survival (RFS), which is the time from transplantation to either relapse or death, whichever comes first.

Summary table

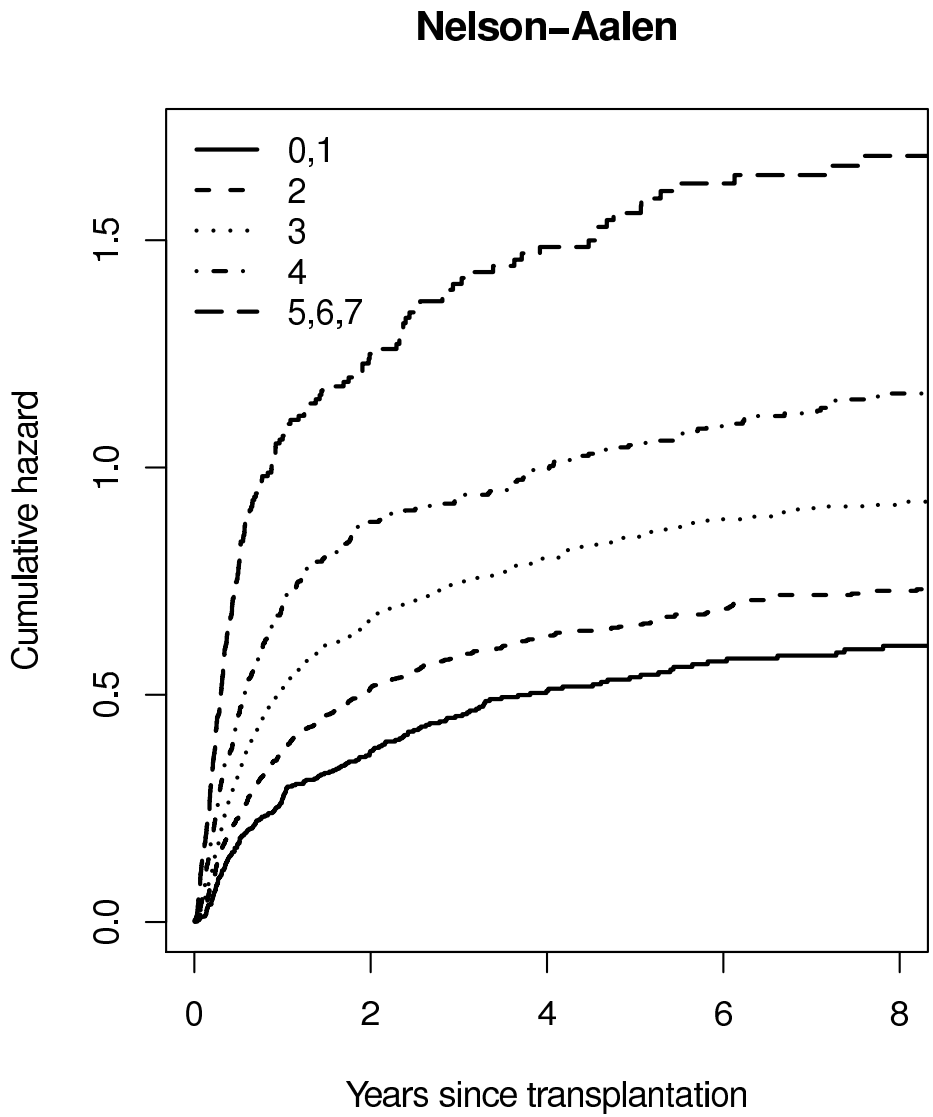
EBMT risk group	Relapse *n* (%)	NRM *n* (%)	Censored *n* (%)	Total *n* (%)
0,1	113 (22.3)	94 (18.6)	299 (59.1)	506 (100)
2	247 (21.3)	323 (27.9)	589 (50.8)	1159 (100)
3	292 (24.0)	404 (33.2)	522 (42.9)	1218 (100)
4	193 (25.9)	300 (40.3)	252 (33.8)	745 (100)
5,6,7	112 (31.6)	169 (47.7)	73 (20.6)	354 (100)

Next, we study RFS in relation to risk group.

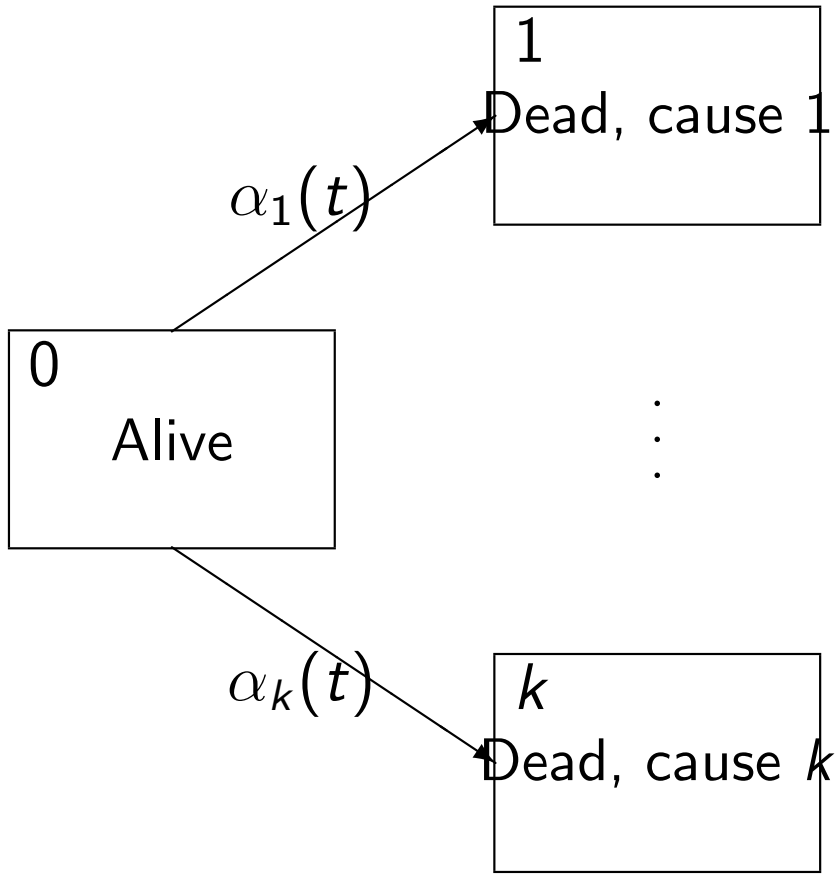
Kaplan-Meier curves (RFS)



