A multistate additive relative survival semi-Markov model

Florence Gillaizeau[*,†,§] & Etienne Dantan[*] & Magali Giral[†,§] & Yohann Foucher[*,§]

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

[*] *EA 4275 - SPHERE - Biostatistique, Pharmacoépidémiologie et Mesures Subjectives en Santé, Université de Nantes, Nantes, France.*
[†] *INSERM UMR1064, Institut Transplantation-Urologie-Néphrologie, Nantes, France.*
[§] *Centre Hospitalier Universitaire de Nantes, Nantes, France.*

3 June 2014

Context

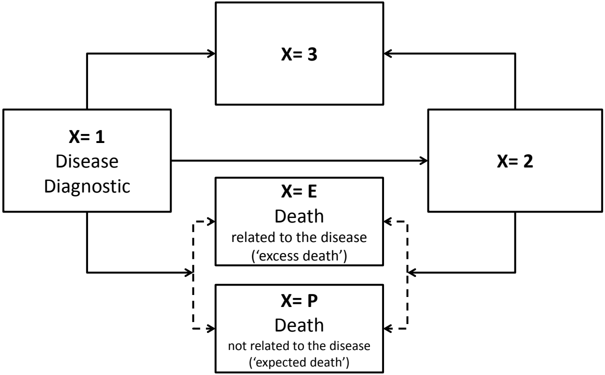
- multiple time-to-events data (disease progression, death) ⇒ multistate models
- association with excess death associated to the disease ⇒ relative survival analysis *
- litterature
  - Belot et al. [2011] : competing risks + relative survival (excess mortality related to colon cancer)
  - Huszti et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

⇒ Gillaizeau et al. [2014] : **semi-Markov additive relative survival model (SMRS)**

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

Representation

FIGURE 1: The SMRS representation for a multistate model including the death related to the disease ($X = E$) and the death not related to the disease ($X = P$). Arrows for the transitions to $X = E$ and $X = P$ are represented with dashed lines since the two states can not be distinguished individually.



Notations

- T : chronological time from baseline
- S : duration (or sojourn time) in a state
- $\mathscr{X}$: finite space of the possible clinical states
- ϵ : set of possible transitions ij with $(i,j) \in (\mathscr{X},\mathscr{X})$, 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

Instantaneous hazard function

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 $Z_{ij} = z_{ij}$:

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

with $\Lambda_{ij}(s|z_{ij}) = \int_0^s \lambda_{ij}(u|z_{ij})\mathrm{d}u$ 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 :

$$\lambda_P(s+a_i|y,g) = \lim_{\Delta s \to 0^+} \frac{P(s+a_i \leq A < s+a_i+\Delta s, X = P | A > s+a_i, y, g)}{\Delta s} \tag{2}$$

$\Rightarrow$ calculated from life tables
(available by calendar year $\times$ birthdate $\times$ gender)

Observed mortality

Instantaneous hazard function :

$$\lambda_{iO}(s|z_{iE}, a_i, y, g) = \lambda_{iE}(s|z_{iE}) + \lambda_P(s + a_i|y, g) \quad (3)$$

Cumulative hazard :

$$\Lambda_{iO}(s|z_{iE}, a_i, y, g) = \Lambda_{iE}(s|z_{iE}) + \Lambda_P(s + a_i|y, g) - \Lambda_P(a_i|y, g) \quad (4)$$

$\Rightarrow \Lambda_P(s + a_i|y, g) - \Lambda_P(a_i|y, g)$ 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 sejourn at least a duration s in state i :

$$S_{i.}(s|z,a_i,y,g) = exp\left[-\left(\sum_{\substack{j:ij\in\epsilon \\ j\neq \text{death}}} \Lambda_{ij}(s|z_{ij})\right) - \Lambda_{iE}(s|z_{iE}) - \Lambda_P(s+a_i|y,g) + \Lambda_P(a_i|y,g)\right] \quad (5)$$

equations (1) + (3) + (5)

