Internal time-dependent covariates and competing risks, an application to bone marrow transplant data

Giuliana Cortese [1], Per K. Andersen [2]

[1] Department of Statistical Sciences, University of Padova
[2] Department of Biostatistics, Copenhagen University

ISCB 30
Prague, August 23-27th, 2009

Outline

- Time-dependent covariates and problems in competing risks regression
- Possible approaches for handling internal time-dependent covariates
- Example: Bone marrow transplant data.

Time-dependent covariates

$$\boldsymbol{X}(t) = \{X(u); 0 \le u \le t\}, \quad t \le \tau, \quad t \in [0, \tau] \quad \tau < \infty$$

T = time to the occurrence of a given event in $[0, \tau]$.

- Internal (random) time-dependent covariates:
  (Kalbfleisch and Prentice, 2002)
  for $u, t: \; u \le t$,

$$\alpha(u; \boldsymbol{X}(u), T \ge u) = P\{u \le T < u + \Delta u | \boldsymbol{X}(u), T \ge u\}$$
$$\ne P\{u \le T < u + \Delta u | \boldsymbol{X}(t), T \ge u\}$$

- Possibility of observing $X(t)$ depends on the individual survival, its path carries information about the time of failure occurrence.

- Examples: disease complications, biochemical or clinical measurements recorded at follow-up visits.

Internal time-dependent covariates

- The model for $\boldsymbol{X}(t)$ depends on parameter of interest and then would need to be specified.

Internal time-dependent covariates

- The model for $\boldsymbol{X}(t)$ depends on parameter of interest and then would need to be specified.
- For estimating hazard functions, inference can still be based on partial likelihood.

Internal time-dependent covariates

- The model for $\boldsymbol{X}(t)$ depends on parameter of interest and then would need to be specified.
- For estimating hazard functions, inference can still be based on partial likelihood.
- Prediction of survival probabilities is not feasible,

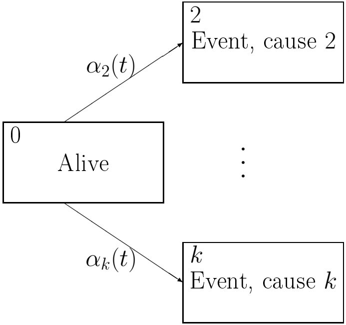
$$P(T > t) = \exp(-\int_0^t \alpha(u)du)$$

depends simultaneously on $\alpha(\cdot)$ and the random development of the covariate $\boldsymbol{X}(t)$.

The competing risks model

- Markov process $\{Z(t), t \in [0, \tau]\}$, with state space $\mathcal{S} = \{0, 2, \ldots, k\}$
- Transition intensities $\alpha_h(t)$ (cause-specific hazard functions)

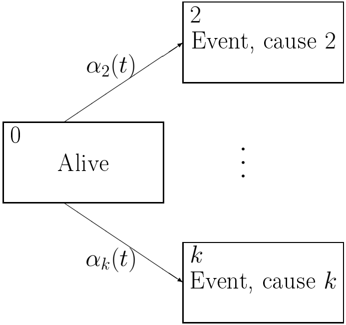
  $h = 2, \ldots, k$



The competing risks model

- Markov process $\{Z(t), t \in [0, \tau]\}$, with state space $\mathcal{S} = \{0, 2, \ldots, k\}$
- Transition intensities $\alpha_h(t)$ (cause-specific hazard functions)

  $h = 2, \ldots, k$



Approaches for regression on X (time independent):

- cause-specific hazards (Andersen et al., 1993)
- subdistribution hazards (Fine & Gray, 1999)

Competing risks regression: Cause-specific hazards

$\mathcal{S} = \{0, 2, 3\} \quad h = 2, 3$

$\boldsymbol{X}(t)$ **external time-dependent covariate**

Cumulative incidence function for cause 2

$$P_{02}(0, t; X(t)) = \int_0^t S(u-; X(u-))\alpha_2(u; X(u))du.$$

$$S(t; X(t)) = \exp\left(-\int_0^t \{\alpha_2(u; X(u)) + \alpha_3(u; X(u))\} \, du\right),$$

