Statistical Methods in Medical Research 2010; **19**: 127–145

A flexible semi-Markov model for interval-censored data and goodness-of-fit testing

Y Foucher Institute for Transplantation and Research in Transplantation and INSERM U643. 30 bd. Jean Monnet, Nantes 44093, France and Department of Biostatistics, Clinical Research Institute, University of Montpellier, 641 av. du doyen G. Giraud, Montpellier 34093, France, **M Giral, JP Soulillou** Institute for Transplantation and Research in Transplantation and INSERM U643. 30 bd. Jean Monnet, Nantes 44093, France and **JP Daures** Department of Biostatistics, Clinical Research Institute, University of Montpellier, 641 av. du doyen G. Giraud, Montpellier 34093, France

Multi-state approaches are becoming increasingly popular to analyse the complex evolution of patients with chronic diseases. For example, the evolution of kidney transplant recipients can be broken down into several clinical states. With this application in mind, we present a flexible semi-Markov model. The distribution functions are fitted to the durations in states and the relevance of the generalised Weibull distribution is shown. The corresponding likelihood function allows for interval censoring, i.e. the times of transitions and the sequences of states are not available during the elapsed times between two visits. The explanatory variables are introduced through the Markov chain and through the probability density functions of durations. A goodness-of-fit test is also defined to examine the stationarity of the semi-Markov model.

1 Introduction

In longitudinal analyses, multi-state models are becoming increasingly popular to deal with the complex evolution of chronic diseases. The use of the Markov chain appears to be very useful for this purpose.[1–3] Another approach consists of modelling the probabilities of transitions according to the times spent in different states, the Markov chain being only associated with the sequence of states.[4] These semi-Markov models (SMM) are adapted to certain diseases, particularly to transplantation.[5] In this paper, we use this approach to study the evolution of kidney transplant recipients (a cohort followed up at Nantes University Hospital in France).

The first methodology issue concerns the use of a continuous marker in order to define the states of disease severity of a patient during his/her follow-up. This marker is available at certain visits, then the transitions between these transient states are interval-censored. This means that the transition times and the sequence of states are unknown during the elapsed time between two consecutive visits. This type of data is often encountered in longitudinal studies. The time of entry into the study and the time of the final event (for example the date of death) are exactly known; contrary to the intermediate

Address for correspondence: Y Foucher, Institute for Transplantation and Research in Transplantation and INSERM U643. 30 bd. Jean Monnet, Nantes 44093, France. E-mail: Yohann.Foucher@univ-nantes.fr

10.1177/0962280208093889

times of transitions, which are interval censored because of the irregularity of the observation process. In this paper, we thus chose to focus the modelling on multi-state data where only intermediate and transient states are interval censored.

Sternberg and Satten[6] proposed non-parametric estimators based on an EM algorithm and showed that a SMM can still be applied to interval-censored data with a unidirectional model without covariates. From a practical standpoint, parametric models are appealing since the parameters can be estimated at a $\sqrt{n}$ rate, leading to a smaller sample size to achieve a given accuracy. Furthermore, the parameters are often interpretable and the covariates can easily be included. As noted by Lindsey,[7] parametric regression models in the presence of heavily interval-censored data are robust. In this paper, we define a SMM based on the generalised Weibull distribution. Convolution products are used to adequately deal with interval-censored transitions, contrary to the recent paper by Foucher *et al.*,[8] in which the interval-censoring concerns only the duration in a given state without considering the censoring of the sequence of consecutive states. Kang and Lagakos[9] also insist on the necessity to take into account all the information available in such a type of incomplete data, that is to say, both the times of transitions and the nature of the states visited are unknown.

Another extension consists in the introduction of covariates through the probability density functions of the times spent in each state and through the Markov chain associated with the sequence of states. The explanatory factors can then be associated either with the speeds or with the trajectories of the process, contrary to most studies where the covariates are only related to the durations.[4,5,10–12]

The second issue in modelling multi-state processes is the goodness-of-fit. Authors generally grant little attention to this point. Most available tests concern Markov models assuming that, for each subject, the number of visits is fixed and that no explanatory variables are measured.[13] The main assumption of the proposed SMM is the stationarity related to the time since the inclusion of patients. Castelli *et al.*[14] recently proposed a goodness-of-fit test of this stationary, but the authors did not deal with the interval censoring. Based on the recent works on the Markov process by Aguirre and Farewell,[15] we herein define a Pearson-type goodness-of-fit statistic in order to determine whether this stationary assumption of the SMM is valid in the presence of interval censored data. However, their method only considers interval-censored times of transition between transient states. Like the modified goodness-of-fit test proposed by Titman and Sharples[16] for Markov or hidden Markov models, the statistic defined in the present paper takes into account the exact times of transition into absorbing states for the semi-Markov process.

We first present the data on which the model is based in Section 2. The modelling of the resulting multi-state data is considered in Section 3. The test of the stationarity is presented in Section 4 and the methods and results are discussed in Section 5.

2 Data and multi-state structure

Such developments were motivated by the analysis of a French prospective study of kidney transplant recipients, extracted from the DIVAT (Données Informatisées et VAlidées en Transplantation) data bank where biological and clinical data of transplant patients have been recorded since 1990 at Nantes University Hospital. Data were computerised

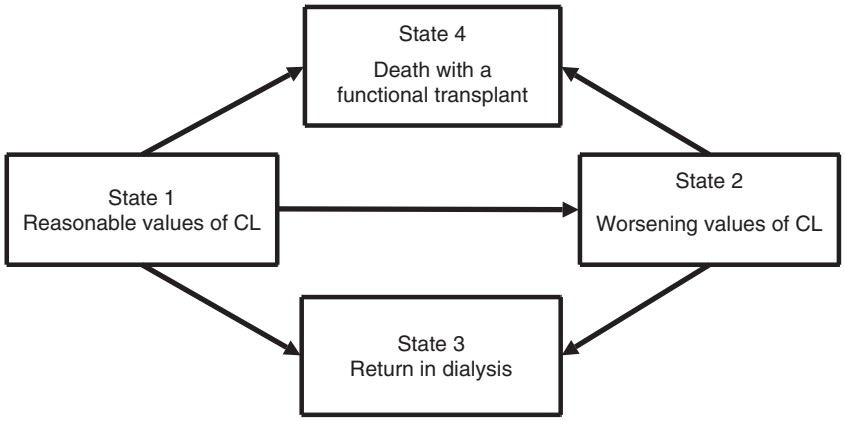

Figure 1 Four-state model for the study of kidney transplant recipient evolution according to a worsening/failure structure.

at each checkup visit and at each anniversary of the transplantation. See the paper by Giral *et al.*[17] for more details. Suppose that the sample consists of n subjects, denoted by h ($h = 1, 2, ..., n$). The follow-up of each patient is a succession of visits at times since the transplantation $\{v_{h,0}, v_{h,1}, ..., v_{h,n_h}\}$. The mean and the median of the number of visits per patient are respectively 30 and 25.