Nelson-Aalen curves (RFS)



The competing risks multi-state model



Basic parameters

Cause-specific hazards $h = 1, 2, \ldots$ (transition intensities):

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

State occupation probabilities:

1. Overall survival function:

$$\begin{aligned} Q_0(t) = S(t) &= P(\text{alive time } t) \\ &= \exp\Big(-\int_0^t \sum_h \alpha_h(u)du\Big). \end{aligned}$$

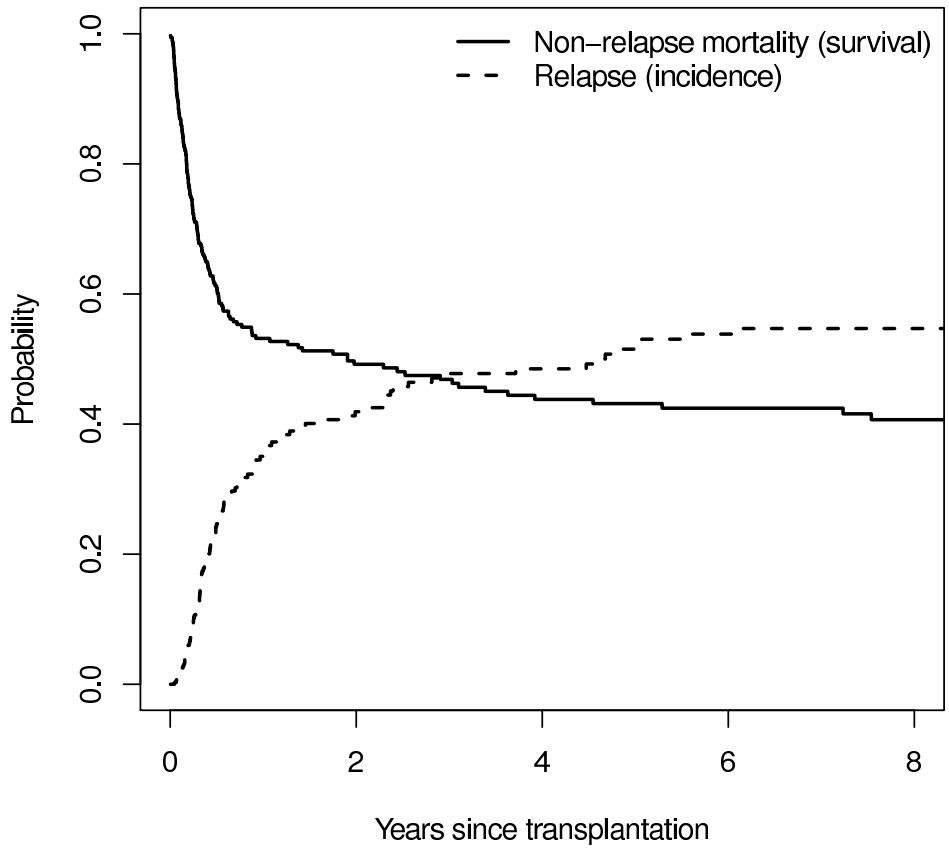
2. Cumulative incidences $h = 1, 2, \ldots$:

$$\begin{aligned} Q_h(t) = F_h(t) &= P(\text{dead from cause } h \text{ before time } t) \\ &= \int_0^t S(u)\alpha_h(u)du. \end{aligned}$$

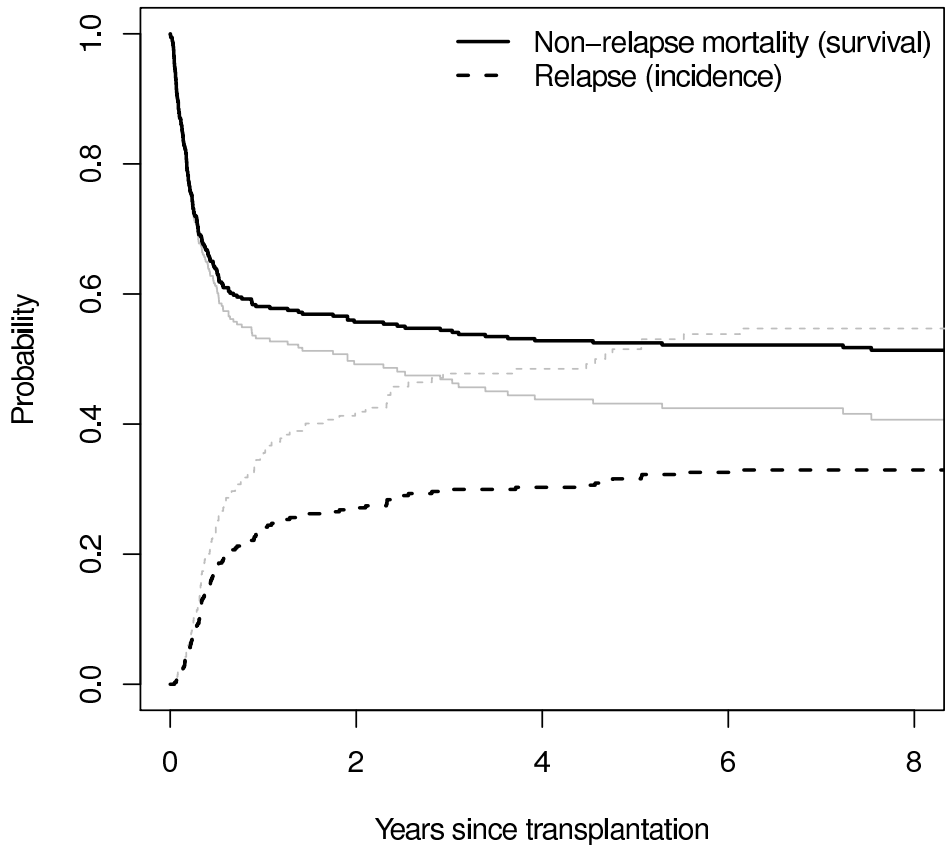
Cumulative incidence. vs. 1-KM

We look at risk group 5,6,7 and compare the 1-Kaplan-Meier estimates with the correct Aalen-Johansen estimates for relapse and for NRM.