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

$$f_{ij}(s|z,a_i,y,g) = \left(\mathbb{1}_{\{j\neq \text{death}\}}\lambda_{ij}(s|z_{ij}) + \mathbb{1}_{\{j=\text{death}\}}\lambda_{iO}(s|z_{iE},a_i,y,g)\right) S_{i.}(s|z,a_i,y,g) \quad (6)$$

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

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

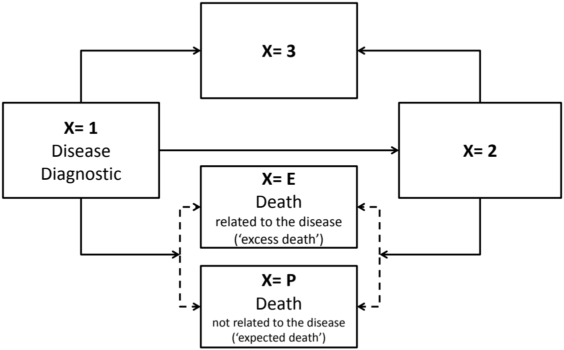
Patient censored in the transient state k (for a duration s_k) at his/her last time of follow-up

$$\prod_{ij \in \epsilon} \{f_{ij}(s_{ij}|z, a_i, y, g)\}^{\delta_{ij}} S_{k.}(s_k|z, a_k, y, g)$$

- $\lambda_{.}(.)$: 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)



Distributions

Simulations based on kidney transplant recipients data :

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

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

Scenarios :

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

5-state SM model

TABLE 1: Estimations of effects with the 5-state SM model (100 simulated samples, N=3000 patients, censoring rate=60%)

Coefficient	Theoretical value[1]	Mean estimate	Absolute bias	RMSE	Empiric SE	Asymptotic SE	Coverage rate (%)
β_{12} Male	0.160	0.165	0.005	0.08	0.08	0.07	94
β_{12} Age	-0.012	-0.012	0.000	0.00	0.00	0.00	97
β_{12} z	0.210	0.211	0.001	0.08	0.08	0.08	94
β_{13} Male	-0.160	-0.194	-0.034	0.15	0.14	0.15	95
β_{13} Age	0.014	0.014	0.000	0.01	0.01	0.01	95
β_{13} z	0.910	0.902	-0.008	0.15	0.15	0.15	99
β_{1E} Male	0.180	0.197	0.014	0.19	0.19	0.18	94
β_{1E} Age	-0.050	-0.050	0.000	0.01	0.01	0.01	96
β_{1E} z	0.600	0.609	0.009	0.20	0.20	0.18	92
β_{23} Male	-0.420	-0.407	0.013	0.24	0.24	0.22	93
β_{23} Age	-0.008	-0.007	0.001	0.01	0.01	0.01	97
β_{23} z	0.400	0.420	0.020	0.24	0.24	0.23	95
β_{2E} Male	-0.150	-0.112	0.038	0.24	0.24	0.23	95
β_{2E} Age	-0.035	-0.035	0.000	0.01	0.01	0.01	91
β_{2E} z	0.740	0.756	0.016	0.26	0.26	0.23	94

RMSE : Root Mean Square Error, SE : Standard Error

[1] Theoretical values for the baseline hazard functions with Weibull distribution (log(scale), log(shape)) : transition 12 (2.5,1.2), transition 13 (5.0,-0.3), transition $1E$ (1.1,0.2), transition 23 (2.8,-0.4), transition $2E$ (1.3,-0.1), right-censoring (0.6,-0.2).
Theoretical values for effects on expected death : β_P Gender=0.4, β_P Age=0.02.

SMRS model

TABLE 2: Estimations of effects with the SMRS model (100 simulated samples, N=3000 patients, censoring rate=60%)

