# Dealing with competing risks in survival analysis

Per Kragh Andersen
Section of Biostatistics
University of Copenhagen, Denmark

On behalf of TG8 - survival analysis

STRATOS meeting, Banff, 2016

## Summary

1. Repetition of a few well-known facts from survival analysis
2. Competing risks
3. Applied papers on competing risks
4. Competing risks and other topic groups
5. Pseudo-observations
6. Concluding remarks

TG8 members: Michal Abrahamowicz, Per Kragh Andersen, Richard Cook, Coraline Danieli, Pierre Joly, Torben Martinussen, Maja Pohar Perme, Jeremy Taylor, Terry Therneau

## Target population and censoring

When doing statistical inference, data are considered a *sample* from some *target population* to which parameters refer.

In *survival analysis*, parameters (like the survival function $S(t)$ or the hazard function $h(t)$) refer to a potentially completely observed population, i.e. one *without censoring*.

The object of survival analysis is then (the ambitious one) of drawing inference for such parameters based on *incomplete data*.

This requires an assumption of *independent censoring*.

## Independent censoring

Independent censoring means that individuals censored at any given time $t$ should not be a biased sample of those who are *at risk* at time $t$.

Stated in other words: the hazard

$$h(t) \approx \mathrm{P}(T \leq t + dt \mid T > t)/dt$$

gives the *event rate* at time $t$, i.e. the failure rate given that the subject is still alive ($T > t$).

Independent censoring then means that the extra information that the subject is not only alive, but also uncensored at time $t$ does not change the hazard.

Typically, independent censoring cannot be tested from the available data - it is a matter of discussion. However, it is crucial that the complete population (without censoring) is *well-defined*.

## Inference for independently censored data

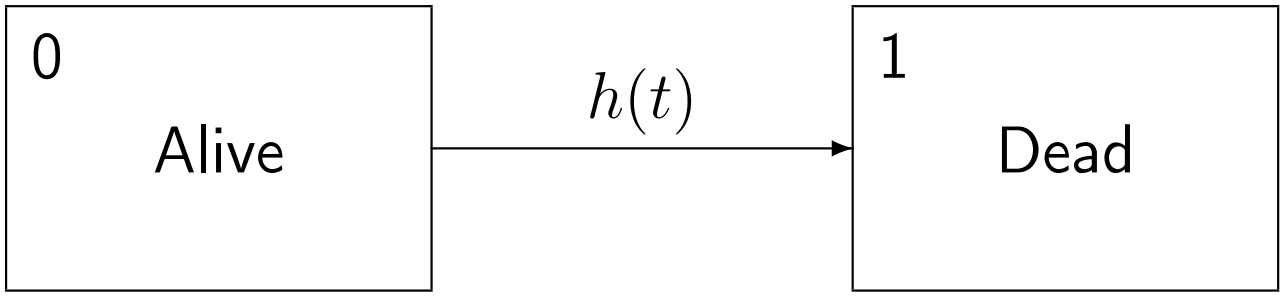
For independent observations $(X_i, \Delta_i, i = 1, \ldots, n)$, with $X_i = \min(T_i, U_i), \Delta_i = I(T_i \leq U_i)$, the likelihood is conveniently expressed via the hazard (and cumulative hazard) functions:

$$L(\theta) = \prod_i h_\theta(X_i)^{\Delta_i} \exp(-H_\theta(X_i)).$$

This leads (with suitable definition of NPMLE) to the *Nelson-Aalen estimator* for $H(\cdot)$ and to the *Cox partial likelihood*.

Using the relation between $S(t)$ and $H(t)$ leads to the *Kaplan-Meier estimator* as a plug-in estimator (because the product-integral relation $S(t) = \exp(-H(t))$ for absolutely continuous distributions becomes a finite product for a discrete distribution).