Kaplan-Meier curves (biased)



Cumulative incidence curves (correct)



Recap of intensity models (cause-specific hazards)

Likelihood

Data: (X_i, D_i), $\quad i = 1, ..., n$ where $D_i = h$, $h = 1, ..., k$ if observed failure from cause h, $D_i = 0$ if censored.
Likelihood:

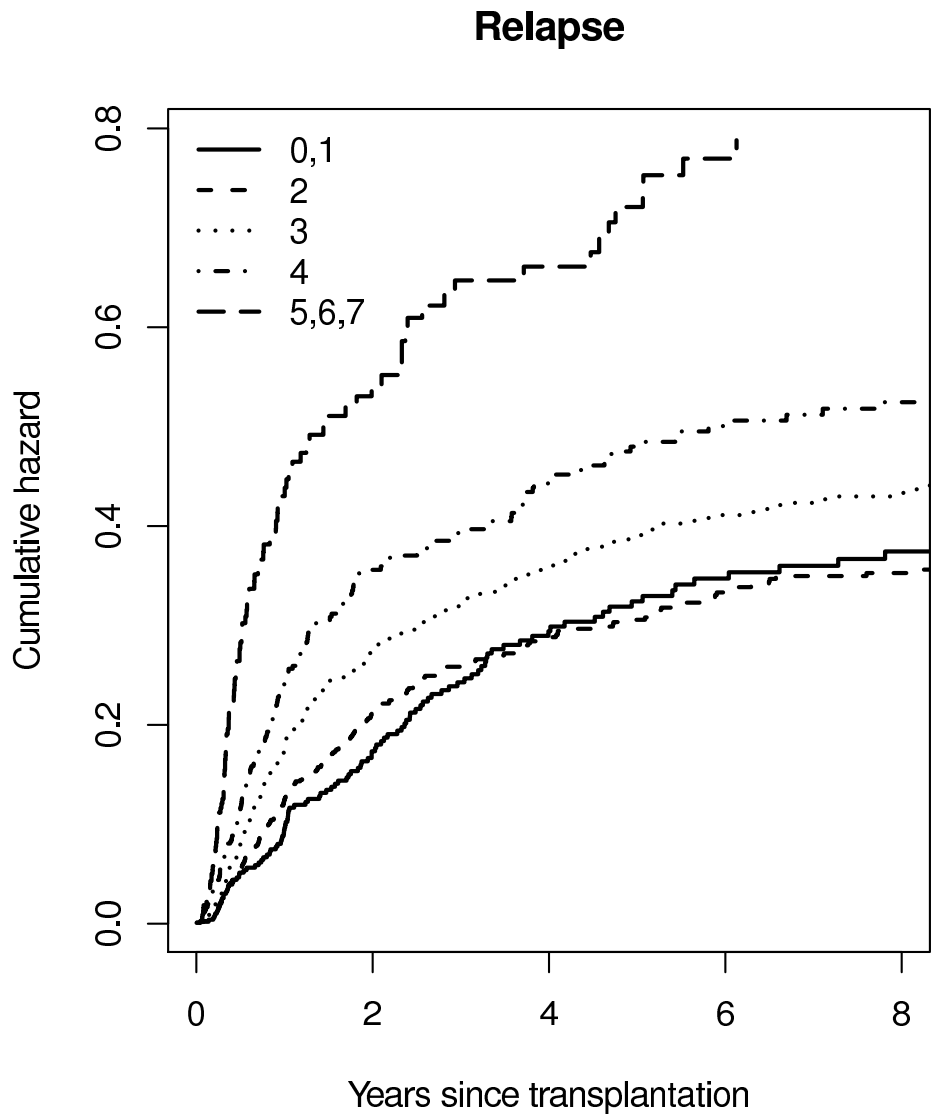
$$
\begin{aligned}
L &= \prod_{i=1}^{n} S(X_i) \prod_{h=1}^{k} (\alpha_h(X_i))^{I(D_i=h)} \\
&= \prod_{i=1}^{n} \left(\exp(-\sum_{h=1}^{k} A_h(X_i))\right) \prod_{h=1}^{k} (\alpha_h(X_i))^{I(D_i=h)} \\
&= \prod_{h=1}^{k} \left(\prod_{i=1}^{n} \exp(-A_h(X_i))(\alpha_h(X_i))^{I(D_i=h)}\right).
\end{aligned}
$$

Inference for cause-specific hazards

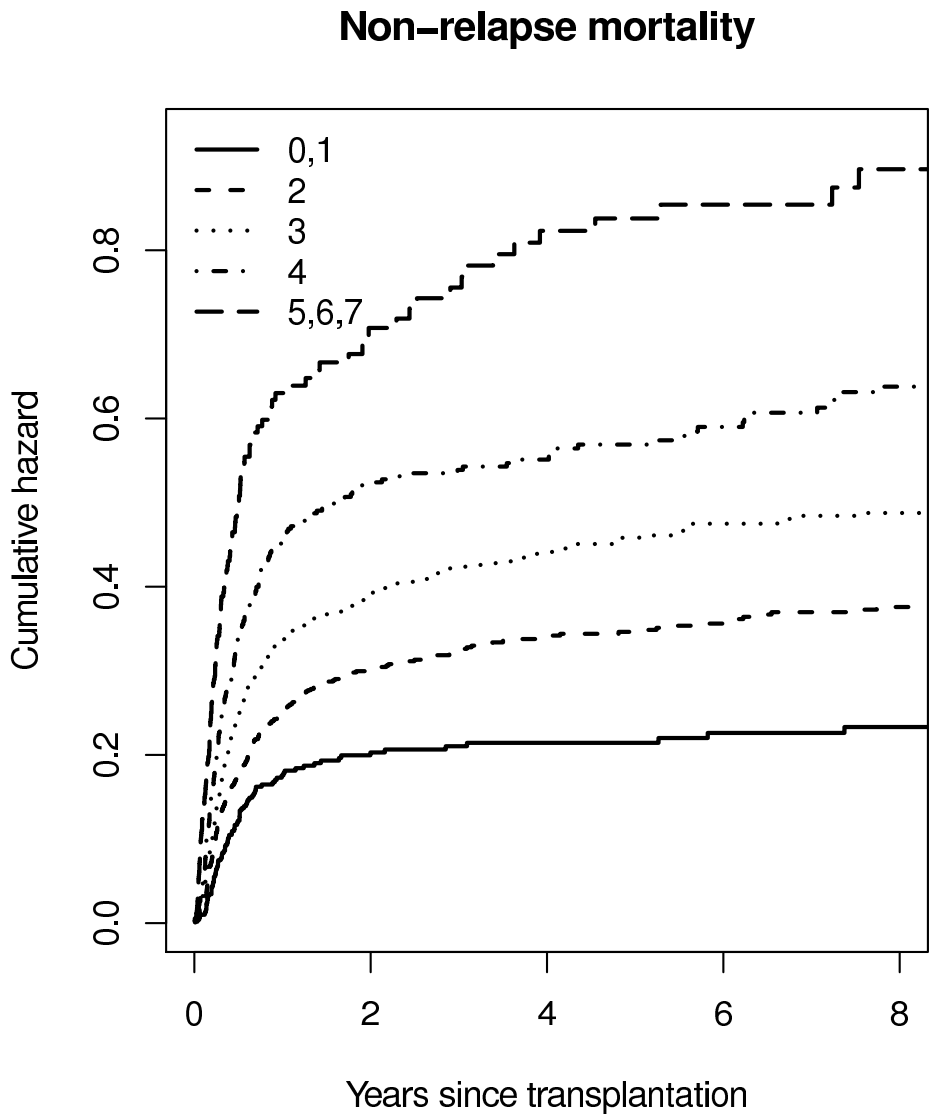
Note:

- Product over causes, h,
- The hth factor is what we would get if only that cause were studied *and all other causes were right-censorings*
- This has nothing to do with 'independence' of causes - it is solely a consequence of the definition of cause-specific hazards as hazards of exclusive events.
- It means that all standard hazard-based models for survival data apply when analyzing cause-specific hazards
  - non-parametric: estimate $A_h(t) = \int_0^t \alpha_h(u)\mathrm{d}u$, $h = 1, \ldots, k$ by Nelson-Aalen estimator, compare using, e.g., logrank tests
  - parametric models
  - Cox regression, (Poisson regression, Aalen model)

Nelson-Aalen curves: relapse



Nelson-Aalen curves: NRM



Cox models for cause-specific hazards

Model for cause h:

$$\alpha_h(t \mid Z) = \alpha_{0h}(t) \exp(\beta_h^\mathsf{T} Z),$$

that is, separate baseline hazards and separate regression coefficients for each cause.

It is technically possible to fit Cox models for cause-specific hazards with

- identical or proportional baselines for some causes
- regression coefficients that are shared between several causes

However, that is rarely relevant!

These features may be more relevant for other multi-state models than the competing risks model.

Cox models for cause-specific hazards

EMBT risk group	Relapse HR (95% ci)	NRM HR (95% ci)
0,1		
2	1.01 (0.81–1.27)	1.57 (1.25–1.97)
3	1.28 (1.03–1.59)	2.01 (1.61–2.52)
4	1.57 (1.25–1.99)	2.68 (2.12–3.37)
5,6,7	2.67 (2.06–3.47)	3.98 (3.09–5.13)

Same rate of relapse in group 2 as in group 0,1.

R code

```
## Nelson-Aalen for relapse and NRM separately

plot(survfit(Surv(days,dc==1)~ factor(riskscore),
data=ebmt), cumhaz=TRUE)

plot(survfit(Surv(days,dc==2)~ factor(riskscore),
data=ebmt), cumhaz=TRUE)

## and jointly

plot(survfit(Surv(days,factor(dc))~ factor(riskscore),
data=ebmt), cumhaz=TRUE, id=id)
```

R code (ctd.)

```
## Cox models for cause-specific hazards separately:

summary(coxph(Surv(days,dc==1)~ factor(riskscore),
method="breslow", data=ebmt))

summary(coxph(Surv(days,dc==2)~ factor(riskscore),
method="breslow", data=ebmt))

## and jointly

summary(coxph(Surv(days,factor(dc) ~ factor(riskscore),
method="breslow", data=ebmt, id=id))
```

Estimation of cumulative incidences from hazards

Estimate $F_h(t \mid Z)$ by plug-in:

$$\widehat{F}_h(t \mid Z) = \int_0^t \widehat{S}(u- \mid Z) d\widehat{A}_h(u \mid Z).$$

Here,

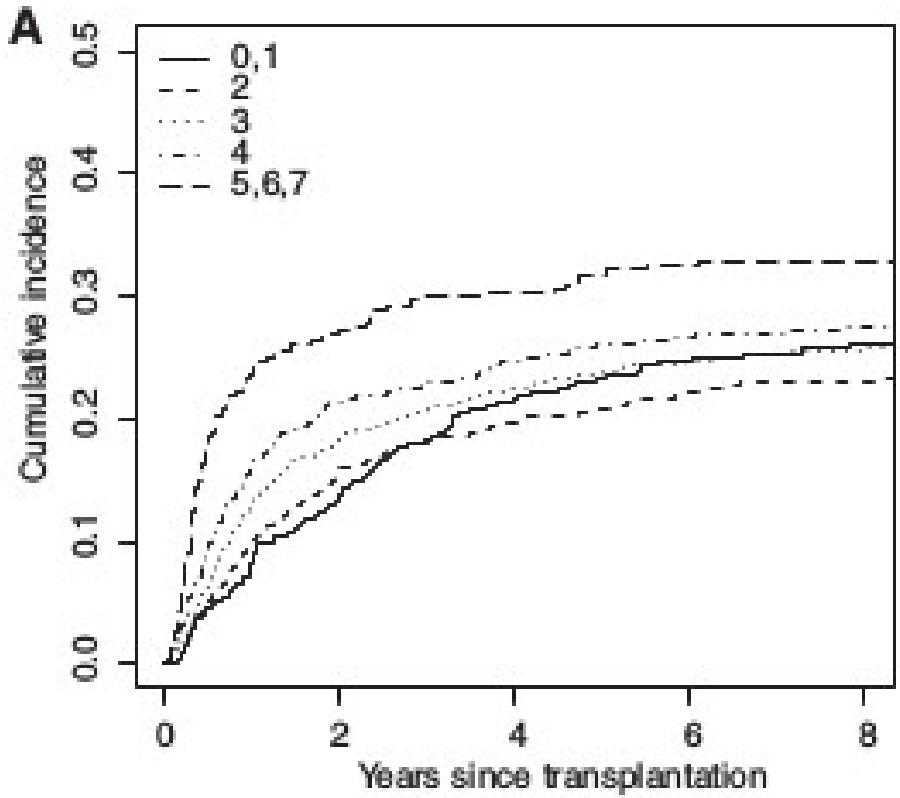
$$\widehat{A}_h(u \mid Z) = \widehat{A}_{h0}(u) \exp(\widehat{\beta}_{h1} Z_1 + ... + \widehat{\beta}_{hp} Z_p)$$

is the cumulative cause-h-hazard estimate from the Cox model and $\widehat{S}(u \mid Z)$ the Cox model based estimator for the overall survival function, e.g.,

