CrossMark

A new joint model of recurrent event data with the additive hazards model for the terminal event time

Xiaoyu Che[1] · John Angus[2]

Abstract Recurrent event data are frequently encountered in clinical and observational studies related to biomedical science, econometrics, reliability and demography. In some situations, recurrent events serve as important indicators for evaluating disease progression, health deterioration, or insurance risk. In statistical literature, non informative censoring is typically assumed when statistical methods and theories are developed for analyzing recurrent event data. In many applications, however, there may exist a terminal event, such as death, that stops the follow-up, and it is the correlation of this terminal event with the recurrent event process that is of interest. This work considers joint modeling and analysis of recurrent event and terminal event data, with the focus primarily on determining how the terminal event process and the recurrent event process are correlated (i.e. does the frequency of the recurrent event influence the risk of the terminal event). We propose a joint model of the recurrent event process and the terminal event, linked through a common subject-specific latent variable, in which the proportional intensity model is used for modeling the recurrent event process and the additive hazards model is used for modeling the terminal event time.

✉ John Angus
  john.angus@cgu.edu

  Xiaoyu Che
  xc2273@cumc.columbia.edu

[1] Center for Infection and Immunity, Columbia University, 722 W. 168th St., Room 1701, New York, NY 10032, USA

[2] Institute of Mathematical Sciences, Claremont Graduate University, 710 N. College Ave., Claremont, CA 91711, USA

1 Introduction

In many studies, the event of interest can be experienced more than once per subject. Such outcomes have been termed as recurrent events, which are commonly encountered in longitudinal follow-up studies. Examples of recurrent events include, among others, bladder tumor recurrence times among patients in a randomized treatment trial (Byar 1980), times to the development of mammary tumors for rats in a carcinogenicity experiment (Gail et al. 1980), infection occurrence times among leukemia patients receiving bone marrow transplants (Prentice et al. 1981), times to valve seat replacement on diesel engines in a service fleet (Lawless and Nadeau 1995), and times to inpatient hospital admissions among intravenous drug users (Wang et al. 2001).

In practice, the data collection will have a fixed termination date, beyond which the occurrence times are censored. Recurrent events occur as naturally ordered multivariate failure time data. As a result, recurrent event data are often analyzed using methods of multivariate survival analysis (e.g., Prentice et al. 1981; Andersen and Gill 1982; Wei et al. 1989). However, applying these methods to recurrent event data requires care, since they are designed for a larger class of general data structures amenable to multivariate survival analysis. Correspondingly, the analysis of recurrent event data continues to be the subject of much methodological research, with interest focused on assessing the effects of covariates on certain features of the recurrent event process.

Let $N_i^*(t) = \int_0^t dN_i^*(u)$ be the number of recurrent events in $(0, t]$ for subject i, where $dN_i^*(t) = N_i^*(t + dt) - N_i^*(t)$ denotes the number of events in the small time interval $(t, t + dt]$. Assume that subject i is observed over the period $[0, C_i]$, and is observed to experience events at times $T_{i1}, \ldots, T_{i,m_i}$, where C_i denotes the follow-up or censoring time and m_i denotes the total number of recurrent events for subject i. For subject i, let $\tilde{T}_{ij} = T_{ij} - T_{i,j-1}$ denote the time elapsed from the $(j-1)$th occurrence of the event to the jth occurrence of the event with $T_{i,0} \equiv 0$, $j = 1, \ldots, m_i$. $N_i^*(t)$ is thus a subject-specific counting process since (1) $N_i^*(t) \geq 0$, (2) $N_i^*(t)$ is integer-valued, (3) for $s < t$, $N_i^*(s) \leq N_i^*(t)$ and (4) for $s < t$, the number of events in $(s, t]$ is given by $N_i^*(t) - N_i^*(s)$ (e.g., Ross 1989).

