

# Competing risks.

7 October 2008, RSS/BIR Biometric Society.

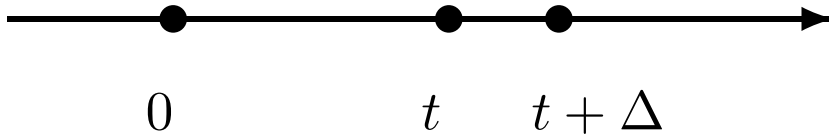
Per Kragh Andersen, Biostatistics, Univ. Copenhagen, Denmark.

Summary:

- Introduction: survival analysis and competing risks as multi-state models
- Example
- Mathematical formulations
- Inference for cause-specific hazards
- Cumulative incidences
- Regression analysis
- Further topics, conclusions, references

## Survival data

- $X$ : time from “zero” to event (death)
- right-censoring at  $C$ : observe  $\tilde{X} = \min(X, C)$ ,  $D = I(\tilde{X} = X)$
- (delayed entry)
- basic quantity:
  - hazard function (= death intensity = mortality rate)
  - $= \alpha(t) \approx \text{Prob}(\text{die before } t + \Delta \mid \text{alive } t) / \Delta$



## Survival data

Survival function:

$$S(t) = \text{Prob}(\text{alive time } t) = \exp\left(-\int_0^t \alpha(u)du\right) = \exp(-A(t)).$$

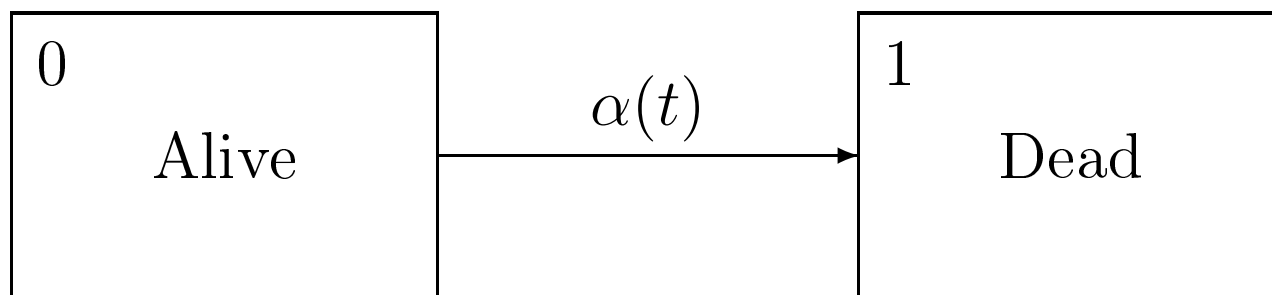
Likelihood based on independent  $(\tilde{X}_i, D_i, i = 1, \dots, n)$  under “independent” censoring:  $\prod_i S(\tilde{X}_i)\alpha(\tilde{X}_i)^{D_i}$ .

Models for  $\alpha(t)$ :

- non-parametric: estimate  $A(t)$  by Nelson-Aalen estimator
- parametric models: Weibull, piecewise exponential, ...
- regression: Cox, Poisson (piecewise exponential), ...
- ...

Simple one-to-one correspondance between the “local” parameter, the hazard,  $\alpha(t)$ , and the “global” parameter, the survival probability,  $S(t)$ .