## Target population




In the target population with complete observation, every one ends up in state 1 and the probability of being in state 1 at time $t$ (the 'failure risk') is given uniquely from the hazard:

$$F(t) = 1 - S(t) = 1 - \exp(-H(t)).$$

## Competing risks

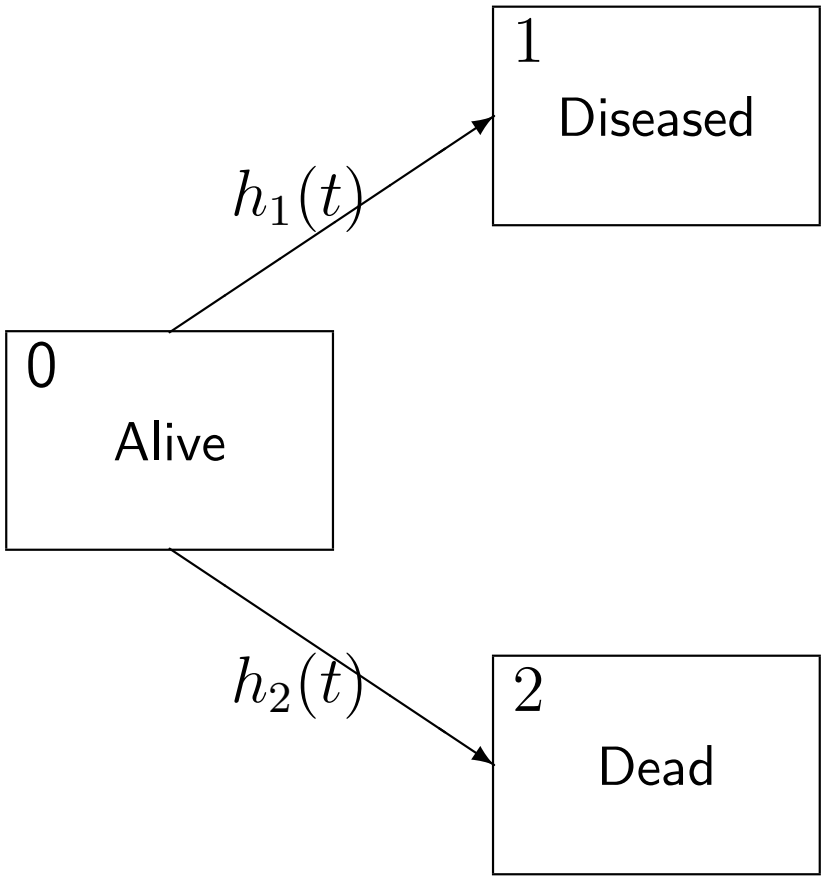
Suppose now that the event of interest is not overall mortality (or a composite end-point that includes death) but rather something like the occurrence of a given disease.

Then, in the target population even with complete follow-up, every one will *not* experience the event: some subjects will die without ever getting the disease.

Deaths without the disease should then *not* be thought of as an 'independent censoring' because a target population 'without censoring' (i.e., one where subjects are not allowed to die without the disease) is completely hypothetical.

We have *competing risks*.

# Target population



## Relationship between rates and risks

The cause-specific hazards $j = 1, 2$ are the transition intensities:

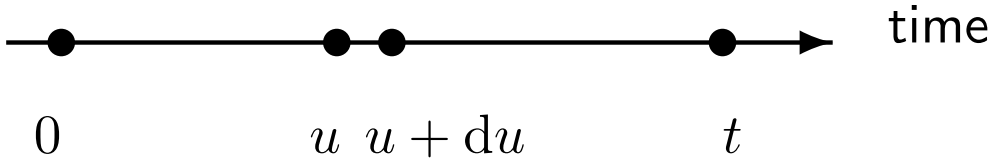
$$h_j(t) \approx \mathrm{P}(\text{state } j \text{ time } t + dt \mid \text{state } \ 0 \ \text{ time } \ t)/dt.$$

The state occupation probabilities include the *overall survival function*:

$$S(t) = \mathrm{P}(\text{alive time } t) = \exp(-(H_1(t) + H_2(t))$$

and the *cumulative incidences* ('risks') for causes $j = 1, 2$:

$$F_j(t) = \mathrm{P}(\text{state } j \text{ time } t) = \int_0^t S(u) h_j(u) du.$$




Both rates are needed to compute one risk!