Recurrent event data are usually analyzed in the context of counting or point process models (Andersen et al. 1993). Following classical survival analysis, these methods are based on modeling the intensity and hazard functions. Because the mean number of events is more easily interpreted than the hazard, some authors propose modeling the mean and rate functions (Pepe and Cai 1993; Lawless and Nadeau 1995; Lin et al. 2000). We now define these terms formally: intensity function, mean function and rate function.

Let $\mathcal{N}_i^*(t) = \{N_i^*(s); 0 \leq s < t\}$ denote the event history of the ith subject up to time t^-. If

$$E\{dN_i^*(t)|\mathcal{N}_i^*(t)\} = \lambda_i(t|\mathcal{N}_i^*(t))dt$$

then $\lambda_i(t|\mathcal{N}_i^*(t))$ is called the intensity function of $N_i^*(t)$. The probability distribution of $\{dN_i^*(t); t \geq 0\}$ can be determined completely in terms of $\lambda_i(t|\mathcal{N}_i^*(t))$ as discussed by Andersen et al. (1993). For Poisson processes, the intensity func-

tion is non-stochastic; that is $\lambda_i(t|\mathcal{N}_i^*(t)) = \lambda_i(t)$, and for renewal processes, $\lambda_i(t|\mathcal{N}_i^*(t)) = h_i(t - T_{i,N_i^*(t^-)})$ where $h_i(\cdot)$ is the hazard function for the inter-event times (which are iid) of the ith subject (Chiang 1968). The recurrent event process can be modeled as a Poisson process when event counts are of interest. The mean function of $N_i^*(t)$ is defined by $\mu_i(t) = E[N_i^*(t)]$. If $E[dN_i^*(t)] = r_i(t)dt$, then $r_i(t)$ is called the rate function of $N_i^*(t)$. When one wants to assess the effect of covariates on the process, analysis of the mean and rate functions is suggested, particularly since assumptions on the intra-subject dependence structure are avoided.

For recurrent event data, there are various models proposed in the survival analysis literature. These include conditional intensity models (Prentice et al. 1981; Andersen and Gill 1982; Chang and Wang 1999; Zeng and Lin 2006), marginal intensity models (Wei et al. 1989; Lee et al. 1992), the frailty model approach (Nielsen et al. 1992; Murphy 1994, 1995; Zeng and Lin 2007), and marginal means and rates models (Pepe and Cai 1993; Lawless and Nadeau 1995; Lin et al. 2000, 2001; Ghosh 2004; Schaubel et al. 2006; Sun and Su 2008; Liu et al. 2010; Sun et al. 2011, 2012). Cook and Lawless (2007) provided a broad review of the existing literature for the analysis of this type of data.

Often, however, there may exist a dependent terminal event that stops the recurrent events and the follow-up. This terminal event is usually correlated with the recurrent events of interest, and this correlation should be accounted for in the analysis. Indeed, this correlation can be counter intuitive. For example, frequent visits to the hospital for treatments can be positively correlated (e.g. aggressive, invasive, or risky therapies) or negatively correlated (physical therapy or dialysis) with a high death rate. Mazroui et al. (2010) emphasized the importance of ascertaining the nature of the dependence between the terminal event and the recurrent event process in the medical context, citing the reason that some therapies may be beneficial in slowing the rate of disease recurrence, but not prolonging survival. Not surprisingly, modeling the dependence between the terminal event and the recurrent events (as opposed to simply studying the recurrent event process) has gained more importance over the past few years.

The existing methods for handling recurrent event data in the presence of a terminal event generally fall into two approaches: frailty methods and marginal methods. Frailty models use random effects to account for the correlation between the recurrent and terminal events (Mazroui et al. 2010; Wang et al. 2001; Huang and Wang 2004; Liu et al. 2004; Ye et al. 2007; Zeng and Lin 2009). Marginal methods focus on the marginal rates of the recurrent and terminal events, leaving the correlation between the recurrent and terminal events unspecified (Cook and Lawless 1997; Ghosh and Lin 2000, 2002; Miloslavsky et al. 2004; Pan and Schaubel 2009; Zeng and Cai 2010).

