Biostatistics (2007), **8**, 4, *pp.* 708–721
doi:10.1093/biostatistics/kxl043
Advance Access publication on January 30, 2007

Joint frailty models for recurring events and death using maximum penalized likelihood estimation: application on cancer events

VIRGINIE RONDEAU*

Institut National de la Santé et de la Recherche Médicale, U875 (Biostatistique), Bordeaux, F-33076, France and Université Victor Segalen Bordeaux 2, Bordeaux, F-33076, France
virginie.rondeau@isped.u-bordeaux2.fr

SIMONE MATHOULIN-PELISSIER

Institut Bergonié—Centre Régional de Lutte Contre le Cancer du Sud-Ouest, Bordeaux F-33076, France

HÉLÈNE JACQMIN-GADDA

Institut National de la Santé et de la Recherche Médicale, U875 (Biostatistique), Bordeaux, F-33076, France and Université Victor Segalen Bordeaux 2, Bordeaux, F-33076, France

VÉRONIQUE BROUSTE, PIERRE SOUBEYRAN

Institut Bergonié—Centre Régional de Lutte Contre le Cancer du Sud-Ouest, Bordeaux F-33076, France

SUMMARY

The observation of repeated events for subjects in cohort studies could be terminated by loss to follow-up, end of study, or a major failure event such as death. In this context, the major failure event could be correlated with recurrent events, and the usual assumption of noninformative censoring of the recurrent event process by death, required by most statistical analyses, can be violated. Recently, joint modeling for 2 survival processes has received considerable attention because it makes it possible to study the joint evolution over time of 2 processes and gives unbiased and efficient parameters. The most commonly used estimation procedure in the joint models for survival events is the expectation maximization algorithm. We show how maximum penalized likelihood estimation can be applied to nonparametric estimation of the continuous hazard functions in a general joint frailty model with right censoring and delayed entry. The simulation study demonstrates that this semiparametric approach yields satisfactory results in this complex setting. As an illustration, such an approach is applied to a prospective cohort with recurrent events of follicular lymphomas, jointly modeled with death.

Keywords: Cancer; Joint frailty models; Penalized likelihood; Recurrent events.

*To whom correspondence should be addressed.

1. Introduction

In many clinical or epidemiological studies, subjects can potentially experience recurrent or repeated events. For instance, patients may experience repeated epileptic seizures or cancer patients may experience recurrent superficial tumors or repeated episodes of hospitalization. Statistical models have been proposed to analyze these recurrent event data (Cook and Lawless, 2002).

Furthermore, the time frame for an individual's repeated event process may depend on other "terminating" events, such as death. Often the recurrence of serious events, such as tumors and opportunistic infections, is associated with an elevated risk of death. In this context, the usual assumption of noninformative censoring of the recurrent event process by death, required by most statistical analyses, can be violated. This dependence should be accounted for in the joint modeling of recurrent events and deaths.

The approach we develop in this paper is motivated by a study of patients with follicular lymphoma (FL) undergoing episodic relapses of FL. FLs account for one-third of non-Hodgkin lymphomas in adults. The prognosis of FL is heterogenous, and numerous treatments may be proposed (Solal-Celigny *and others*, 2004). The course of this disease is usually characterized by a response to initial treatment, followed by relapses, sometimes associated with high-grade non-Hodgkin lymphomas. After the initial treatment, each patient was monitored regularly for routine visits, and the presence of FL relapses was notified at each visit. Estimation of the risk of recurrence allows for better planning of follow-up schedules after diagnosis or first treatment and permits clinicians to determine therapeutic approaches based on the patient's risk of relapse. Furthermore, FL relapses may increase the risk of death. As a result, there is an association between the FL relapses process and the survival process, which precludes the use of standard analyses of recurrent events. Specifically, those subjects experiencing FL relapses at the highest rate are typically observed for shorter periods of observation due to mortality. In this work, we will thus consider the FL relapses and the terminal event process jointly, in a joint frailty model setting.

Li and Lagakos (1997) considered the marginal approach of Wei–Lin–Weissfeld (1989). They assumed the terminating event as a censoring event for each recurrent event, or they treated the failure time for each recurrence as the first occurrence of the recurring event or terminating event, whichever came first. However, these marginal models do not specify the dependence between recurrent events and death. Ghosh and Lin (2003) proposed a joint marginal formulation for the distributions of the recurrent event process and the dependent censoring time.

Some methods based on counts have been proposed. Lancaster and Intrator (1998) considered joint parametric modeling of repeated inpatient episodes (via a Poisson model) and survival time of a panel of patients over 15 months. The model induced a correlation between hospitalization and death via a person-specific frailty term. Sinha and Maiti (2004) used a more general joint model for panel count data and a dependent termination, using a Bayesian approach.

Huang and Wolfe (2002) proposed to take into account the informative censoring in clustered data. Liu *and others* (2004) proposed a joint semiparametric model for the intensity functions of both recurrent events and death by a shared gamma frailty model. In these models, the frailty effect on recurrent events and death rates is not the same. In these approaches, estimation is carried out through a Monte Carlo expectation maximization (EM) algorithm, which could be time consuming. Furthermore, these methods cannot be used to correctly estimate hazard functions, which often have a meaningful interpretation in epidemiological studies. Most of the time, the baseline intensity estimate is based on Breslow's estimate leading to a piecewise-constant baseline hazard function or unspecified baseline hazard function.

In this paper, we propose a nonparametric penalized likelihood method for estimating hazard functions in a general joint frailty model for recurrent events and terminal events, with both right-censored survival data and delayed entries. This approach is of interest for several reasons. First, it makes it possible to deal with informative censoring for recurrent event data; in addition, it also allows joint treatment of 2 processes which evolve with time leading to more accurate estimates. This work extends the previous work by giving smoothed estimates of the 2 hazard functions which represent incidence and mortality

rates in epidemiology. It is natural in epidemiology to impose a continuous hazard function with small local variations.

