Multivariate frailty models for two types of recurrent events with a dependent terminal event: Application to breast cancer data

Yassin Mazroui*[,1,2], **Simone Mathoulin-Pélissier**[3,4], **Gaetan MacGrogan**[3], **Véronique Brouste**[3], and **Virginie Rondeau**[1]

[1] INSERM, ISPED, Centre INSERM U-897-Epidemiologie-Biostatistique, 146 rue Léo Saignat, 33076, Bordeaux Cedex, France
[2] Université Bordeaux Segalen, ISPED, Centre INSERM U-897-Epidemiologie-Biostatistique, Bordeaux, F-33000, France
[3] Unité de recherche et d'épidemiologie cliniques - Institut Bergonié, Bordeaux, F-33000, France
[4] Université Bordeaux Segalen, ISPED, INSERM CIC-EC7, Bordeaux, F-33000, France

Received 2 October 2012; revised 17 April 2013; accepted 21 April 2013

Individuals may experience more than one type of recurrent event and a terminal event during the life course of a disease. Follow-up may be interrupted for several reasons, including the end of a study, or patients lost to follow-up, which are noninformative censoring events. Death could also stop the follow-up, hence, it is considered as a dependent terminal event. We propose a multivariate frailty model that jointly analyzes two types of recurrent events with a dependent terminal event. Two estimation methods are proposed: a semiparametrical approach using penalized likelihood estimation where baseline hazard functions are approximated by M-splines, and another one with piecewise constant baseline hazard functions. Finally, we derived martingale residuals to check the goodness-of-fit. We illustrate our proposals with a real dataset on breast cancer. The main objective was to model the dependency between the two types of recurrent events (locoregional and metastatic) and the terminal event (death) after a breast cancer.

Keywords: Breast cancer; Frailty models; Penalized likelihood; Recurrent events.

Additional supporting information may be found in the online version of this article at the publisher's web-site

1 Introduction

Relapses and death are often the events of interest for long-term diseases. In statistics, relapses are treated as recurrent events and death as a terminal event. These last decades, researchers have developed methods for analyzing such events separately or jointly. Proportional hazard (PH) models (see Cox, 1972), were first developed for an event of interest that could happen once. This model implicitly assumes a homogeneous population. Frailty models (see Clayton, 1978; Vaupel et al., 1979), which are extensions of PH models, aim to account for potential heterogeneity caused by unmeasured prognostic factors and interrecurrent event time dependency.

A frailty is an unobserved random proportionality factor that modifies the hazard function of an individual, or of a group of individuals. Frailty models (Duchateau and Janssen, 2008) can also be used with univariate data, for instance to analyze the hazard of a single event per subject but this means that frailty accounts for a latent heterogeneity among individuals. In the context of multivariate data,

*Corresponding author: e-mail: Yassin.Mazroui@isped.u-bordeaux2.fr

such as recurrent events/clustered data, the frailty is specific to individuals/groups and accounts for the interrecurrence/intercluster dependency. The random effect explains the dependence in the sense that, given the frailty, the events are independent. In other words, the event times are conditionally independent given frailties. We focus here on recurrent events and hence we consider individual frailties.

The occurrence of an event may be due to biological weakening or damage, which is why the notion of frailty was introduced. A review of the existing methods for the analysis of recurrent events has been detailed in Cook and Lawless (2007). These methods assume that the follow-up period is independent of the underlying recurrent process. Later, joint modeling approaches relying on frailty models (see Lancaster and Intrator, 1998; Huang and Wolfe, 2002), or marginal models (see Schaubel and Cai, 2005) considered occurrences of events as informative censoring or terminal event as a dependent processes. Several authors such as Liu et al. (2004), Rondeau et al. (2007), Huang and Liu (2007), Mazroui et al. (2012) have proposed a more flexible approach allowing frailty to differ between the risk of recurrences and the risk of a terminal event.

Nevertheless, for some diseases, there are several, different but related types of recurrent events of interest. A class of semiparametric marginal models (see Cai and Schaubel, 2004) has been proposed for multiple type recurrent events without considering a terminal event. A generalization of the approach of Wang et al. (2001) has been presented in Zhu et al. (2010) with an arbitrary structure for both the relationship between the recurrent events and the terminal event and the effect of covariates on the terminal event. A semiparametric additive model was proposed in Zeng and Cai (2010) where the dependency between recurrent events and the terminal event is nonparametric. In these recent approaches, dependency between events is considered, but not of interest.

Multiple event modeling could also be done using multistate models such as in Broët et al. (1999), Putter et al. (2006), De Bock et al. (2009). However, in most of these approaches, only one part of the information is used, as for instance only the first locoregional relapse and not the successive ones. A drawback considering multistate models is that they do not provide direct information about the dependence between events of the same type or not. Clearly, the dependence between events is of interest in this article. We can also notice that multistate models are often not adapted for multiple recurrent event data because this would mean that too many states should be considered and then computational issues would certainly appear.

In this article, we develop a multivariate frailty model for recurrent events in the presence of a dependent terminal event, with right-censored survival data. Hence, the relationship between disease recurrences and survival can be assessed with random effects, frailties, or other dependency parameters that add more flexibility. Our model considers the natural history of the disease beyond the first diagnosis and also accounts for covariate effects. For instance, after a first breast cancer, the evolution toward locoregional recurrences or metastases and their association with death is studied. 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). The proposed approach can deal with both timescales. In some papers such as in Liu et al. (2004), Huang and Liu (2007), the Monte Carlo EM algorithm is used to estimate the hazard functions and the parameters, which could be time-consuming. Furthermore, these methods, proposed to estimate the cumulative hazard functions, cannot directly estimate the smooth hazard function, which often has a meaningful interpretation in epidemiological studies. We propose two estimation methods: maximization of likelihood for models with a parametric (piecewise constant) baseline hazard function and maximization of the penalized likelihood for models with baseline hazard functions approximated by M-splines (see Rondeau et al., 2007).

Therefore, the aim of this article is to provide methods for jointly analyzing two types of recurrent events with a terminal event, to measure potential dependency between each type of event and to estimate the influence of prognostic factors. This work was motivated by the analysis of different types of recurrences, locoregional or metastatic, and death for patients after a breast cancer diagnosis. Researchers have shown that the risk of death is increased after a metastatic relapse (see O'Shaughnessy, 2005). Further, after a locoregional relapse, the risk of metastatic recurrence is also increased (see Montagna et al., 2011). The dependency between locoregional relapse and the risk of death is observed

for young women (see Elkhuizen et al., 1998), but it is not for women of any age (see Monteiro Grillo et al., 2005). Wapnir et al. (2006) showed also that locoregional relapses after breast-conserving surgery are associated with increased risk of distant disease and death. These papers used the Cox model with one event of interest, death or distant metastasis. Although popular, this model is quite limited to establish such associations. In fact, by using it, the authors did not use the whole available information.

In the next section, we introduce the multivariate frailty model and propose two estimating procedures for the model parameters. Simulations studies to evaluate the proposed method in finite samples are conducted in Section 3. We illustrate the method in Section 4 by applying it to a breast cancer relapse study. Finally, Section 5 is devoted to concluding remarks and discussion.

2 Joint modeling framework for different types of events

In this section, we present the multivariate frailty model and the estimation methods as it was suggested in Huang and Wolfe (2002).

2.1 The model

For each individual i, $i = 1, \ldots, N$, we consider the two types of recurrent event times $X_{ij}^{(l)}$, $j = 1, \ldots, n_i^{(l)}$ since the initiation of the processes; $l \in \{1, 2\}$ indicates the type of recurrent event. The number $n_i^{(l)}$ of observations for recurrent events of type l is a random variable. Each individual is censored by the terminal event time $T_i^* = \min(C_i, D_i)$ which could be a noninformative censoring C_i or the death D_i. Censoring events other than death are treated as independent events, that is, as noninformative censoring. The considered event time vectors are $T_{ij}^{(l)} = \min(X_{ij}^{(l)}, C_i, D_i)$, $j = 1, \ldots, n_i^{(l)}$. We consider event times since the diagnosis, that is why, for each individual i, T_{i0} is assumed equal to 0. We denote the event indicators $\delta_{ij}^{(l)} = I(T_{ij}^{(l)} = X_{ij}^{(l)})$, and the death indicator $\delta_i^* = I(T_i^* = D_i)$. The recurrent gap times $S_{ij}^{(l)}$ represent the duration between two consecutive events of type l: $S_{ij}^{(l)} = T_{ij}^{(l)} - T_{i(j-1)}^{(l)}$. We actually observe $\{T_{ij}^{(l)}$ (or $S_{ij}^{(l)}$ in the case of gap times); T_i^*; $\delta_{ij}^{(l)}, l \in \{1, 2\}$; $\delta_i^*\}$. Let $N_i^{R(l)^*}(t)$ count the number of recurrent events of type l for individual i over the interval $(0, t]$, $i = 1, \ldots, N$. Because of censoring, it is impossible to observe the true number of recurrent events experienced by the individual i. Actually, we observe the processes $N_i^{R(l)}(t) = N_i^{R(l)^*}(\min(T_i^*, t))$ which count the observed numbers of recurrent events of type l. Similarly, denote by $N_i^{D^*}(t) = I(D_i \le t)$ and $N_i^D(t) = I(T_i^* \le t, \delta_i^* = 1)$ the actual and the observed death indicator by time t, respectively. Furthermore, let $Y_i(t) = I(t \le T_i^*)$ denote whether or not the individual i is at risk of an event at time t. Over the small interval $[t, t + dt)$, the number of recurrent events that occur for subject i is $dN_i^{R(l)^*}(t) = N_i^{R(l)^*}((t + dt)^-) - N_i^{R(l)^*}(t^-)$ and the number of observed recurrent events is $dN_i^{R(l)}(t) = Y_i(t) dN_i^{R(l)^*}(t)$, $l \in \{1, 2\}$. We can notice that $n_i^{(l)} = N_i^{R(l)}(T_i^*)$. The history of the i-th process up to time t is denoted by: $\mathcal{H}_{it} = \sigma\{Y_i(h), N_i^{R(l)}(h), l \in \{1, 2\}, N_i^D(h), Z_i(h), 0 \le h \le t\}$, $i = 1, \ldots, N$, where $Z_i(h)$ is a vector of possibly time-dependent covariates. We denote the filtration by $\mathcal{F}_{it} = \sigma\{H_{it}, u_i, v_i\}$, $i = 1, \ldots, N$ such as in Liu et al. (2004). The underlying intensity processes are jointly dependent through two correlated random effects u_i, v_i which account for the nonobserved heterogeneity, the interrecurrence dependencies and the dependency between different event types. We assume that the two recurrent and the terminating processes are continuous, which means that two different events cannot happen at the same time. In the case of simultaneous death and recurrent event types, we consider that the recurrent event happens first in the small interval $[t, t + dt)$. We also assume that the actual recurrent event processes $N_i^{R(l)}(t)$, $l \in \{1, 2\}$, remain at their last value after death time D_i but can increase after the censoring time C_i. That means death precludes the observation of new recurrent events. On the contrary, a censoring, such as end of study or lost to