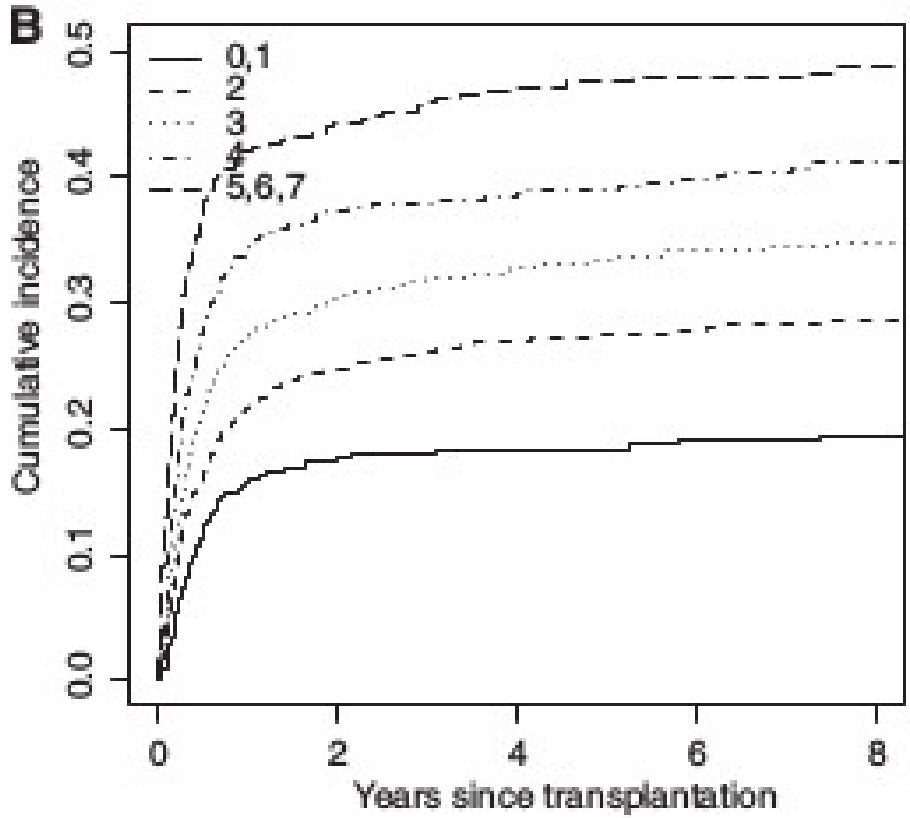
$$\widehat{S}(u \mid Z) = \exp\left(- \sum_h \widehat{A}_h(u \mid Z) \right),$$

or, preferably, the corresponding product-integral estimator.

