A SEMI-MARKOV MODEL FOR INTERVAL-CENSORED DATA

ANALYSIS OF THE EVOLUTION OF KIDNEY TRANSPLANT RECIPIENTS

Y. Foucher, M. Giral, JP. Soulillou, JP. Daures

yohann.foucher@iurc.montp.inserm.fr
University of Montpellier 1, France

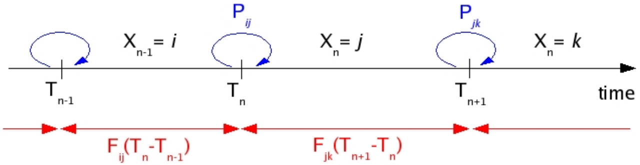
ISCB - Geneva 2006

Outline

Introduction

- Multistate approaches are becoming increasingly used for the analysis of longitudinal data.
- Semi-Markov models explicitly define distributions of waiting times.
- In the follow-up of patients, transition times are known to have occurred in some interval.
- Objective : The development of a flexible semi-Markov model which allow for interval censoring.

Definitions (1)



- $\longrightarrow$ Probability of jumping from the State *i* to the State *j*.
- $\longleftrightarrow$ Staying times $T_{n+1} - T_n \rightsquigarrow F_{ij}(T_{n+1} - T_n)$.

Definitions (2)

Embedded Markov chain

$$P_{ij} = P(X_{n+1} = j | X_n = i)$$

- If state i is not persistent then $P_{ij} \geq 0$ and $P_{ii} = 0$.
- If state i is persistent then $P_{ij} = 0$ and $P_{ii} = 1$.

Definitions (2)

Embedded Markov chain

$$P_{ij} = P(X_{n+1} = j | X_n = i)$$

- If state i is not persistent then $P_{ij} \geq 0$ and $P_{ii} = 0$.
- If state i is persistent then $P_{ij} = 0$ and $P_{ii} = 1$.

Distribution of waiting times

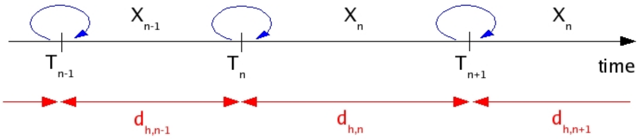
- $F_{ij}(x) = P(T_{n+1} - T_n \leq x | X_{n+1} = j, X_n = i)$.
- $F_{ij}(x) = F^{(ij)}(x, \varphi_{ij})$.

$$\Longrightarrow S_{ij}(x),\ f_{ij}(x) \text{ et } \lambda_{ij}(x)$$

Loglikelihood (1)

Contribution of a transition exactly observed $\delta^E_{h,r}$

Let $d_{h,r} = T_{h,r+1} - T_{h,r}$, the waiting time in the state $X_{h,r}$ before jumping to the state $X_{h,r+1}$.

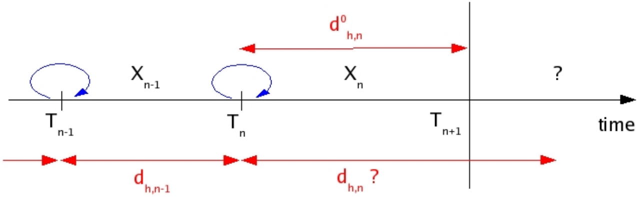


$$
\begin{aligned}
& lim_{\Delta d \to 0^+} \frac{P(d_{h,r} < x < d_{h,r} + \Delta d, X_{h,r+1} = j | X_{h,r} = i)}{\Delta d} \\
= \quad & lim_{\Delta d \to 0^+} \frac{P(d_{h,r} < x < d_{h,r} + \Delta d | X_{h,r+1} = j, X_{h,r} = i)}{\Delta d} \\
& \times P(X_{h,r+1} = j | X_{h,r} = i) = P_{ij} f_{ij}(d_{h,r})
\end{aligned}
$$

Loglikelihood (2)

Contribution of a right-censored transition $\delta^R_{h,r}$

Let $d^0_{h,r}$ be a value such that if $d^0_{h,r} < d_{h,r}$ then $d^0_{h,r}$ is observed and $d_{h,r}$ is not.