## Inference for independently censored data

For independent observations $(X_i, \Delta_i \cdot D_i, i = 1, \ldots, n)$, with $X_i = \min(T_i, U_i), \Delta_i = I(T_i \leq U_i), D_i =$final state, the likelihood may again be expressed via the cause-specific hazard (and cumulative hazard) functions:

$$L(\theta) = \prod_i h_{1,\theta}(X_i)^{\Delta_i I(D_i=1)} h_{2,\theta}(X_i)^{\Delta_i I(D_i=2)} \exp(-H_{1\theta}(X_i) - H_{2\theta}(X_i))$$

which factorizes:

$$L(\theta) = \prod_i h_{1,\theta}(X_i)^{\Delta_i I(D_i=1)} \exp(-H_{1\theta}(X_i))$$

$$\times \prod_i h_{2,\theta}(X_i)^{\Delta_i I(D_i=2)} \exp(-H_{2\theta}(X_i)).$$

## Consequences of likelihood factorization

- Each factor has the form it would have if only cause $j$ events were events and events of the other type were treated as (independent) censorings
- In particular, the *Nelson-Aalen estimator* for $H_j(\cdot), j = 1, 2$ and the *Cox partial likelihood* for each cause separately work perfectly well
- Cumulative incidences may be estimated by plug-in (the 'Aalen-Johansen estimator'), while the '1-Kaplan-Meier estimator' for cause $j$ estimates

$$\int_0^t \exp(-H_j(u))h_j(u)du$$

  and is (upwards) biased for $F_j(t)$
- First of all, it leads to *confusion* and *misunderstanding*

## Confusion

Frequent question:

"why can we use the Cox model with competing risks but not the Kaplan-Meier estimator?"

Questions like that have lead to *numerous* (more or less) *pedagogical* papers in the applied (mostly medical) literature trying to explain these difficulties. These include:


Satagopan et al. (2004). *Br. J. Cancer*, 91, 1229-1235.

Southern et al. (2006). *J. Clin. Epidemiol.*, 59, 1110-1114.

Kim (2007). *Clin. Cancer Res.*, 13, 559-565.

Biau et al. (2007). *Clin. Orthopaed. and Rel. Res.*, 462, 229-233.

Scrucca et al. (2007). *Bone Marrow Transpl.*, 40, 381-387.

Zhang et al. (2008). *Expert Rev. Clin. Pharmacol.* 1, 391 .

Dignam et al. (2008). *J. Clin. oncol.*, 26, 4027-4034.

Shiels et al. (2009). *J. Clin. Epidemiol.*, 63, 459-467.

Wolbers et al. (2009). *Epidemiol.*, 20, 555-561.

Lau et al. (2009). *Amer. J. Epidemiol.*, 170, 244-256.

Allignol et al. (2011). *BMC Med. Res. Methodol.*, 11, 86.

Andersen et al. (2012). *Int. J. Epidemiol.*, 41, 861-870.

Austin et al. (2016). Circulation, 133, 601-619.

Much of the confusion seems to arise from not clearly distinguishing between censoring and competing risks (defining the target population) and between 'rates' and 'risks'.

# Rates vs. risks?

Competing risks ‘analogy’:

## Gladiators

Suppose that a gladiator may lose via two quite different mechanisms: lions or fellow gladiators.

When training a gladiator, he should both be prepared to face a lion or a fellow gladiator, and special skills may be needed to face a lion (even in the presence of the competing risk) and, similarly, special skills may be needed to beat a fellow gladiator. The cause-specific hazards describe how these mechanisms depend on properties and equipment of the gladiator.

For Ceasar to predict the number of remaining gladiators still around at time $t$, both risks must be considered. He needs the cumulative incidences given the distribution of properties and equipment of the population of gladiators.