To analyze recurrent event data, the focus can be placed on time-between-events (i.e. gap times) or time-to-events (i.e. calendar times) models (see Duchateau *and others*, 2003). These 2 timescales which are 2 important aspects of the data can be linked to the semi-Markov models in which the transition probability between 2 states depends only on the waiting times, whereas in the nonhomogenous Markov models, this transition depends only on the time since inclusion in the study. The proposed approach can deal with both situations and is illustrated in the article.

The paper is organized as follows: In Section 2, we describe the joint frailty model. The construction of the full penalized log-likelihood is explained in Section 3. Results from a detailed simulation study are reported in Section 4. The model is applied to the analysis of episodic relapses of FL and death in Section 5. Finally, Section 6 presents a concluding discussion.

2. JOINT MODEL FOR RECURRENT EVENTS AND A TERMINATING EVENT

2.1 *The model*

We denote for subject i ($i = 1, \ldots, N$), X_{ij} the jth recurrent times ($j = 1, \ldots, n_i$), C_i the censoring times (not by death), and D_i the death times. We first consider X_{ij} as a time to event. $T_{ij} = \min(X_{ij}, C_i, D_i)$ corresponds to each follow-up time, and δ_{ij} is a binary indicator for recurrent events which is 0 if the observation is censored or if the subject died and 1 if X_{ij} is observed $\left(\delta_{ij} = I_{(T_{ij}=X_{ij})}\right.$, where $I_{(\cdot)}$ denotes the indicator function$\left.\right)$. Similarly, we denote T_i^* the last follow-up time for subject i, which is either a time of censoring or a time of death ($T_i^* = \min(C_i, D_i)$), and $\delta_i^* = I_{(T_i^*=D_i)}$. What we actually observe is $(T_{ij}, \delta_{ij}, \delta_i^*)$. We also use the theory of multivariate counting processes (Andersen *and others*, 1993; Liu *and others*, 2004). Let $N_i^{\mathrm{R}^*}(t)$ define the actual number of recurrent events in $(0, t]$ for the ith individual. Because of censoring, it is impossible to observe $N_i^{\mathrm{R}^*}(\cdot)$. Rather, we observe the process $N_i^{\mathrm{R}}(t) = N_i^{\mathrm{R}*}(\min(T_i^*, t))$ which counts the observed number of recurrent events, which may be less than $N_i^{\mathrm{R}^*}(t)$. Similarly, denote by $N_i^{\mathrm{D}}(t) = I\ (T_i^* \leqslant t, \delta_i^* = 1)$ the observed death indicator and $N_i^{\mathrm{D}^*}(t) = I(D_i \leqslant t)$ the actual death indicator. Furthermore, define $Y_i(t) = I(T_i^* \geqslant t)$, the "at-risk" process which indicates whether the subject is still under observation at time t or not. The number of recurrent events that occur for subject i over the small interval $[t, t + \mathrm{d}t)$ is $\mathrm{d}N_i^{\mathrm{R}^*}(t) = N_i^{\mathrm{R}^*}((t + \mathrm{d}t)^-) - N_i^{\mathrm{R}^*}(t^-)$, and we have $\mathrm{d}N_i^{\mathrm{R}}(t) = Y_i(t)\mathrm{d}N_i^{\mathrm{R}^*}(t)$.

We consider $\mathcal{F}_t$ the σ-algebra generated by the whole observed data and the unobserved frailty ω (defined later), $\mathcal{F}_t = \sigma\{Y_i(u), N_i^R(u), N_i^D(u), Z_i(u), 0 \leqslant u \leqslant t, \omega_i, i = 1, \ldots, n\}$ which represents process history of subject i up to time t, the filtration is the family $(\mathcal{F}_t)_{t\geqslant 0}$ and with $Z_i(t)$ the covariate process.

The following assumptions are made:

1) We assume continuous recurrent, terminating, and censoring processes so that recurrent events and death cannot happen at the same time. We adopt the convention that death happens first in the small interval $[t, t + \mathrm{d}t)$. For 2 subjects in the application study who died on the same day as their FL relapsed, they only count for terminal events, not for recurrent event.
2) $N_i^{\mathrm{R}^*}(t)$ is constant after time D_i but can increase after C_i. That means death precludes the observation of new FL relapses, but on the contrary censoring (as lost of follow-up) does not interrupt the occurrence of new relapses, they are simply not observed.
3) We define $Y_i(t)r_i(t)$ the intensity of the recurrent event process at time t in the filtration $(\mathcal{F}_t)_{t\geqslant 0}$, given the covariate process, the frailty, and the condition $D_i \geqslant t$ (being alive just before time t),

using

$$r_i(t)\mathrm{d}t = \mathrm{d}R_i(t) = P\left(\mathrm{d}N_i^{\mathrm{R}^*}(t) = 1 | Z_i(t), \omega_i, D_i \geqslant t\right).$$

We then wish to describe the FL relapse rate among the patients currently alive. We assume as a characterization of the independent censoring $P(\mathrm{d}N_i^{\mathrm{R}}(t) = 1|\mathcal{F}_{t^-}) = Y_i(t)\mathrm{d}R_i(t) = Y_i(t)\, r_i(t)\mathrm{d}t$.

4) Similarly, we define the death intensity process $Y_i(t)\lambda_i(t)$ at time t, given the covariates, the frailty, using

$$\lambda_i(t)\mathrm{d}t = \mathrm{d}\Lambda_i(t) = P\left(\mathrm{d}N_i^{\mathrm{D}^*}(t) = 1 | Z_i(t), \omega_i, D_i \geqslant t\right).$$

Independent censoring for death then requires $P(\mathrm{d}N_i^{\mathrm{D}}(t) = 1|\mathcal{F}_{t^-}) = Y_i(t)\mathrm{d}\Lambda_i(t) = Y_i(t)\lambda_i(t)\mathrm{d}t$.

