A multistate additive relative survival semi-Markov model

Florence Gillaiseau$^{1,2,3}$ & Etienne Dantan$^1$
& Magali Giral$^{2,3}$ & Yohann Foucher$^{1,3}$

E-mail: florence.gillaizeau@univ-nantes.fr

1 EA 4275 - SPHERE - Biostatistics, Pharmacoepidemiology and Subjective Measures in Health Sciences, Nantes University, Nantes, France.
2 INSERM UMR1064, Institute of Transplantation, Urology and Nephrology, Nantes, France.
3 Nantes University Hospital, Nantes, France.

27 August 2014
Motivating example: kidney transplantation

X = 1
Kidney Transplantation

X = 2
Acute Rejection Episode

X = 3
Return to dialysis

X = E
Death related to the ESRD ('excess death')

X = P
Death not related to the ESRD ('expected death')

ESRD = End Stage Renal Disease
The idea

- multiple time-to-events data (disease progression, death) ⇒ multistate models
- association with excess death associated to the disease ⇒ relative survival analysis *
- literature
  - Belot et al. [2011] : competing risks + relative survival (excess mortality related to colon cancer)
  - Huszti et al. [2012] : Markov NH + relative survival (excess mortality related to colon cancer in an illness-death model)
- Gillaiseau et al. [2014] : semi-Markov additive relative survival model (SMRS)

* Hakulinen and Tenkanen [1987], Esteve et al. [1990], Perme et al. [2012]
Additive relative survival

ESRD=End Stage Renal Disease

X= 1
Kidney Transplantation

X= 2
Acute Rejection Episode

X= 3
Return to dialysis

X= P
Death not related to the ESRD (‘expected death’)

X= E
Death related to the ESRD (‘excess death’)

X= E
Death related to the ESRD (‘excess death’)
Notations

- $T$ : chronological time from baseline
- $S$ : duration (or sojourn time) in a state
- $\mathcal{X}$ : finite space of the possible clinical states
- $\epsilon$ : set of possible transitions $ij$ with $(i,j) \in (\mathcal{X},\mathcal{X})$, with $i$ transient state with $j \neq i$
- $X_m$ : state of the patient after the $m$-th transition occurring at time $T_m$, with $T_0 < T_1 < \ldots < T_m$ ($T_0 = 0$ and $X_0 = 1$)
- $Z$ : overall vector of patient characteristics
- $Z_{ij}$ : subvector of characteristics specifically associated to the transition $ij$
semi-Markovian property

transition intensities between two states depend on the duration in the current state

- instantaneous hazard function specific from state $X_m = i$ to the state $X_{m+1} = j$ after a duration $s$, given patient characteristics $Z_{ij} = z_{ij}$: 

$$
\lambda_{ij}(s|z_{ij}) = \lim_{\Delta s \to 0^+} \frac{P(s \leq T_{m+1} - T_m < s + \Delta s, X_{m+1} = j | T_{m+1} - T_m > s, X_m = i, z_{ij})}{\Delta s}
$$

with $\Lambda_{ij}(s|z_{ij}) = \int_0^s \lambda_{ij}(u|z_{ij}) \, du$ the corresponding cumulative hazard function.
Expected mortality

- $X = E$: death related to the disease
- $X = P$: death related to other causes
- $A$: random variable for patient’s age at death
- $a_i$: patient age observed at entry in state $i$
- $y$: patient’s birthyear
- $g$: patient’s gender

Instantaneous hazard function for the mortality not related to the disease after a duration $s$ in the state $i$:

$$
\lambda_P(s + a_i | y, g) = \lim_{\Delta s \to 0^+} \frac{P(s + a_i \leq A < s + a_i + \Delta s, X = P | A > s + a_i, y, g)}{\Delta s}
$$

$\Rightarrow$ calculated from life tables
   (available by calendar year × birthdate × gender)
Observed mortality

Instantaneous hazard function:

\[ \lambda_{iO}(s|z_{iE}, a_i, y, g) = \lambda_{iE}(s|z_{iE}) + \lambda_P(s + a_i|y, g) \]

Cumulative hazard:

\[ \Lambda_{iO}(s|z_{iE}, a_i, y, g) = \Lambda_{iE}(s|z_{iE}) + \Lambda_P(s + a_i|y, g) - \Lambda_P(a_i|y, g) \]

⇒ \( \Lambda_P(s + a_i|y, g) - \Lambda_P(a_i|y, g) \) represents the cumulative hazard of death between age \( a_i \) and \( a_i + s \) in the general population.
Marginal survival function and subdensity

Probability for a patient to stay at least a duration $s$ in state $i$:

$$S_i(s|z, a_i, y, g) = \exp \left[ - \left( \sum_{j: ij \in e \atop j \neq \text{death}} \Lambda_{ij}(s|z_{ij}) - \Lambda_{iE}(s|z_{iE}) - \Lambda_P(s+a_i|y,g) + \Lambda_P(a_i|y,g) \right) \right]$$