follow-up, does not interrupt the occurrence of new recurrent events that are simply not observed. The recurrent event intensity processes at time t are (for $l \in \{1, 2\}$): $Y_i(t)r_i^{(l)}(t)dt = P(dN_i^{R(l)}(t) = 1|\mathcal{F}_{it^-})$, where $r_i^{(l)}(t)dt = P(dN_i^{R(l)^*}(t) = 1|Z_i(t), u_i, v_i, D_i > t^-)$. The death intensity process at time t is $Y_i(t)\lambda_i(t)dt = P(dN_i^D(t) = 1|\mathcal{F}_{it^-})$, where $\lambda_i(t)dt = P(dN_i^{D^*}(t) = 1|Z_i(t), u_i, v_i, D_i > t^-)$. Finally, we model the intensity functions of counting processes for the two types of recurrent events and the terminal event observed processes given that the individual is still alive.

The multivariate frailty model for two types of recurrent events with a terminal event is (in the calendar or time-to-event timescale):

$$\begin{cases} r_i^{(1)}(t|u_i, v_i) = r_0^{(1)}(t)\exp\left(\boldsymbol{\beta}_1^{'}Z_i(t) + u_i\right) = \exp(u_i)r_i^{(1)}(t) & \text{(rec. of type 1)} \\ r_i^{(2)}(t|u_i, v_i) = r_0^{(2)}(t)\exp\left(\boldsymbol{\beta}_2^{'}Z_i(t) + v_i\right) = \exp(v_i)r_i^{(2)}(t) & \text{(rec. of type 2)} \\ \lambda_i(t|u_i, v_i) = \lambda_0(t)\exp\left(\boldsymbol{\beta}_3^{'}Z_i(t) + \alpha_1 u_i + \alpha_2 v_i\right) = \exp(\alpha_1 u_i + \alpha_2 v_i)\lambda_i(t) & \text{(death),} \end{cases} \tag{1}$$

where $r_0^{(l)}(t)$, $l \in 1, 2$, and $\lambda_0(t)$ are, respectively, the recurrent and terminal event baseline hazard functions, and $\boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \boldsymbol{\beta_3}$ the regression coefficient vectors associated with $Z_i(t)$, the covariate vector. The covariates could be different for the different event hazard functions and may be time-dependent. We consider that death stops new occurrences of recurrent events of any type, hence given $t > D$, $dN^{R(l)*}(t)$, $l \in 1, 2$, takes the value 0. Thus, the terminal and the two recurrent event processes are not independent or even conditional upon frailties and covariates. We consider the hazard functions of recurrent events among individuals still alive. The three components in the above multivariate frailty model are linked together by two Gaussian and correlated random effects u_i, v_i:

$$(u_i, v_i)^T \sim \mathcal{N}\left(\mathbf{0}, \begin{pmatrix} \theta & \rho\sqrt{\theta\eta} \\ \rho\sqrt{\theta\eta} & \eta \end{pmatrix}\right).$$

The random effect u_i is shared by the hazard of the recurrent events of type 1 and the hazard of the terminal event. Hence, the variance of u_i (θ) specifies the dependency between occurrences of the recurrent events of type 1 and the terminal event and also the interrecurrence dependency. Similarly, the variance of v_i (η) specifies the dependency between occurrences of the recurrent events of type 2 and the terminal event and also the interrecurrence dependency. The parameter α_l, $l \in 1, 2$, assesses the sign and the strength of the dependency between the l-th type of recurrent events and the terminal event, and informs whether the terminal event and these recurrent events are really dependent. A high value of α_1 (resp. α_2) and a high value of the variances θ (resp. η) illustrates a strong dependency between recurrent events of type 1 (resp. 2) and terminal events. A high absolute value of the correlation coefficient $\rho = \text{Corr}(u_i, v_i)$ corresponds to a strong dependency between the two types of recurrent events. Note that $\rho \in [-1, 1]$. A correlation coefficient $\rho = 0$ means that the occurrences of these two recurrent events for the same subject are independent of each other. If the α_1 (resp. α_2) and/or the variance θ (resp. η) are not significantly different from 0 then this means that the terminal event and the recurrent events of type 1 (resp. 2) are independent. In the gap-time timescale, the multivariate joint frailty model for two types of recurrent events and terminal event is

$$\begin{cases} r_i^{(1)}(t|u_i, v_i) = r_0^{(1)}\left(t - T_{iN_i^{(1)}(t-)}\right)\exp\left(\boldsymbol{\beta}_1^{'}Z_i(t) + u_i\right) & \text{(rec. of type 1)} \\ r_i^{(2)}(t|u_i, v_i) = r_0^{(2)}\left(t - T_{iN_i^{(2)}(t-)}\right)\exp\left(\boldsymbol{\beta}_2^{'}Z_i(t) + v_i\right) & \text{(rec. of type 2)} \\ \lambda_i(t|u_i, v_i) = \lambda_0(t)\exp\left(\boldsymbol{\beta}_3^{'}Z_i(t) + \alpha_1 u_i + \alpha_2 v_i\right) & \text{(death).} \end{cases} \tag{2}$$

The time $t - T_{iN_i^{(l)}(t^-)}$ corresponds to the duration between the time t and the previous recurrent event time. We can notice that, conditionally on frailties, on α_1, α_2 and on covariates $Z_i(t)$, recurrent gap times $S_{ij}^{(l)} = T_{ij}^{(l)} - T_{i(j-1)}^{(l)}$ are independent, with $T_{i0} = 0$, the time at entry into the study.

2.2 Estimation

Maximization of the log-likelihood and maximization of the penalized log-likelihood are the estimation methods proposed to estimate the parameters $\Phi = (r_0^{(1)}(\cdot), r_0^{(2)}(\cdot), \lambda_0(\cdot), \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \theta, \eta, \rho, \alpha_1, \alpha_2)$. The first one is used for parametric baseline hazard functions (piecewise constant or Weibull) and the second one is used when baseline hazard functions are semiparametric and approximated by M-splines. We directly use $\widehat{H}^{-1}$ as a variance–covariance matrix estimator, where H is minus the converged Hessian of the penalized log-likelihood or the log-likelihood, depending on the estimation method.

We also use the robust sandwich estimator $\widehat{H^{-1}IH^{-1}}$, I is the Fisher information matrix of the nonpenalized likelihood. $\widehat{H^{-1}IH^{-1}}$ gives almost same results as $\widehat{H}^{-1}$. Furthermore, to deal with the constraint on the correlation parameter $\rho \in [-1, 1]$, we estimate b where $\rho = \frac{2\exp(b)}{1+\exp(b)} - 1$, $b \in \mathcal{R}$. A positivity constraint is imposed on the baseline hazard function parameters $(r_0^{(l)}, \lambda_0)$ on the variances of the random effects (θ, η). The standard errors for the parameters estimated with constraints were computed by the Δ-method (see Knight, 2000).