Aalen-Johansen curves: relapse

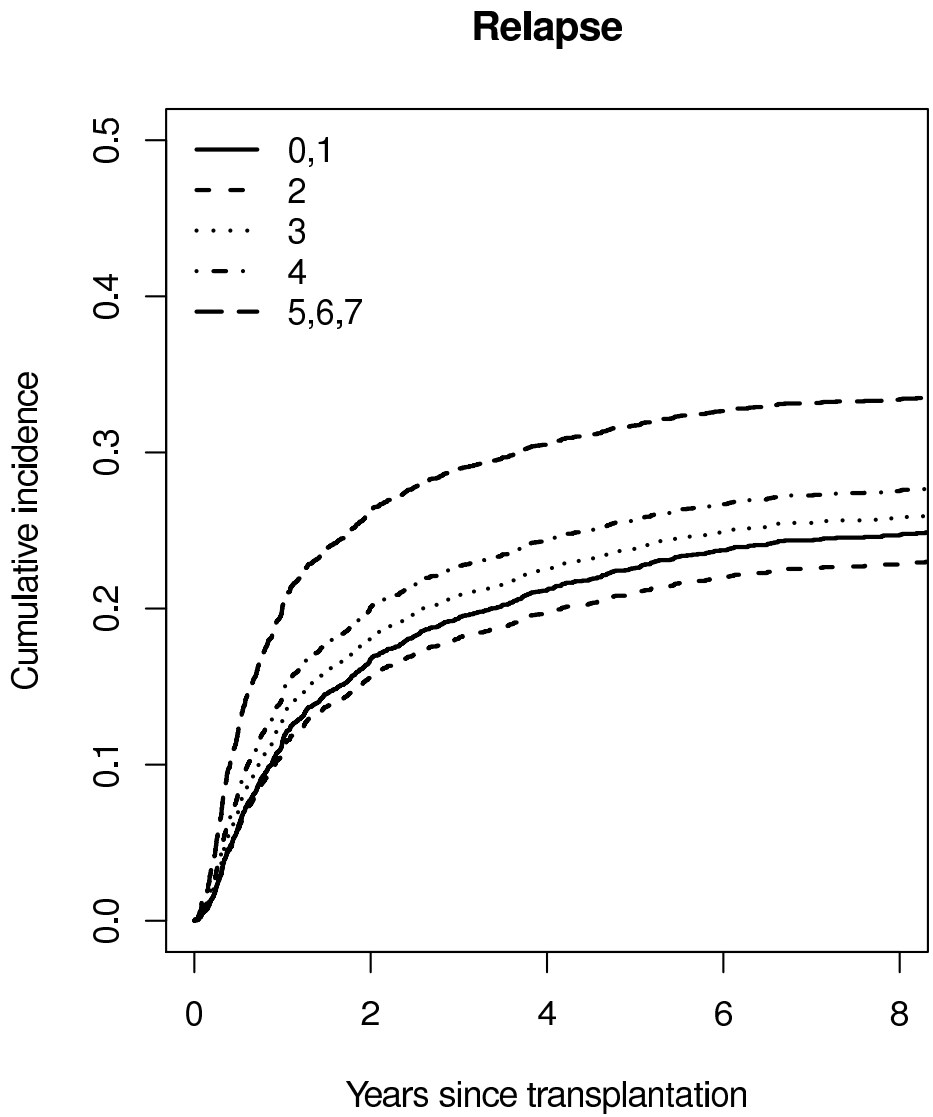


Aalen-Johansen curves: NRM

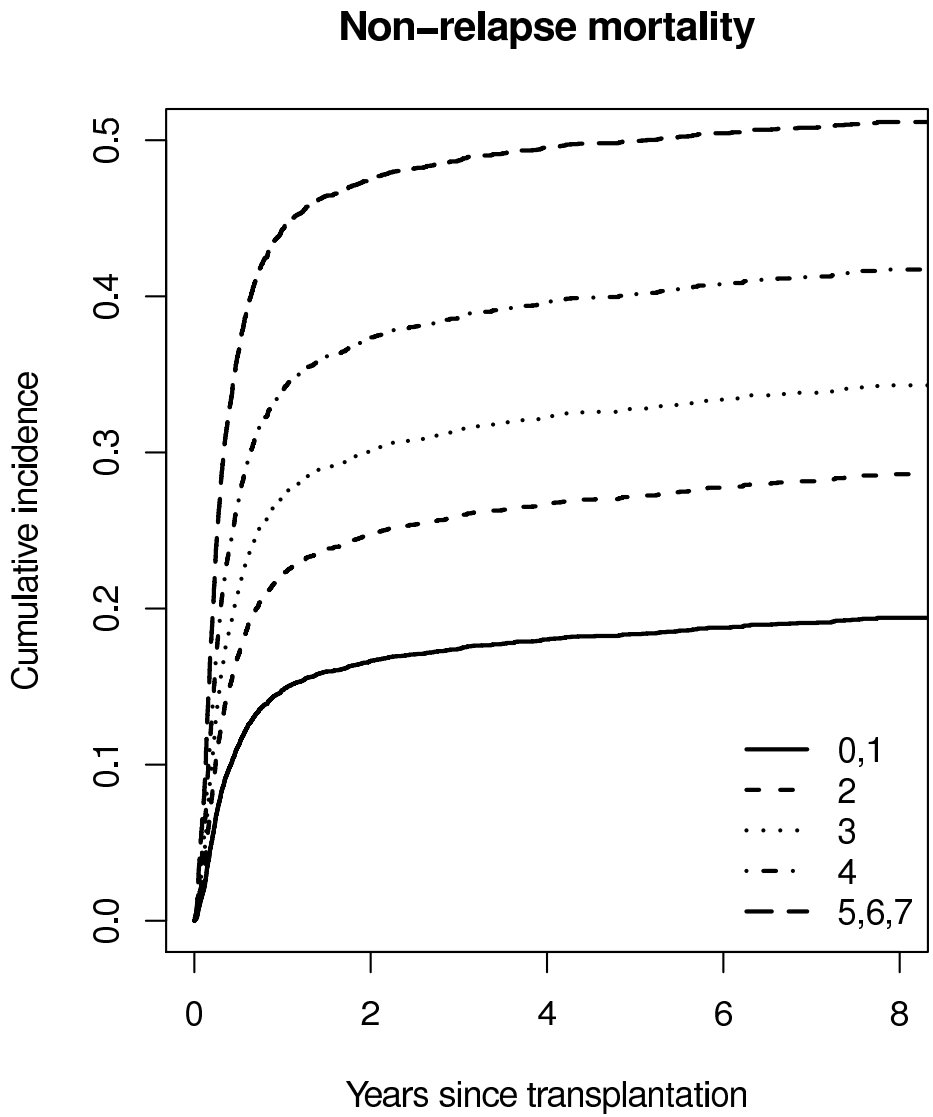


We now estimate the cumulative incidences for relapse and NRM for each of the 5 EMBT risk groups based on Cox models for the two cause-specific hazards.

Relapse


Relapse
0,1
2
3
4
5,6,7
Cumulative incidence
0.0
0.1
0.2
0.3
0.4
0.5
0
2
4
6
8
Years since transplantation

NRM


Non–relapse mortality
Cumulative incidence
0.0
0.1
0.2
0.3
0.4
0.5
0
2
4
6
8
Years since transplantation
0,1
2
3
4
5,6,7

Cumulative incidences from cause-specific Cox models

Important to notice:

- The Cox models impose a simple structure between covariates and *rates*.
- Due to the non-linear relationship between rates and risks, this simple relationship does not carry over to the cumulative incidences.
- In particular, the way in which a covariate affects a rate can be different from the way in which it affects the corresponding risk: this will depend on how it affects the rates for the competing causes.
- EBMT example: group 2 vs. 0,1, relapse

R code

```
fitebmt<-coxph(Surv(days,factor(dc) ~ factor(riskscore),
method="breslow", data=ebmt, id=id)

preddata <-
data.frame(riskscore=factor(c("0-1","2","3","4","5-6-7")))

plot(survfit(fitebmt,newdata=preddata))
```

Direct marginal analysis: The Fine-Gray model

Cumulative incidence regression models

The fact that plugging-in cause-specific hazard models does not provide parameters that in a simple way describe the relationship between covariates and cumulative incidences has led to the development of direct regression models for the cumulative incidences.
The most widely used such model is the *Fine-Gray* model. Recall from a Cox model for all-cause mortality that:

$$\log(-\log(1 - F(t \mid Z))) = \log(A_0(t)) + \beta^{\mathsf{T}} Z.$$

Fine & Gray (1999, *JASA*) studied the similar model for a cumulative incidence:

$$\log(-\log(1 - F_h(t \mid Z))) = \log(\widetilde{A}_{0h}(t)) + \widetilde{\beta}_h^{\mathsf{T}} Z.$$

The Fine-Gray model

This is a model for the hazard for the improper random variable

$$T_h^* = T \cdot I(D = h) + \infty \cdot I(D \neq h),$$

i.e., for

$$\tilde{\alpha}_h(t) = -\frac{\mathrm{d}}{\mathrm{d}t} \log(1 - F_h(t)).$$

Thus, the transformation which for all-cause mortality takes us from cumulative risk to hazard is used for a cumulative incidence in a competing risks model.