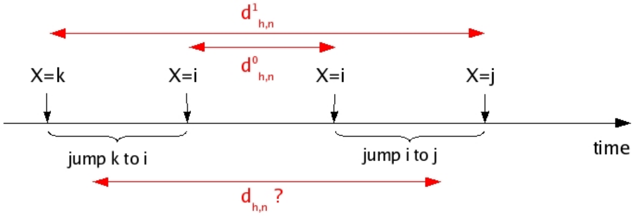


$$
\begin{aligned}
& P(d_{h,r} > d^0_{h,r} | X_{h,r} = i) \\
= & \sum_{j \neq i} P(X_{h,r+1} = j | X_{h,r} = i) P(d_{h,r} > d^0_{h,r} | X_{h,r+1} = j, X_{h,r} = i) \\
= & \sum_{j \neq i} P_{ij} \int_{d^0_{h,r}}^{\infty} f_{ij}(u) du = \sum_{j \neq i} P_{ij} S_{ij}(d^0_{h,r})
\end{aligned}
$$

Loglikelihood (3)

Contribution of a interval-censored transition $\delta^{I}_{h,r}$

Let $d^1_{h,r}$ be a value such that if $d^1_{h,r} > d_{h,r}$ then $d^1_{h,r}$ is observed and $d_{h,r}$ is not.



$$
\begin{aligned}
P(d^0_{h,r} < \quad x \quad < d^1_{h,r}, X_{h,r+1} = j | X_{h,r} = i) \\
&= P(X_{h,r+1} = i | X_{h,r} = i) \int_{d^0_{h,r}}^{d^1_{h,r}} f_{ij}(u)du \\
&= P_{ij}\left(\int_0^{d^1_{h,r}} f_{ij}(u)du - \int_0^{d^0_{h,r}} f_{ij}(u)du\right) \\
&= P_{ij}(F_{ij}(d^1_{h,r}) - F_{ij}(d^0_{h,r})) = P_{ij}(S_{ij}(d^0_{h,r}) - S_{ij}(d^1_{h,r}))
\end{aligned}
$$

Loglikelihood (4)

Contribution of an initial observation for the subject h

By defining $z_{h,0j}$, the vector of covariates associated with the initial state j for the h^{th} subject, the usual multinomial logistic regression can be written as :

$$P(X_{h,1} = j) = \frac{exp(\gamma_{0j} + \beta_{0j} z_{h,0j})}{\sum_{k=1}^{c} exp(\gamma_{0k} + \beta_{0k} z_{h,0k})} \quad \text{for } j = 1, .., c$$

with $\gamma_{0c} = 0$ and $\beta_{0c} = 0$, in order to obtain $\sum_{j=1}^{c} \pi_{0j} = 1$.

Loglikelihood (5)

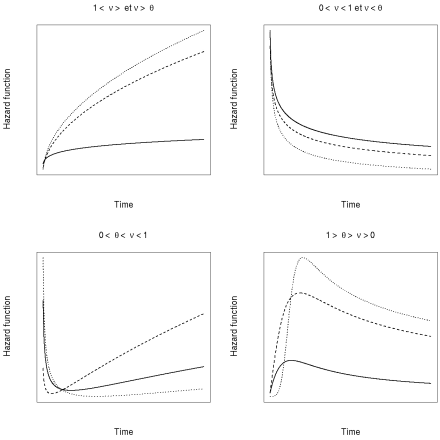
$$
\begin{aligned}
ln\mathcal{L} \;=\; & \sum_h \Bigg\{ \gamma_{0X_{h,1}} + \beta_{0X_{h,1}} z_{h,0X_{h,1}} - ln(\sum_{i=1}^{c} exp(\gamma_{0i} + \beta_{0i} z_{h,0X_{h,1}})) \\
& + \sum_{ij} \sum_{X_{h,r}=i, X_{h,r+1}=j} \Bigg\{ \delta^{E}_{h,r} \left[ln\, P_{ij} + ln\, S_{ij}(d_{h,r}) + ln\, \lambda_{ij}(d_{h,r}) \right] \\
& + \delta^{I}_{h,r} \left[ln\, P_{ij} + ln(S_{ij}(d^0_{h,r}) - S_{ij}(d^1_{h,r})) \right] \Bigg\} \\
& + \sum_i \sum_{X_{h,r}=i} \Bigg\{ \delta^{R}_{h,r} \left[ln(\sum_{j \neq i} P_{ij} S_{ij}(d^0_{h,r})) \right] \Bigg\} \Bigg\}
\end{aligned}
$$

where $\gamma_{0c} = \beta_{0c} = 0$.

Modelling assumptions (1)

Generalised Weibull distribution ($\nu_{ij}, \sigma_{ij}, \theta_{ij} > 0$)

- Hazard, $\lambda_{ij}(x) = \frac{1}{\theta_{ij}}\left(1 + (\frac{x}{\sigma_{ij}})^{\nu_{ij}}\right)^{\frac{1}{\theta_{ij}}-1} \frac{\nu_{ij}}{\sigma_{ij}}\left(\frac{x}{\sigma_{ij}}\right)^{\nu_{ij}-1}$
- Survival, $S_{ij}(x) = exp\left(1 - (1 + (\frac{x}{\sigma_{ij}})^{\nu_{ij}})^{\frac{1}{\theta_{ij}}}\right)$



Modelling assumptions (2)

Incorporation of covariates (PH)

- Proportional Hazard (PH) assumption.

$$S_{ij}(x, \eta_{h,ij}) = S_{0,ij}(x)^{exp(\eta_{h,ij})}$$

$$\lambda_{ij}(x, \eta_{h,ij}) = \lambda_{0,ij}(x) exp(\eta_{h,ij})$$

- Respect of the PH assumption.

  plotting $log(-log(S_{ij}(x)))$ against the survival time x.