2.2.1 Maximum likelihood estimation

The marginal likelihood is calculated by integrating the conditional likelihood (detailed in the web Supporting Information). The details are given in the Supporting Information. The full likelihood does not adopt a closed form because the integrals do not have an analytical solution. The integrals are evaluated using the Gauss–Hermite quadrature. Considering a calendar timescale, the i-th's individual contribution to the likelihood is

$$
\begin{aligned}
L_i(\Phi) = {} & \frac{\lambda\left(T_i^*\right)^{\delta_i^*} \prod_{j=1}^{n_i^{(1)}} r_i^{(1)}\left(T_{ij}^{(1)}\right)^{\delta_{ij}^{(1)}} \prod_{j=1}^{n_i^{(2)}} r_i^{(2)}\left(T_{ij}^{(2)}\right)^{\delta_{ij}^{(2)}}}{2\pi\sqrt{\theta\eta}\sqrt{(1-\rho^2)}} \\
& \times \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\Bigg[-\exp(u_i)\sum_{j=1}^{n_i^{(1)}}\int_{T_{i(j-1)}^{(1)}}^{T_{ij}^{(1)}} r_i^{(1)}(t)dt - \exp(v_i)\sum_{j=1}^{n_i^{(2)}}\int_{T_{i(j-1)}^{(2)}}^{T_{ij}^{(2)}} r_i^{(2)}(t)dt \\
& - \exp(\alpha_1 u_i + \alpha_2 v_i)\int_0^{T_i^*}\lambda_i(t)dt \\
& + \left(\frac{-u_i^2/\theta + 2\rho u_i v_i/\sqrt{\theta\eta} - v_i^2/\eta}{2(1-\rho^2)} + \left(n_i^{(1)} + \delta_i^*\alpha_1\right)u_i + \left(n_i^{(2)} + \delta_i^*\alpha_2\right)v_i\right)\Bigg] du_i dv_i.
\end{aligned}
$$

In the previous expression, we consider recurrent event times from the beginning of the study (calendar time). We do not consider left-truncation (delayed entries) which are meaningless in the application, because we are more interested in the time-to-event since diagnosis and then for all individuals $T_{i0} = 0$. In the gap timescale formulation, the likelihood expression is the same as for the calendar timescale except that $T_{ij}^{(l)}$, $l \in 1, 2$, is replaced by $S_{ij}^{(l)}$ and $\int_{T_{i(j-1)}^{(l)}}^{T_{ij}^{(l)}}$ by $\int_0^{S_{ij}^{(l)}}$.

2.2.2 Baseline hazard functions approximated by M-splines

B-splines are piecewise polynomial functions that are combined linearly to approximate a function on an interval. M-splines, which are a variant of B-splines, are well adapted to model baseline hazard functions. Indeed, M-splines are nonnegative functions that can approximate the hazard functions $\lambda_0(\cdot), r_0(\cdot)^{(l)}, l = \{1, 2\}$, easy to integrate or differentiate. Calculations of the M-splines basis are detailed in Ramsay (1988) and Joly (1996). We denote $t_1, \ldots, t_Q$, the sequence of the spline knots and the order of the splines is $d = 4$ (cubic splines). Based on these Q knots, m spline functions $M_i(\cdot)$ associated to coefficients η_i are defined from the Ramsay's equations (Ramsay, 1988). The number of parameters to estimate is $m = Q + 2$ (because $m = Q + d - 2$). To be sure to have positive hazard functions, the spline coefficients η_{Di}, $\eta_{Ri^{(l)}}$ must be positive. If we denote $\tilde{r_0}^{(l)}(\cdot)$ an approximation to the maximum penalized likelihood estimator (MPnLE) $\hat{r_0}^{(l)}(\cdot)$, the approximation error can be made as small as desired by increasing the number of knots Q. The more knots we use, the closer the approximation is to the true hazard function. We started with a small number of knots and increased the number of knots until the graph of the baseline hazard function remained unchanged. A trade-off have to be between the graphs by eye, the likelihood cross-validation (LCVa) criterion and the computational issues due to a high number of parameters to estimate. The approximations $\widetilde{\lambda}_0(\cdot), \widetilde{r}_0(\cdot)^{(l)}$ of the baseline hazard function estimators $\hat{\lambda}_0(\cdot), \hat{r}_0(\cdot)^{(l)}$ are, respectively, $\widetilde{\lambda}_0(\cdot) = \sum_{i=1}^{m} \eta_{Di} M_i(\cdot)$ and $\widetilde{r}_0(\cdot)^{(l)} = \sum_{i=1}^{m} \eta_{Ri}^{(l)} M_i(\cdot)$. Such approximation allows flexible shapes of the hazard functions while reducing the number of parameters. In our approach, although there are three different hazard functions, we use the same basis of splines for each function but the spline coefficients are different for the distinct functions. An approximation for the pointwise confidence intervals at 95% of $\widetilde{\lambda}_0(\cdot)$ is provided by $\widetilde{\lambda}_0(\cdot) \pm 1.96\sqrt{\mathbf{M}(\cdot)^{\mathbf{T}} \mathbf{I}_{\hat{\eta}_{\mathbf{D}}}^{-\mathbf{1}} \mathbf{M}(\cdot)}\}$, and $\widetilde{r}_0^{(l)}(\cdot) \pm 1.96\sqrt{\mathbf{M}(\cdot)^{\mathbf{T}} \mathbf{I}_{\hat{\eta}_{\mathbf{R}}^{(\mathbf{l})}}^{-\mathbf{1}} \mathbf{M}(\cdot)}\}$, $l \in \{1, 2\}$, where $M(\cdot) = (\mathrm{M}_1(\cdot), \ldots, \mathrm{M}_{\mathrm{m}}(\cdot))$ is the M-splines vector, $I_{\hat{\eta}_R^{(l)}}^{-1} = \frac{\partial^2 pl(\hat{\Phi})}{\partial \eta_R^2}$, $I_{\hat{\eta}_D}^{-1} = \frac{\partial^2 pl(\hat{\Phi})}{\partial \eta_D^2}$. A function approximated by splines could have local fluctuations, in particular when the number of knots is high. A solution to limit these local variations is to penalize the likelihood of the model to have smooth estimation of the baseline hazard functions.

2.2.3 Maximum of the penalized log-likelihood estimation

To avoid wide local variations of baseline hazard functions $\lambda_0(t)$ and $r_0^l(t)$ ($l \in \{1, 2\}$), we introduce smoothness with a penalized log-likelihood of the form:

$$pl(\Phi) = l(\Phi) - \kappa_1 \int_0^\infty r_0^{(1)''}(t)^2 dt - \kappa_2 \int_0^\infty r_0^{(2)''}(t)^2 dt - \kappa_3 \int_0^\infty \lambda_0''(t)^2 dt,$$

$$l(\Phi) = \sum_{i=1}^{N} \log L_i(\Phi), \tag{3}$$

where $l(\Phi)$ is the full log-likelihood defined previously and $\kappa_1, \kappa_2, \kappa_3$ are the positive smoothing parameters which control the trade-off between the data fit and the smoothness of the functions. The penalization is on hazard function fluctuations represented by $\int_0^\infty \lambda_0''(t)^2 dt$ and $\int_0^\infty r_0^{(l)''}(t)^2 dt$ which are the roughness terms based on the second derivative of the baseline hazard functions (approximated by a combination of polynomial functions of order 1). Maximization of $pl(\Phi)$ defines the MPnLEs $\hat{\Phi}$. For a fixed value of the smoothing parameters, the maximization of the penalized likelihood provides estimators for Φ. The smoothing parameters are obtained by fitting shared frailty models, one for each recurrent event type and a Cox model for the terminal event, using the cross-validation estimation method with the `R` package `Frailtypack` (see Rondeau et al., 2012; Rondeau and Gonzalez, 2005).

The estimated parameters were obtained by the robust Marquardt algorithm (see Marquardt, 1963), which is a combination between a Newton–Raphson algorithm and a steepest descent algorithm. This

algorithm has the advantage of being more stable than the Newton–Raphson algorithm while preserving its fast convergence property. The vector Φ of the parameters is updated until the convergence using the following recurring expression:

$$\Phi^{(r+1)} = \Phi^{(r)} - \delta \left(\widetilde{H}^{(r)}\right)^{-1} \Delta \left(L\left(\Phi^{(r)}\right)\right). \tag{4}$$

Step δ has been set equal to 1 by default but can be modified to ensure that the likelihood is improved at each iteration. The matrix $\widetilde{H}$ is a diagonal-inflated Hessian matrix to ensure positive definiteness. The term $\Delta(L(\Phi^{(r)}))$ corresponds to the penalized log-likelihood gradient at the r-th iteration. The iterations stop when the difference between two consecutive log-likelihoods is small ($<10^{-3}$), the coefficients are stable ($<10^{-3}$), and the gradient is small enough ($<10^{-3}$). The first and second derivatives are calculated using the finite differences method. After the convergence, the standard errors of the estimates are directly obtained from $\widehat{H}^{-1}$.

2.2.4 Parametric baseline hazard functions

We propose multivariate frailty models with parametric baseline hazard functions, especially piecewise constant or Weibull functions. The main drawback is that such parametric functions could not be, in some cases and specially for the Weibull function, flexible enough compared to an approximation by M-splines. The advantage with such considerations is that we reduce the number of parameters to estimate. There is no need to estimate smoothing parameters and we directly use the maximum likelihood estimation method.

2.3 Goodness-of-fit

2.3.1 The approximated LCVa criterion

The approximated LCVa criterion proposed in Commenges et al. (2007), LCVa, is used as criterion for model choice: LCVa $= \frac{trace(\widehat{H^{-1}I})-l(\widehat{\Phi})}{\sum_{i=1}^{N} n_i}$, where $n_i = n_i^{(1)} + n_i^{(2)} - 1$ is the number of observations for the i-th subject. It allows to compare a model estimated with maximum likelihood method and a model estimated with maximum penalized likelihood method. In the case of maximum penalized likelihood estimation, the AIC is not applicable. Note that for maximum likelihood estimation $trace(\widehat{H^{-1}I}) = np$ (np: number of parameters in the model), and LCVa $= \frac{np-l(\widehat{\Phi})}{\sum_{i=1}^{N} n_i}$.