$\Rightarrow$ density function specific to transition $ij$, after a duration $s$:

$$f_{ij}(s|z, a_i, y, g) = \left( 1_{\{j \neq \text{death}\}} \lambda_{ij}(s|z_{ij}) + 1_{\{j = \text{death}\}} \lambda_{iO}(s|z_{iE}, a_i, y, g) \right) S_i(s|z, a_i, y, g)$$
Contribution to the likelihood

- \( s_{ij} \): duration time in state \( i \) before transition to state \( j \)
- \( \delta_{ij} = 1 \) if the transition \( ij \) is observed, \( \delta_{ij} = 0 \) otherwise

Patient in an absorbing state at his/her last time of follow-up

\[
\prod_{ij \in \epsilon} \{f_{ij}(s_{ij}|z, a_i, y, g)\}^{\delta_{ij}}
\]

Patient censored in the transient state \( k \) (for a duration \( s_k \)) at his/her last time of follow-up

\[
\prod_{ij \in \epsilon} \{f_{ij}(s_{ij}|z, a_i, y, g)\}^{\delta_{ij}} S_k(s_k|z, a_k, y, g)
\]

- \( \lambda.(.) \): parametric PH models with time-fixed covariates
- estimations: maximization of the likelihood function + Hessian matrix (Nelder and Mead algorithms)
Performances of 2 models

- SMRS model
- 5-state SM model (causes of death known)

\[ X = 1 \]
Kidney Transplantation

\[ X = 2 \]
Acute Rejection Episode

\[ X = 3 \]
Return to dialysis

\[ X = E \]
Death related to the ESRD ('excess death')

\[ X = P \]
Death not related to the ESRD ('expected death')

ESRD = End Stage Renal Disease
Simulations based on kidney transplant recipients data:

- year of entry into the study $\sim U([1998;2010])$
- gender $g$: men $\sim B(0.61)$
- explicative variable $z$ $\sim B(0.30)$
- age $a$ at baseline $\sim$ truncated $\mathcal{N}$ (from 18 to 80 years old) with parameters varying according to $g$ and $z$

The 5 sojourn time distributions $\sim \mathcal{W}$ depending on $(a, g, z)$.

Scenarios:

- 3 sample sizes ($N=500, N=1000, N=3000$ subjects)
- 3 censoring rates (15%, 30%, 60%)
### Simulation study

#### Objective

Data generation

#### Results

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Theoretical value</th>
<th>Mean estimate</th>
<th>Absolute bias</th>
<th>Coverage rate(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_{12} ) Male</td>
<td>0.160</td>
<td>0.167</td>
<td>0.007</td>
<td>93.67</td>
</tr>
<tr>
<td>( \beta_{12} ) Age</td>
<td>-0.012</td>
<td>-0.012</td>
<td>0.000</td>
<td>96.00</td>
</tr>
<tr>
<td>( \beta_{12} ) ( z )</td>
<td>0.210</td>
<td>0.216</td>
<td>0.006</td>
<td>94.67</td>
</tr>
<tr>
<td>( \beta_{13} ) Male</td>
<td>-0.160</td>
<td>-0.180</td>
<td>-0.020</td>
<td>93.67</td>
</tr>
<tr>
<td>( \beta_{13} ) Age</td>
<td>0.014</td>
<td>0.014</td>
<td>0.000</td>
<td>94.33</td>
</tr>
<tr>
<td>( \beta_{13} ) ( z )</td>
<td>0.910</td>
<td>0.912</td>
<td>0.002</td>
<td>97.00</td>
</tr>
<tr>
<td>( \beta_{1E} ) Male</td>
<td>0.180</td>
<td>0.191</td>
<td>0.011</td>
<td>96.67</td>
</tr>
<tr>
<td>( \beta_{1E} ) Age</td>
<td>-0.050</td>
<td>-0.050</td>
<td>0.000</td>
<td>96.67</td>
</tr>
<tr>
<td>( \beta_{1E} ) ( z )</td>
<td>0.600</td>
<td>0.590</td>
<td>-0.010</td>
<td>94.00</td>
</tr>
<tr>
<td>( \beta_{23} ) Male</td>
<td>-0.420</td>
<td>-0.413</td>
<td>0.007</td>
<td>96.33</td>
</tr>
<tr>
<td>( \beta_{23} ) Age</td>
<td>-0.008</td>
<td>-0.008</td>
<td>0.000</td>
<td>97.00</td>
</tr>
<tr>
<td>( \beta_{23} ) ( z )</td>
<td>0.400</td>
<td>0.408</td>
<td>0.008</td>
<td>96.00</td>
</tr>
<tr>
<td>( \beta_{2E} ) Male</td>
<td>-0.150</td>
<td>-0.122</td>
<td>0.028</td>
<td>96.00</td>
</tr>
<tr>
<td>( \beta_{2E} ) Age</td>
<td>-0.035</td>
<td>-0.035</td>
<td>0.000</td>
<td>92.33</td>
</tr>
<tr>
<td>( \beta_{2E} ) ( z )</td>
<td>0.740</td>
<td>0.748</td>
<td>0.008</td>
<td>94.67</td>
</tr>
</tbody>
</table>
### Variability