In order to obtain a homogeneous sample, our data set concerns 839 patients older than 18 years and having received a kidney transplant between January 1996 and September 2006. The follow-up started 3 months after the date of transplantation (all patients had a visit at 3 months). The creatinine clearance (CL) is then regarded as stabilised. CL, calculated using the modification of diet in renal disease formula,[18] is recorded at each visit and is used as a marker of the clinical aggravation towards the final events. We consider that all patients enter into an initial state where the risk of failure is low (state 1). For a subject, if her/his value of CL decreases by more than 30% of the maximum observed value, we consider that she/he transits into a state where the risk of failure is high (state 2). Absolute values of CL are not used to classify health-states, since preliminary analyses have shown that it is more informative to consider relative changes.[19] From state 2, no return into state 1 is considered, the medical assumption is of a non-reversible aggravation. Of course, the hypothesis of no backward transition between states 2 and 1 is a substantial assumption. If the state 2 is observed for a patient, the number of the following state will be greater than or equal to 2, whatever the CL level. In other words, we suggest that the recovery of the kidney after such a shock is not possible. However, the possible recovery should constitute a stronger assumption from a clinical point of view.

Two final events are recorded with their exact date of occurrence: the return to dialysis (state 3) and the death of the patient (state 4). In contrast to the transient states of disease severity, state 3 and state 4 are absorbing. The $3 \rightarrow 4$ transition is impossible since the follow-up of patients in the DIVAT data bank is stopped if the patient returns to dialysis. According to this structure, which is presented in Figure 1, the visit v_{h,n_h} relates to a final event or right-censoring with a functional transplant. Table 1 offers some additional

Table 1 Description of the sample according to the possible observed trajectories

Observed states	Size	Percentages	Mean[a]	Median[a]
1	537	64,0 %	3.79	3.48
1; 2	190	22,7 %	4.05	3.63
1; 2; 3	61	7,3 %	3.63	3.61
1; 2; 4	18	2,1 %	2.83	2.51
1; 3	16	1,9 %	1.75	1.39
1; 4	17	2,0 %	2.74	1.18
total	839	100,0 %	3.77	3.44

[a]Mean/Median follow-up times in years.

descriptions of the data set. For instance, 190 patients are right-censored in state 2 with a mean follow-up time since the transplantation of 4 years.

Even though the description of this structure concerns the analysis of kidney transplant recipients, this type of model can also be adapted to other longitudinal data analyses. In many applications, the transient states of disease severity may be interval censored and the date of final failures exactly known.

Nine explanatory variables have been retained, the reference group is specified in the brackets: donor and recipient gender (women), cold ischemia time (<24 h), donor and recipient age at the time of transplantation (<55 years), number of HLA-incompatibilities (<4), induction treatment (no Simulect), panel reactive antibody (0%) and delayed graft function (<6 days). These thresholds were chosen by the clinicians according to the literature.

3 The semi-Markov model

3.1 The semi-Markov framework

Let $\{X_{h,r}, r = 1, \ldots, n_h\}$ be the observed sequence of states at each visit for the hth subject ($h = 1, \ldots, n$), where n_h is the number of visits for this subject. By definition, all patients begin in state 1, i.e. $X_{h,1} = 1$. If the end of the follow-up corresponds to return to dialysis, then $X_{h,n_h} = 3$. If the final event consists in the death of the patient, then $X_{h,n_h} = 4$. Otherwise, $\{X_{h,r}, r = 1, \ldots, n_h\} \in \{1, 2\}$. From $\{X_{h,r}, r = 0, \ldots, n_h\}$, the sequence of the distinct observed states can directly be deduced: $\{W_{h,r}, r = 0, \ldots, m_h\}$ where m_h is the number of observed transitions. By definition, $W_{h,0} = 1$. This sequence forms a Markov chain. Assuming the stationarity of the process with the time since the transplantation, the probabilities of jumping from state $W_{h,r} = i$ to state $W_{h,r+1} = j$, associated with this chain, can be written as:

$$P_{ij} = P(W_{h,r+1} = j | W_{h,r} = i) \tag{1}$$

with the constraint

$$\sum_j P_{ij} = 1. \tag{2}$$

If state i is not persistent, then $P_{ij} \geq 0$ for $i \neq j$ and $P_{ij} = 0$ for $i = j$. Otherwise, if state i is a final event, then $P_{ij} = 0$ for $i \neq j$ and $P_{ij} = 1$ for $i = j$. The semi-Markov property is assumed for the process, that is to say, time is reset to zero at each transition. The probability density function (PDF) of the duration in state $W_{h,r} = i$, before jumping to state $W_{h,r+1} = j$ (for $i \neq j$), is given by:

$$f_{ij}(d_{h,r}) = \lim_{\Delta d \to 0^+} P(d_{h,r} < D_{h,r} < d_{h,r} + \Delta d | W_{h,r+1} = j, W_{h,r} = i)/\Delta d \tag{3}$$

in which $D_{h,r}$ is the time spent in state $W_{h,r}$. As is usual in survival analysis, we deduce from f_{ij} the corresponding survival, hazard and cumulative hazard functions, S_{ij}, λ_{ij} and Λ_{ij} respectively. One can notice that, for a particular multi-state structure in which there is one initial state and a few absorbing states (without transient states), the proposed SMM model generalises the traditional competing risk model (CRM). Indeed, for this type of CRM, the density function specific to the transition $i \to j$ is based on the following joined probability : $\lim_{\Delta d \to 0^+} P(d_{h,r} < D_{h,r} < d_{h,r} + \Delta d, W_{h,r+1} = j | W_{h,r} = i)/\Delta d$. This quantity is equal to $P_{ij} f_{ij}(d_{h,r})$. Our separate modelling of the trajectories (P_{ij}) and of the times of transitions (f_{ij}) provides a high degree of flexibility and is very convenient for result interpretations. For instance, P_{ij} directly represents the prediction of the proportion of transition $i \to j$ if all patients were followed-up until their final failure. Moreover, and as we will see later, it is possible to incorporate covariates either through P_{ij} or through f_{ij}.

We based our modelling strategy on the generalised Weibull distribution with

$$\lambda_{ij}(d_{h,r}) = \theta_{ij}^{-1}(1 + (d_{h,r}/\sigma_{ij})^{\nu_{ij}})^{1/\theta_{ij}-1}(\nu_{ij}/\sigma_{ij})(d_{h,r}/\sigma_{ij})^{\nu_{ij}-1} \tag{4}$$

for all $d_{h,r} \geq 0$, $\nu_{ij} > 0$, $\sigma_{ij} > 0$ and $\theta_{ij} > 0$. This class of distribution has interesting properties.[20] Depending on the parameter values, the hazard function can be either constant, monotone (increasing or decreasing), $\cup$−shaped or $\cap$−shaped. For the non-monotonic functions, it is interesting to calculate the times corresponding to the minimum or the maximum hazard rate. If $0 < \theta_{ij} < \nu_{ij} < 1$, then the hazard function of the $i \to j$ transition decreases from ∞ to its minimum value at the time

$$c_{ij} = \sigma_{ij}\left(\frac{\theta_{ij} - \nu_{ij}\theta_{ij}}{\nu_{ij} - \theta_{ij}}\right)^{1/\nu_{ij}} \tag{5}$$

and then increases to ∞, it is then $\cup$−shaped. If $\theta_{ij} > \nu_{ij} > 1$, then the hazard rate increases from 0 to its maximum value at the time c_{ij} and then decreases to 0, it is then $\cap$−shaped. We can easily simplify this distribution using the likelihood ratio statistic (LRS). For example, if θ_{ij} is fixed at 1, the Weibull formulation is obtained. Moreover, if ν_{ij} equals 1, the hazard function is constant (exponential distribution).