Following the model of Liu and Wolfe (2004), the joint model for the hazard functions for recurrent event ($r_i(\cdot)$) and death ($\lambda_i(\cdot)$) is

$$\begin{cases} r_i(t|\omega_i) = \omega_i r_0(t)\exp(\beta_1' Z_i(t)) = \omega_i r_i(t), \\ \lambda_i(t|\omega_i) = \omega_i^{\alpha}\lambda_0(t)\exp(\beta_2' Z_i(t)) = \omega_i^{\alpha}\lambda_i(t). \end{cases} \tag{2.1}$$

The effect of the explanatory variables is assumed to be different for recurrent and death times. The parameters β_1 and β_2 are interpretable in terms of the instantaneous probability of occurrence of the recurrent events and the terminal event, respectively, conditional on the subject's past event history and on the subject being alive. The model and the estimation can deal with external time-dependent covariates in the sense of Kalbfleisch and Prentice (2002, p. 197). The previous number of recurrent events can also be considered as an internal time-dependent covariate that requires the survival of the individual for its existence, and its path thus carries direct information on the time to failure.

The random effects ω_i (frailties) are assumed independent. The gamma frailty density is adopted here with unit mean and variance θ. The dependence between T_i^* and T_{ij} conditional on $Z_i(t)$ is solely due to the fact that the unobserved ω_i affects both the recurrent times and the death times. The common frailty parameter ω_i will take into account the heterogeneity in the data, associated with unobserved covariates.

In the traditional model, the assumption is that $\alpha = 0$ in (2.1), that is $\lambda_i(t)$ does not depend on ω_i, and thus death (or the terminal event process) is not informative for the recurrent event rate $r_i(t)$, that is the 2 rates $\lambda_i(t)$ and $r_i(t)$ are not associated, conditional on covariates. When $\alpha = 1$, the effect of the frailty is identical for the recurrent events and the terminating event. When $\alpha > 1$, the recurrent rate and the death rate are positively associated; higher frailty will result in higher risk of recurrence and higher risk of death.

In the gap timescale formulation, T_{ij} is replaced by $S_{ij} = T_{ij} - T_{ij-1}$ with $T_{i0} = 0$ for the recurrent hazard functions, and the corresponding joint model is

$$\begin{cases} r_i(s|\omega_i) = \omega_i r_0(s)\exp(\beta_1' Z_i(t)) = \omega_i r_i(s), \\ \lambda_i(t|\omega_i) = \omega_i^{\alpha}\lambda_0(t)\exp(\beta_2' Z_i(t)) = \omega_i^{\alpha}\lambda_i(t). \end{cases}$$

2.2 *Inference in the joint frailty model*

We show the expression of the full log-likelihood for calendar times (or time-to-events) and explain how to deduce it for gap times (or time-between-events). Using the time-to-events timescale, it is easy to incorporate time-varying covariates, and the likelihood must incorporate delayed entries. The length of

the time-at-risk period is the same for the 2 timescales; however, in the calendar-time formulation, the start of the at-risk period is not reset to 0 but to the actual time since entry to the study.

Contrary to the shared gamma frailty models (Rondeau *and others*, 2003), the full log-likelihood of the joint frailty model does not take a simple form because the integrals do not have a close form. Thus, using other distributions for the frailty, such as log normal or positive stable, will not induce more difficulties. Moreover, Pickles and Crouchley (1995) suggest that the results should not be sensitive to the choice of the frailty distribution.

We denote $\phi = (r_0(\cdot), \lambda_0(\cdot), \boldsymbol{\beta}, \alpha, \theta)$. The construction of the log-likelihood is detailed in Appendix A.1. We obtain the following expression of the full marginal log-likelihood in the calendar timescale:

$$l(\phi) = \sum_i \left\{ \sum_j \delta_{ij} \log r_i(T_{ij}) + \delta_i^* \log \lambda_i(T_i^*) - \log \Gamma(1/\theta) - \frac{1}{\theta} \log \theta \right.$$
$$\left. + \log \int_0^\infty \omega^{(N_i^R(T_i^*)+\alpha\delta_i^*+1/\theta-1)} \exp\left(-\omega \int_0^{T_i^*} \mathrm{d}R_i(t) - \omega^\alpha \int_0^{T_i^*} \mathrm{d}\Lambda_i(t) - \frac{\omega}{\theta} \right) \mathrm{d}\omega \right\} \quad (2.2)$$

with $T_{i0} = 0$ and $T_{in_i} = T_i^*$ (for each subject, we assume that the last observation time is a censoring time or a death time and not a relapse time), $\Lambda_i(t) = \int_0^t \lambda_i(u)\partial\, \mathrm{d}u$ the cumulative hazard function for death, with $\Lambda_i(\cdot|\omega) = \omega^\alpha \Lambda_i(\cdot)$, and $R_i(t) = \int_0^t r_i(u)\mathrm{d}u$ the cumulative hazard function for recurrent events, with $R_i(\cdot|\omega) = \omega R_i(\cdot)$.

In the gap timescale formulation, the likelihood expression is the same except that T_{ij} is replaced by $S_{ij} = T_{ij} - T_{ij-1}$ giving the expression

$$l(\phi) = \sum_i \left\{ \sum_j \delta_{ij} \log r_i(S_{ij}) + \delta_i^* \log \lambda_i(T_i^*) - \log \Gamma(1/\theta) - \frac{1}{\theta} \log \theta \right.$$
$$\left. + \log \int_0^\infty \omega^{(N_i^{\mathrm{R}}(T_i^*)+\alpha\delta_i^*+1/\theta-1)} \exp\left(-\omega \sum_{j=1}^{n_i} \int_0^{S_{ij}} \mathrm{d}R_{ij}(s) - \omega^\alpha \int_0^{T_i^*} \mathrm{d}\Lambda_i(t) - \frac{\omega}{\theta} \right) \mathrm{d}\omega \right\}.$$

3. THE SEMIPARAMETRIC PENALIZED LIKELIHOOD APPROACH

We introduced a semiparametric penalized likelihood approach to estimate the different parameters $\boldsymbol{\beta}$, α, θ, and the baseline hazard function $r_0(t)$ for recurrent events or $\lambda_0(t)$ for death times.