2.3.2 Martingale residuals

The use of martingale residuals have been proposed for model checking in survival data (see Commenges and Rondeau, 2000). It enables us to check whether the model assess the precision of the prediction correctly the number of observed events. Martingale residuals applied to our proposed model become

$$\begin{cases} M_i^{R(1)}(t) = N_i^{R(1)}(t) - Y_i(t)\exp(\widehat{u}_i)\displaystyle\int_0^t \widehat{r_i^{(1)}}(u)du & \text{(rec. of type 1)} \\ M_i^{R(2)}(t) = N_i^{R(2)}(t) - Y_i(t)\exp(\widehat{v}_i)\displaystyle\int_0^t \widehat{r_i^{(2)}}(u)du & \text{(rec. of type 2)} \\ M_i^{D}(t) = N_i^{D}(t) - Y_i(t)\exp(\widehat{\alpha}_1\widehat{u}_i + \widehat{\alpha}_2\widehat{v}_i)\displaystyle\int_0^t \widehat{\lambda}_i(u)du & \text{(death).} \end{cases} \tag{5}$$

The martingale residuals are evaluated at the end of follow-up, that is, $t = T_i^*$. The parameters $\widehat{\Phi} = \{\widehat{r_0^{(1)}}, \widehat{r_0^{(2)}}, \widehat{\lambda}_0, \widehat{\boldsymbol{\beta}}_1, \widehat{\boldsymbol{\beta}}_2, \widehat{\boldsymbol{\beta}}_3, \widehat{\theta}, \widehat{\eta}, \widehat{\rho}, \widehat{\alpha}_2, \widehat{\alpha}_1\}$ used here are the estimated parameters, obtained after the maximization of the log-likelihood or the penalized log-likelihood. For the calculation, we need

individual estimates of the random effects $\widehat{u}_i$, $\widehat{v}_i$. We used empirical Bayesian estimators obtained by maximizing the posterior probability density function of the random effects (details can be found in the Supporting Information). Then we use the Lowess function of R software to have smooth curves for these martingale residuals.

3 Simulation

A simulation study of the proposed multivariate frailty model was performed to evaluate the performance of the estimators and to compare it with separate shared frailty models with one random effect specific for each type of event. In order to investigate the effect of increased sample size on estimator performance, we generated 500 replicate datasets and considered three sample sizes ($N = 250$, $N = 500$, and $N = 1000$) with a variable number of events per subject. The results are reported in the next subsection. The empirical standard errors and the estimator of standard error ($\sqrt{\hat{H}^{-1}}$) are also reported. We treated the right-censored cases only and used a gap-time timescale representation.

3.1 Generating data

We generate recurrent event gap times and death times from a multivariate frailty model:

$$\begin{cases} r_i^{(1)}\left(s_{ij}^{(1)}|u_i, v_i\right) = r_0^{(1)}\left(s_{ij}^{(1)}\right)\exp(\beta_1 Z_{i1} + \beta_2 Z_{i2} + u_i) & = \exp(u_i) r_i^{(1)}\left(s_{ij}^{(1)}\right) \\ r_i^{(2)}\left(s_{ij}^{(2)}|u_i, v_i\right) = r_0^{(1)}\left(s_{ij}^{(2)}\right)\exp(\beta_3 Z_{i1} + \beta_4 Z_{i2} + \beta_5 Z_{i3} + v_i) & = \exp(v_i) r_i^{(2)}\left(s_{ij}^{(2)}\right) \\ \lambda_i\left(t_i^*|u_i\right) = \exp(\alpha_1 u_i + \alpha_2 v_i)\lambda_0\left(t_i^*\right)\exp(\beta_6 Z_{i1}) & = \exp(\alpha_1 u_i + \alpha_2 v_i)\lambda_i\left(t_i^*\right) \end{cases} . \quad (6)$$

For each subject i ($i = 1, \ldots, N$), we generated the two correlated or not normally distributed random variables (u_i, v_i) with $\theta = 0.5$ and $\eta = 0.5$ as random effect variances and different correlation coefficients ρ. A right-censoring variable ($C_i = 5$) was set at a fixed value in order to have on average a third of censored data. The binary explanatory variables Z_{i1}, Z_{i2}, and Z_{i3} were generated from a Bernoulli distribution with $p = 0.5$. We set $\beta_1 = 0.5$, $\beta_2 = 0.7$, $\beta_3 = 0.6$, $\beta_4 = -0.5$, $\beta_5 = 0.3$, and $\beta_6 = 0.4$.

- First, we generated an exponential death time D_i using the hazard function $\lambda_i(t_i^*|u_i, v_i)$ with $\lambda_0(t) = 0.25$. If $D_i \leq C_i$ then the terminal event time is a death time ($\delta_i^* = 1$ and $T_i^* = D_i$). If $D_i > C_i$ then it is a censoring time ($\delta_i^* = 0$ and $T_i^* = C_i$).
- Then, we generated the recurrent gap times $s_{ij}^{(l)}$, $l = 1, 2$ using the two hazard functions $r_i^{(l)}(s_{ij}^{(l)}|u_i, v_i)$ with constant baseline hazard functions $r_0^{(1)} = 0.5$, $r_0^{(2)} = 0.65$. The corresponding observed calendar times for recurrent events are $T_{ij}^{(l)} = \min(D_i, C_i, \sum_{k=1}^{j} S_{ik}^{(l)})$. If $T_{ij}^{(l)} < T_i^*$ then we observed a recurrent event of type l and $\delta_{ij}^{(l)} = I(S_{ij}^{(l)} = s_{ij}^{(l)})$ and the data generation continues until $T_{ij}^{(l)} < T_i^*$. If $T_{ij}^{(l)} > T_i^*$ then we consider that subject is censored at T_i^* and $\delta_{ij}^{(l)} = 0$ and then the data generation stops.

3.2 Scenarios

We considered different settings for the parameters α_1, α_2, and ρ. The objective was to compare the parameter estimates between the proposed models and the reduced models when these three different events are dependent or not totally dependent. Note that the reduced models for recurrent events were shared frailty models and a Cox model for the terminal event. In the first setting, we generated data with a significant and positive dependency between recurrent events and terminal event ($\rho = 0.5$,

$\alpha_1 = 1, \alpha_2 = 1$). In the second setting, we generated independent random effects ($\rho = 0$). The two types of recurrent events were independent but were positively linked to the terminal event ($\alpha_1 = 1, \alpha_2 = 1$). In the third setting, the two recurrent events were negatively dependent ($\rho = -0.5$). The first one was negatively associated to the terminal event hazard function ($\alpha_1 = -1$), whereas the second one was positively linked to the terminal event hazard function ($\alpha_2 = 0.5$).

3.3 Results of the simulation study

The results for sample size of $N = 1000$ are presented in Table 1, the results for the sample size of $N = 250$ and $N = 500$ are reported in the Supporting Information. The death rate is around two-thirds. The average number of observed recurrent events (the first and second types) per subject ranges from 1.62 to 2.08 in the conducted simulation studies. Around a third of the subjects did not have any recurrent events.

First, it can be seen that, in the three settings, parameter estimates of our proposed models (piecewise constant and M-splines hazard functions) are well estimated. The magnitudes of the empirical biases of the estimates are negligible. Small biases (around 5%) are observed for $\hat{\theta}$ in setting I for the proposed model with M-splines hazard function, and $\hat{\alpha_1}$ in setting II. The coverage probabilities are close to the nominal level 95%. We observe only minor biases, regarding the coverage probabilities, for $\hat{\theta}$ in settings I and II, $\hat{\eta}$ in setting II and $\hat{\alpha_1}$ in setting I. We also used reduced model estimators assuming that the two types of recurrent events and the terminal event did not share any parameters, that is, $\rho = 0$ and $\alpha_1 = \alpha_2 = 0$. Recurrent event data were fitted with shared frailty models and the terminal event with a Cox model (results shown on the right side of Table 1). Using reduced models, we obtained biased parameter estimates, especially for the variance parameters $\hat{\theta}$, $\hat{\eta}$, and for the regression coefficient $\hat{\beta}_6$ associated with the terminal event, which are underestimated. These biases are due to the presence of dependencies between the two recurrent events and the terminal event.

Reported in the Supporting Information, we tried additional scenarios, close to the application dataset, that is, with few events and a high correlation coefficient and we noticed convergence issues of the model for such extreme case of data. However, we reached the same conclusion as these three scenarios.

4 Application to breast cancer patients

This study aimed to estimate the prognostic factors (hazard ratios) associated with the occurrences of breast cancer locoregional relapses, metastatic relapses, and death. It also aimed to study the dependencies between these three events of interest. Several papers such as Montagna et al. (2011) showed that after a locoregional relapse, there is a high risk of experiencing a metastatic event. O'Shaughnessy (2005) showed a strong link between a metastatic relapse and death. The link between locoregional relapses and death has been established for young women in Elkhuizen et al. (1998). This association is not obvious for women of any age in Monteiro Grillo et al. (2005). We consider here two different types of recurrent events which could be associated. Moreover, death is considered as a dependent terminating event for the relapses, it is necessary to analyze these recurrent events jointly with the terminal event to make valid inferences. The use of the proposed multivariate frailty model is justified with such data. The coefficients α_1 and α_2 indicate the sign of the association whether a type of recurrent event, locoregional or/and metastatic relapses, and death are significantly positively or negatively associated. The variances of the random effects (u_i, v_i) measure also the dependencies between the two types of recurrent events and death and also whether there are interrelapse dependencies. With such an approach, we are able to assess the association between breast cancer locoregional relapses, metastatic relapses, and death, and secondly the interrecurrence dependency.

Table 1 Results of simulation study ($N = 1000$).