## Conclusion

- Cause-specific hazards describe the mechanisms by which subjects may fail.
- Cumulative incidences describe the fraction of the population that fails from given causes.
- Both are useful (and both are needed?) for a complete description of the competing risks situation.

Latouche A., Allignol A., Beyersmann J., Labopin M., Fine J.P.: A competing risks analysis should report results on all cause-specific hazards and cumulative incidence functions. *J. Clin. Epidemiol.* (2013) **66**, 648-653.

## Competing risks and TG8

Competing risks is, therefore, cruicial for STRATOS TG8 and we have to find our niche in this vast literature.

The above-mentioned papers correspond to 'STRATOS-level 1'.

Can/should we contribute more to that?

Or should we focus on level 2?

(I am quite sceptical towards STRATOS-level 3.)

## Competing risks and other topic groups

TG1 (missing data) There is a steadily increasing literature on how to do multiple imputation in competing risks analysis with missing data (e.g., Bartlett & Taylor, *Biostatistics*, 2016; Delord & Genin, *J. Statist. Comp. Sim.*, 2016).

Methods depend on whether covariates or cause of failure is missing. Important to distinguish between censoring and failure from a competing cause.

TG2 (variable selection and functional forms of a dose-response relationship) Special regression models are typically used for survival analysis (possibly with competing risks) though working with models with a linear predictor ($LP$) is quite similar for different types of outcome variable.

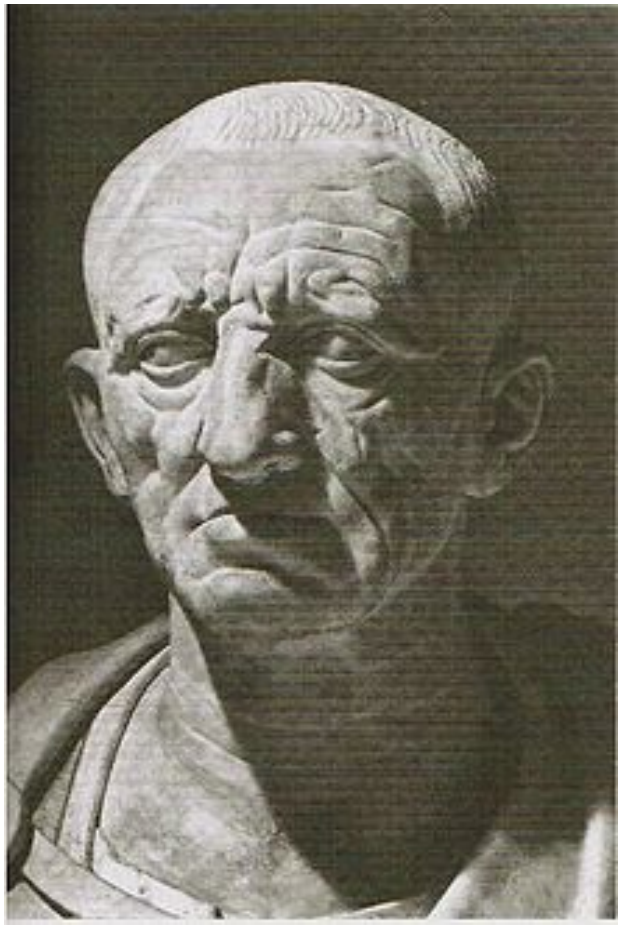
More below.

TG6 (evaluating diagnostic tests and prediction models) When assessing predictive accuracy, special care is needed for right-censored outcomes, including situations with competing risks.

More below.