The cumulative incidence $F_h(t)$ is often denoted a *sub-distribution function*, and the resulting $\widetilde{\alpha}_h(t)$ is a *sub-distribution hazard*. Thus, the Fine-Gray model is a proportional sub-distribution hazards model.

The Fine-Gray model

A problem is that, while the hazard function has the useful 'rate' interpretation:

$$\alpha(t) \approx P(\text{die before } t + dt \mid \text{ alive } t)/dt, \quad dt > 0 \text{ small},$$

and so has the cause-specific hazard:

$$\alpha_h(t) \approx P(\text{die from cause } h \text{ before } t+dt \mid \text{ alive } t)/dt, \quad dt > 0 \text{ small},$$

the sub-distribution hazard has *not*. Thus

$$\widetilde{\alpha}_h(t) \approx P(\text{die from cause } h \text{ before } t + dt \mid$$
either alive at t *or* dead from a competing cause by $t)/dt, \quad dt >$ 0 small.

The Fine-Gray model

The model for the sub-distribution hazard is:

$$\widetilde{\alpha}_h(t \mid Z) = \widetilde{\alpha}_{0h}(t) \exp(\widetilde{\beta}_{h1} Z_1 + ... + \widetilde{\beta}_{hp} Z_p),$$

but, while a 'sub-distribution hazard' sounds like a hazard, it is not! Therefore, the resulting parameters $\exp(\widetilde{\beta}_h)$ in the Fine-Gray model have a rather indirect interpretation as 'sub-distribution hazard ratios'.

Anyway, the model is being used quite a bit and is, indeed, useful by giving parameters that directly link the cumulative incidence to covariates.

Estimation with complete data

With no censoring, Fine and Gray defined the cause h 'risk set'

$$\widetilde{R}_h(t) = \{i : (T_i \geq t) \text{ or } (T_i \leq t, D_i \neq h)\},$$

and $\widetilde{\beta}_h$ is estimated from the partial likelihood score equation

$$U_h(\widetilde{\beta}_h) = \sum_i I(D_i = h)\big(Z_i - \frac{\sum_{\ell \in \widetilde{R}_h(T_i)} Z_\ell \exp(\widetilde{\beta}_h^{\mathsf{T}} Z_\ell)}{\sum_{\ell \in \widetilde{R}_h(T_i)} \exp(\widetilde{\beta}_h^{\mathsf{T}} Z_\ell)}\big) = 0$$

corresponding to replacing times of failure from causes other than h by $+\infty$.
Using counting process notation

$$U_h(\widetilde{\beta}_h) = \sum_i \int_0^\infty \big(Z_i - \frac{\sum_\ell \widetilde{Y}_{h\ell}(t) Z_\ell \exp(\widetilde{\beta}_h^{\mathsf{T}} Z_\ell)}{\sum_\ell \widetilde{Y}_{h\ell}(t) \exp(\widetilde{\beta}_h^{\mathsf{T}} Z_\ell)}\big) dN_{hi}(t) = 0$$

with $\widetilde{Y}_{hi}(t) = 1 - N_{hi}(t-)$, the indicator for no cause h failure by time t.

Estimation with censored data

With known (e.g., 'administrative') censoring (at C_i), the cause h risk set is replaced by

$$\widetilde{R}_h(t) = \{i : (T_i \wedge C_i \geq t) \text{ or } (T_i \leq t, D_i \neq h, C_i \geq t)\},$$

i.e., $\widetilde{Y}_{hi}(t)$ is replaced by $\widetilde{Y}_{hi}(t)I(C_i \geq t)$, the indicator for no cause h failure *and* no censoring by time t.
With general censoring, an Inverse Probability of Censoring Weighted (IPCW) 'score' equation is used and to do this, a model for censoring is needed.
In the simplest case, one uses the 'Kaplan-Meier for un-censoring', that is, estimating $G(t) = P(C > t)$ (in this analysis 'failures are censorings').
If censoring depends on covariates, then a model for $G(t \mid Z) = P(C > t \mid Z)$ is needed for the weights, e.g., a Cox model.

Estimation with censored data

For this to work, we define the weights

$$w_i(t) = I(C_i \geq T_i \wedge t) \frac{\widehat{G}(t)}{\widehat{G}(\widetilde{T}_i \wedge t)},$$

and Fine and Gray showed that if (in the simplest case) C_i is independent of (T_i, D_i) and Z_i, then the 'score' equation

$$\widetilde{U}_h(\widetilde{\beta}_h) = \sum_i \int_0^\infty (Z_i - \frac{\sum_\ell w_\ell(t) \widetilde{Y}_{h\ell}(t) Z_\ell \exp(\widetilde{\beta}_h^\mathsf{T} Z_\ell)}{\sum_\ell w_\ell(t) \widetilde{Y}_{h\ell}(t) \exp(\widetilde{\beta}_h^\mathsf{T} Z_\ell)}) w_i(t) dN_{hi}(t) = 0$$

is an *unbiased estimating equation* yielding consistent estimates of $\widetilde{\beta}_h$.

Estimation with censored data

The resulting weights are as follows:

$t, \widetilde{T}_i$	Status	$I(C_i \geq T_i \wedge t)$	$\widetilde{Y}_{hi}(t)$	$w_i(t)$
$t \leq \widetilde{T}_i$	$D_i = 0$	1	1	1
	$D_i = h$	1	1	1
	$D_i \neq 0, h$	1	1	1
$t > \widetilde{T}_i$	$D_i = 0$	0	1	0
	$D_i = h$	1	0	$\widehat{G}(t)/\widehat{G}(\widetilde{T}_i)$
	$D_i \neq 0, h$	1	1	$\widehat{G}(t)/\widehat{G}(\widetilde{T}_i)$

After an observed time of failure (from a cause $\neq h$), a subject gets a smaller and smaller weight as time passes (and it is, therefore, less and less likely that the subject would still be uncensored).