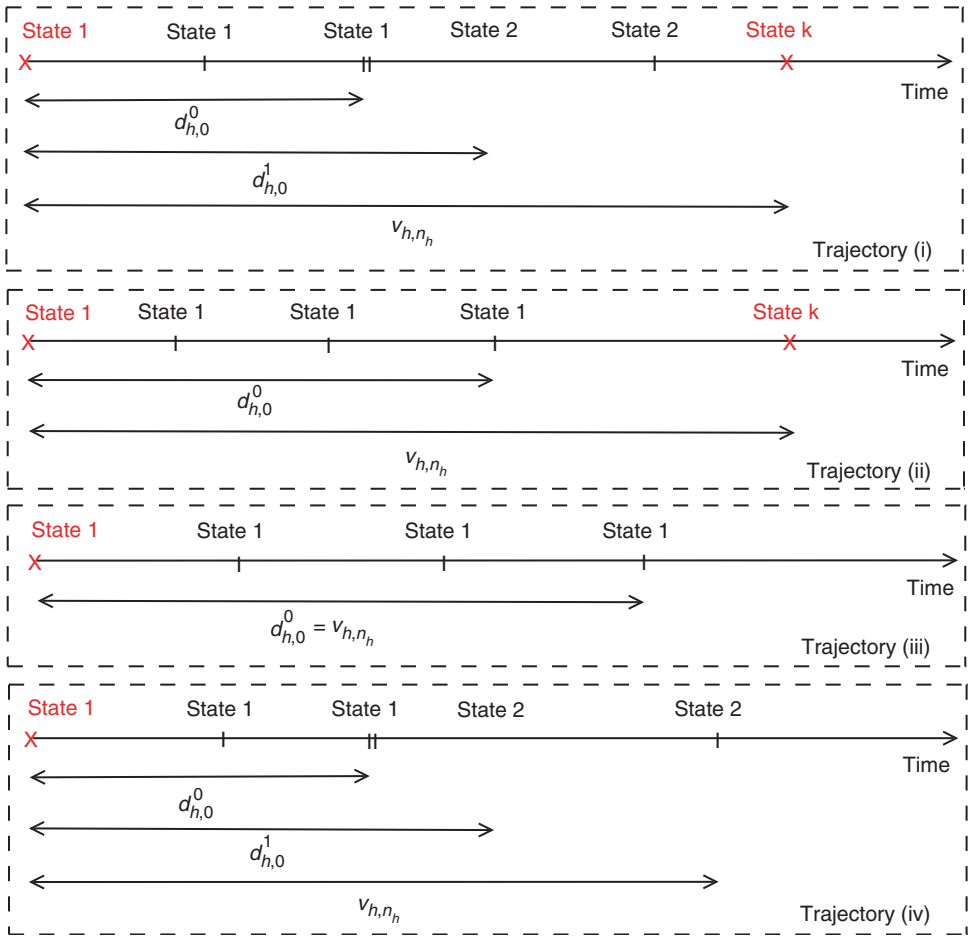

Figure 2 Possible trajectories of a kidney transplant recipient. | | | Observed states of disease severity during visits; X X X Events for which the time of entrance since the transplantation is exactly known.

3.2 Likelihood function

In order to compute the likelihood function, we need to identify the different observed trajectories according to Figure 1. Such trajectories are described in Figure 2. Demonstrations of the following developments are included in the Appendix.

Trajectory (i) – The individual h is observed in both states of disease severity before the occurrence of a final event k ($k = 3, 4$). The duration in the first state, $d_{h,0}$, is included in the interval $]d^0_{h,0}, d^1_{h,0}]$. The event k appears a time v_{h,n_h} after the transplantation. Let $C_{h,1}$ represent this individual likelihood contribution.

$$C_{h,1} = P_{12}P_{2k}\int_{d^0_{h,0}}^{d^1_{h,0}} f_{12}(u)f_{2k}(v_{h,n_h} - u)\mathrm{d}u \tag{6}$$

Trajectory (ii) – The individual h is not observed in state 2 before the occurrence of a final event k ($k = 3, 4$). Let $C_{h,2}$ be this contribution. Figure 1 shows that the patient may directly jump from state 1 to state k, but he/she may have passed through state 2. This observation illustrates the interval censoring: the durations and the sequence of states.

$$C_{h,2} = P_{1k} f_{1k}(v_{h,n_h}) + P_{12} P_{2k} \int_{d^0_{h,0}}^{v_{h,n_h}} f_{12}(u) f_{2k}(v_{h,n_h} - u) \mathrm{d}u \tag{7}$$

Trajectory (iii) – The individual h is right-censored in state 1 at the time of her/his last visit v_{h,n_h}. Let $C_{h,3}$ denote this type of observation.

$$C_{h,3} = \sum_{j=2}^{4} P_{1j} S_{1j}(v_{h,n_h}) \tag{8}$$

Trajectory (iv) – The individual h is right-censored in state 2 at the time of his/her last visit v_{h,n_h}. The contribution of such trajectory is

$$C_{h,4} = P_{12} \int_{d^0_{h,0}}^{d^1_{h,0}} f_{12}(u) \Big\{ \sum_{j=3}^{4} P_{2j} S_{2j}(v_{h,n_h} - u) \Big\} \mathrm{d}u \tag{9}$$

Making the product of these individual contributions, we obtain the likelihood function of the observed sample. We used the quasi-Newtonian algorithm to maximise this function and to compute the Hessian matrix. This optimisation was performed using the R software with the *optim* function.

The integral calculations are numerically approximated using the 10-points Gauss–Legendre quadrature. In order to evaluate the accuracy of the method, we also used a 20-points quadrature. One should notice that the previous integrals may include singularities:

(a) Expressions (6) and (9) present a singularity at $u = 0$, where, $d^0_{h,0} = 0$ and $\nu_{12} < 1$. Tacking $s = (u/(\sigma_{12}))^{\nu_{12}}$ will remove this singularity.
(b) Expression (6) gives a singularity at $u = v_{h,n_h}$, where, $d^1_{h,0} = v_{h,n_h}$ and $\nu_{2k} < 1$.
(c) Expression (7) gives a singularity at $u = v_{h,n_h}$, where, $\nu_{2k} < 1$ and $d^0_{h,0} > 0$. Tacking $s = ((v_{h,n_h} - u)/\sigma_{2k})^{\nu_{2k}}$ will remove the singularities in cases (b) and (c).
(d) Expression (6) gives singularities at both $u = 0$ and $u = v_{h,n_h}$, where, $\nu_{12} < 1$, $\nu_{2k} < 1$, $d^0_{h,0} = 0$ and $d^1_{h,0} = v_{h,n_h}$.
(e) Expression (7) gives singularities at both $u = 0$ and $u = v_{h,n_h}$, where, $\nu_{12} < 1$, $\nu_{2k} < 1$ and $d^0_{h,0} = 0$. Cases (d) and (e) are more difficult to solve. One way of removing the singularities is to split the integral into two parts, the first part from 0 to $0.5 v_{h,n_h}$ and the second part from $0.5 v_{h,n_h}$ to v_{h,n_h}. Each part has only one singularity and previous substitutions can then be applied.