In most situations, it is reasonable to expect smooth baseline hazard functions, piecewise constant modeling for the hazard functions being often unrealistic. To introduce such *a priori* knowledge, we penalize the likelihood by a term which has large values for rough functions (O'Sullivan, 1988; Joly *and others*, 1998). The roughness penalty function is represented by the sum of 2 squared norms of the second derivative of the hazard functions (O'Sullivan, 1988). The penalized log-likelihood is thus defined as

$$\mathrm{pl}(r_0(\cdot), \lambda_0(\cdot), \boldsymbol{\beta}, \alpha, \theta) = l(\phi) - \kappa_1 \int_0^\infty r_0''^2(t)\mathrm{d}t - \kappa_2 \int_0^\infty \lambda_0''^2(t)\mathrm{d}t, \quad (3.1)$$

where $l(\lambda_0(\cdot), \boldsymbol{\beta}, \alpha, \eta)$ is the full log-likelihood defined in (2.2) and $\kappa \geqslant 0$ is a positive smoothing parameter which controls the trade-off between the data fit and the smoothness of the functions. Maximization of (3.1) defines the maximum penalized likelihood estimators (MPnLE) $\hat{r}_0(t)$, $\hat{\lambda}_0(t)$, $\hat{\boldsymbol{\beta}}$, $\hat{\alpha}$, and $\hat{\theta}$. We

directly use $\hat{H}^{-1}$ as a variance estimator, where H is minus the converged Hessian of the penalized log-likelihood. Furthermore, to deal with the constraint on the variance component ($\theta > 0$), we used a squared transformation and the standard error of θ was computed by the Δ-method (Knight and Xekalaki, 2000).

The estimators $\hat{r}_0(t)$ and $\hat{\lambda}_0(\cdot)$ cannot be calculated explicitly but can be approximated on the basis of splines. Splines are piecewise polynomial functions that are combined linearly to approximate a function on an interval. We use cubic M-splines, which are a variant of cubic B-splines (for more details, see Ramsay, 1988). M-splines are nonnegative and easy to integrate or differentiate. As we use cubic spline (or of order 4), the second derivative of r or λ is approximated by a linear combination of piecewise polynomial of order 2. This approximation allows flexible shapes of the hazard functions while reducing the number of parameters. If we denote $\tilde{r}(\cdot)$ an approximation to the MPnLE $\hat{r}(\cdot)$, the approximation error can be made as small as desired by increasing the number of knots. In our approach, although there are 2 different hazard functions (for recurrent events and death), we use the same basis of splines for each function, but the spline coefficients are different for the distinct functions.

We have previously shown that to obtain a good estimation of the theoretical hazard function, the more knots we used, the closer the MPnLE was to the true hazard function (Rondeau *and others*, 2003). The smoothing parameters can be chosen by maximizing a likelihood cross-validation criterion as in Joly *and others* (1998). Another approach consists in fixing the number of degrees of freedom to estimate the hazard function, as has been previously described (Rondeau *and others*, 2003; Gray, 1992). We thus use the relation linking the model degrees of freedom (mdf) and the smoothing parameter κ to evaluate the smoothing parameter: $\text{mdf} = \text{trace}([\hat{H}]^{-1}\hat{I})$ (with I the Hessian matrix of the log-likelihood computed at the MPnLE). Indeed, it is easier to specify a number of degrees of freedom to estimate a given curve, rather than to specify a smoothing parameter.

We proposed to directly maximize the observed log-likelihood (3.1) using a modified robust Marquardt (1963) optimization algorithm, which is a combination between the Newton–Raphson algorithm and the steepest descent algorithm. This algorithm is more stable than the Newton–Raphson algorithm (Fletcher, 2000) but preserves its fast convergence property near the maximum. The integrations in the full log-likelihood expression in (2.2) were evaluated using Gaussian quadrature. Laguerre polynomials with 20 points were used to treat the integration $[0, \infty)$.

4. Simulations

A simulation study of the joint frailty model was performed to evaluate the performance of the estimators and to compare a joint frailty model with a single/reduced frailty model. In order to investigate the effect of increased sample size on estimator performance, we considered 2 sample sizes with a variable number of subjects and a variable number of recurrent events by subject. There were 200 or 500 subjects and 1000 simulated data sets for each case. For each simulation run, the joint frailty model (2.1) was used. We treated the right-censored case only and used a calendar timescale representation.

For each subject i:

1) we generated the random variables ω_i, $i = 1, \ldots, N$, i.i.d. $\Gamma(1/\theta; 1/\theta)$ with $\theta = 0.5$, the variance of the random effect.
2) a fixed right-censoring variable was used, $C_i = 0.8$ ($i = 1, \ldots, N$).
3) we generated an exponential death time D_i using $\lambda_i(t|\omega_i) = \omega_i^{\alpha}\lambda_0(t)\exp(\beta_1^* Z_{1i})$ with $\lambda_0(t) = 2.0$ and $\delta_i^* = 1$ if $D_i < C_i$.
4) we generated the gap times X_{ik} using $r_i(t|\omega_i) = \omega_i r_0(t)\exp(\beta_1 Z_{1ij} + \beta_2 Z_{2ij})$ with an exponential $r_0(t) = 1.0$; the corresponding observed calendar times are $T_{ij} = \min\left(C_i, D_i, \sum_{k=1}^{j} X_{ik}\right)$, $\delta_i = 1$ if $T_{ij} = \sum_{k=1}^{j} X_{ik}$ with $T_{i0} = 0$. This simulation scheme is valid since $r_0(t)$ is constant.

To summarize, if the observed time is a recurrent event time $T_{ij} = \sum_{k=1}^{j} X_{ik}$ and $\delta_{ij} = 1$, the data generation continues; if the observed time is a censoring time $T_{ij} = C_i$, $\delta_{ij} = 0$ and $T_i^* = C_i$, $\delta_i^* = 0$; or if the observed time is a death time $T_{ij} = D_i$, $\delta_{ij} = 0$ and $T_i^* = D_i$, $\delta_i^* = 1$, the data generation stops.

Death times and recurrent event times have in common only one explanatory variable Z_{1ij}. The binary explanatory variables Z_{1ij} and Z_{2ij} were generated from a Bernoulli distribution with $P(Z = 1) = 0.5$. We set $\beta_1 = 1.0$, $\beta_2 = -0.5$, and $\beta_1^* = 0.7$. We consider 3 settings for α, setting I corresponds to $\alpha = 0.5$, setting II $\alpha = -0.5$, and setting III $\alpha = 0$.