	Proposed model (piecewise constant) $[u_i, v_i, \alpha_1 u_i + \alpha_2 v_i]$				Proposed model (M-splines) $[u_i, v_i, \alpha_1 u_i + \alpha_2 v_i]$				Reduced (M-splines) $[u_i, v_i, -]$			
	Mean	S.E. emp.	E.S.E. $(\sqrt{\widehat{H^{-1}}})$	C.P.	Mean	S.E. emp.	E.S.E. $(\sqrt{\widehat{H^{-1}}})$	C.P.	Mean	S.E. emp.	E.S.E. $(\sqrt{\widehat{H^{-1}}})$	C.P.
Setting I: $\rho = 0.5, \alpha_1 = \alpha_2 = 1$												
First type of recurrent events												
$\beta_1 = 0.5$	0.500	0.065	0.066	94.8	0.506	0.071	0.067	93.2	0.468	0.065	0.065	92.6
$\beta_2 = 0.7$	0.699	0.059	0.061	95.0	0.694	0.063	0.061	95.0	0.716	0.065	0.066	94.8
Second type of recurrent events												
$\beta_3 = 0.6$	0.598	0.069	0.070	96.2	0.609	0.071	0.071	94.8	0.557	0.068	0.067	89.6
$\beta_4 = -0.5$	−0.503	0.062	0.063	95.2	−0.486	0.061	0.063	95.8	−0.507	0.067	0.067	94.4
$\beta_5 = 0.3$	0.308	0.064	0.062	94.4	0.298	0.064	0.062	93.8	0.311	0.069	0.066	94.2
Terminal event (death)												
$\beta_6 = 0.4$	0.403	0.113	0.116	96.2	0.394	0.115	0.114	94.0	0.134	0.109	0.104	22.2
$\theta = 0.5$	0.479	0.056	0.056	92.2	0.473	0.059	0.059	91.0	0.355	0.054	0.050	21.2
$\eta = 0.5$	0.493	0.067	0.061	93.2	0.496	0.065	0.066	94.2	0.338	0.059	0.053	18.8
$\rho = 0.5$	0.495	0.063	0.065	95.0	0.530	0.063	0.066	93.6	–	–	–	–
$\alpha_1 = 1$	1.035	0.170	0.163	94.2	0.959	0.191	0.179	91.4	–	–	–	–
$\alpha_2 = 1$	1.035	0.169	0.167	94.6	0.990	0.187	0.186	95.0	–	–	–	–
Setting II: $\rho = 0, \alpha_1 = 1, \alpha_2 = 1$												
First type of recurrent events												
$\beta_1 = 0.5$	0.501	0.068	0.066	94.4	0.502	0.068	0.066	95.0	0.483	0.069	0.066	93.4
$\beta_2 = 0.7$	0.705	0.065	0.063	94.2	0.704	0.066	0.063	93.8	0.724	0.071	0.067	92.0
Second type of recurrent events												
$\beta_3 = 0.6$	0.598	0.070	0.069	95.0	0.598	0.070	0.070	95.4	0.575	0.071	0.068	91.6
$\beta_4 = -0.5$	−0.506	0.064	0.065	95.8	−0.505	0.064	0.065	95.6	−0.513	0.069	0.068	94.2
$\beta_5 = 0.3$	0.305	0.061	0.064	96.0	0.306	0.062	0.064	95.8	0.311	0.067	0.067	95.2
Terminal event (death)												
$\beta_6 = 0.4$	0.404	0.103	0.105	95.0	0.408	0.104	0.106	95.4	0.147	0.085	0.085	26.4
$\theta = 0.5$	0.481	0.050	0.052	92.6	0.477	0.053	0.054	91.2	0.435	0.053	0.051	71.2
$\eta = 0.5$	0.494	0.059	0.056	92.8	0.493	0.062	0.060	92.2	0.429	0.061	0.055	68.4
$\rho = 0$	−0.011	0.075	0.074	94.4	−0.014	0.076	0.075	94.8	–	–	–	–
$\alpha_1 = 1$	1.052	0.143	0.141	94.2	1.064	0.163	0.161	96.6	–	–	–	–
$\alpha_2 = 1$	1.029	0.151	0.143	95.2	1.042	0.173	0.163	96.8	–	–	–	–

(*continued*)

Table 1 Continued

	Proposed model (piecewise constant) $[u_i, v_i, \alpha_1 u_i + \alpha_2 v_i]$				Proposed model (M-splines) $[u_i, v_i, \alpha_1 u_i + \alpha_2 v_i]$				Reduced (M-splines) $[u_i, v_i, -]$			
	Mean	S.E. emp.	E.S.E. $(\sqrt{\widehat{H^{-1}}})$	C.P.	Mean	S.E. emp.	E.S.E. $(\sqrt{\widehat{H^{-1}}})$	C.P.	Mean	S.E. emp.	E.S.E. $(\sqrt{\widehat{H^{-1}}})$	C.P.
Setting III: $\rho = -0.5, \alpha_1 = -1, \alpha_2 = 0.5$												
First type of recurrent events												
$\beta_1 = 0.5$	0.503	0.067	0.063	94.4	0.502	0.066	0.063	94.6	0.521	0.094	0.060	89.4
$\beta_2 = 0.7$	0.700	0.061	0.058	93.2	0.698	0.061	0.058	93.0	0.681	0.110	0.064	90.8
Second type of recurrent events												
$\beta_3 = 0.6$	0.597	0.068	0.069	96.0	0.598	0.068	0.070	96.0	0.573	0.069	0.068	93.2
$\beta_4 = -0.5$	−0.502	0.064	0.065	94.4	−0.502	0.064	0.065	94.8	−0.511	0.070	0.068	94.2
$\beta_5 = 0.3$	0.303	0.064	0.064	95.8	0.293	0.088	0.090	95.6	0.309	0.070	0.067	94.0
Terminal event (death)												
$\beta_6 = 0.4$	0.405	0.104	0.101	93.4	0.406	0.106	0.102	94.4	0.133	0.088	0.075	26.6
$\theta = 0.5$	0.487	0.046	0.047	94.0	0.486	0.047	0.048	93.0	0.439	0.067	0.066	64.2
$\eta = 0.5$	0.499	0.056	0.056	94.8	0.501	0.060	0.060	95.4	0.428	0.058	0.055	67.0
$\rho = -0.5$	−0.503	0.058	0.058	95.6	−0.505	0.057	0.058	95.6	–	–	–	–
$\alpha_1 = -1$	−1.019	0.159	0.152	94.6	−1.046	0.178	0.169	95.4	–	–	–	–
$\alpha_2 = 0.5$	0.520	0.135	0.130	94.2	0.518	0.140	0.133	94.4	–	–	–	–

[a]Mean, mean of the estimates; S.E. emp., empirical standard errors mean; E.S.E., estimated standard errors mean; C.P., coverage probability.

4.1 Rationale

Breast cancer is the most frequently diagnosed cancer and the leading cause of cancer death among females worldwide, accounting for 23% of the total cancer cases and 14% of cancer deaths (see Jemal et al., 2011). The elucidation of prognostic factors and their influence are critical for making progress in disease prevention and treatment. The occurrences of locoregional and metastatic relapses may provide information about the deterioration of a patient's health. These recurrent events and death can be partly explained with prognostic factors such as biological measurements, or characteristics of the individuals, associated with the tumor or the environment. Several prognostic factors for which we do not have measurements in the clinical dataset should be added to explain the evolution of the disease. A model that can deal with nonobserved prognostic factors (heterogeneity of data), correlated recurrent event times with a dependent terminal event time, is needed. This leads us to the proposed multivariate frailty model.

4.2 Data

Our data were provided by the Institut Bergonié, the Regional Comprehensive Cancer Center, Bordeaux, France. A total of 1161 patients were included in the study, after a first diagnosis, between 1989 and 1993 and followed until 2010. The median follow-up period established with the reverse Kaplan–Meier method was 13.9 years with a range of five months to 20 years. Patients were included after the diagnosis, and are considered at risk of relapse or death since the date of breast-conserving surgery. Age, tumor size, grade, peritumoral vascular invasion (PVI), lymph node involvement (node: N+ or N−), positive status of hormone receptors (HR+ or HR−), HER2+ and proliferation index with Mib1 (Pro-mib1) were considered. Ninety-one patients were excluded because of missing data. A patient was considered N+ if at least one of her axillary lymph nodes was metastatic. A tumor was considered HR+ if more than 10% of tumor cells were positive by immunohistochemical analysis. HER2 or human epidermal growth factor receptor-2 is a protein and if its rate is abnormally raised at the surface of the tumoral cells then the tumor is classified HER2+. A brief description of the data is provided in the Supporting Information. A total of 718 (67.1%) patients did not have a relapse, of which 107 (10%) died. A total of 273 (25.5%) patients experienced a metastatic relapse, 177 (16.5%) had one locoregional relapse, and 10 (< 1%) had two locoregional relapses. In fact, 129 (12.1%) patients had a metastatic relapse before dying, whereas 9 (< 1%) patients had a locoregional relapse before death. Moreover, 78 (7.3%) patients had a locoregional relapse, then a metastatic relapse and then died. During the follow-up, 332 (31.0%) patients had died, and the remaining patients were alive (censored) at the end of the follow-up. We assumed this censoring was independent of any type of relapse or death. Most patients, followed over a long time, died after a distant metastasis. Their number is almost 14 times bigger than the number of patients died after a locoregional relapse. This confirms that distant metastasis is a lethal event. We applied the proposed model to analyze the different recurrences and survival among these patients.

4.3 Results with multivariate and reduced models

We fitted a multivariate frailty model with baseline hazard functions approximated by M-splines (with five equidistant knots), a multivariate frailty models with piecewise constant baseline hazard functions (with nine intervals) and three separate reduced models with baseline hazard function approximated by M-splines (with five equidistant knots). Maximum of the penalized likelihood estimation method is used for the models of which the baseline hazard functions are approximated by M-splines. Otherwise, maximum likelihood estimation method is used for the models with piecewise constant baseline hazard function. First we aimed to compare the results obtained by fitting the multivariate frailty models using two different estimation methods. Second, we aimed to compare the results obtained using a

Table 2 Analysis with multivariate frailty and reduced models of the locoregional and metastatic recurrences and death for breast cancer patients.