## Survival data as a two-state model.



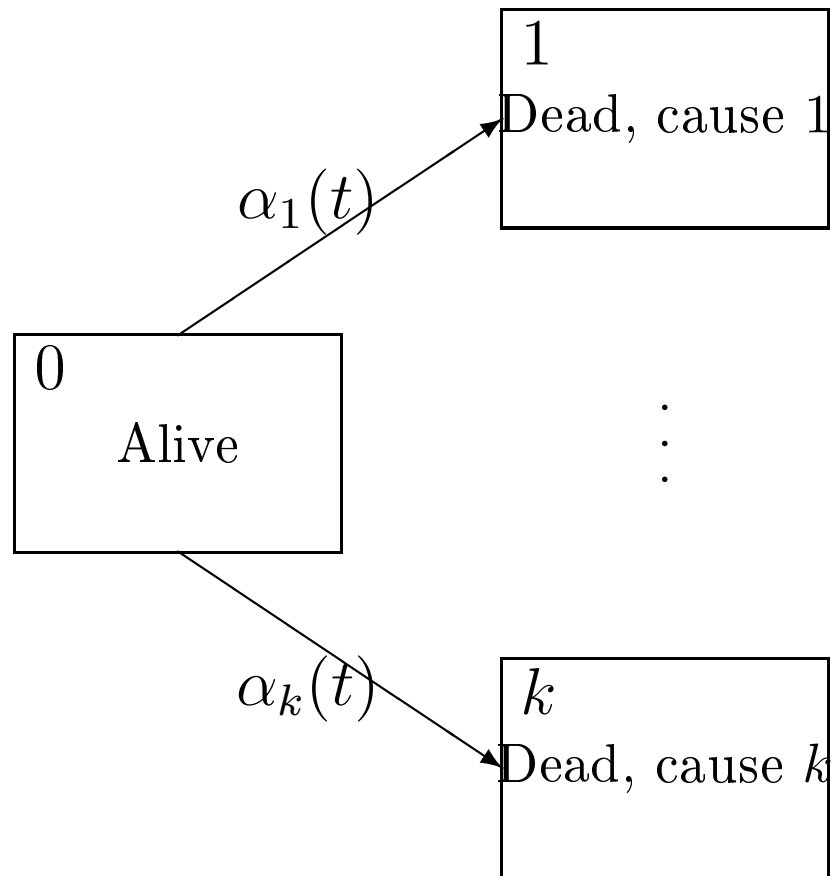
Transition intensity

$$\alpha(t) \approx \text{Prob}(\text{state 1 time } t + \Delta \mid \text{state 0 time } t) / \Delta$$

State occupation probabilities

$$S(t) = \text{Prob}(\text{state 0 time } t), \quad F(t) = 1 - S(t).$$

## Competing risks as a multi-state model.



## Parameters.

Transition intensities: Cause-specific hazards,  $j = 1, \dots, k$ :

$$\alpha_j(t) \approx \text{Prob}(\text{state } j \text{ time } t + \Delta \mid \text{state } 0 \text{ time } t) / \Delta,$$

State occupation probabilities:

$$S(t) = \text{Prob}(\text{state } 0 \text{ time } t),$$

$$F_j(t) = \text{Prob}(\text{state } j \text{ time } t), j = 1, \dots, k.$$

$$S(t) + \sum_{j=1}^k F_j(t) = 1.$$

## **Bone marrow transplantation.**

1715 leukemia patients with BMT: (Szydlo et al., 1997)

- 537 ALL, 340 AML, 838 CML
- 1026 early stage, 410 intermediate stage, 279 advanced stage
- 1224 HLA-identical sibling, 383 HLA-matched unrelated donor, 108 HLA-mismatched unrelated donor
- 311 patients relapsed, 557 died in remission.

Purpose:

- Study risk factors for relapse and death in remission

## Mathematical formulation (1).

$X$ , survival time,  $D$  cause of failure, joint distribution:

$$P(X \leq t, D = j).$$

Cause-specific hazards:

$$\alpha_j(t) = \lim_{\Delta \rightarrow 0} \frac{P(X \leq t + \Delta, D = j \mid X \geq t)}{\Delta}.$$

Marginal distribution of  $X$ : survival function

$$S(t) = P(X > t) = \exp\left(-\sum_j \int_0^t \alpha_j(u) du\right).$$

Total hazard is  $\alpha(t) = \sum_{j=1}^k \alpha_j(t)$ .

## Mathematical formulation (2).

Latent failure times

$$X_1^L, \dots, X_k^L,$$

observe  $\min\{X_1^L, \dots, X_k^L\}$  and corresponding  $D$ .

Joint survival distribution:

$$Q(t_1, \dots, t_k) = P(X_1^L > t_1, \dots, X_k^L > t_k).$$

Relations between the two formulations:

$$S(t) = Q(t, \dots, t),$$

$$\alpha_j(t) = - \left. \frac{\partial \log Q(t_1, \dots, t_k)}{\partial t_j} \right|_{t_1 = \dots = t_k = t}.$$

## Parameter identification.

Which parameters may be identified from the competing risks data  $(X_i, D_i), i = 1, \dots, n$ ?