Estimating equations (in general)

Let $\widehat{\theta}$ be the solution to the unbiased estimating equation $U(\theta) = 0$.
Taylor expansion of $U(\cdot)$ around the true value, θ_0 yields:

$$U(\theta) = U(\theta_0) + U'(\theta^*)(\theta - \theta_0)$$

with θ^* on the line segment between θ and θ_0. Inserting $\widehat{\theta}$ and re-arranging we get:

$$n^{1/2}(\widehat{\theta} - \theta_0) \approx -U'(\theta_0)^{-1}(n^{-1/2}U(\theta_0)).$$

A CLT for $n^{-1/2}U(\theta_0)$ (sum of independent terms with mean zero) gives a CLT for $n^{1/2}(\widehat{\theta} - \theta_0)$ and the robust ('sandwich') variance estimate is

$$(U'(\widehat{\theta})^{-1})^{\mathsf{T}}(\sum_i U_i(\widehat{\theta})U_i(\widehat{\theta})^{\mathsf{T}})U'(\widehat{\theta})^{-1}.$$

Estimating the cumulative baseline sub-distribution hazard

There is a 'Breslow-type' estimator for $\widetilde{A}_{0h}(t)$.

With the weights $w_i(t)$, the estimator is

$$\widehat{\widetilde{A}}_{0h}(t) = \sum_i \int_0^t \frac{w_i(u) dN_{hi}(u)}{\sum_\ell w_\ell(u) \widetilde{Y}_{h\ell}(u) \exp(\widetilde{\beta}_h^\mathsf{T} Z_\ell)}.$$

Fine and Gray (1999) provided asymptotics for the estimator and discussed covariate-specific estimated cumulative incidences based on the model:

$$\widehat{F}_h(t \mid Z) = 1 - \exp(-\widehat{\widetilde{A}}_{0h}(t) \exp(\widetilde{\beta}_h^\mathsf{T} Z)).$$

The Fine-Gray model

The Fine-Gray model provides parameters describing the relationship between the covariates and the cause h risk. For example, for a binary covariate Z_1 with an estimated regression coefficient $\widetilde{\beta}_{h1} > 0$ it follows that for all values, Z_2^0, for the other covariates in the model we have that

$$\widehat{F}_h(t \mid Z_1 = 1, Z_2^0) > \widehat{F}_h(t \mid Z_1 = 0, Z_2^0).$$

The positive regression coefficient has the *qualitative* meaning that individuals with $Z_1 = 1$ have a uniformly increased cause h cumulative incidence compared to those with $Z_1 = 0$.
However, the resulting estimates $\exp(\widetilde{\beta}_h)$ are sub-distribution hazard ratios, so the *quantitative* meaning of the regression coefficient is not simple.
The model is related to the Gray (1988, *Ann. Statist.*) test for comparison of cumulative incidences and, indeed, the Fine-Gray model provides useful significance tests.

Fine-Gray models for EBMT data

EMBT risk	Relapse			NRM		
group	$\widetilde{\beta}$	SD	exp($\widetilde{\beta}$) (95% ci)	$\widetilde{\beta}$	SD	exp($\widetilde{\beta}$) (95% ci)
0,1						
2	-0.068	0.111	0.93 (0.75–1.16)	0.443	0.116	1.56 (1.24–1.96)
3	0.072	0.108	1.07 (0.87–1.33)	0.661	0.114	1.94 (1.55–2.42)
4	0.161	0.117	1.17 (0.93–1.48)	0.906	0.118	2.48 (1.96–3.12)
5,6,7	0.439	0.135	1.55 (1.19–2.02)	1.185	0.131	3.27 (2.53–4.22)

Somewhat lower risk of relapse in group 2 than in group 0,1.

Note that Cox models for cause-specific hazards and Fine-Gray models for cumulative incidences are *mathematically incompatible*. This, however, does not prevent people from quoting results from both (and since both models may fit data reasonably well, this may not be a major problem).

Other link functions than $\log(-\log)$ may be analysed, e.g., using *pseudo-values*.

R-code for Aalen-Johansen and Fine-Gray

```
plot(survfit(Surv(days,factor(dc))~
factor(riskscore), data=ebmt))

## The finegray function creates a new data set for a
## cause and fitting a certain Cox model to it
## gives the Fine-Gray model.

reldat<-finegray(Surv(days,factor(dc))~ .,
data=ebmt, etype=1)
coxrel<-coxph(Surv(fgstart,fgstop,fgstatus)~
factor(riskscore), data=reldat, weight=fgwt)
nrmdat<-finegray(Surv(days,factor(dc))~ .,
data=ebmt, etype=2)
coxnrm<-coxph(Surv(fgstart,fgstop,fgstatus)~
factor(riskscore), data=nrmdat, weight=fgwt)
```

Plug-in estimation based on Fine-Gray models

Even though regression parameters from Fine-Gray models have undesirable interpretations, plug-in estimation of cumulative incidences based on the model may be useful.

Also, summarizing such curves using the g-formula is useful. This is implemented in the `mets` package in `R`.

Note, however, that predicted curves for different causes may not sum to 'one minus a proper survival function'.

```
preddata<-data.frame(riskscore=
factor(c("0-1","2","3","4","5-6-7")))

plot(survfit(coxrel,newdata=preddata))
plot(survfit(coxnrm,newdata=preddata))
```

Summary (competing risks)

- In studies of all-cause mortality, risks (probabilities, cumulative incidences) can be computed from rates (hazards) and vice versa - in other words the two functions contain *equivalent* information
- In studies of events which will not eventually happen for every one in the population, this is no longer the case and death (and maybe other events) are *competing risks* which need to be addressed
- In such cases, the risk of a given cause depends on the rates for *all* competing causes
- Therefore, using '1-Kaplan-Meier for a single cause' as a risk estimator is (upward) biased
- The magnitude of the bias depends on the frequency of the competing events

Summary (ctd.)

- A rather simple, unbiased estimator for the risk exists - the 'Aalen-Johansen' estimator
- Effects of covariates on rates (cause-specific hazards) may be (qualitatively) different from their effects on the risks (cumulative incidences)
- Rates may be analysed using standard hazard based methods from survival analysis (Nelson-Aalen, Cox, Poisson, logrank, ...)
- Risks may be analysed by 'plugging-in' results from such hazard models or directly using, e.g., the Fine-Gray model
- Interpretation of coefficients from a Fine-Gray model is not appealing, but prediction from the model is useful
- Other link functions may be used based on pseudo-values