According to our database and the number of observations per subject, only four patients are concerned. These patients correspond to the trajectory (9) with $d^0_{h,0} = 0$. For these contributions, we have made the appropriate substitution. If many of the observed trajectories was associated with singularities, the Gauss–Legendre quadrature would not have been the best approach. The adaptive numerical integration (function *integrate* in R) would be preferable.

3.3 Incorporation of explanatory factors

In SMM, most authors include covariates through the PDF of the times spent in states, using the proportionality assumption. Many recent works in traditional survival analysis have shown that the assumption of proportionality does not hold in many cases and may lead to serious bias.[21] The same issue arises in multi-state processes. We consider that the covariates are fixed in time and that the corresponding regression parameters depend on the durations. The hazard function of durations is defined by:

$$\lambda_{ij}(d_{h,r}, \mathbf{z}_{h,ij}) = \lambda_{0,ij}(d_{h,r})\exp(\boldsymbol{\gamma}_{ij}(d_{h,r})\mathbf{z}_{h,ij}) \tag{10}$$

in which $\lambda_{0,ij}$ is the baseline hazard function at the duration $d_{h,r}$ specific to the transition from state i to state j. We assume that this function corresponds to a generalised Weibull distribution. $\boldsymbol{\gamma}_{ij}(d_{h,r})^T = (\gamma^1_{ij}(d_{h,r}), \gamma^2_{ij}(d_{h,r}), ..., \gamma^{m_{ij}}_{ij}(d_{h,r}))$ is the vector of time-dependent regression parameters, associated with the m_{ij} time fixed covariates $\mathbf{z}_{h,ij} = (z^1_{h,ij}, z^2_{h,ij}, ..., z^{m_{ij}}_{h,ij})$ of the hth subject. Notice that the effect of the explanatory variables can change from one transition to another. In order to fit time varying effects, we adopt the method of Stablein, Carter and Nova[22] by introducing some time interaction terms. Although the linear term would eventually be sufficient to allow for a time-varying relationship, the quadratic term allows for the modelling of a surface that rises and falls, constituting a non-monotonic time dependence. This method of modelling time-dependent effects is particularly adapted to the following application since the related hazard functions seem to be exponentially distributed. Moreover, the inclusion in the exponential function of interactions with durations does not change the positivity of the hazard function, regardless of the value of regression parameters. One can also notice that the regression parameters are always interpretable as hazard ratios.

In order to compute the likelihood, we need the survival function associated with (10). No analytical solution is available. Using again the 10-points Gauss–Legendre quadrature, $\exp(-\Lambda_{0,ij}(d_{h,r}, \mathbf{z}_{h,ij}))$ is numerically approximated. Log-minus-log survival plots are used to determine if baseline hazard functions are proportional across groups of categorical variables.

Parallel to this approach, we propose to include the explanatory factors through the Markov chain (1). The idea is to take into account heterogeneity in the sequences of the process. Using regression models for categorical dependent variables with more than two response categories,[23] the probabilities of jumping from the state $W_{h,r} = i$ to the

state $W_{h,r+1} = j$, can be written as:

$$P_{ij}(\boldsymbol{\psi}_{h,ik}, k \neq i) = \exp(\boldsymbol{\beta}_{ij}\psi_{h,ij}) / \sum_{k \neq i} \exp(\boldsymbol{\beta}_{ik}\psi_{h,ik}) \tag{11}$$

where $\boldsymbol{\beta}_{ij}^T = (\beta_{ij}^0, \beta_{ij}^1, ..., \beta_{ij}^{u_{ij}})$ is the vector of regression parameters associated with the u_{ij} covariates $\boldsymbol{\psi}_{h,ij} = (1, \psi_{ij}^1, ..., \psi_{ij}^{u_{ij}})$. The constraint (2) imposes a reference state. From state 1, a patient may transit to states 2, 3 or 4. By convention, we fix $\beta_{14} = 0$. Identically, $\beta_{24} = 0$.

3.4 Results

Explanatory factors and PDF parameters are first tested in univariate ($p \leq 0.20$) and the assumption of proportionality is graphically examined. Among the 90 possible parameters for the complete model without interaction, 42 are retained. We have plotted the logarithm of the estimated cumulative hazard function against the duration, for each modality of each covariate and for all transitions. If the resulting curves are non-parallel, then we model the effect of this covariate on this transition by adding interactions with the duration. After the descending procedure ($p \leq 0.05$), 11 factors are finally retained with a log likelihood equals -1550.71, regardless of the number of points in the Gauss–Legendre quadrature. The estimations of regression parameters are given in Table 2. Notice that in the case of convergence problems associated with the 42-parameters model, it would have been more adequate to adopt an ascending method to construct the multivariate model. In our database, the $1 \rightarrow 2$ transitions are interval censored into a short time period with a median equal to 2.5 months. Moreover, the proportion of observations for which the trajectory of the subject is unknown (contributions 7) is relatively small, representing only 3.9% of the total number of contributions. Problems of fitting may have been greater if less observations were made per patient, i.e. higher interval censored periods and a higher proportion of unknown trajectories.

Concerning the trajectories, the $1 \rightarrow 2$ transition is more likely in patients with a delayed graft function greater than or equal to 6 days compared to the others. In contrast, receiving a kidney from a donor greater than 55 years of age is a risk factor for death, given that the initial state is occupied. Finally, kidney transplant recipients with more than 4 HLA-incompatibilities (A+B+DR) constitute a risk factor for returning to dialysis, given that the second state is occupied.

The effects of covariates on the durations in the different states can also be examined. Given that state 2 follows, the time spent in the initial state is shorter for men treated by Simulect or receiving a transplant from a donor greater than 55 years of age. Given that the subject returned to dialysis from the initial state, this transition seems to be accelerated when the cold ischemia time is >24 h. For subjects in the worsened state who returned to dialysis, the transition is faster for transplantations with high HLA-incompatibilities. Again for this transition, individuals with panel reactive antibody (> 0 %) seem to transit quicker than the others. However, this effect decreases linearly according to time spent in the worsened state. Finally, regarding patient death, the

Table 2 Regression parameters of the semi-Markov model

Transition	Covariate	Coef.	SE	*P*-value
Regression parameters associated with the trajectories, $P_{ij}()$.				
$1 \rightarrow 2$	Intercept	1.87	0.68	0.0064
$1 \rightarrow 2$	Delayed graft function	0.73	0.34	0.0302
$1 \rightarrow 2$	Donor age	−2.03	0.69	0.0035
$1 \rightarrow 3$	Intercept	−2.81	0.53	0.0001
$2 \rightarrow 3$	Intercept	1.13	0.34	0.0008
$2 \rightarrow 3$	Incompatibilities A+B+DR	0.93	0.47	0.0473
Regression parameters associated with the durations, $f_{ij}()$.				
$1 \rightarrow 2$	Induction treatment	0.36	0.14	0.0094
$1 \rightarrow 2$	Recipient gender	−0.26	0.13	0.0542
$1 \rightarrow 2$	Donor age	0.96	0.23	0.0001
$1 \rightarrow 3$	Cold ischemia time	5.02	1.20	0.0001
$2 \rightarrow 3$	Incompatibilities A+B+DR	0.88	0.28	0.0019
$2 \rightarrow 3$	Panel reactive antibody	1.09	0.35	0.0016
$2 \rightarrow 3$	Panel reactive antibody $\times\ d^{a}$	−0.48	0.22	0.0308
$2 \rightarrow 4$	Delayed graft function	2.03	0.60	0.0008
$2 \rightarrow 4$	Recipient gender	1.54	0.64	0.0165
$2 \rightarrow 4$	Recipient gender $\times\ d^{a}$	−4.30	1.19	0.0003
$2 \rightarrow 4$	Recipient gender $\times\ d^{2a}$	1.30	0.32	0.0001

[a]Time interaction with the duration d in the state.

duration in state 2 appears to be shorter for men with a long delayed graft function. The regression coefficient associated with gender varies across the time spent in the worsened state.