(or from similar right-censored data  $(\tilde{X}_i, D_i), i = 1, \dots, n$  where either  $X_i = \tilde{X}_i$  and  $D_i \in \{1, \dots, k\}$  or  $X_i > \tilde{X}_i$  and  $D_i = 0$ .)

Likelihood:  $\prod_i S(\tilde{X}_i) \prod_j (\alpha_j(\tilde{X}_i))^{I(D_i=j)}$ .

From this we may identify the cause specific hazards  $\alpha_j(t)$  but not the whole joint distribution  $Q(\cdot)$  of the “latent failure times”  $X_1^L, \dots, X_k^L$ .

For instance NOT the “marginal distribution” of  $X_j^L$ :

$$P(X_j > t_j) = Q(0, \dots, 0, t_j, 0, \dots, 0) = S_j(t_j)$$

with (“net”) hazard function  $h_j(t) = -\partial \log S_j(t) / \partial t$ .

## “Independent” competing risks.

Definition:  $X_1^L, \dots, X_k^L$  are independent, i.e.

$$Q(t_1, \dots, t_k) = \prod_j S_j(t_j)$$

or (weaker): marginal (or “net”) and cause-specific (“crude”) hazards are identical:  $\alpha_j(t) = h_j(t)$ .

Since  $S_j(t)$  and  $h_j(t)$  cannot be identified from the data (without further, unidentifiable conditions) these assumptions are unverifiable.

Likewise: the question of “what would happen if certain causes were removed” (“partial crude hazards”) is quite hypothetical in most biological settings.

(Sensitivity analysis?)

Possible exception: failure of technical systems due to components in “unrelated parts” of the system.

## “Counterexample”.

Kalbfleisch and Prentice (2002). Let  $k = 2$  and:

$$Q(t_1, t_2) = \exp(1 - \alpha_1 t_1 - \alpha_2 t_2 - \exp(\alpha_{12}(\alpha_1 t_1 + \alpha_2 t_2))).$$

Cause-specific hazards:  $\alpha_j(t) = \alpha_j(1 + \alpha_{12} \exp(\alpha_{12}(\alpha_1 + \alpha_2)t))$ .

If  $\alpha_{12} = 0$  then risks 1 and 2 are “independent”. However, likelihood would be the same if the model was

$$Q^*(t_1, t_2) = \exp(1 - \alpha_1 t_1 - \alpha_2 t_2) \\ \times \exp\left(-\frac{\alpha_1 e^{\alpha_{12}(\alpha_1 + \alpha_2)t_1} + \alpha_2 e^{\alpha_{12}(\alpha_1 + \alpha_2)t_2}}{\alpha_1 + \alpha_2}\right)$$

risks are independent (also for  $\alpha_{12} \neq 0$ ); cause-specific hazards are the same (but marginal hazards are different).

## Identifiable probabilities.

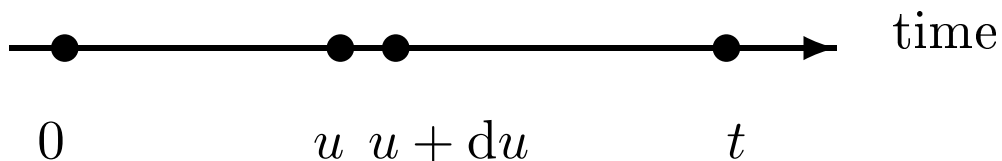
These are the state occupation probabilities in the competing risks multi-state model. That is,

the overall survival function:

$$S(t) = \exp\left(-\sum_j \int_0^t \alpha_j(u) du\right),$$

and the *cumulative incidences*:

$$F_j(t) = \int_0^t S(u-) \alpha_j(u) du, j = 1, \dots, k.$$



## Inference for cause-specific hazards.

Likelihood:

$$\begin{aligned} & \prod_{i=1}^n S(\tilde{X}_i) \prod_{j=1}^k (\alpha_j(\tilde{X}_i))^{I(D_i=j)} \\ &= \prod_{i=1}^n \left( \exp\left(-\sum_{j=1}^k A_j(\tilde{X}_i)\right) \right) \prod_{j=1}^k (\alpha_j(\tilde{X}_i))^{I(D_i=j)} \\ &= \prod_{j=1}^k \left( \prod_{i=1}^n \exp(-A_j(\tilde{X}_i)) (\alpha_j(\tilde{X}_i))^{I(D_i=j)} \right). \end{aligned}$$