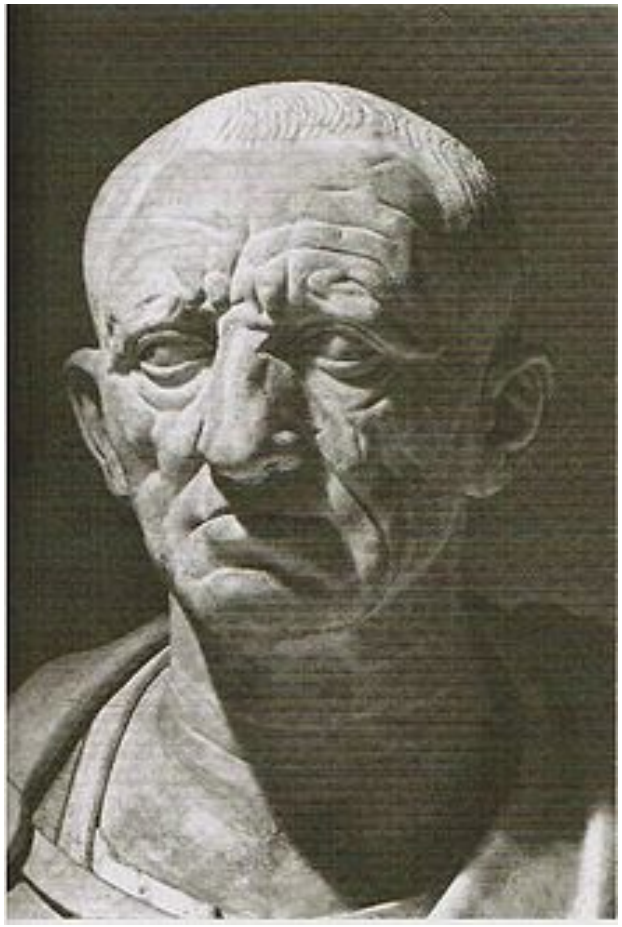
TG7 (causal inference) Both when using IPTW and when using the g-formula, special techniques are needed in the presence of competing risks.

More below.



Marcus Porcius Cato ("Cato the elder", $\approx$ 234 BC - 149 BC): Roman soldier, consul, senator etc.

"Ceterum censeo Carthaginem esse delendam"



Marcus Porcius Cato ("Cato the elder", $\approx$ 234 BC - 149 BC): Roman soldier, consul, senator etc.

"Moreover, I advise that Carthage must be destroyed"

## Pseudo-observations

"Moreover, I think that pseudo-observations are useful in survival analysis"

Consider censored data and a mean value parameter $\theta = \mathrm{E}(f(T))$.

Possible mean value parameters, $\theta$ of interest are:

- survival probability $S(t) = \mathrm{E}(I(T_i > t))$,
- $t-$restricted mean $\mu_t = \mathrm{E}(T_i \wedge t)$,
- cumulative incidence in competing risks

  $F_j(t) = \mathrm{E}(I(T_i \leq t, D_i = j))$.

Regression analysis for $\mathrm{E}(f(T) \mid Z)$ may be performed using *pseudo-observations* as targets in a GEE:

$$\theta_i = n\widehat{\theta} - (n-1)\widehat{\theta}^{-i}.$$

Here, $\widehat{\theta}$ is a consistent estimator for the marginal mean (e.g., the Aalen-Johansen estimator for the cumulative incidence) and $\widehat{\theta}^{-i}$ is the same estimator applied to the sample of size $n-1$ obtained by eliminating subject $i$.

With no censoring, $\theta_i$ in the examples is simply, respectively:

- $I(T_i > t)$,
- $T_i \wedge t$,
- $I(T_i \leq t, D_i = j)$.