Concerning the PDF parameters, presented in Table 3, the duration hazard function of the $1 \rightarrow 2$ transition is $\cup$−shaped with a minimum about 4 years post transplantation. Using the LRS, the other transitions seem to be exponentially distributed, i.e. the baseline hazard functions are constant over the durations. If the $1 \rightarrow 2$ transition had been exponentially distributed, the modelling of a Markov versus a semi-Markov process would have been more suitable. However, the semi-Markov property we describe here is interesting for several reasons. Firstly, even if the $2 \rightarrow 3$ and $2 \rightarrow 4$ transitions times are exponentially distributed, the resulting distribution of duration in state 2 is a mixture of exponential distributions. Secondly, again concerning this duration in state 2, we have identified significant interactions between panel reactive antibody, recipient gender and durations. Thus, the specific hazard functions of the $2 \rightarrow 3$ and $2 \rightarrow 4$ are not constant according to durations for patients with positive panel reactive antibody and for men, respectively. Finally, in order to compute the probability of entering into an absorbing state according to chronological time since the transplantation, convolution products similar to expressions (6) and (7) use the density f_{12}, which is not constant over time regardless of the covariate values. Thus, the results obtained are globally different from those obtained using a Markov model.

Table 3 PDF parameters related to the durations in states

	σ_{ij}		ν_{ij}		θ_{ij}	
Transition	Estim.	SE	Estim.	SE	Estim.	SE
$1 \rightarrow 2$	68.80	76.53	0.77	0.05	0.25	0.20
$1 \rightarrow 3$	42.84	47.34	1	.	1	.
$1 \rightarrow 4$	109.83	73.00	1	.	1	.
$2 \rightarrow 3$	12.66	3.20	1	.	1	.
$2 \rightarrow 4$	6.54	4.13	1	.	1	.

The values of all the parameters are presented in Appendix B (Table B1) for the integral calculation made using 20-points Gauss–Legendre quadrature. One can see that the supplementary number of points is not associated with a significant difference in terms of parameter estimation, compared to the 10-points approximation.

4 The goodness-of-fit test

4.1 Definition of the statistic

In SMM, the estimated transition probabilities depend on the sequence of observed states, on the time spent in states and on the covariate pattern. In Section 3, the LRS is used to test the significance of parameters associated with duration distribution and of regression coefficients. Explanatory variables are also graphically examined. One of the assumptions of the modelling is the stationarity (1) regarding the chronological time, i.e. the time since the transplantation. The null hypothesis (H_0) assumes that the semi-Markov is stationary. The alternative hypothesis (H_1) assumes that the model is not stationary. To test if the stationarity is valid, the behaviour of the model across chronological time should be examined. We propose a Pearson-type goodness-of-fit statistic, in which observations are grouped according to K categories of transitions and to L intervals of chronological times with boundaries $\{t_0, ..., t_l, ..., t_L\}$.

In contrast to the $1 \rightarrow 2$ transitions, the times of occurrence of the final events k ($k = 3, 4$) are exactly known. This property of the data makes the classification of these transitions according to chronological time easier. In other words, classification of the $1 \rightarrow 2$ transitions according to chronological time is not possible, since all the durations of these transitions are interval-censored. We thus consider two types of observations ($K = 2$): the jump to state 3 ($e \rightarrow 3$, with $e = 1, 2$) and the jump to state 4 ($e \rightarrow 4$, with $e = 1, 2$).

Concerning the classification of transitions according to chronological time, we define $L = 5$ intervals limited using the quantiles of the times of occurrence of a final event. These choices lead to balanced number of observed transitions.

Let $e_{l,k}$ and $o_{l,k}$ be respectively the expected and observed number of transitions in cell (l, k). The Pearson type goodness-of-fit statistic, to examine the adequacy of the

stationary SMM, is

$$G = \sum_{l=1}^{L} \sum_{k=3}^{4} (o_{l,k} - e_{l,k})^2 / e_{l,k} \tag{12}$$

This goodness-of-fit test bears some resemblances to the test for lifetables proposed by Lawless,[24] which have been applied to testing the fit of the survival curve for a multi-state model with a single absorbing state.[25] In these works, the statistic is based on the number of entries into the absorbing state in time periods conditional on an individual being in a transient state at the start of that period, meaning an individual contributes a series of binary observations (whether they survived or were absorbed in each period). In contrast, the statistic in this paper categorises observations according to which absorbing state they entered and the time period, meaning an individual contributes a single multinomial observation.

Respecting the expression (7) and taking into account the right-censoring, the expected number of transitions into the state k ($k = 3, 4$) between t_{l-1} and t_l ($l = 1, ..., L$) equals

$$\begin{aligned} e_{l,k} = \sum_{R(t_{l-1})} \Big\{ & P_{1k}(\boldsymbol{\psi}_{h,12}, \boldsymbol{\psi}_{h,13}) \int_{t_{l-1}}^{\min(v_{h,n_h}, t_l)} f_{1k}(x, \mathbf{z}_{h,1k}) \mathrm{d}x \\ & + P_{12}(\boldsymbol{\psi}_{h,12}, \boldsymbol{\psi}_{h,13}) P_{2k}(\boldsymbol{\psi}_{h,23}) \int_{t_{l-1}}^{\min(v_{h,n_h}, t_l)} \int_{0}^{x} f_{12}(u, \mathbf{z}_{h,12}) \\ & \times f_{2k}(x-u, \mathbf{z}_{h,2k}) \mathrm{d}u \mathrm{d}x \Big\} \end{aligned} \tag{13}$$

where v_{h,n_h} is the possible time of right-censoring of the subject h and $R(t_{l-1})$ represents the sum over all the individuals h which are not censored at the time t_{l-1} since the transplantation. The number of final events is obtained by counting their occurrence in an interval:

$$o_{l,k} = \sum_{R(t_{l-1})} I_{\{v_{h,n_h} \leq t_l \text{ and } w_{h,n_h} = k\}}$$

where $I_{\{a\}} = 1$ if the condition a is respected and 0 otherwise.