Note:

- Product over causes,  $j$ ,
- The  $j$ th factor is what we would get if only that cause was studied *and all other causes were right-censorings*

## Inference for cause-specific hazards.

- 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  $A_j(t) = \int_0^t \alpha_j(u)du, j = 1, \dots, k$  by Nelson-Aalen estimator, compare using, e.g. logrank tests
  - parametric models
  - Cox regression, Poisson regression
  - ...

## Bone marrow transplantation: Cox regression models for cause-specific hazards

Covariate	Relapse		Death	
	$\hat{\beta}$	(SE)	$\hat{\beta}$	(SE)
HLA-id. sibling	0	-	0	-
HLA-matched donor	0.011	0.15	0.811	0.097
HLA-mismatched donor	-0.944	0.36	1.118	0.14
ALL	0	-	0	-
AML	-0.271	0.15	-0.195	0.14
CML	-0.721	0.16	0.291	0.117
Early stage	0	-	0	-
Intermed. stage	0.640	0.15	0.474	0.10
Advanced stage	1.848	0.15	0.781	0.13
Karnofsky > 90	-0.118	0.14	-0.504	0.11

## Cumulative incidences.

Recall:

$$\begin{aligned} F_j(t) &= \int_0^t S(u-) \alpha_j(u) du, j = 1, \dots, k \\ &= \int_0^t \exp\left(-\sum_{h=1}^k A_h(u-)\right) \alpha_j(u) du, j = 1, \dots, k. \end{aligned}$$

Note that  $F_j(t)$ , via  $S(u) = \exp(-\sum_{h=1}^k A_h(u-))$ , depends on the cause-specific hazards for *all causes*: think of cancer and cardio-vascular mortality in smokers.

That is, the simple one-to-one correspondance between the “rate”,  $\alpha_j(t)$ , and the “risk”,  $F_j(t)$ , which we are used to from simple survival analysis does no longer hold when competing risks are operating.

This is the key to understanding competing risks!

## Rates and risks.

Comparison of rates (cause-specific hazards) and risks (cumulative incidences differ):

$\alpha_j^{(1)}$	$\alpha_j^{(2)}$	$\alpha_h^{(1)}$	$\alpha_h^{(2)}$	$\frac{\alpha_j^{(2)}}{\alpha_j^{(1)}}$	$\frac{F_j^{(2)}(t_1)}{F_j^{(1)}(t_1)}$	$\frac{F_j^{(2)}(t_2)}{F_j^{(1)}(t_2)}$
0.1	0.2	1.0	2.0	2	1.33287*	1.11080*
0.1	0.2	2.0	1.0	2	2.78712	3.23094
0.1	0.1	1.0	2.0	1	0.68902	0.58025
0.1	0.1	2.0	1.0	1	1.45134	1.72340
0.1	0.2	1.0	1.0	2	1.92038	1.87474
0.1	0.1	1.0	1.0	1	1.00000	1.00000
0.1	0.2	0.1	0.2	2	1.81873**	1.67032**
0.1	0.2	0.2	0.1	2	2.00000	2.00000
0.1	0.1	0.1	0.2	1	0.95321	0.91238
0.1	0.1	0.2	0.1	1	1.04909	1.09604
0.1	0.2	0.1	0.1	2	1.90642	1.82475
0.1	0.1	0.1	0.1	1	1.00000	1.00000

## Inference for cumulative incidences.

Estimate for  $F_j(t)$ : plug-in.