Coefficient	Theoretical value[1]	Mean estimate	Absolute bias	RMSE	Empiric SE	Asymptotic SE	Coverage rate (%)
β_{12} Male	0.160	0.165	0.005	0.08	0.08	0.07	94
β_{12} Age	-0.012	-0.012	0.000	0.00	0.00	0.00	97
β_{12} z	0.210	0.211	0.001	0.08	0.08	0.08	94
β_{13} Male	-0.160	-0.193	-0.033	0.15	0.14	0.15	95
β_{13} Age	0.014	0.014	0.000	0.01	0.01	0.01	95
β_{13} z	0.910	0.902	-0.008	0.15	0.15	0.15	98
β_{1E} Male	0.180	0.205	0.025	0.23	0.23	0.23	95
β_{1E} Age	-0.050	-0.049	0.001	0.01	0.01	0.01	95
β_{1E} z	0.600	0.609	0.009	0.26	0.26	0.23	96
β_{23} Male	-0.420	-0.407	0.013	0.24	0.24	0.22	93
β_{23} Age	-0.008	-0.007	0.001	0.01	0.01	0.01	97
β_{23} z	0.400	0.420	0.020	0.24	0.24	0.23	95
β_{2E} Male	-0.150	-0.108	0.042	0.28	0.28	0.28	97
β_{2E} Age	-0.035	-0.035	0.000	0.01	0.01	0.01	94
β_{2E} z	0.740	0.770	0.030	0.32	0.32	0.28	92

RMSE : Root Mean Square Error, SE : Standard Error

[1] Theoretical values for the baseline hazard functions with Weibull distribution (log(scale), log(shape)) : transition 12 (2.5,1.2), transition 13 (5.0,-0.3), transition 1E (1.1,0.2), transition 23 (2.8,-0.4), transition 2E (1.3,-0.1), right-censoring (0.6,-0.2).
Theoretical values for effects on expected death : β_P Gender=0.4, β_P Age=0.02.

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
- Model needs extensions :
  - time-dependent variables
  - non-proportional hazards
  - interval-censored data
- Package R : eSemiMarkov (`www.divat.fr`)
- Application to data from kidney transplant recipients (DIVAT cohort)

Application data (goodness-of-fit)

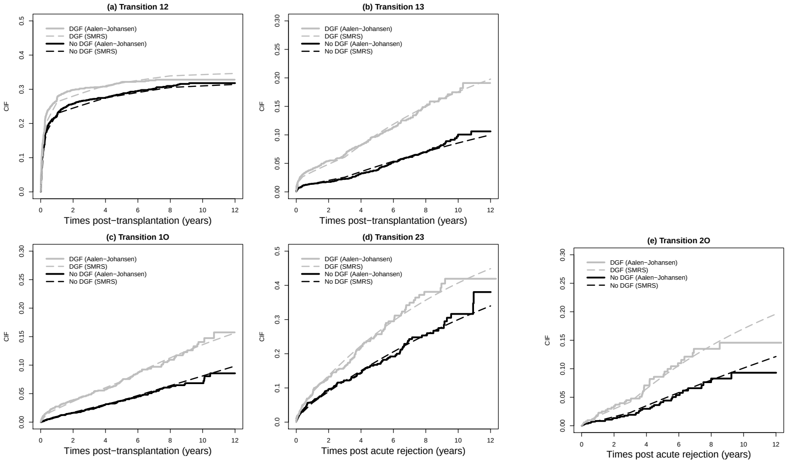


Figure 2: CIF estimates provided by the SMRS model (dashed lines) and the Aalen-Johansen nonparametric estimator (solid lines) on data from kidney transplant recipients for the 5 possible transitions and according to the presence of an explicative variable.

References

T. Hakulinen and L. Tenkanen. Regression analysis of relative survival rates.
J R Stat Soc Ser C Appl Stat, 36 :309–317, 1987

J. Esteve, E. Benhamou, M. Croasdale, and L. Raymond. R : elements for further discussion.
Stat Med, 9(5) :529–538, May 1990

M. P. Perme, J. Stare, and J. Estève. On estimation in relative survival.
Biometrics, 68(1) :113–120, March 2012

A. Belot, L. Remontet, G. Launoy, V. Jooste, and R. Giorgi. Competing risk models to estimate the excess mortality and the first recurrent-event hazards.
BMC Med Res Methodol, 11 :78, 2011

E. Huszti, M. Abrahamowicz, A. Alioum, C. Binquet, and C. Quantin. Relative survival multistate markov model.
Stat Med, 31(3) :269–286, February 2012

Merci de votre attention