4.2 Distribution of the statistic

As written in Section 4.1, the $1 \rightarrow 2$ transitions are not included in the statistic (12), since only the times of entering in absorbing states are exactly known. Thus, the proposed statistic does not include the right-censored contributions. For this reason, and regardless of the asymptotic circumstances, the approximation using a χ^2 distribution is not adequate. One method for estimating the distribution of the statistic (12) is to generate B independent bootstrap samples from the parameters of the SMM under H_0 and to calculate the goodness-of-fit statistic for each sample. This semi-parametric bootstrap procedure can be divided into the following steps:

(i) Generation of B bootstrap samples, constituted of n subjects: Each individual h^* ($h^* = 1, \ldots, n$) is observed during visits at the times since the transplantation $\{v^*_{h^*,0}, v^*_{h^*,1}, \ldots, v^*_{h^*,n_{h^*}}\}$. This vector corresponds to the observed times of visit in the

initial sample. Since we do not simulate these times, the bootstrap is named semi-parametric. All patients begin in the state 1, $W^*_{h^*,0} = 1$.

(ii) Simulation of the trajectory of each patient h^* from the estimated parameters of the SMM under H_0: From the estimated regression coefficients associated with the Markov chain, $\{\hat{\boldsymbol{\beta}}_{1j}, j = 2, 3\}$, the second bootstrap state, $W^*_{h^*,1}$, is defined using the multinomial distribution with parameters $\{P_{1j}(\boldsymbol{\psi}_{h^*,12}, \boldsymbol{\psi}_{h^*,13}), j = 2, 3, 4\}$. If $W^*_{h^*,1} = 2$, then the value of $W^*_{h^*,2}$ is obtained using a binomial distribution with probabilities $\{P_{2j}(\boldsymbol{\psi}_{h^*,23}), j = 3, 4\}$. Given these bootstrap sequence of states, the duration in state 1 for the subject h, $D^*_{h^*,0}$, follows the PDF $f_{1W^*_{h^*,1}}$, which takes into account the covariates $\mathbf{z}_{h,1W^*_{h^*,1}}$. Thus, the inverse of the corresponding distribution function can be used to obtain a simulated duration from a random uniform variable $U(0, 1)$. Identically, if $W^*_{h^*,1} = 2$, then the duration in this can be simulated from $f_{2W^*_{h^*,2}}$.

(iii) Estimation of the SMM for each bootstrap sample: The contributions are obtained as follows.

- If $d^*_{h^*,0} > v^*_{h^*,n_{h^*}}$, then this subject h^* is censored in state 1 and her/his contribution is equivalent to (8), replacing v_{h,n_h} by $v^*_{h^*,n_{h^*}}$.
- Otherwise, if $v^*_{h^*,r-1} < d^*_{h^*,0} \leq v^*_{h^*,r}$ $(r = 1, ..., n_{h^*})$ and $W^*_{h^*,1} = k$ $(k = 3, 4)$, the contribution corresponds to (7), where, v_{h,n_h} becomes $d^*_{h^*,0}$ and $d^0_{h,0}$ becomes $v^*_{h^*,r-1}$.
- Otherwise, if $v^*_{h^*,r-1} < d^*_{h^*,0} \leq v^*_{h^*,r}$ $(r = 1, ..., n_{h^*})$, $W^*_{h^*,1} = 2$ and $d^*_{h^*,0} + d^*_{h^*,1} > v^*_{h^*,n_{h^*}}$, then the contribution equals (9), in which the values $\{d^0_{h,0}, d^1_{h,0}, v_{h,n_h}\}$ become $\{v^*_{h^*,r-1}, v^*_{h^*,r}, v^*_{h^*,n_{h^*}}\}$.
- Otherwise, if $v^*_{h^*,r-1} < d^*_{h^*,0} \leq v^*_{h^*,r}$ $(r = 1, ..., n_{h^*})$, $W^*_{h^*,1} = 2$ and $d^*_{h^*,0} + d^*_{h^*,1} \leq v^*_{h^*,n_{h^*}}$, then the contribution is equivalent to (6), in which the values $\{d^0_{h,0}, d^1_{h,0}, v_{h,n_h}\}$ become $\{v^*_{h^*,r-1}, v^*_{h^*,r}, d^*_{h^*,0} + d^*_{h^*,1}\}$.

(iv) Calculation of the distribution of the bootstrap statistics: For each bootstrap sample, the statistics G^*_b $(b = 1, ..., B)$ is calculated, using the expressions (12) and (13). The p-value, i.e. the error probability of rejecting H_0 if this hypothesis is valid, equals $B^{-1} \sum_{b=1}^{B} I_{\{G^*_b \geq G\}}$, where $I_{\{G^*_b \geq G\}}$ equals 1 if the bootstrap statistics G^*_b is greater than or equal to G, and 0 otherwise.

4.3 Application to data

The expected and observed counts of transitions into a final state are shown in the Contingency Table 4. The first two rows correspond to the observed and expected returns to dialysis and death between 0.011 and 0.689 years post transplantation. For these two types of events, the respective cells contribute 5.19 and 10.12% of the value of the goodness-of-fit statistic, with $G = 14.12$. The two cells in bold contribute 46.52% and correspond to the returns to dialysis. For reasons of computing time,

Table 4 Contingency table for the observed and expected counts of final events

Chronological time (in years)		Transition		Percentage	
		$e \to 3$	$e \to 4$	$e \to 3$	$e \to 4$
]0.011; 0.689]	Observed	12	8	5.19%	10.12%
	Expected	9.38	5.25		
]0.689; 2.168]	Observed	13	8	**21.21%**	2.89%
	Expected	20.91	10.02		
]2.168; 3.826]	Observed	16	5	15.73%	11.73%
	Expected	23.17	8.81		
]3.826; 5.213]	Observed	17	4	2.14%	4.45%
	Expected	14.87	6.18		
]5.213; 9.158]	Observed	14	7	**25.31%**	0.24%
	Expected	23.08	7.51		

These cells correspond to the two higher contributions.

400 bootstrap samples were carried out. The quantiles of the cumulative probability function of the bootstrap statistic in {0.75, 0.50, 0.25, 0.10, 0.05, 0.01} are respectively {9.79, 12.77, 15.26, 18.27, 20.95, 27.37}. A total of 159 bootstrap statistics are greater than or equal to G, corresponding to a p-value equal to 0.3975. Thus, the fitted stationary SMM seems to be adequate to the kidney transplant data.

5 Conclusions

We have considered a flexible multi-state model in order to analyse inherently complex longitudinal data. This model has been applied to the follow-up of kidney transplant recipients. The generalised Weibull PDF appears to be suitable. The use of a multinomial and multivariate logistic approach, in order to introduce explanatory factors in the Markov chain, also lead to a more precise modelling.

This parametric regression allows for the analysis of interval-censored data. More precisely, as is often the case in longitudinal data, the date of the final and absorbing events are exactly known (the death or the return to dialysis in our application) and only the transition times of the intermediate and transient states are interval censored. Between two consecutive visits, the information regarding a patient is unavailable. Using convolution products, the proposed estimation procedure deals with all the possible trajectories of a patient during this interval and can be easily adapted to other applications.

This model is fitted using standard likelihood estimation. The LRS can then be applied to the parameters selection strategy. This relates to the parsimony of the distribution functions and the selection of the significant covariates.