Variables	Modalities	Proposed (piecewise constant)		Proposed (M-splines)		Reduced (M-splines)	
		HR	(95% CI)	HR	(95% CI)	HR	(95% CI)
For locoregional relapses							
Age (ref: >55 years)	≤40 years	2.86	(1.76–4.64)	2.91	(1.80–4.71)	2.59	(1.77–3.93)
	>40 years and ≤55 years	1.32	(0.94–1.86)	1.32	(0.94–1.86)	1.46	(1.07–1.99)
PVI (ref: No)	Yes	1.63	(1.15–2.30)	1.65	(1.16–2.33)	1.49	(1.10–2.02)
Grade (ref: I)	II	2.05	(1.33–3.17)	2.10	(1.36–3.24)	1.81	(1.21–2.69)
	III	2.45	(1.49–4.03)	2.51	(1.53-4.13)	1.84	(1.18–2.92)
Lymph node (ref: N−)	N+	1.61	(1.15–2.25)	1.62	(1.16–2.26)	1.41	(1.04–1.91)
HER2+ (ref: No)	Yes	1.83	(1.18–2.82)	1.83	(1.18–2.84)	1.59	(1.09–2.34)
Tumor size (ref: <20 mm)	≥20 mm	1.59	(1.11–2.27)	1.59	(1.11–2.27)	1.34	(0.98–1.86)
For metastatic relapses							
Age (ref: >55 years)	≤40 years	2.81	(1.31–6.03)	2.91	(1.37–6.15)	1.85	(1.29–2.67)
	>40 years and ≤55 years	0.80	(0.49–1.29)	0.79	(0.48–1.28)	1.09	(0.84–1.42)
PVI (ref: No)	Yes	2.19	(1.32–3.63)	2.23	(1.34–3.69)	1.41	(1.09–1.82)
Grade (ref: I)	II	2.79	(1.53–5.09)	2.90	(1.59–5.28)	1.63	(1.16–2.31)
	III	4.56	(2.26–9.20)	4.98	(2.47–10.04)	1.89	(1.29–2.78)
Lymph node (ref: N−)	N+	3.47	(2.15–5.59)	3.60	(2.24–5.79)	1.99	(1.53–2.59)
HER2+ (ref: No)	Yes	2.19	(1.10–4.34)	2.31	(1.15–4.64)	1.55	(1.11–2.15)
Tumor size (ref: <20 mm)	≥20 mm	3.95	(2.37–6.60)	3.87	(2.30–6.52)	1.76	(1.36–2.28)
For death							
Age (ref: >55 years)	≤40 years	1.22	(0.39–3.80)	1.39	(0.48–4.07)	0.94	(0.64–1.38)
	>40 years and ≤55 years	0.29	(0.14–0.59)	0.29	(0.14–0.59)	0.66	(0.51–0.84)
PVI (ref: No)	Yes	3.22	(1.48–7.00)	3.18	(1.54–6.57)	1.40	(1.11–1.77)
Grade (ref: I)	II	4.59	(1.79–11.74)	4.65	(2.00–10.82)	1.70	(1.24–2.31)
	III	14.78	(4.76–45.90)	14.88	(5.62–39.36)	2.36	(1.67–3.33)
Lymph node (ref: N−)	N+	4.33	(2.06–9.08)	4.14	(2.11–8.12)	1.74	(1.37–2.19)
HER2+ (ref: No)	Yes	2.10	(0.75–5.85)	2.11	(0.76–5.88)	1.33	(0.97–1.83)
Tumor size (ref: <20 mm)	≥20 mm	5.92	(2.53–13.86)	5.57	(2.56–12.09)	1.61	(1.27–2.04)
$\theta = \text{var}(u_i)$ (SE)		1.10	(0.11)	1.12	(0.10)		–
$\eta = \text{var}(v_i)$ (SE)		7.39	(0.63)	7.54	(0.63)		–
α_1 (SE)		−0.25	(1.45)	−0.11	(1.48)		–
α_2 (SE)		1.66	(0.59)	1.57	(0.58)		–
ρ (SE)		0.99	(0.01)	0.99	(0.01)		–
LCVa[a]		0.160		0.144			–

[a]LCVa or approximated likelihood cross-validation criterion (see Section 2.3.1).

multivariate frailty model and the results obtained using three separate reduced models. The results are presented in Table 2.

According to all the fitted models, the proposed and the reduced models, the risk of locoregional, metastatic recurrences, or death increased for patients with PVI and for patients with at least one positive lymph node (N+) or for patients with a grade greater than *I*. For all models fitted, the risk of death was not significantly different for patients younger than 40 years versus older than 55 years,

but it was significantly lower for patients aged between 40 and 55 years compared to those older than 55 years. The risk of metastatic or locoregional recurrences increased for patients younger than 40 years compared to those older than 55 years. Tumor size (>20 mm) has a significant effect for the risks of these three events. The reduced model did not detect the effect of tumor size for the risk of locoregional recurrences. The risk of locoregional and metastatic recurrences was higher for HER2+ patients, but no association was found with the risk of death.

The proposed models showed that the risk of locoregional, metastatic recurrences, or death increased for patients with tumors greater than 20 mm. The reduced models detect only a significant effect of tumor size for the risk of metastatic recurrences and death; it was not significant for the risk of locoregional recurrences. This illustrates that if the dependencies between these three events of interest is not taken into account, we may omit a significant association. We further observe an underestimation of the regression coefficient in the reduced models, especially for the death hazard function, which was expected given the simulations.

The variables HR+ and Pro-mib1 were significantly associated ($p < 0.05$, using a Wald test) with the risks of locoregional relapse, metastatic relapse, and death for all models fitted, reduced, and proposed. We decided not to keep them in the global analysis. We noticed that the variable HR+ did not satisfy the hazard proportionality assumption, using the Cox model. This could explain why it is not significant for the complete follow-up period. The time-varying effect of the variable HR+ has also been pointed out in Dignam et al. (2009) and Hilsenbeck et al. (1998). This PHs assumption is satisfied when we divided the follow-up period into two periods: 0–5 years and 5–20 years. For the first period, the effect of HR+ is significant and strengthening and for the second period it was significant and weakening for any type of event. The other covariates including Pro-mib1 satisfied the hazard proportional assumption. For the proposed models, the parameters η and α_2 are both high and significantly different from 0, this means that there was a positive and strong dependency between the risk of metastatic recurrences and the risk of death. The parameter θ was significantly different from 0, the parameter α_1 was not. This means that there was interrecurrence dependency and no association between the risk of death and the risk of locoregional recurrences. The correlation coefficient ρ was also high and significantly different from 0, illustrating a strong dependency between the risk of locoregional recurrences and the risk of metastatic recurrences. The dependencies obtained here confirmed what we found in O'Shaughnessy (2005), Elkhuizen et al. (1998), Montagna et al. (2011), and Monteiro Grillo et al. (2005). Since metastatic relapse is an event which often happens before dying, it is not surprising to notice from this model that the association between locoregional relapses and death is not significant.

A correlation coefficient at 0.99 could be quite unexpected. Hence, we estimated random effects $\hat{u}_i$, $\hat{v}_i$ using empirical Bayesian estimators (see Section 2.3.2), and the empirical correlation coefficient was very close to 1. Furthermore, we checked parameter identifiability in the presence of a high correlation coefficient by doing simulations similar to the first scenario with $\rho = 0.9$. We obtained unbiased estimates for the correlation coefficient and satisfactory results for the other parameters. This means that there were no identifiability issues for a high correlation coefficient. We conclude that both proposed models have correctly converged and that these results come from the data we used. Furthermore, in our application, few repeated occurrences of both types, locoregional relapses and distant metastasis, were observed. Hence, the random effects u_i, v_i reflect more the link between locoregional relapses and death or between distant metastasis and death, than a within individual dependence. In this particular setting, it can be hard to obtain an independence between the two outcomes (locoregional relapses and distant metastasis). This can explain the very high value of the correlation coefficient, ρ, between the random effects. However, these results should be viewed with caution, given the small number of recurrent events.

The approximated LCVa criterion indicates that the proposed model with baseline hazard functions approximated by M-splines fits the data better than the model with piecewise constant baseline hazard functions. Baseline hazard functions for both fitted multivariate frailty models (piecewise constant and M-splines, from Table 2) are represented in Fig. 1.