We note that Huang and Wang (2004) just used the proportional hazards model for the terminal event. It is known that sometimes the proportional hazards model may not fit failure time data well. When this is the case, one alternative is the additive hazards model. The latter describes a different aspect of the relationship between the survival time and covariates and in many situations could be more plausible than the proportional hazards model (Lin and Ying 1994). For example, in public health studies, the risk difference described by the additive hazards model is the quantity used more often than the risk ratio described by the proportional hazards model. This is discussed further in Breslow and Day (1987). Dunson and Herring (2005) proposed

a new Bayesian model selection and averaging procedure which can be used to choose between proportional and additive models.

In this paper, we propose a joint model with the recurrent event process and the terminal event linked through a common subject-specific latent variable, with the proportional intensity model used for modeling the recurrent event process and the additive hazards model used for modeling the terminal event time. This model is flexible in that no parametric assumptions on the distributions of censoring times and latent variables are made. In Sect. 2, we describe the proposed models. An estimating procedure and asymptotic properties for the parameter estimators are established in Sect. 2.1. Section 2.2 reports some results from simulation studies conducted for evaluating the finite sample performance of the proposed method. In Sect. 3, the methodology is applied to hospitalization data for heart failure patients from the clinical data repository at University of Virginia Health System. All technical proofs are included in the Appendix.

2 Model specification

Let $N^*(t)$ denote the number of recurrent events over the time interval $(0, t]$, and let X be a p-dimensional vector of covariates. A non negative latent variable v is associated with the recurrent process, satisfying the property that $E(v|X) = 1$. Let D be the terminating event time (e.g., death) and C be the follow-up or censoring time. Write $Y = C \wedge D$ and $\delta = I(D \leq C)$, where $a \wedge b = \min(a, b)$, and $I(\cdot)$ is the indicator function. The deterministic time $\tau > 0$ signifies the end of the observation period, so that $C \leq \tau$ with probability 1. Due to censoring, $N^*(\cdot)$ is not fully observed, and the number of observed events is denoted by $N(t) = N^*(t \wedge Y)$. We denote by m the total number of recurrent events to occur; that is, $m = N(Y)$. The random object $(N^*(\cdot), X, C, D, Y, v, \delta, m)$ will signify a model for a subject chosen at random, and subscripts assigned to distinguish subjects. Thus, for a random sample of n subjects, $\left(N_i^*(\cdot), X_i, C_i, D_i, Y_i, v_i, \delta_i, m_i\right), i = 1, 2, \ldots, n$ will be the associated independent and identically distributed random objects. Note that for this sequence, only $(N_i(\cdot), Y_i, \delta_i, X_i, m_i), \ i = 1, \ldots, n$, is observable.

For the recurrent event process, we assume that conditional on X_i and a latent variable v_i, $N_i^*(\cdot)$ is a Poisson process with intensity function

$$E(dN_i^*(t)|v_i, X_i) = d\Lambda_i(t) = v_i \exp\{\gamma_0' X_i\} d\Lambda_0(t) \tag{1}$$

where $\Lambda_0(t)$ is the unspecified deterministic baseline cumulative intensity function and γ_0 is the regression parameter. Note that the intensity function $E(dN_i^*(t)|v_i, X_i)$ in model (1) is not conditional on the event history. It is valid in our case because, due to the memory-less property of the Poisson Process, the intensity function is equivalent to the rate function of the recurrent event process. The latent variable v_i is treated as a nuisance parameter and no parametric assumptions are imposed. Since any other positive value of the mean of v_i can be absorbed into $\Lambda_0(t)$, for identifiability of model (1), we assume that v_i is non negative and has conditional mean 1 given X_i. Thus, v_i

plays the role of a frailty parameter in the recurrent event process, with the intensity of recurrent events increasing with v_i.