For simple non-parametric inference, plugging the Nelson-Aalen estimator into  $F_j(t)$  gives the estimator

$$\widehat{F}_j(t) = \int_0^t \widehat{S}(u-) d\widehat{A}_j(u),$$

where  $\widehat{S}$  is the Kaplan-Meier estimator for the overall survival function,  $S$ .

This is a simple special case of the general *Aalen-Johansen estimator* for non-homogeneous Markov processes.

A variance estimator is also available.

## Sidetrack.

In the competing risks model, what is the interpretation of

$$\bar{F}_j(t) = 1 - \exp\left(-\int_0^t \alpha_j(u) du\right)?$$

It is: Prob(Dead from cause  $j$  before  $t$ ) **IF** all other  $\alpha_h(t) = 0$ , i.e. if the competing risks did not exist!

It can, therefore, only be interpreted in a hypothetical population where mortality from causes other than cause  $j$  have been eliminated (and where the mortality from cause  $j$  is still given by the same  $\alpha_j(t)$  - "independent competing risks").

This is an untestable assumption and the estimator  $\widehat{\bar{F}}_j(t)$  should not be used.

It has, in fact, been used extensively in, e.g. clinical cancer studies: “Relapse survival curve”.

The paradox is that  $\hat{A}_j(t)$ , the cumulative cause-specific hazard, **is not** problematic (except for the fact that a cumulative hazard is hard to interpret), but presenting this as a “Kaplan-Meier-type” estimator **is** problematic since this does not have a probability interpretation.

The magnitude of this problem, obviously, depends on the magnitude of the competing risk; but note that we always have:  $\bar{F}_j(t) \geq F_j(t)$ .

The quantity  $\bar{F}_j(t)$  corresponds to the *partial* Markov process where all transition intensities other than  $\alpha_j(t)$  are set to 0.

## Censoring in survival studies

When, in survival studies, we draw the Kaplan-Meier estimator only the death intensity is taken into account - NOT the censoring intensity. This makes sense if **BOTH**: (I), the population without censoring makes sense **AND**: (II), censoring is “*independent*”.

Example: event = death due to cancer, consider censoring due to:

- end of study
- emigration
- loss to follow-up
- death due to traffic accidents
- death due to cardiovascular diseases

The magnitude of the first problem (I) depends on the magnitude of the competing risk.