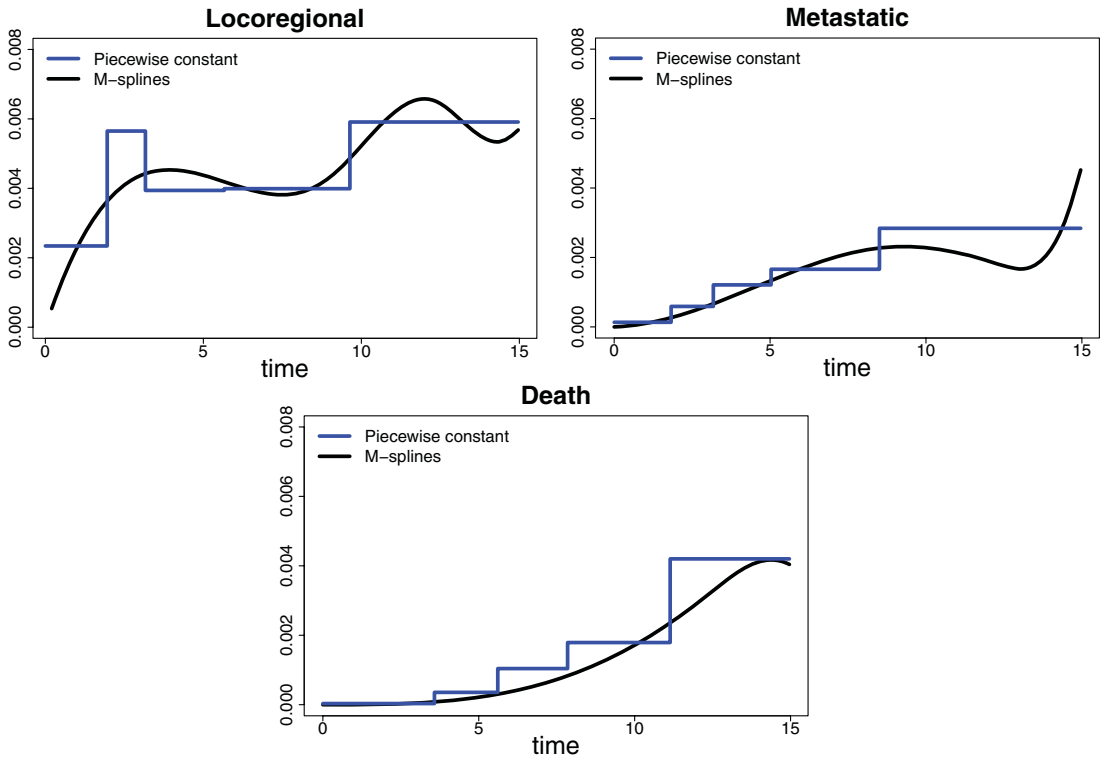

Figure 1 Baseline hazard functions for locoregional, metastatic, and death processes after a breast cancer obtained by fitting proposed multivariate frailty models.

Regarding the martingale residuals, we can see in the legends of the graphs (see Fig. 2) that the means of the martingale residuals for all event processes are close to 0. Moreover, the smoothing of these residuals obtained with the `R` software Lowess function is close to the X-axis. We can conclude that both proposed models fit the data well. In particular, the models correctly predict, on average, correctly the number of observed events during follow-up.

5 Conclusion and discussion

In this work, we presented a multivariate frailty model with two correlated random effects to simultaneously model two types of recurrent events with a dependent terminal event. The proposed model for possibly right-censored data was able to express the dependency among the two types of recurrent events, but it could also deal with the association between recurrent and terminal events. Simulation studies indicated that the approach works well for practical situations and was better than using three separate reduced models (shared frailty models for the recurrent events and a Cox model for death). This approach accounts for between events dependencies, interrecurrence dependencies, and nonobserved heterogeneity. As observed in Putter and van Houwelingen (2011) in the context of shared frailty models, the latent frailties can pick up anything that improves the fit of the model. This could be true heterogeneity but also lack of fit of the model without frailties. For instance, the lack of fit of the PHs models can be confused with the presence of heterogeneity, that is, it is difficult to tell whether there is true heterogeneity or whether the PHs assumption is violated. We also did the PHs assumption conditionally on frailties in our proposed model.

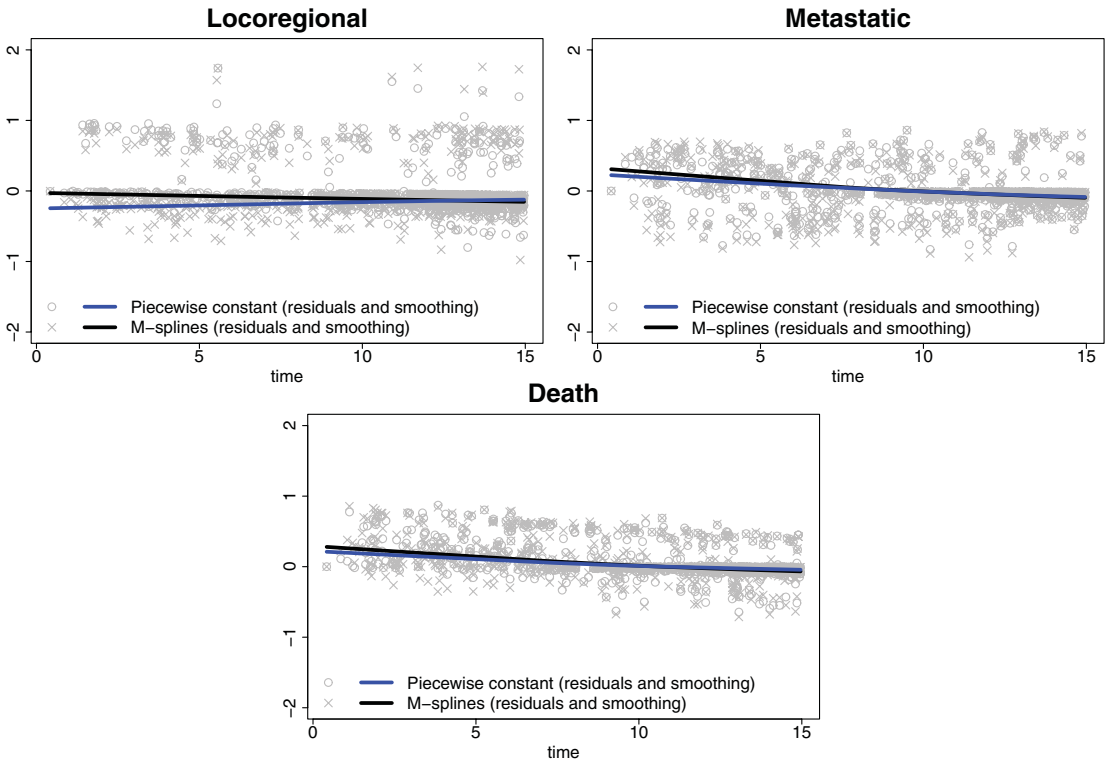

Figure 2 Martingale residuals for locoregional, metastatic, and death processes against follow-up time.

One advantage of the above-mentioned approach is that different covariate effects may be assessed for the two types of recurrence hazard functions or death hazard function; these covariates may be time-dependent or time-independent. In total, there are many reasons to use multivariate frailty models for three survival endpoints, including giving a general idea of the possible dependencies, correcting for regression coefficient bias in survival analysis. Furthermore, they also provide information on whether one or both types of relapses can be used as surrogate endpoints for overall survival (see Rondeau et al., 2011; Burzykowski et al., 2001).

In the application of these models, we concluded that the risk of locoregional recurrences is associated with the risk of metastatic recurrences and the risk of metastatic recurrences is also associated with death. The risk of death is not directly associated with the risk of locoregional recurrence. However, Wapnir et al. (2006) showed a dependence between locoregional relapses and death. A major issue of debate in this area of research is whether for death following locoregional relapse, the locoregional relapse is causal (produces additional risk of metastatic spread) versus incidental (in fact, a marker or flag for high-risk disease that may have disseminated already at the time of diagnosis; Fisher et al., 1991; Fortin et al., 1999; Vicini et al., 2003). For example, the reason early locoregional relapse associated with death is that it is indeed a flag for aggressive (Schmoor et al., 2000; Tanis et al., 2012), already disseminated disease (and also greater risk of locoregional relapse concurrently), whereas a late locoregional relapse may be just a de novo recurrence found early due to increased screening vigilance, and any disseminated disease has just begun and will not increase mortality risk any time soon. We fitted a Cox model on the risk of death adjusted for the timing of the first locoregional relapse in three categories: patients having no locoregional relapse (as reference), patients having a locoregional relapse before five years of follow-up (the median of the locoregional relapse delays), and

patients having a locoregional relapse after five years of follow-up. An early locoregional recurrence significantly increased the risk of death, whereas a later one was not significantly associated with death. Furthermore, we tried to analyze jointly locoregional relapse and death by fitting a joint frailty model (Rondeau et al., 2007) using the `R` package `Frailtypack` (Rondeau et al., 2012). We noticed that locoregional relapses and death are related events. We did the same joint analysis with distant metastasis and death and we conclude that these two events are also associated. We may conclude that if we omit to consider the distant metastasis events in the analysis, we can conclude that locoregional relapses and death are highly associated, which is not totally true. Again, this shows the value of the proposed model which can deal with the whole data information and provide direct dependencies between events.

Convergence of the proposed model could be difficult to obtain in the case of data with few events. Nevertheless, convergence could be reached when decreasing the number of parameters to estimate, that is, the number of parameters used to estimate baseline hazard functions. This work gives smoothed estimates of hazard functions that represent incidence and mortality rates in epidemiology, which are meaningful for clinicians. It is natural in epidemiology to impose a continuous hazard function with small local variations. We also proposed to use the maximum likelihood estimation method when the baseline hazard functions are parametrized as piecewise constant functions. In several articles such as Liu et al. (2004), Huang and Liu (2007), the estimation method used is a variant of the EM algorithm which does not provide a direct variance estimator of random effects, nor a smoothed estimation of the hazard functions.

We also proposed appropriate martingale residuals that can be used in graphical examinations of the fit of the multivariate frailty model with the data. One use of these residuals is to examine the possible nonlinear effect of a variable: the residuals are computed in a model not including the variable of interest; then they are plotted against this variable. A smoothing can be performed to make the interpretation easier.

According to few number of recurrent events observed in the application, several multistate models could be investigated. Contrary to the frailty modeling, multistate modeling does not directly provide information about the dependence between events of interest. Moreover, to use the same information as in our approach, the number of intensities which should be modeled to consider all potential states would be high and may lead to computational issues.