In order to determine whether the stationary assumption of the SMM is valid, we proposed the Pearson-type goodness-of-fit test. Because the theoretical distribution of this statistic is intractable, a bootstrap algorithm was proposed. This test does not provide evidence of lack of fit. To sum up, this paper offers three extensions relating to the existing works in SMM: the approach for addressing covariates, the modelling of interval-censoring and the testing of stationarity.

A way to improve this statistic should take into account the $1 \to 2$ transition directly in the contingency table. In our application, all the transitions from state 1 into state 2

are interval-censored with different intervals for each subject. Thus, the classification of the number of $1 \rightarrow 2$ transitions in some intervals of time since the transplantation is difficult. However, even if the statistic is based on the number of final failures, the $1 \rightarrow 2$ transition enters into the computation of the expected number of final events [see Equation (13)].

As we mentioned previously, the date of entry into an absorbing state, v_{h,n_h}, is exactly known. However, the estimation of the model can be adapted if this date is interval censored. For example, we can reconsider the individual likelihood contribution of a subject h respecting the trajectory *(i)*, but where the date of entry into the state $W_{h,2} = k$ occurs between v^0_{h,n_h} and v^1_{h,n_h}. In this case, the Equation (6) can be extend as follows:

$$P_{12}P_{2k}\int_{d^0_{h,0}}^{d^1_{h,0}}\int_{v^0_{h,n_h}}^{v^1_{h,n_h}} f_{12}(u)f_{2k}(w-u)\mathrm{d}w\mathrm{d}u$$

One can see that this expression is equal to (6) when the difference between v^0_{h,n_h} and v^1_{h,n_h} tends to be null. The double integral calculations are always feasible using the Gauss–Legendre quadrature. For a higher dimension of integrals, for instance if three consecutive transitions were interval censored for a given subject (2 transient and 1 absorbing states), this numerical approximation is no longer feasible since the computation time increases too much and since removing singularities becomes harder.

The main limit is the assumption of a unidirectional model. Indeed, if returns are included (i.e. the transition from state 2 to state 1). The recent paper of Kang and Lagakos[9] considers the case where backward transitions are permitted. This excellent work constitutes a way of extending our proposed SMM, since the authors do not include covariate effects in their modelling.

Acknowledgements

We thank the Roche Laboratory for its support of the DIVAT data bank and Dr J. Ashton-Chess for editing the manuscript.

References

1 Lagakos S, Sommer C, Zelen M. Semi-Markov models for partially censored data. *Biometrika* 1978; **65**: 311–17.
2 Kay R. A Markov model for analyzing cancer markers and disease states in survival studies. *Biometrics* 1986; 7: 855–66.
3 Andersen P. Multistate models in survival analysis: a study of nephropathy and mortality in diabetes. *Statistics in Medicine* 1988; 7: 661–70.
4 Perez-Ocon R, Ruiz-Castro J. *Semi-Markov models and applications*. Dordrecht, Netherlands: Kluwer Academic Publishers, 1999.
5 Dabrowska D, Sun G, Horowitz M. Cox Regression in a Markov Renewal Model : an application to the analysis of bone transplant

data. *Journal of the American Statistical Association* 1994; **89**: 867–77.
6 Sternberg M, Satten G. Discrete time nonparametric estimation for chain of events data subject to interval censoring and truncation. *Biometrics* 1999; **55**: 514–22.
7 Lindsey J. A study of interval censoring in parametric regression models. *Lifetime Data Analysis* 1998; **4**: 329–54.
8 Foucher Y, Giral M, Soulillou J, Daures J. A semi-Markov model for multistate and interval-censored data with multiple terminal events. Application in renal transplantation. *Statistics in Medicine* 2007; **26**: 5381–93.
9 Kang M, Lagakos S. Statistical methods for panel data from a semi-Markov process, with application to HPV. *Biostatistics* 2007; **8**: 252–64.
10 Escolano S, Golmard J, Korinek A, Mallet A. A multi-state model for evolution of intensive care unit patiets: prediction of nosocomial infections and deaths. *Statistics in Medicine* 2000; **19**: 3465–82.
11 Huang Y. Censored regression with the multistate accelerated sojourn times model. *Journals of the Royal Statistical Society B* 2002; **64**: 17–29.
12 Foucher Y, Mathieu E, Saint-Pierre P, Durand J, Daures J. A Semi-Markov model based on generalized Weibull distribution with an illustration for HIV disease. *Biometrical Journal* 2005; **47**: 825–33.
13 Kalbfleisch J, Lawless J. The analysis of panel data under a Markov assumption. *Journal of the American Statistical Association* 1985; **80**: 863–71.
14 Castelli C, Combescure C, Foucher Y, Daures J. Cost-effectiveness analysis in colorectal cancer using a semi-Markov model. *Statistics in Medicine* 2007; **26**: 5557–71.
15 Aguirre-Hernandez R, Farewell V. A Pearson-type goodness-of-fit test for stationary and time-continuous Markov regression models. *Statistics in Medicine* 2002; **21**: 1899–911.
16 Titman A, Sharples L. A general goodness-of-fit test for Markov and hidden Markov models. *Statistics in Medicine* 2007; DOI:10.1002/sim.3033.
17 Giral M, Bertola J, Foucher Y, Villers D, Bironneau E, Blanloeil Y, et al. Effect of brain-dead donor resuscitation on delayed graft function: results of a monocentric analysis. *Transplantation* 2007; **83**: 1174–81.
18 Levey A, Bosch J, Lewis J, Greene T, Rogers N, Roth D. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of Diet in Renal Disease Study Group. *Annals of Internal Medicine* 1999; **130**: 461–70.
19 Foucher Y. Modèles semi-markoviens: Application à l'analyse de l'évolution de pathologies chroniques (PhdThesis). Université Montpellier 1. http://tel.archives-ouvertes.fr/tel-00267683/fr/; 2007.
20 Bagdonavicius V, Nikulin M. *Accelerated life models*. Chapman and Hall/CRC, London; 2002.
21 Orbe J, Ferreira E, Nunez-Anton V. Comparing proportional hazards and accelerated failure time models for survival analysis. *Statistics in medicine* 2002; **21**: 3493–3510.
22 Stablein D, Carter W, Novak J. Analysis of survival data with nonproportional hazard functions. *Controlled Clinical Trials* 1981; **2**: 149–159.
23 McCullagh P. *Generalized linear models*. Chapman and Hall, London; 1983.
24 Lawless J. *Statistical models and methods for lifetime data*. Wiley, New York; 1982.
25 Perez-Ocon R, Ruiz-Castro J, Gamiz-Perez M. A piecewise Markov process for analysing survival from breast cancer in different risk groups. *Statistics in Medicine* 2001; **20**: 109–22.