# BMT study: cumulative incidences for relapse.

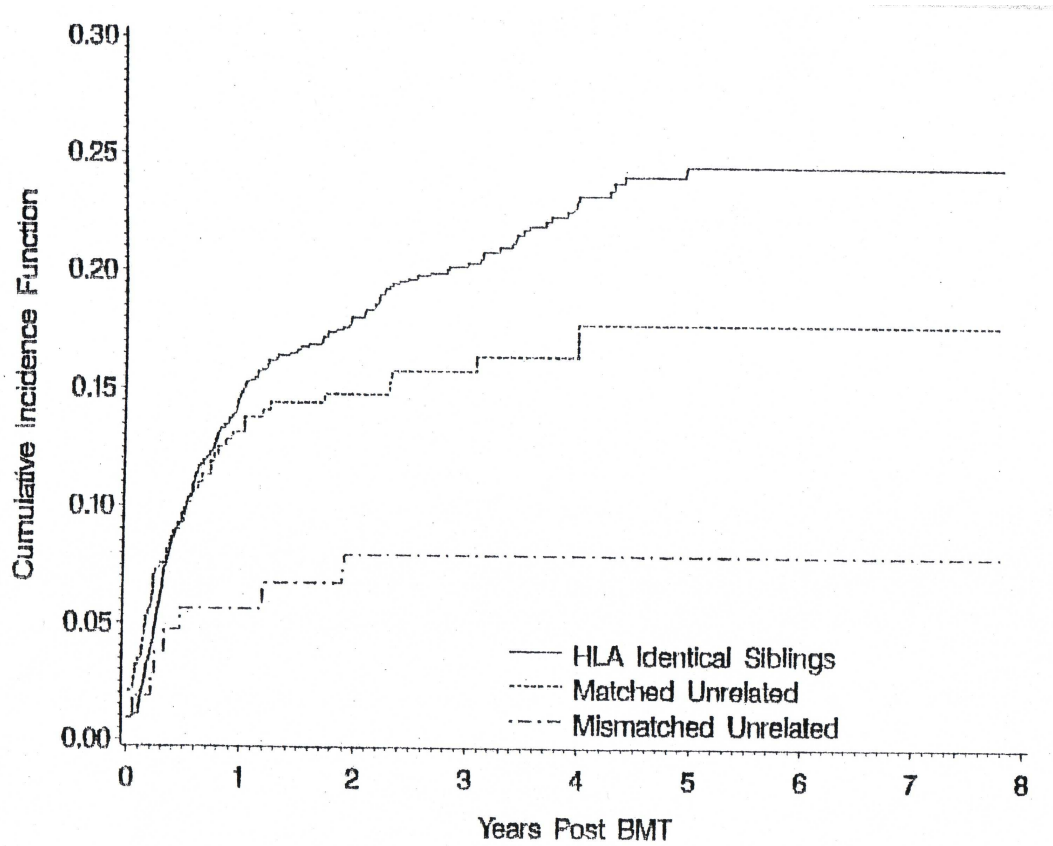


Figure 1. Cumulative incidence of relapse.

## **Time-dependent covariates.**

As in many standard hazard-based models for survival data, time-dependent covariates are easily accommodated in models for cause-specific hazards.

However, as in standard survival models, the one-to-one relation between hazard and probability only holds when time-dependent covariates are exogenous or deterministic. If covariates are endogeneous (“truly random”) then the survival probability depends on the possible future development of the time-dependent covariate.

## Regression models for competing risks

- Models for rates (cause-specific hazards):
  - all well-known hazard regression models from survival analysis
- Models for risks (cumulative incidences):
  - plug in models for rates: no simple covariate effects, but prediction of cumulative incidence for given covariates is possible (e.g., psychiatric studies). Standard errors available via the delta-method (Cox model: Andersen, Hansen and Keiding (1992, *SJS*), Cheng, Fine and Wei (1998, *Biometrics*); additive hazard model: Shen and Cheng (1999, *Biometrics*); flexible “Cox-Aalen” model: Scheike and Zhang (2003, *Biometrics*)).

– Direct regression models

- \* Fine & Gray (1999, *JASA*), cloglog model: (Sub-distribution) hazard for  $X_j^* = X \cdot I(D = j) + \infty \cdot I(D \neq j)$  is  $\tilde{\alpha}_j(t) = -\frac{d}{dt} \log(1 - F_j(t))$ , model:  $\log(\tilde{\alpha}_j(t | Z)) = \log(\tilde{\alpha}_{j0}(t)) + \beta^\top Z$  where  $\beta$  is estimated by partial likelihood with no or known censoring and by an IPCW score equation with general censoring. Related to the Gray (1988, *Ann. Statist.*) test for comparison of cumulative incidences.
- \* Fine (1999, *JRSS B*, 2001, *Biostatistics*), general links
- \* pseudo-observations (Andersen, Klein et al., *Biometrika*, 2003; *Biometrics*, 2005; *SJS*, 2007),
- \* Direct binomial regression using inverse probability of censoring weights (Scheike & Zhang, 2007, *SJS*, Scheike, Zhang & Gerds *Biometrika*)
- \* (The latter two work for more general multi-state models.)

## Direct models for probabilities:

Without censoring, the counting process  $N_{ij}(t) = I(X_i \leq t, D_i = j)$  is observed and its values can be used as outcome variables in a regression model.

### 1. Pseudo-observations:

$\hat{\theta}(t) = \hat{F}_j(t)$  estimator based on entire sample,

$\hat{\theta}_{-i}(t)$  estimator based on data obtained by deleting  $i$ .

Pseudo-observation no.  $i$  is then given by

$$\hat{\theta}_i(t) = n \cdot \hat{\theta}(t) - (n - 1) \cdot \hat{\theta}_{-i}(t).$$

For selected (all?) time points,  $t_1, \dots, t_m$ , these are used as outcome variables in standard regression models using GEE with some choice of link function, e.g. cloglog, and sandwich estimator for variances.

Note that censoring must be independent of covariates.