$\boldsymbol{X}(t)$ **internal time-dependent covariate**

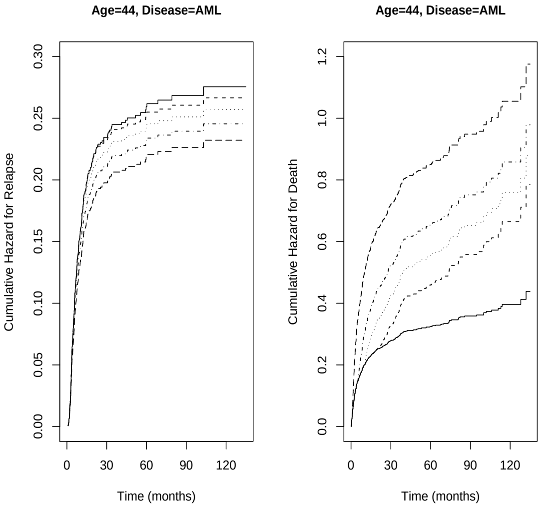
* $\alpha_h(t; X(t-))$ and hazard ratios can be estimated.
* $P_{0h}(0, t; X(t))$ and $S(t; X(t))$ can no longer be interpreted as probabilities
  (Kalbfleisch and Prentice, 2002).
* This problem arises also in the subdistribution hazards approach

Bone Marrow Transplant Data (from European CIBMTR studies)

- **Data set**: 2009 patients who received bone marrow transplant from 1995 to 2004.
- **Follow-up period**: time since transplant to death or censoring.
- **Events**: Relapse, Death.
- **Internal time-dependent covariate**: Development of Graft versus Host Disease (GvHD).
- **Aim of the study**: investigate the effect of GvHD on the cumulative risk for Death and Relapse.
- **Prognostic factors**: (measured at time of diagnosis)
  age,
  disease (AML versus ALL).

Bone Marrow Transplant Data: Competing risks model

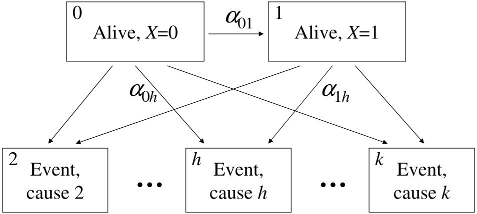
Relapse **Death**



Broken lines: Presence of GvHD developed at times $t = 1, 5, 10, 20$

Solid lines: No appearance of GvHD over follow-up time.

An extended model with transient states (from Andersen, 1986)

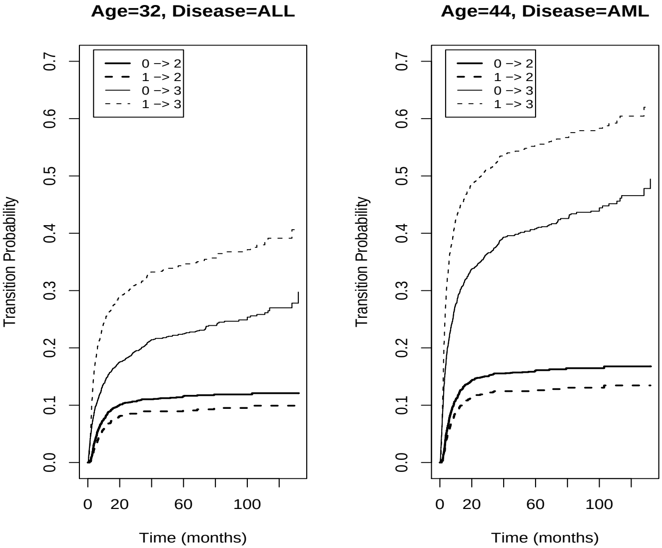


$X(t) \in \{0,1\}, \quad \boldsymbol{X}(t) = I(T_1 > t),\ T_1 =$ time of entrance in state 1.

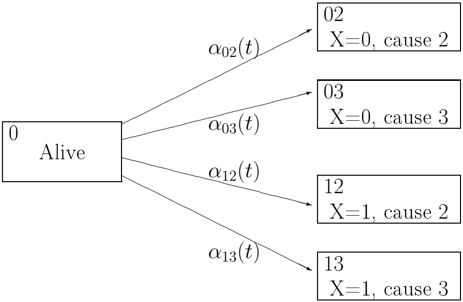
- $\mathcal{S}$ **is enlarged** to $\mathcal{S}' = \{0, 1, 2, \ldots, k\}$ and $Z(t)$ is modified so that $Z'(t) = X(t+)$ in $[0, T)$.
- **Test for Markovian assumption**:
  Consider a semi-Markov model with $\alpha_{1h}(t, t - T_1)$, $h = 2, \ldots, k$ and $t - T_1 =$ sojourn time in state 1.
- It needs a model for the development of $X(t)$ over time.