Kidney transplant recipients (1)

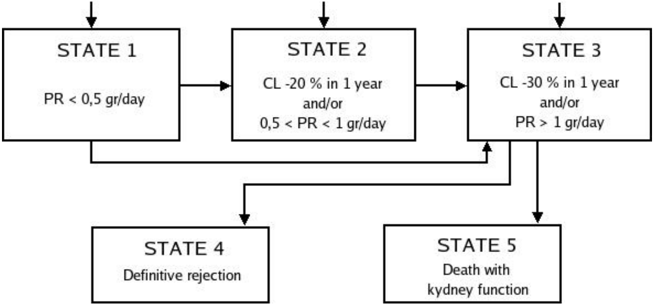
Data description

- Prospective study of kidney transplant recipients (DIVAT).
- 997 patients and 1980 exact or censored transitions.
- Data were computerized at each checkup visit.
- 5 explanatory variables have been retained :
  - gender (men = 1 ; women = 0),
  - cold ischemia time (1 if $\geq$ 16 hours and 0 otherwise),
  - year of the transplantation (1 if $<$ 1998 and 0 otherwise),
  - recipient age at the time of transplantation (1 if $\geq$ 55 years of age and 0 otherwise),
  - delayed graft function (1 if $\geq$ 6 days and 0 otherwise).

Kidney transplant recipients (2)

Multistate structure

- 3-gravity states with two markers :
  - Creatinine clearance (CL)
  - Proteinuria (PR)
- 2-terminal states : chronic rejection of the kidney and death of the patient.



Results (1)

Covariates associated with the initial probabilities

Transition	Covariate	Estim.	SE	p-value
$0 \rightarrow 1$	Intercept	2.85	0.19	0.0001
$0 \rightarrow 1$	Recipient Gender	-0.39	0.17	0.0226
$0 \rightarrow 1$	Delayed graft function	-0.53	0.17	0.0014
$0 \rightarrow 2$	Intercept	-0.67	0.44	0.1258
$0 \rightarrow 2$	Cold ischemia time	1.13	0.44	0.0092

Results (2)

Covariates associated with the intensities of transition

Transition	Covariate	Estim.	SE	RR	p-value
$1 \rightarrow 2$	Year of transplant	-0.80	0.12	0.45	0.0001
$1 \rightarrow 3$	Recipient Gender	0.29	0.15	1.34	0.0484
$1 \rightarrow 3$	Year of transplant	-1.20	0.21	0.30	0.0001
$2 \rightarrow 3$	Year of transplant	-0.54	0.12	0.59	0.0001
$3 \rightarrow 5$	Recipient age	1.48	0.39	4.41	0.0001

Results (3)

Parameters of the waiting times distributions

Transition	σ_{ij} Estim.	σ_{ij} ET	ν_{ij} Estim.	ν_{ij} ET	θ_{ij} Estim.	θ_{ij} ET
$1 \rightarrow 2$	36.14	31.97	0.53	0.03	0.24	0.09
$1 \rightarrow 3$	34.11	65.20	0.52	0.05	0.19	0.15
$2 \rightarrow 3$	33.40	31.34	0.56	0.03	0.30	0.13
$3 \rightarrow 4$	10.16	1.56	1.49	0.11	.	.
$3 \rightarrow 5$	18.48	47.62	1.14	0.23	1.46	3.75

Concluding remarks

Summary of the results

- Multinomial logistic regression usefull in order to identify covariates associated with the initial probabilities.
- Parcimony of the generalized Weibull distribution ($\bigcup -$ or $\bigcap -$ shape).

Concluding remarks

Summary of the results

- Multinomial logistic regression usefull in order to identify covariates associated with the initial probabilities.
- Parcimony of the generalized Weibull distribution ($\bigcup -$ or $\bigcap -$ shape).

Limits of the model and work in progress

- Delate the transition $1 \rightarrow 3$, even if this transition is informative for clinicians.

$$P_{12}P_{23}\int_0^{d_{h,r}} f_{12}(x)f_{23}(d_{h,r}-x)dx$$

- Estimate the cut-off of the markers in order to determine the best states of gravity.
- Construction of an hidden semi-Markov model in order to take into account the short-term fluctuation.