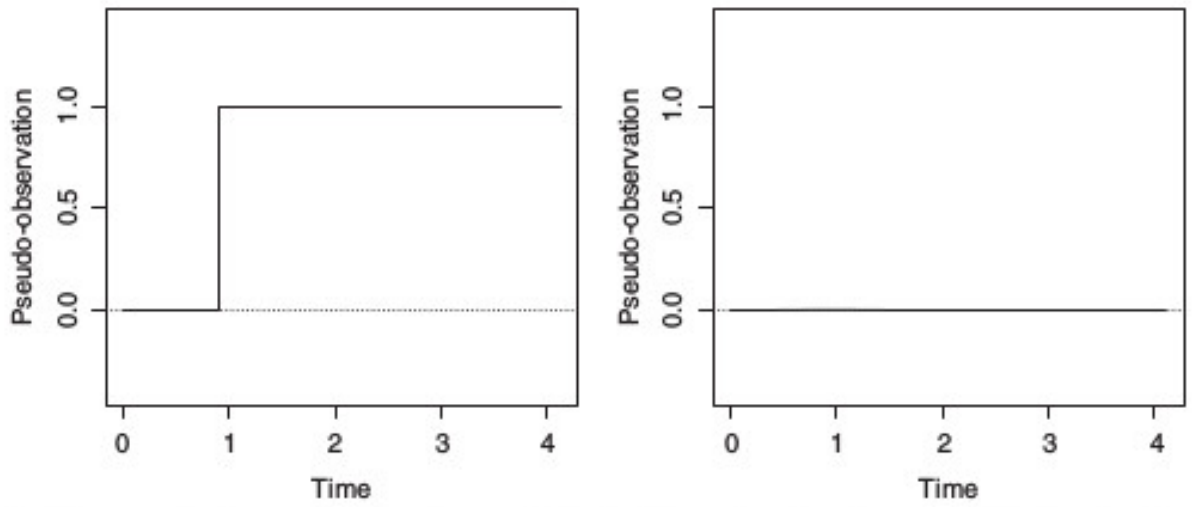



**Figure 6** The pseudo-observations for the cumulative incidence function (for cause 1) in time for the case of no censoring. (a) Individual dying from cause 1; (b) Individual dying from cause 2.

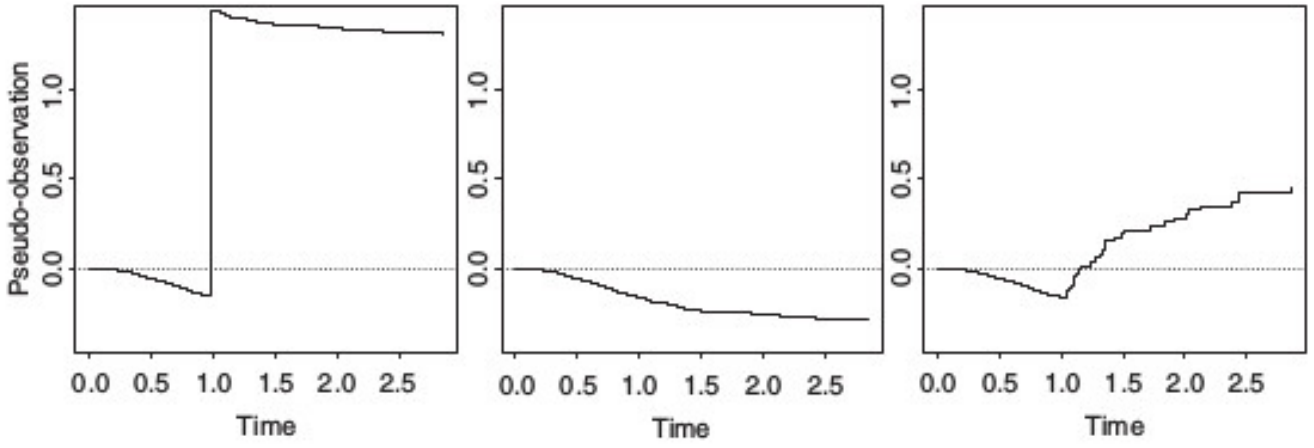



**Figure 7** The pseudo-observations for the cumulative incidence function (for cause 1) in time for a censored data set. (a) Individual failing from cause 1; (b) Individual failing from cause 2; (c) Censored individual.