We used cubic splines to approximate each hazard function. The number of equidistant knots was 5 for all simulations. For the first replicate of each simulation (i.e. for a given θ and sample size), we estimated κ using the cross-validation method, the same κ was used for the other 999 generated data sets. We eliminated the rare cases (less than 5%) when convergence or numerical problems occurred in the estimation of the parameters.

4.1 *Results*

The death rate ranges from 27.1% to 49.4%. The average number of observed recurrent events by subject ranges from 0.60 to 1.52 in the conducted simulation studies with a maximum fixed at 24. Between 40.5% and 70.0% of the subjects did not have a recurrent event. The results of simulation studies using a penalized likelihood estimation are summarized in Tables 1–5 of the supplementary material (available at *Biostatistics* online, http://www.biostatistics.oxfordjournals.org). The regression coefficients from the joint model were very well estimated in the 3 settings. We observe in the first setting ($\alpha = 0.5$), a bias on the regression coefficients using the simple, shared frailty model instead of the joint model. The bias on the estimates of the variance of the random effects (θ) was very small in the joint model. In setting II ($\alpha = -0.5$), that is with a negative association between recurrent events and death, we observed a significant bias using the simple frailty model ($\hat{\theta} = 0.365$ with $N = 500$ and $\hat{\theta} = 0.359$ with $N = 200$). This demonstrates that ignoring the dependence between the terminal and the recurrent events can lead to erroneous results. It can be seen that in the 3 settings for the joint model, $\hat{\alpha}$ was unbiased. As expected, in all simulations the estimates for the standard errors were smaller for $N = 500$ than for $N = 200$. In setting III ($\alpha = 0$), the 2 models (joint and reduced) are valid and give similar results.

We also increased the degree of dependency between the recurrent events and death with $\alpha = 1.0$ (results not shown). We observed larger differences between the joint and the reduced models. The results increasing the number of recurrent events by subject are summarized in the supplementary material available at *Biostatistics* online. We obtained a clear improvement in the estimations.

We evaluated the estimation of the survival functions and the hazard functions for the recurrent events using the mean integrated squared error (Härdle, 1990). More details and results are described in Table 5 of the supplementary material available at *Biostatistics* online. We observed that our penalized likelihood estimation gives good estimates for the survival and hazard functions. This also illustrates that better estimations are obtained using the joint model instead of the reduced model.

5. FL, RELAPSES AND DEATH

The scope of our investigation was to estimate a joint model to describe the risk factors associated with the recurrences of FLs and death, taking into account the informative censoring by death. If the death times depend on the recurrent event times, it is necessary to use a joint model to make valid inferences. Another important point was to study whether the subjects who are at higher risk of FL recurrences tend to be at an elevated risk of death or inversely at a lower risk of death. This approach allows us to quantify the association between the recurrent events of FL and death. From 1965 to 2000, 409 patients with FL (190

males, 46.5%) were monitored at Institut Bergoni, a regional comprehensive cancer center in south west France. All the patients were prospectively included by one research assistant in a clinical, histological, and biologic database. A FL recurrence was defined as the first clinical sign of FL. Patients came to the hospital for a routine visit every 4 months for 3 years, every 6 months for 2 years, once a year for 5 years, and then every 2–3 years. Some other spontaneous inter-visits could take place. The FL recurrence was or was not detected at each visit.

Information on patient gender, age, and the number of recurrent events or deaths is given in Table 1. A total of 249 (60.9%) patients died during the follow-up and 49.1% of subjects did not have a recurrent event. For 2 subjects who died on the same day as their recurrence, they only count for terminal events, not for recurrent event. The follow-up period thus varied between 11 days and 30 years. The median follow-up of surviving patients was 9.8 years. Table 1 shows that the older subjects ($\geqslant$60 years) have fewer recurrences but more deaths. This would suggest that older subjects could die before developing a recurrence.

The number of recurrences ranges from 0 to 4, averaging 0.71 per patient. Episodes were categorized into 1, 2, 3, 4, or 5 corresponding to the number of observation times for each subject. The fifth episode number corresponds to a censoring time or death.

Figure 1 presents the survival functions following successive recurrences of FL. This figure does not illustrate clear trends in the evolution of the risk of recurrence.

We modeled the joint distribution of the inpatient recurrences and the survival times (model (2.1)) using the fact that we wished to describe the relapse's rate among patients currently alive. The person-specific frailty term represents the effect of unmeasured factors on the chances of both recurrence and death. The time variable ("gap time") was the time since the latest episode. We expected that the hazard rate would not change substantially over time but as a function of the time since the last event. The covariates included in the analyses were the number of prior episodes (as an internal time-dependent covariate), gender, age at diagnosis (60 years or older versus younger than 60 years), the tumor burden with the Ann Arbor stage (III–IV versus I–II), the number of nodal areas involved ($\geqslant$4 versus $<$4), and the initial treatment at diagnosis classified as "any type of radiotherapy" versus "chemotherapy alone, another treatment, or no treatment." We did not adjust for the serum lactate dehydrogenase levels (for tumor aggressiveness) or for the hemoglobin levels (consequences of the lymphoma on the host) even if

Table 1. *Number of FL recurrences and death according to age and gender*

	No. of patients	No. of deaths	No. of recurrences since diagnosis				
			0	1	2	3	4
Male	190	116	94	76	14	5	1
(%)	100	61.1	49.5	40.0	7.4	2.6	0.5
Female	219	133	104	74	30	6	5
(%)	100	60.7	47.5	33.8	13.7	2.7	2.2
Diagnosis age $<$ 60 years	208	96	96	79	25	6	2
(%)	100	46.2	46.2	37.9	12.0	2.9	0.9
Diagnosis age $\geqslant$ 60 years	201	153	102	71	19	5	4
(%)	100	76.1	50.7	35.3	9.5	2.5	2.0
Total	409	249	198	150	44	11	6
(%)	100	60.8	48.4	36.7	10.8	2.7	1.5