Many other applications could be adapted to the proposed multivariate model, for instance when we want to study simultaneously three types of competing causes of failure (e.g., competing risk of death) or when we want to study the evolution of a patient's state over time, after an admission in an intensive care unit (ICU), to study infections, discharge of alive patients or death in ICU. In analyses of the natural history of cancer, there is great interest in dynamic prediction of death, that is, in the computation of the predictive distribution of death at a certain moment in time, given the history of events (local or distant relapses) and covariates until that moment. These predictions and a measure of their accuracy are in progress, they may provide valuable insight for future research.

6 Software

All of our proposals will be very soon implemented in the `R` software with the freely available package `Frailtypack` (see Rondeau et al., 2012; Rondeau and Gonzalez, 2005). This package can also be used to fit a variety of joint frailty models or other frailty models for recurrent or clustered time-to-event data with several different options for the baseline risk functions. The package can be downloaded from the Comprehensive R Archive Network accessible via http://cran.r-project.org/package=frailtypack.

Acknowledgments We thank the Institut Bergonié for providing the data, Pippa McKelvie-Sebileau for medical editorial services, and the "Conseil Régional d'Aquitaine" for funding this work. We also thank the reviewers and the associate editor for their relevant and useful comments.

Conflict of interest
The authors have declared no conflict of interest.

References

Broët, P., de la Rochefordiere, A., Scholl, S., Fourquet, A., Rycke, Y., Pouillart, P., Mosseri, V. and Asselain, B. (1999). Analyzing prognostic factors in breast cancer using a multistate model. *Breast Cancer Research and Treatment* **54**, 83–89.

Burzykowski, T., Molenberghs, G., Buyse, M., Geys, H. and Renard, D. (2001). Validation of surrogate end points in multiple randomized clinical trials with failure time end points. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **50**, 405–422.

Cai, J. and Schaubel, D. (2004). Marginal means/rates models for multiple type recurrent event data. *Lifetime Data Analysis* **10**, 121–138.

Clayton, D. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* **65**, 141–151.

Commenges, D., Joly, P., Gégout-Petit, A. and Liquet, B. (2007). Choice between semi-parametric estimators of Markov and non-Markov multistate models from generally coarsened observations. *Scandinavian Journal of Statistics* **34**, 33–52.

Commenges, D. and Rondeau, V. (2000). Standardized martingale residuals applied to grouped left truncated observations of dementia cases. *Lifetime Data Analysis* **6**, 229–235.

Cook, R. and Lawless, J. (2007). *The Statistical Analysis of Recurrent Events*. Springer Verlag, New-York.

Cox, D. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)* **34**, 187–220.

De Bock, G., Putter, H., Bonnema, J., Van Der Hage, J., Bartelink, H. and Van De Velde, C. (2009). The impact of loco-regional recurrences on metastatic progression in early-stage breast cancer: a multistate model. *Breast Cancer Research and Treatment* **117**, 401–408.

Dignam, J., Dukic, V., Anderson, S., Mamounas, E., Wickerham, D. and Wolmark, N. (2009). Hazard of recurrence and adjuvant treatment effects over time in lymph node-negative breast cancer. *Breast Cancer Research and Treatment* **116**, 595–602.

Duchateau, L. and Janssen, P. (2008). *The Frailty Model*. Springer, New York.

Elkhuizen, P., van de Vijver, M., Hermans, J., Zonderland, H., van de Velde, C. and Leer, J. (1998). Local recurrence after breast-conserving therapy for invasive breast cancer: high incidence in young patients and association with poor survival. *International Journal of Radiation Oncology Biology Physics* **40**, 859–867.

Fisher, B., Anderson, S., Fisher, E., Redmond, C., Wickerham, D., Wolmark, N., Mamounas, E., Deutsch, M. and Margolese, R. (1991). Significance of ipsilateral breast tumour recurrence after lumpectomy. *The Lancet* **338**, 327–331.

Fortin, A., Larochelle, M., Laverdière, J., Lavertu, S. and Tremblay, D. (1999). Local failure is responsible for the decrease in survival for patients with breast cancer treated with conservative surgery and postoperative radiotherapy. *Journal of Clinical Oncology* **17**, 101–109.

Hilsenbeck, S., Ravdin, P., de Moor, C., Chamness, G., Osborne, C. and Clark, G. (1998). Time-dependence of hazard ratios for prognostic factors in primary breast cancer. *Breast Cancer Research and Treatment* **52**, 227–237.

Huang, X. and Liu, L. (2007). A joint frailty model for survival and gap times between recurrent events. *Biometrics* **63**, 389–397.

Huang, X. and Wolfe, R. (2002). A frailty model for informative censoring. *Biometrics* **58**, 510–520.

Jemal, A., Bray, F., Center, M., Ferlay, J., Ward, E. and Forman, D. (2011). Global cancer statistics. *CA: A Cancer Journal for Clinicians* **61**, 69–90.

Joly, P. (1996). *Estimation de la fonction de risque dans un contexte général de troncature et de censure: application à l'estimation de l'incidence de la démence*. Ph.D. thesis, Université de Bordeaux.

Knight, K. (2000). *Mathematical Statistics, Texts in Statistical Science*. Chapman and Hall/CRC, London.

Lancaster, T. and Intrator, O. (1998). Panel data with survival: hospitalization of HIV-positive patients. *Journal of the American Statistical Association* **93**, 4653.

Liu, L., Wolfe, R. 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. *Journal of the Society for Industrial and Applied Mathematics* **11**, 431–441.

Mazroui, Y., Mathoulin-Pelissier, S., Soubeyran, P. and Rondeau, V. (2012). General joint frailty model for recurrent event data with a dependent terminal event: application to follicular lymphoma data. *Statistics in Medicine* **31**, 1162–1176.

Montagna, E., Bagnardi, V., Rotmensz, N., Viale, G., Renne, G., Cancello, G., et al. (2011). Breast cancer subtypes and outcome after local and regional relapse. *Annals of Oncology* **23**, 324–331.

Monteiro Grillo, I., Jorge, M., Marques Vidal, P., Ortiz, M. and Ravasco, P. (2005). The effect of locoregional recurrence on survival and distant metastasis after conservative treatment for invasive breast carcinoma. *Clinical Oncology* **17**, 111–117.

O'Shaughnessy, J. (2005). Extending survival with chemotherapy in metastatic breast cancer. *The Oncologist* **10**, 20–29.

Putter, H., van der Hage, J., de Bock, G., Elgalta, R. and van de Velde, C. (2006). Estimation and prediction in a multistate model for breast cancer. *Biometrical Journal* **48**, 366–380.

Putter, H. and van Houwelingen, H. (2011). Frailties in multistate models: are they identifiable? Do we need them? *Statistical Methods in Medical Research*, Nov 23 [Epub ahead of print].

Ramsay, J. (1988). Monotone regression splines in action. *Statistical Science* **3**, 425–441.

Rondeau, V. and Gonzalez, J. (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.

Rondeau, V., Mathoulin-Pelissier, S., Jacqmin-Gadda, H., Brouste, V. and Soubeyran, P. (2007). Joint frailty models for recurring events and death using maximum penalized likelihood estimation: application on cancer events. *Biostatistics* **8**, 708–721.

Rondeau, V., Mazroui, Y. and Gonzalez, J. (2012). Frailtypack: an r package for the analysis of correlated survival data with frailty models using penalized likelihood estimation or parametrical estimation. *Journal of Statistical Software* **47**.

Rondeau, V., Pignon, J. and Michiels, S. (2011). A joint model for the dependence between clustered times to tumour progression and deaths: a meta-analysis of chemotherapy in head and neck cancer. *Statistical Methods in Medical Research*, Oct 23 [Epub ahead of print]

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

Schmoor, C., Sauerbrei, W., Bastert, G., Schumacher, M., et al. (2000). Role of isolated locoregional recurrence of breast cancer: results of four prospective studies. *Journal of Clinical Oncology* **18**, 1696–1708.

Tanis, E., van de Velde, C., Bartelink, H., van de Vijver, M., Putter, H. and van der Hage, J. (2012). Locoregional recurrence after breast-conserving therapy remains an independent prognostic factor even after an event free interval of 10 years in early stage breast cancer. *European Journal of Cancer* **48**, 1751–1756.

Vaupel, J., Manton, K. and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* **16**, 439–454.

Vicini, F., Kestin, L., Huang, R. and Martinez, A. (2003). Does local recurrence affect the rate of distant metastases and survival in patients with early-stage breast carcinoma treated with breast-conserving therapy? *Cancer* **97**, 910–919.

Wang, M., Qin, J. and Chiang, C. (2001). Analyzing recurrent event data with informative censoring. *Journal of the American Statistical Association* **96**, 1057–1065.

Wapnir, I., Anderson, S., Mamounas, E., Geyer Jr, C., Jeong, J., Tan-Chiu, E., Fisher, B. and Wolmark, N. (2006). Prognosis after ipsilateral breast tumor recurrence and locoregional recurrences in five national surgical adjuvant breast and bowel project node-positive adjuvant breast cancer trials. *Journal of Clinical Oncology* **24**, 2028–2037.

Zeng, D. and Cai, J. (2010). A semiparametric additive rate model for recurrent events with an informative terminal event. *Biometrika* **97**, 699–712.

Zhu, L., Sun, J., Tong, X. and Srivastava, D. (2010). Regression analysis of multivariate recurrent event data with a dependent terminal event. *Lifetime Data Analysis* **16**, 478–490.