We specify the additive hazards model for the terminal event time D_i as

$$\alpha_i(t) = \alpha_0(t) + \eta_0' X_i + \theta_0 v_i \tag{2}$$

for $t \geq 0$, where $\alpha_0(\cdot)$ is an unspecified deterministic baseline hazard function, η_0 is a p-dimensional vector of unknown regression parameters, and θ_0 is an unknown parameter. Note that (2) defines the conditional (on v_i and X_i) hazard function when the quantity on the right is non negative. This hazard function is 0 otherwise.

It is assumed that conditional on X_i and v_i, D_i, C_i and $N_i^*(\cdot)$ are mutually independent. Thus, (1) and (2) link the recurrent event process with the terminal event for a given individual i through the random effect v_i. The parameter θ_0 conveys information about the correlation (conditional on X_i) between the conditional hazard function for the terminal event time, and the conditional intensity function of the recurrent event process for a given individual. Specifically, if we denote (with t and X_i held fixed) $f_i(v_i) = v_i \exp\{\gamma_0' X_i\} d\Lambda_0(t)/dt$ and $g_i(v_i) = \max\{0, \alpha_0(t) + \eta_0' X_i + \theta_0 v_i\}$, then f_i and g_i are both non-decreasing in v_i when $\theta_0 \geq 0$, and when $\theta_0 \leq 0$, f_i is non-decreasing in v_i while g_i is non-increasing in v_i. Therefore, the conditional (on X_i) covariance of $f(v_i)$ and $g(v_i)$ is non-negative if $\theta_0 \geq 0$, and is non-positive if $\theta_0 \leq 0$. It follows that the conditional recurrent event intensity and conditional hazard function for the terminal event time for individual i are non-negatively correlated when $\theta_0 \geq 0$, and non-positively correlated when $\theta_0 \leq 0$. Thus, testing the null hypothesis $\theta_0 = 0$ versus the alternatives $\theta_0 > 0$ or $\theta_0 < 0$ will provide evidence on how the recurrent event process influences the risk of the terminal event; if the null hypothesis is rejected in favor of $\theta_0 > 0$, then higher intensities of recurrent events are associated with an increased risk of the terminal event occurring, while if the null hypothesis is rejected in favor of $\theta_0 < 0$, then higher intensities of recurrent events are associated with lower risk of the terminal event occurring.

2.1 Estimation of regression parameters

The model we have outlined is semiparametric, since it involves both finite and infinite dimensional unknowns. Since we make no assumptions about the form of the distribution of the latent variable or follow-up/censoring times, we use estimating-equation procedures to form our parameter estimators. Estimators formed in this way can be semiparametric efficient under certain regularity conditions. We do not attempt such a discussion here, opting instead for our minimal assumptions and the tractability of the estimating-equation technique.

Note that we are using the same informative censoring model that was studied by Wang et al. (2001), who focused on the estimation of parameters in model (1). They treated the nonparametric distributions of the censoring and latent variable as nuisance parameters, and were able to formulate consistent, asymptotically normal estimators for the regression parameters γ_0. Thus, our focus is mainly on estimation results for model (2), with distributional results being asymptotic. Our purpose is to demon-

strate the importance of taking into account the association between the recurrent event process and the terminal event process when drawing inferences on the covariate effects on the terminal event. In Sect. (3), we examine this issue using marginal p-values (based on asymptotic distributions) computed using data from hospitalization data for heart failure patients from the clinical data repository at University of Virginia Health System. For this, we find it sufficient here to derive the joint asymptotic distribution of the parameter estimators in model (2) only, rather than to solve the more complicated problem of doing this for the joint distribution of all the parameter estimators in both models (1) and (2). It will be seen in this section that we use estimating equations for all the parameters in both models, but only solve the equations and develop the asymptotic distributions for the parameters from model (2). For the marginal asymptotic distribution needed for model (1), we use the results from Wang et al. (2001) in our analysis in Sect. (3) and in our simulation results. For analyses requiring simultaneous inferences on the parameters in both models (1) and (2), it will be necessary to derive the joint asymptotic distribution of the parameter estimators for both models. For now, we leave this as a open question for future research.