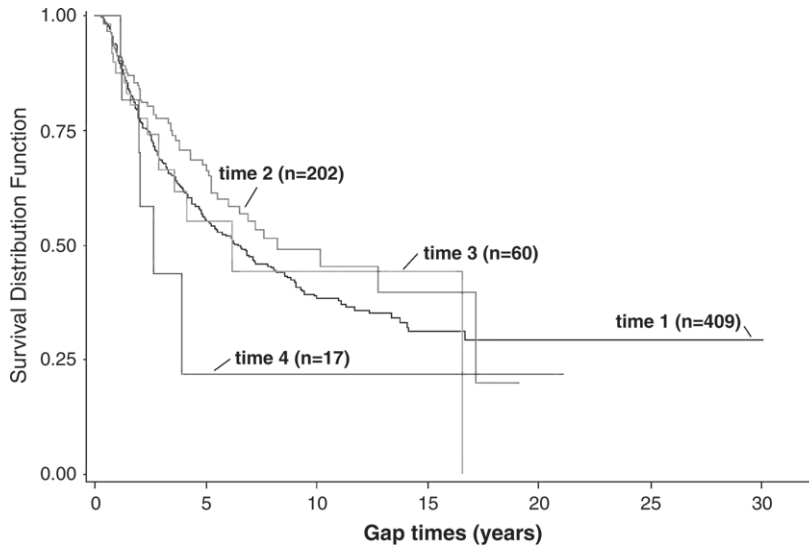

Fig. 1. Survival functions for successive FL recurrences.

they are also involved in the FL International Prognostic Index (Solal-Celigny *and others*, 2004), because there were too many missing data.

The statistical softwares used were R and the library "frailtypack" (version 2.0-0) with the function "frailtyPenal" for the shared frailty models (Rondeau and Gonzalez, 2005), and a fortran program was developed for the joint modeling and will be inserted in the frailtypack. Penalized likelihood maximization was used. In the reduced models, κ_1 and κ_2 were evaluated using the cross-validation method, thereafter this value was used in the joint model.

Table 2 presents the results using adjusted joint models and reduced shared frailty models. The rate of recurrence increased with age (age $\geqslant$ 60 years, RR = 1.93, 95% confidence interval [CI] (1.39–2.67)) was higher for women and was associated with the stage of the tumor. These 3 effects were underestimated using the reduced shared frailty model instead of the joint model. It is clear that ignoring the dependence between the terminal event and recurrent events resulted in significant biases in the independent shared frailty model compared to the joint model. For instance, the effect of gender was greater using the joint frailty model compared to the reduced shared frailty model (1.53 versus 1.16). As a result, some covariates can be incorrectly observed as nonsignificant variables using a simple, reduced shared frailty model which does not take into account the informative censoring by death. Age, gender, and the stage of the tumor were also identified as significant prognostic factors.

The positive value of $\alpha = 2.17$ in the joint model indicates that the incidence of recurrences is positively associated with death after controlling for the number of past events. Patients with a large frailty value tend to have a high rate of recurrence after any episode, whatever the number of past relapses is. The same positive association was also obtained without the adjustment for the number of past relapses ($\alpha = 1.63$).

The number of previous episodes influenced the risk of recurrence or death given the frailty, however, it was significantly associated with a decreased risk of recurrence (RR = 0.60, 95% CI (0.49–0.73)) and a decreased risk of death (RR = 0.22, 95% CI (0.15–0.31)). This protective effect of the number of recurrences on the risk of death (RR = 0.22) could be explained by the probable existence of at least 2 different types of FL based on clinical observation: FL with large tumor mass and FL with several small and disseminated nodes. Patients of the first group often behave more aggressively with higher risk

Table 2. *Analysis of the recurrences and death for FLs using gap times*

Covariate	Joint model		Reduced model	
	RR	95% CI	RR	95% CI
For recurrences				
Sex				
Men	1		1	
Women	1.53	(1.13–2.06)	1.16	(0.89–1.50)
Age				
Younger than 60 years	1		1	
60 years or older	1.93	(1.39–2.67)	1.45	(1.09–1.94)
Original Ann Arbor stage				
I–II	1		1	
III–IV	1.43	(1.05–1.97)	1.22	(0.93–1.61)
Number of prior episodes	0.60	(0.49–0.73)	0.78	(0.55–1.09)
For survival				
Sex				
Men	1		—	
Women	2.92	(1.69–5.02)		
Age				
Younger than 60 years	1		—	
60 years or older	8.79	(4.85–15.94)		
Original Ann Arbor stage				
I–II	1		—	
III–IV	3.68	(2.12–6.41)		
Number of prior episodes	0.22	(0.15–0.31)	—	
θ(SE)	1.19 (0.09)		0.41 (0.34)	
α(SE)	2.17 (0.22)			

RR, relative risk; CI, confidence interval; SE, standard error of the mean.

of treatment failure and death in the short term, while the others generally have a slow progression, a good response to treatment, and are more often in remission (partial or complete). These last patients will correspond to patients with a higher risk of recurrences but with a longer survival. However, this assumption remains to be formerly validated.

These models attempt to capture the effect of process history through a single covariate, which indicates the number of previous recurrences of FL occurred by time t. An alternative would be to consider a model without this variable but with $r_0(t)$ replaced by $r_{0j}(t)$. This stratified analysis can be easily conducted with $j = 2$ or 3 using the MPnLE but can become less tractable with more recurrences by subject simply because the number of parameters to estimate will increase with the number of different baseline hazard functions. We did not perform a stratified analysis. Models including the number of prior episodes as category variables (3 binary variables for 5 classes) confirmed the above estimations.

The variance of the frailties is a measure of the heterogeneity of the observations. The recurrence rate varied greatly among patients ($\hat{\theta} = 1.19$ in the joint model), even after adjustment for the individual variables. We observed a greater heterogeneity using the joint frailty model.