## 2: Direct binomial regression:

Let

$$V_{ji} = X_i I(C_i > X_i, D_i = j) + \infty(I(D_i \neq j) + I(C_i < X_i)).$$

Then

$$E \left( \frac{I(V_{ji} \leq t)}{S_C(V_{ji})} \right) = F_j(t),$$

and inverse probability of censoring weights (IPCW) can be used in the regression.

1. and 2. shown by Graw, Gerds & Schumacher (2009) to be equivalent:

$$\hat{\theta}_i(t) = \frac{I(X_i \leq t, D_i = j, X_i \leq C_i)}{S_C(X_i)} + o_P(1).$$

## Bone marrow transplantation.

Regression models for cumulative incidences (adjusted for disease, stage, Karnofsky) at 10 time points.

Relapse Covariate	Fine-Gray		Pseudo: cloglog		Pseudo: logit	
	$\hat{\beta}$	(SE)	$\hat{\beta}$	(SE)	$\hat{\beta}$	(SE)
HLA-id. sibling	0	-	0	-	0	-
HLA-matched donor	-0.32	0.16	-0.37	0.16	-0.45	0.19
HLA-mismatched donor	-1.37	0.38	-1.61	0.45	-1.88	0.51

Death in remission Covariate	Fine-Gray		Pseudo: cloglog		Pseudo: logit	
	$\hat{\beta}$	(SE)	$\hat{\beta}$	(SE)	$\hat{\beta}$	(SE)
HLA-id. sibling	0	-	0	-	0	-
HLA-matched donor	0.76	0.10	0.75	0.10	0.95	0.12
HLA-mismatched donor	1.15	0.13	1.23	0.14	1.64	0.21

## Further topics.

- Missing failure type information, e.g. Goetghebeur and Ryan (*Biometrika*, 1995).
- Relative survival:  $\alpha_i(t) = \alpha_i^*(t) + \alpha(t)$ .
- Pattern mixture regression model, e.g. Larson and Dinse (*Appl. Stat.*, 1985):

$$\text{logit}P(D = j | Z) = a_j + b_j^\top Z, \text{cloglog}P(X > t | Z, D = j) = \log(\alpha_{0j}(t)) + \beta_j^\top Z.$$

- Parameters in common for several cause-specific hazards, e.g.

$$\alpha_h(t | Z) = \alpha_{0h}(t) \exp(\beta_{h1}Z_1 + \beta Z_2), h = 1, 2$$

(Andersen, Borgan, Gill and Keiding, 1993, Springer; Lunn and McNeil, 1995, *Biometrics*).

- Random effects (frailty) models

## More on identifiability.

We saw previously that the joint survival function for the latent failure times was not identifiable from the competing risks data.

Two remarks to that effect:

1. Heckman and Honoré (*Biometrika*, 1989) showed that, with quantitative covariates and marginal proportional hazards models, the joint distribution *is* identifiable under suitable regularity conditions.
2. Zheng and Klein (*Biometrika*, 1995) showed that, with an assumed form of the joint distribution (an assumed “copula”), the marginal distribution of an  $X_j^L$  is identifiable. They used this to estimate in the presence of dependent censoring.

## Software.

- Models for cause-specific hazards: `coxph` in R, PHREG in SAS, ...
- Fine-Gray model, Gray's test: `cmprsk` package in R
- Plug-in estimation of cumulative incidence based on Cox models for cause-specific hazards: SAS MACRO by Rosthøj, Andersen and Abildstrom (*Comp. Progr. Meth. Biomed.*, 2004). Also John Klein's homepage at Medical College of Wisconsin, Milwaukee.
- Computation of pseudo-values: SAS MACRO and R function by Klein, Gerster, Andersen, Tarima, Pohar Perme (*Comp. Progr. Meth. Biomed.*, 2008).
- ...