Appendix

Appendix A : Proof of likelihood contributions

Trajectory (i) – The contribution is the joint probability that the subject h jumped into state 2 after a duration in state 1 between $d^0_{h,0}$ and $d^1_{h,0}$ and that she/he returned to dialysis ($k = 3$) or died ($k = 4$) at time v_{h,n_h} post transplantation.

$$
\begin{aligned}
C_{h,1} = &\lim_{\Delta d \to 0^+} \Big\{ P(v_{h,n_h} - d_{h,0} < d_{h,1} < v_{h,n_h} - d_{h,0} + \Delta d,\, W_{h,2} = k, \\
&\quad d^0_{h,0} < d_{h,0} < d^1_{h,0},\, W_{h,1} = 2 | W_{h,0} = 1)/\Delta d \Big\} \\
= &\lim_{\Delta d \to 0^+} \Big\{ P(v_{h,n_h} - d_{h,0} < d_{h,1} < v_{h,n_h} - d_{h,0} + \Delta d,\, W_{h,2} = k | \\
&\quad d^0_{h,0} < d_{h,0} < d^1_{h,0},\, W_{h,1} = 2,\, W_{h,0} = 1) \\
&\quad \times P(d^0_{h,0} < d_{h,0} < d^1_{h,0},\, W_{h,1} = 2 | W_{h,0} = 1)/\Delta d \Big\} \\
= &\lim_{\Delta d \to 0^+} \Big\{ P(W_{h,2} = k | W_{h,1} = 2) P(v_{h,n_h} - d_{h,0} < d_{h,1} < v_{h,n_h} - d_{h,0} \\
&\quad + \Delta d | W_{h,2} = k,\, W_{h,1} = 2)/\Delta d \Big\} \\
&\quad \times P(W_{h,1} = 2 | W_{h,0} = 1) P(d^0_{0,h} < d_{h,0} < d^1_{h,0} | W_{h,1} = 2,\, W_{h,0} = 1) \\
= &P_{12} P_{2k} \int_{d^0_{h,0}}^{d^1_{h,0}} f_{12}(u) f_{2k}(v_{h,n_h} - u) \mathrm{d}u
\end{aligned}
$$

Trajectory (ii) – Because the state 2 is not observed, we must take into account that the individual transited directly between state 1 and state k at time v_{h,n_h} after the transplantation or that her/he had two consecutive transitions $1 \to 2 \to k$ in the interval $]d^0_{h,0}, v_{h,n_h}]$. Respecting the last development, we obtain:

$$
\begin{aligned}
C_{h,2} = &\lim_{\Delta d \to 0^+} \Big\{ P(v_{h,n_h} < d_{h,0} < v_{h,n_h} + \Delta d,\, W_{h,1} = k | W_{h,0} = 1)/\Delta d \Big\} \\
&+ \lim_{\Delta d \to 0^+} \Big\{ P(v_{h,n_h} - d_{h,0} < d_{h,1} < v_{h,n_h} - d_{h,0} + \Delta d,\, W_{h,2} = k, \\
&\quad d^0_{h,0} < d_{h,0} < v_{h,n_h},\, W_{h,1} = 2 | W_{h,0} = 1)/\Delta d \Big\} \\
= &\, P_{1k} f_{1k}(v_{h,n_h}) + P_{12} P_{2k} \int_{d^0_{h,0}}^{v_{h,n_h}} f_{12}(u) f_{2k}(v_{h,n_h} - u) \mathrm{d}u
\end{aligned}
$$

Trajectory (iii) – The individual h is right-censored in state 1 after a time v_{h,n_h} post transplantation.

$$
\begin{aligned}
C_{h,3} &= P(d_{h,0} > v_{h,n_h} | W_{h,0} = 1) \\
&= \sum_{j=2}^{4} P(W_{h,1} = j | W_{h,0} = 1) P(d_{h,0} > v_{h,n_h} | W_{h,1} = j,\, W_{h,0} = 1) \\
&= \sum_{j=2}^{4} P_{1j} S_{1j}(v_{h,n_h})
\end{aligned}
$$

Trajectory (iv) – The individual h is right-censored in state 2 after a time v_{h,n_h} post transplantation and her/he jumped into state 2 after a duration in state 1 between $d^0_{h,0}$ and $d^1_{h,0}$.

$$
\begin{aligned}
C_{h,4} &= P(d_{h,1} > v_{h,n_h} - d_{h,0},\, W_{h,1} = 2,\, d^0_{h,0} < d_{h,0} < d^1_{h,0} | W_{h,0} = 1) \\
&= P(d_{h,1} > v_{h,n_h} - d_{h,0} | W_{h,1} = 2,\, d^0_{h,0} < d_{h,0} < d^1_{h,0},\, W_{h,0} = 1) \\
&\quad \times P(d^0_{h,0} < d_{h,0} < d^1_{h,0},\, W_{h,1} = 2 | W_{h,0} = 1) \\
&= P(W_{h,1} = 2 | W_{h,0} = 1) P(d^0_{h,0} < d_{h,0} < d^1_{h,0} | W_{h,1} = 2,\, W_{h,0} = 1) \\
&\quad \times \sum_{j=3}^{4} P(W_{h,2} = j | W_{h,1} = 2) P(d_{h,1} > v_{h,n_h} - d_{h,0} | W_{h,2} = j,\, W_{h,1} = 2) \\
&= P_{12} \int_{d^0_{h,0}}^{d^1_{h,0}} f_{12}(u) \Big\{ \sum_{j=3}^{4} P_{2j} S_{2j}(v_{h,n_h} - u) \Big\} \mathrm{d}u
\end{aligned}
$$

Appendix B : 20-points Gauss–Legendre quadrature

Table B1 Comparison of the parameter estimations according to the number of points in the Gauss–Legendre quadrature

Transition	Parameters	Estim. (10 points)	Estim. (20 points)
	Parameters of the baseline hazard functions		
$1 \rightarrow 2$	σ_{12}	68.80	68.75
$1 \rightarrow 2$	ν_{12}	0.77	0.05
$1 \rightarrow 2$	θ_{12}	0.25	0.25
$1 \rightarrow 3$	σ_{13}	42.84	42.76
$1 \rightarrow 4$	σ_{14}	109.83	109.87
$2 \rightarrow 3$	σ_{23}	12.66	12.66
$2 \rightarrow 4$	σ_{24}	6.54	6.83
	Regression parameters associated with the trajectories, $P_{ij}()$.		
$1 \rightarrow 2$	Intercept	1.87	1.87
$1 \rightarrow 2$	Delayed graft function	0.73	0.73
$1 \rightarrow 2$	Donor age	−2.03	−2.03
$1 \rightarrow 3$	Intercept	−2.81	−2.81
$2 \rightarrow 3$	Intercept	1.13	1.13
$2 \rightarrow 3$	Incompatibilities A+B+DR	0.93	0.93
	Regression parameters associated with the durations, $f_{ij}()$.		
$1 \rightarrow 2$	Induction treatment	0.36	0.36
$1 \rightarrow 2$	Recipient gender	−0.26	−0.26
$1 \rightarrow 2$	Donor age	0.96	0.96
$1 \rightarrow 3$	Cold ischemia time	5.02	5.02
$2 \rightarrow 3$	Incompatibilities A+B+DR	0.88	0.88
$2 \rightarrow 3$	Panel reactive antibody	1.09	1.09
$2 \rightarrow 3$	Panel reactive antibody $\times\ d^{a}$	−0.48	−0.47
$2 \rightarrow 4$	Delayed graft function	2.03	2.06
$2 \rightarrow 4$	Recipient gender	1.54	1.56
$2 \rightarrow 4$	Recipient gender $\times\ d^{a}$	−4.30	−4.30
$2 \rightarrow 4$	Recipient gender $\times\ d^{2a}$	1.30	1.30

[a]Time interaction with the duration d in the state.