Figure 2 illustrates the hazard of recurrence using the joint or the shared frailty model. We did not present the hazard function after 15 years because of the lack of information in the data set after this

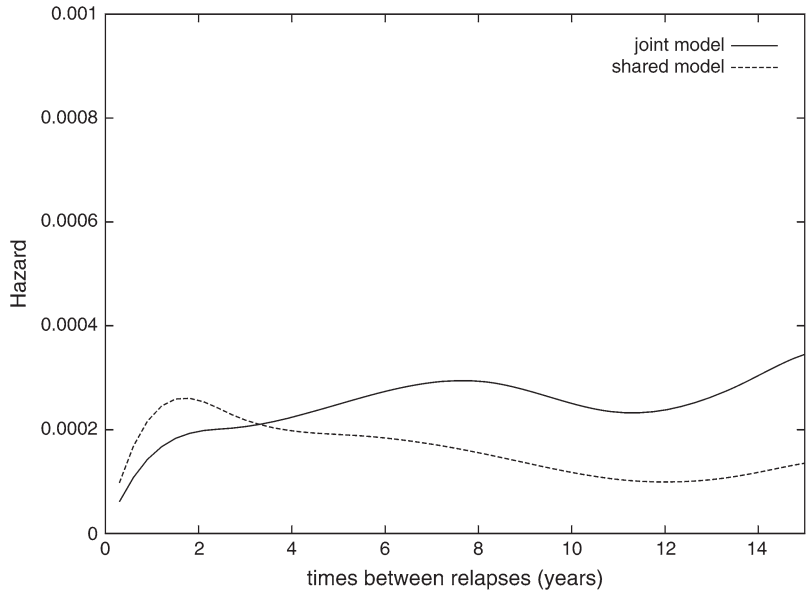

Fig. 2. Joint modeling and reduced shared frailty modeling for FL recurrence hazards.

period. We observe that the hazard function was underestimated when using the shared frailty model because this model does not correctly model death. Indeed, when $\alpha > 0$, the frail subjects with higher failure risk are also frail subjects for death and are more likely to die before we observe their failures. The recurrence risk is then underestimated. In contrast, when $\alpha < 0$ the recurrence risk is overestimated.

The analyses were solely based on the gap timescales, and we studied how the hazard rate evolves after an event has taken place; time is reset to zero after a recurrence. Hence, one neglects the recurrence history when describing the inter-recurrences time. If it is expected that the recurrence rate changes as a function of time since inclusion in the cohort, the analysis can be based on the calendar timescales, that is the time since inclusion in the cohort. As the model can be set up in a counting process framework, it is easy to incorporate time-varying covariates and delayed entry. The joint model using the calendar timescale led to equivalent results but slightly smaller gender effect (RR = 1.34, 95% CI (0.98–1.84)) or stage effect (RR = 1.37, 95% CI (0.99–1.90)), and these covariates were no longer significant.

6. CONCLUSION

This paper proposed a method of estimation in joint modeling for 2 survival processes which enables us to study the joint evolution over time of recurrent events and death and gives unbiased and efficient parameters. The most commonly used estimation procedure in the joint models for survival events is the EM algorithm. The strength of this article is that it shows how MPnLE can be applied to nonparametric estimation of the continuous baseline hazard functions in a joint frailty model with right-censored data and delayed entry. The method of estimation proposed and the program used also have the advantage of not being time consuming even for large applications. For instance, the joint model presented in Table 2 used 50 s of CPU time. Valid and rapid inferences under minor assumptions are then obtained.

A major advantage of joint frailty models is their ability to analyze simultaneously the recurrent events data and a terminating event that can be associated and to assess their degree of dependence. We have shown by simulation that using a reduced shared frailty model instead of a joint frailty model when

there is a significant dependence between the 2 processes leads to unreliable estimates, with regression factors falsely nonsignificant or with an underestimation of the recurrence risk. This implies that the noninformative censoring of the recurrent event process by death needs to be taken into account in survival analysis to obtain accurate inferences. In general, omission of important features of dependence in the data from the models we estimate results in biased and inefficient estimates. On the other hand, if no association exists between the 2 processes, a more restricted model might be acceptable, such as a reduced shared frailty model.

The marginal model has already been proposed and compared to the frailty model to deal with the dependence between recurring events and death (Schaubel and Cai, 2005). This joint frailty approach compared to the marginal approach has the advantage to quantify this dependence. The frailty model is implicitly conditional on the previously described filtration $(\mathcal{F}_t)_{t\geqslant 0}$ and the frailty term; marginal models are in this sense marginal as opposed to conditional and can be seen as having averaged over all possible filtrations. Furthermore, the regression coefficients of the frailty model are interpreted conditionally, given the unobserved frailty, and do not have a clear interpretation marginally since the marginal RR does not equal $\exp(\beta)$.

We applied our approach to the joint modeling of FL recurrences and death, and we found a positive association between these 2 processes. The censoring by death was informative for the risk of recurrences, and this was taken into account in the joint modeling. There are cases when a history of higher rates of recurrent events implies an expected delay in the favorable termination event such as cure or discharge from hospital. The flexible model that we used, introduced by Liu *and others* (2004), can accommodate this kind of negative ($\alpha < 0$) relationship between recurrent event history and risk of termination.

Other approaches allowing for additional correlation structures on the random effects may provide valuable insight for future research.

Acknowledgments

Conflict of Interest: None declared.

APPENDIX

A.1 *Construction of the full log-likelihood for the joint frailty model (2.1) with calendar timescale*

We denote T_{ij} the jth follow-up time for subject i and δ_{ij} the failure indicator for the recurrent events. Similarly, we define $T_i^* = \min(D_i, C_i)$ the last follow-up time for subject i and the death indicator $\delta_{ij}^* = I(D_i < C_i)$.