## Conclusions.

- Competing risks are frequently seen in biomedical (and other) applications of survival analysis.
- The identifiable parameters are cause-specific hazards and cumulative incidences (and overall survival function).
- In standard survival analysis, rates (hazards) and risks (failure probabilities) are equivalent
- In the competing risks model, this is no longer the case: for each failure type, the failure risk depends on all failure rates

- Modelling cause-specific hazards using techniques from standard survival analysis is quite straightforward
- The cause-specific hazards, however, only provide a *local* (in time) description of the operation of the causes
- To obtain a *global* description (cumulated over time), the competing risks must be specified (and accounted for) unless they are small
- Questions like independence of risks and what would happen if certain causes were removed may be important and interesting but their answers rely on unverifiable assumptions. However, sensitivity analysis may give useful insight.

## References.

- OO Aalen (1976): *Scand. J. Statist.* **3**, 15.
- OO Aalen (1978). *Ann. Statist.* **6**, 534.
- OO Aalen, S Johansen (1978). *Scand. J. Statist.* **5**, 141.
- PK Andersen, SZ Abildstrom, S Rosthøj (2002). *Statist. Meth. Med. Res.* **11**, 203.
- PK Andersen, Ø Borgan, RD Gill, N Keiding (1993). *Statistical Models Based on Counting Processes*. Springer
- PK Andersen, LS Hansen, N Keiding (1991). *Scand. J. Statist.* **18**, 153.
- PK Andersen N Keiding (2002). *Statist. Meth. Med. Res.* **11**, 91.
- PK Andersen, JP Klein, S Rosthøj (2003). *Biometrika* **90**, 15.
- PK Andersen, JP Klein (2005). *Biometrics* **61**, 223.
- PK Andersen, JP Klein (2007). *Scand. J. Statist.* **34**, 3.
- SC Cheng, JP Fine, LJ Wei (1998). *Biometrics* **54**, 219.
- DR Cox (1959). *J. Roy. Statist. Soc. B* **21**, 411.
- M Crowder (2001). *Classical Competing Risks*, Chapman and Hall/CRC.
- PW Dickman, A Slogett, M Hills et al. (2004). *Statist. Med.* **40**, 326.
- JP Fine, RJ Gray (1999). *J. Amer. Statist. Assoc.* **94**, 496.

- JP Fine (1999). *J. Roy. Statist. Soc. B* **61**, 817.
- JP Fine (2001). *Biostatistics* **2**, 85.
- JJ Gaynor, EJ Feuer, CC Tan et al. (1993). *J. Amer. Statist. Assoc.* **88**, 400.
- E Goethghebeur, L Ryan (1995). *Biometrika* **82**, 821.
- TA Gooley, W Leisenring, J Crowley et al. (1999). *Statist. Med.* **18**, 695.
- F Graw, TA Gerds, M Schumacher (2009). *LIDA*, to appear.
- RJ Gray (1988). *Ann. Statist.* **16**, 1141.
- JJ Heckman, BE Honoré (1989). *Biometrika* **76**, 325.
- JD Kalbfleisch, RL Prentice (2002): *The Statistical Analysis of Failure Time Data, 2nd ed.*. Wiley.
- JP Klein, M Gerster, PK Andersen et al. (2008). *Comp. Progr. Meth. Biomed.*
- MG Larson, GE Dinse (1985). *Appl. Statist.* **34**, 201.
- DY Lin (1997). *Statist. Med.* **16**, 901.
- M Lunn, D McNeil (1995). *Biometrics* **51**, 524.
- M Pintilie (2006). *Competing Risks. A Practical Perspective*, Wiley.
- RL Prentice, JD Kalbfleisch, AV Peterson et al. (1978). *Biometrics* **34**, 541.
- S Rosthøj, PK Andersen, SZ Abildstrom (2004). *Comp. Progr. Meth. Biomed.* **74**, 69.

TH Scheike, M-J Zhang (2003). *Biometrics* **59**, 1036.

TH Scheike, M-J Zhang (2007). *Scand. J. Statist.* **34**, 17.

TH Scheike, M-J Zhang, TA Gerds (2008). *Biometrika* (in press).

Y Shen, SC Cheng (1999). *Biometrics* **55**, 1093.

R Szydlo, JM Goldman, JP Klein et al. (1997). *J. Clin. Oncol.* **15**, 1767.

AA Tsiatis (1975). *Proc. Nat. Acad. Sci, USA* **72**, 20.

M Zheng, JP Klein (1995). *Biometrika* **82**, 127.