Bone Marrow Transplant Data: The extended model

1 = GvHD appearence, 2 = Relapse, 3 = Death



An alternative extended competing risks model



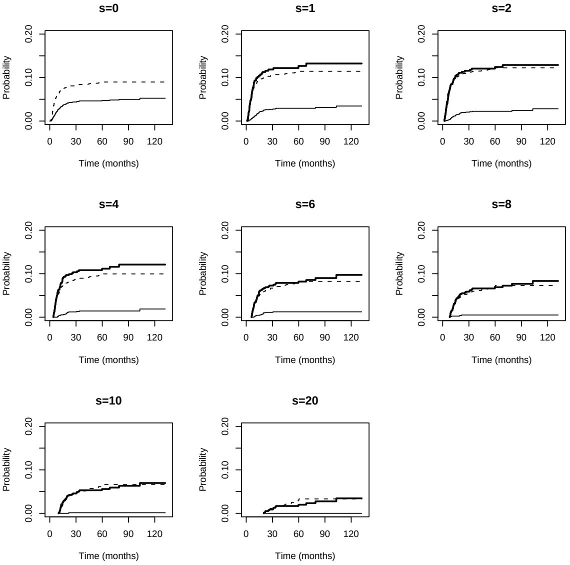
$X(t) \in \{0,1\}$, $\boldsymbol{X}(t) = I(T_1 > t)$, T_1 = time of entrance in state 1.

Conditional cumulative incidences for different suitable times s:

$$P_{0h}(T \leq t, Z(T) = h \mid T \geq s, X(s)), \qquad h \in \{02, 03, 12, 13\},$$

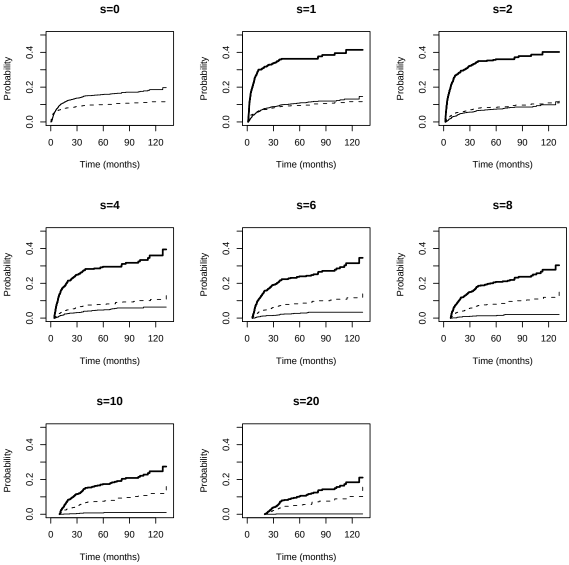
$X(s)$ time-constant on $[s, t]$.

Relapse



Dashed line: $P_{0,02}(s, t \,|X(s) = 0\,)$, **Thick solid line**: $P_{0,12}(s, t \,|X(s) = 1\,)$,
Solid line: $P_{0,12}(s, t \,|X(s) = 0\,)$

Death



Dashed line: $P_{0,03}(s, t\,|X(s) = 0)$, **Thick solid line**: $P_{0,13}(s, t\,|X(s) = 1)$,
Solid line: $P_{0,13}(s, t\,|X(s) = 0)$