## Properties of pseudo-observations

Graw, Gerds and Schumacher (*LIDA*, 2009) showed (for the competing risks cumulative incidence function) that if censoring is independent of covariates then, uniformly in $i = 1, \ldots, n$ as $n \to \infty$:

$$\mathrm{E}(F_{ji}(t) \mid Z_i) = \mathrm{E}(I(T_i \leq t, D_i = j) \mid Z_i) + o_P(1).$$

(For covariate-dependent censoring - see Binder, Gerds and Andersen, *LIDA*, 2014)

Later asymptotic results by Jacobsen and Martinussen, (*Scand. J. Statist.*, 2016) and Overgaard, Parner and Pedersen (submitted), show that the estimated SD using the sandwich estimator when pseudo-values are used as targets in GEE may be (slightly) conservative.

## Variable selection and functional forms

Competing risks regression for the cumulative incidence (at one or a few time points) may be performed using pseudo-observations (Klein & Andersen, *Biometrics*, 2005; Gerds et al., *Stat. in Med.*, 2012):

$$\sum_i \left( \frac{\partial}{\partial \beta} g^{-1}(LP_i) \right)^{\mathsf{T}} V_i^{-1} (F_{ji} - g^{-1}(LP_i)) = 0.$$

Pros:

- Easy to apply different link functions, $g$ not only the cloglog link
- Scatterplots and residual plots are available (Pohar Parme & Andersen, *Stat. in Med.*, 2008)

Cons:

- Not fully efficient
- Not (yet) possible to do for all time points simultaneously

## Evaluation of prediction models

Pseudo-observations may be used when estimating the Brier score for a competing risks regression model for the cumulative incidence (Cortese et al., *Stat.in Med.*, 2013). The estimated value for complete data:

$$\widehat{B}_j(s,t) = \frac{1}{m_s} \sum_{i \in R(s)} \left( I(X_i(t) = j) - r_j(t \mid Z_s) \right)^2$$

(where $r_j(t \mid Z_s)$ is the predictor used at time $s$) is then replaced by

$$\widetilde{B}_j(s,t) = \frac{1}{m_s} \sum_{i \in R(s)} \left( F_{ji}(t \mid s)(1 - 2r_j(t \mid Z_s)) + r_j(t \mid Z_s)^2 \right)$$

## Evaluation of prediction models

Pseudo-observations may also be used for making a calibration plot for a prediction model $r_j(t \mid z)$ (Gerds et al., *Stat. in Med.*, 2014).

The calibration curve is the graph of $p \to C(p, t, r)$:

$$C_j(p, t, r) = \mathrm{E}_{X,Z}\big(I(X(t) = j) \mid Z \in G_r(t; p)\big)$$

where $G_r(t; p)$: set of covariates $Z$ for which $r$ predicts an event probability of $p$ at $t$:

$$G_r(t; p) = \{z \in \mathbf{R}^d : r_j(t \mid z) = p\}.$$

The calibration curve may be estimated by:

$$\widehat{C}_{j,a_n}(p, t, r) = \frac{1}{n_{a_n}} \sum_{i=1}^{n} F_{ji}(t) K_{a_n}(p, r_j(t \mid Z_i)).$$

## Causal inference

The 'g-formula' for a completely observed putcome, $Y$ is:

$$\frac{1}{n}\sum_{i=1}^{n}\big(\widehat{\mathrm{E}}(Y_i \mid A_i := 1, Z_i) - \widehat{\mathrm{E}}(Y_i \mid A_i := 0, Z_i)\big).$$

The propensity score weighted outcome to be used in a marginal structural model is:

$$\widetilde{Y}_i = \frac{A_i Y_i}{\widehat{e}(Z_i)} + \frac{(1-A_i)Y_i}{(1-\widehat{e}(Z_i))}$$

(where $e(Z) = \mathrm{P}(A = 1 \mid Z)$ is the propensity score). For a time-to-event outcome, e.g. the competing risks cumulative cause $j$ incidence at time point $t$, Andersen et al. (submitted) showed that the outcome $Y_i$ in these expressions may be replaced by the pseudo-value $F_{ji}(t)$.

## Concluding remarks

- Competing risks is a central topic for TG8
- (However, there are many other topics of interest for TG8)
- Many examples of overlap between TG8 and other TG's
- Here, exemplified via competing risks
- Pseudo-values are useful (and fun!)