The marginal contribution to the likelihood $L_i(r_0(\cdot), \lambda_0(\cdot), \boldsymbol{\beta}, \alpha, \theta) = L_i(\phi)$ for subject i and for $j = 1, \ldots, n_i$ is $L_i(\phi) = \int_\omega L_i(\phi|\omega) f(\omega) \mathrm{d}\omega$

1) The conditional distribution of the survival times given ω_i is the product of the individual contributions:

$$L_i(\phi|\boldsymbol{\omega}_i) = \prod_{j=1}^{n_i} \left[\mathrm{d}R_i(T_{ij}|\omega_i)^{\delta_{ij}} \times \exp\left(-\omega \sum_{j=1}^{n_i} \int_{T_{ij-1}}^{T_{ij}} \mathrm{d}R_i(t) \right) \right]$$
$$\times \mathrm{d}\Lambda_i(T_i^*|\omega_i)^{\delta_i^*} \times \exp\left(-\omega^\alpha \int_0^\infty Y_i(t) \mathrm{d}\Lambda_i(t) \right).$$

2) The probability density function for the random effects ω is $f(\omega) = \frac{\omega^{(1/\theta-1)} \exp(-\omega/\theta)}{\Gamma(1/\theta)\theta^{1/\theta}}$.

3) Using the previous expressions, the ith marginal contribution to the likelihood is obtained by integrating out the random effects:

$$L_i(\phi) = \frac{\prod_{j=1}^{n_i} \mathrm{d}R_i(T_{ij})^{\delta_{ij}} \times \mathrm{d}\Lambda_i(T_i^*)^{\delta_i^*}}{\Gamma(1/\theta)\theta^{1/\theta}} \times \int_0^\infty \omega^{\left(N_i^R(T_i^*)+\alpha\delta_i^*+\frac{1}{\theta}-1\right)} \times \exp\left(-\omega \sum_{j=1}^{n_i} \int_{T_{ij-1}}^{T_{ij}} \mathrm{d}R_i(t) - \omega^\alpha \int_0^\infty Y_i(t)\mathrm{d}\Lambda_i(t) - \frac{\omega}{\theta}\right) \mathrm{d}\omega.$$

We then obtain the expression (2.2) of the full log-likelihood by using

$$l(\phi) = \log \prod_{i=1}^{N} L_i(\phi).$$

REFERENCES

ANDERSEN, P. K., BORGAN, O., GILL, R. D. AND KEIDING, N. (1993). *Statistical Models Based on Counting Processes*. New York: Springer.

COOK, R. J. AND LAWLESS, J. F. (2002). Analysis of repeated events. *Statistical Methods in Medical Research* **11**, 141–166.

DUCHATEAU, L., JANSSEN, P., KEZIC, I. AND FORTPIED, C. (2003). Evolution of recurrent asthma event rate over time in frailty models. *Applied Statistics* **52**, 355–363.

FLETCHER, R. (2000). *Practical Methods of Optimization*, 2nd edition. New York: John Willey & Sons.

GHOSH, D. AND LIN, D. Y. (2003). Semiparametric analysis of recurrent events data in the presence of dependent censoring. *Biometrics* **59**, 877–885.

GRAY, R. J. (1992). Flexible methods for analyzing survival data using splines, with applications to breast cancer prognosis. *Journal of the American Statistical Association* **87**, 942–951.

HÄRDLE, W. (1990). *Smoothing Techniques with Implementation in S*. Berlin: Springer.

HUANG, X. AND WOLFE, R. A. (2002). A frailty model for informative censoring. *Biometrics* **58**, 510–520.

JOLY, P., COMMENGES, D. AND LETENNEUR, L. (1998). A penalized likelihood approach for arbitrarily censored and truncated data: application to age-specific incidence of dementia. *Biometrics* **54**, 185–194.

KALBFLEISCH, J. D. AND PRENTICE, R. L. (2002). *The Statistical Analysis of Failure Time Data*. New York: Wiley.

KNIGHT, K. AND XEKALAKI, E. (2000). *Mathematical Statistics*, chapters 3 and 5, New York: Chapman & Hall.

LANCASTER, T. AND INTRATOR, O. (1998). Panel data with survival: hospitalization of HIV-positive patients. *Journal of the American Statistical Association* **93**, 46–53.

LI, Q. H. AND LAGAKOS, S. W. (1997). Use of the Wei-Lin-Weissfeld method for the analysis of a recurring and a terminating event. *Statistics in Medicine* **16**, 925–940.

LIU, L., WOLFE, R. A. AND HUANG, X. (2004). Shared frailty models for recurrent events and a terminal event. *Biometrics* **60**, 747–756.

MARQUARDT, D. (1963). An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal on Applied Mathematics* **11**, 431–441.

O'SULLIVAN, F. (1988). Fast computation of fully automated log-density and log-hazard estimators. *SIAM Journal of Science and Statistical Computation* **9**, 363–379.

PICKLES, A. AND CROUCHLEY, R. (1995). A comparison of frailty models for multivariate survival data. *Statistics in Medicine* **14**, 1447–1461.

RAMSAY, J. O. (1988). Monotone regression splines in action. *Statistical Science* **3**, 425–461.

RONDEAU, V., COMMENGES, D. AND JOLY, P. (2003). Maximum penalized likelihood estimation in frailty models. *Lifetime Data Analysis* **9**, 139–153.

RONDEAU, V. AND GONZALEZ, J. R. (2005). frailtypack: a computer program for the analysis of correlated failure time data using penalized likelihood estimation. *Computer Methods and Programs in Biomedicine* **80**, 154–164.

SCHAUBEL, D. E. AND CAI, J. (2005). Analysis of clustered recurrent event data with application to hospitalization rates among renal failure patients. *Biostatistics* **6**, 404–419.

SINHA, D. AND MAITI, T. (2004). A Bayesian approach for the analysis of panel-count data with dependent termination. *Biometrics* **60**, 34–40.

SOLAL-CELIGNY, P., ROY, P., COLOMBAT, P., WHITE, J., ARMITAGE, J. O., ARRANZ-SAEZ., R., AU, W. Y., BELLEI, M., BRICE, P., CABALLERO, D. (2004). Follicular lymphoma international prognostic index. *Blood* **104**, 1258–1265.

WEI, L. J., LIN, D. Y. AND WEISSFELD, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of the American Statistical Association* **84**, 1065–1073.

[*Received April 11, 2006; first revision September 18, 2006; second revision December 11, 2006; accepted for publication December 20, 2006*]