The landmark approach

$\mathcal{S} = \{0, 2, 3\}$

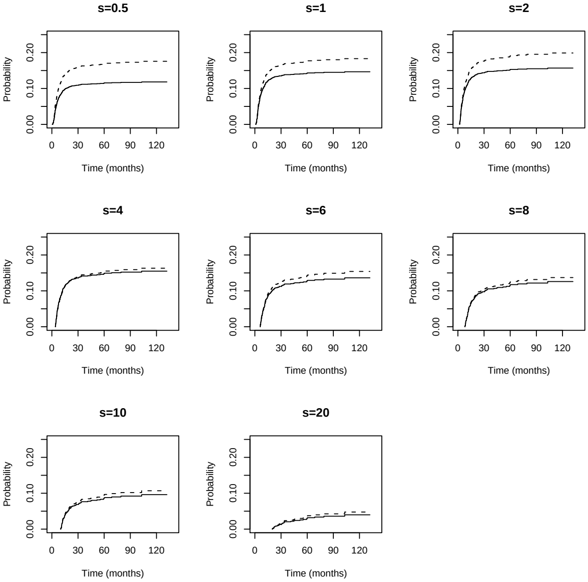
Adopt "landmark analysis" (van Houwelingen, 2007) for the competing risks model:

$$P(T \leq t, Z(T) = 2 \,|\, T \geq s, X(s)),$$

given a number of landmark time points s.

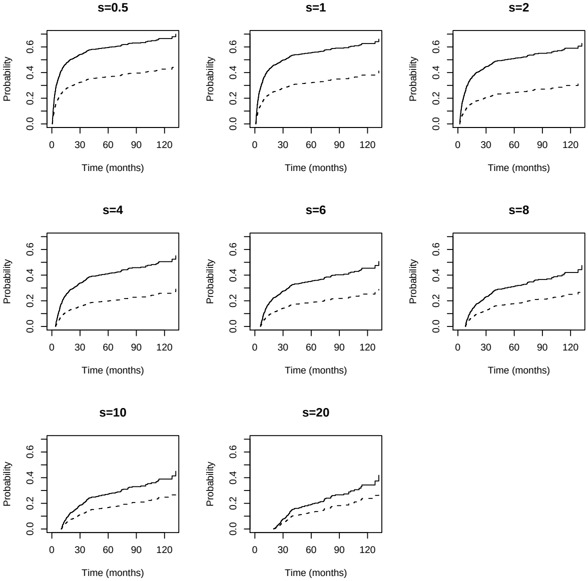
- restricted samples of patients still at risk at each s,
- $X(s)$ time-constant for each s, but it is allowed to vary between different s.
- comparison of $P(\cdot\,|X(s) = 0)$ and $P(\cdot\,|X(s) = 1)$.
- It is also feasible for discrete or continuos internal covariates.

Relapse



Dashed lines: no GvHD at s, **Solid lines**: GvHD is present at s.

Death



Dashed lines: no GvHD at s, **Solid lines**: GvHD is present at s.

Final remarks and Outlook

- Extended competing risks models: only applicable for categorical *internal* covariates.
  Landmark approach is also valid for discrete or continuous covariates.

- Software: R code is available at the webpage `http://homes.stat.unipd.it/gcortese`.

- Use time-varying regression coefficient for $X(t)$.

- Multiplicative-additive hazards model (Scheike & Zhang, 2002)

  $$\alpha_{02}(t) = \alpha_{02}^{(0)}(t) \exp\left\{V^T \gamma\right\} \quad \text{and} \quad \alpha_{12}(t) = \left[\alpha_{02}^{(0)}(t) + \beta(t)\right] \exp\left\{V^T \gamma\right\},$$

  with additional covariates V.

- Aiming at regression in the landmark approach with

  $\alpha(t|s) = \alpha_0(t|s) \exp\left\{X(t)^T \beta_L(t,s)\right\}$ for landmark points s.

References

Cortese G., Andersen PK. (2009). Competing risks and time-dependent covariates, *Biometrical Journal*, In press.

Kalbfleisch JD., Prentice RL. (2002). *The Statistical Analysis of Failure Time Data*, Wiley, New York.

Andersen PK., Borgan Ø., Gill RD., Keiding N. (1993). *Statistical Models Based on Counting Processes*, Springer-Verlag, New York.

Fine JP., Gray RJ. (1999). A Proportional Hazards Model for the Subdistribution of a Competing Risk, *Journal of the American Statistical Association*, **94**:496-509.

van Houwelingen HC. (2007). Dynamic Prediction by Landmarking in Event History Analysis, *Scandinavian Journal of Statistics*, **34**:70-85.

Scheike TH., Zhang M-J. (2002). An additive-multiplicative Cox-Aalen model, *Scandinavian Journal of Statistics*, **28**:1328-1355.