300 simulated samples, N=3000 patients, censoring rate=60%

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Theoretical value</th>
<th>RMSE</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>5 SM</td>
<td>SMRS</td>
<td>5 SM</td>
</tr>
<tr>
<td>$\beta_{12}$ Male</td>
<td>0.160</td>
<td>0.077</td>
<td>0.077</td>
<td>0.077</td>
</tr>
<tr>
<td>$\beta_{12}$ Age</td>
<td>-0.012</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
</tr>
<tr>
<td>$\beta_{12}$ z</td>
<td>0.210</td>
<td>0.080</td>
<td>0.080</td>
<td>0.080</td>
</tr>
<tr>
<td>$\beta_{13}$ Male</td>
<td>-0.160</td>
<td>0.158</td>
<td>0.158</td>
<td>0.157</td>
</tr>
<tr>
<td>$\beta_{13}$ Age</td>
<td>0.014</td>
<td>0.007</td>
<td>0.006</td>
<td>0.006</td>
</tr>
<tr>
<td>$\beta_{13}$ z</td>
<td>0.910</td>
<td>0.155</td>
<td>0.155</td>
<td>0.156</td>
</tr>
<tr>
<td>$\beta_{1E}$ Male</td>
<td>0.180</td>
<td>0.170</td>
<td>0.235</td>
<td>0.170</td>
</tr>
<tr>
<td>$\beta_{1E}$ Age</td>
<td>-0.050</td>
<td>0.007</td>
<td>0.010</td>
<td>0.007</td>
</tr>
<tr>
<td>$\beta_{1E}$ z</td>
<td>0.600</td>
<td>0.181</td>
<td>0.235</td>
<td>0.181</td>
</tr>
<tr>
<td>$\beta_{23}$ Male</td>
<td>-0.420</td>
<td>0.224</td>
<td>0.224</td>
<td>0.224</td>
</tr>
<tr>
<td>$\beta_{23}$ Age</td>
<td>-0.008</td>
<td>0.009</td>
<td>0.009</td>
<td>0.009</td>
</tr>
<tr>
<td>$\beta_{23}$ z</td>
<td>0.400</td>
<td>0.232</td>
<td>0.232</td>
<td>0.233</td>
</tr>
<tr>
<td>$\beta_{2E}$ Male</td>
<td>-0.150</td>
<td>0.229</td>
<td>0.283</td>
<td>0.228</td>
</tr>
<tr>
<td>$\beta_{2E}$ Age</td>
<td>-0.035</td>
<td>0.010</td>
<td>0.013</td>
<td>0.010</td>
</tr>
<tr>
<td>$\beta_{2E}$ z</td>
<td>0.740</td>
<td>0.251</td>
<td>0.293</td>
<td>0.251</td>
</tr>
</tbody>
</table>
Conclusion

- Good performances of the SMRS model
  - as good as the SM model where the causes of death are known
  - similar results for other simulation scenarios

- Application to data from kidney transplant recipients (DIVAT cohort, N=5943)

<table>
<thead>
<tr>
<th>Model</th>
<th>Transition</th>
<th>Coefficient</th>
<th>HR</th>
<th>[95%CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMRS model</td>
<td>Age&lt;35 years</td>
<td>0.06</td>
<td>[0.01;0.31]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age 35 to 50 years</td>
<td>0.34</td>
<td>[0.19;0.61]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age 50 to 65 years</td>
<td>0.55</td>
<td>[0.32;0.93]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Male recipient</td>
<td>0.78</td>
<td>[0.53;1.16]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Delayed Graft Function</td>
<td>3.02</td>
<td>[1.96;4.64]</td>
<td></td>
</tr>
<tr>
<td>4-state SM</td>
<td>Age&lt;35 years</td>
<td>0.06</td>
<td>[0.03;0.13]</td>
<td></td>
</tr>
<tr>
<td>model</td>
<td>Age 35 to 50 years</td>
<td>0.22</td>
<td>[0.15;0.31]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age 50 to 65 years</td>
<td>0.45</td>
<td>[0.33;0.60]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Male recipient</td>
<td>1.14</td>
<td>[0.89;1.48]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Delayed Graft Function</td>
<td>1.93</td>
<td>[1.52;2.44]</td>
<td></td>
</tr>
</tbody>
</table>

(transplantation to death related to the disease)
Use and perspectives

- Package R : eSemiMarkov (www.divat.fr)
- Model needs extensions :
  - time-dependent variables
  - non-proportional hazards
  - interval-censored data


---

**Thanks for your attention**