Let $N^D(t) = I(Y \leq t, \delta = 1)$, $\Delta(t) = I(t \leq Y)$, and τ be the end time of the study. Estimation equations involving the parameters of interest can be written down easily based on the following population equations, where $Q(t)$, $t \geq 0$, is a (possibly data-dependent) weight function:

$$E\left[dN^D(t) - \left\{\eta'X + \theta v + \alpha_0(t)\right\}\Delta(t)dt\right] = 0$$

$$E\left[\int_0^\tau Q(t)X\left\{dN^D(t) - \{\eta'X + \theta v + \alpha_0(t)\}\Delta(t)dt\right\}\right] = 0$$

$$E\left[\int_0^\tau Q(t)v\left\{dN^D(t) - \{\eta'X + \theta v + \alpha_0(t)\}\Delta(t)dt\right\}\right] = 0$$

Note that if v can be observed, then using the generalized estimating equation approach (Liang and Zeger 1986), the parameters for the terminal event time (2) can be estimated as follows. Adding subject-specific subscripts, and since $\left(N_i^D(\cdot), \Delta_i, X_i, v_i\right), i = 1, 2, \ldots, n$ are independent and identically distributed, the estimating equations would be:

$$\sum_{i=1}^n dN_i^D(t) - \{\eta'X_i + \theta v_i + \alpha_0(t)\}\Delta_i(t)dt = 0,$$

$$\sum_{i=1}^n \int_0^\tau Q(t)X_i\left\{dN_i^D(t) - \{\eta'X_i + \theta v_i + \alpha_0(t)\}\Delta_i(t)dt\right\} = 0,$$

$$\sum_{i=1}^n \int_0^\tau Q(t)v_i\left\{dN_i^D(t) - \{\eta'X_i + \theta v_i + \alpha_0(t)\}\Delta_i(t)dt\right\} = 0.$$

However, in practice, v_i can not be observed. For this, define $\Lambda_i^*(t) = \Lambda_0(t)e^{\gamma_0'X_i}$ and $V_i^* = m_i\Lambda_i^*(Y_i)^{-1}$ for subject i. Note that $E(V_i^*|X_i, Y_i, \delta_i, v_i) = v_i$ and

$$E(V_i^*(m_i - 1)\Lambda_i^*(Y_i)^{-1}|X_i, Y_i, \delta_i, v_i) = v_i^2.$$

Thus, for given γ_0 and $\Lambda_0(t)$, we can estimate $\alpha_0(t)$, η_0 and θ_0 using the following three unbiased estimating equations:

$$\sum_{i=1}^n dN_i^D(t) - \{\eta' X_i + \theta V_i^* + \alpha_0(t)\}\Delta_i(t)dt = 0, \tag{3}$$

$$\sum_{i=1}^n \int_0^\tau Q(t)X_i \Big\{dN_i^D(t) - \{\eta' X_i + \theta V_i^* + \alpha_0(t)\}\Delta_i(t)dt\Big\} = 0, \tag{4}$$

$$\sum_{i=1}^n \int_0^\tau Q(t)V_i^* \Big\{dN_i^D(t) - \{\eta' X_i + \theta(m_i - 1)\Lambda_i(Y_i)^{-1} + \alpha_0(t)\}\Delta_i(t)dt\Big\} = 0. \tag{5}$$

Of course, γ_0 and $\Lambda_0(t)$ are unknown, but they can be consistently estimated by fitting the model (1). Following Wang et al. (2001), given v_i and X_i, the recurrent event process is a non homogeneous Poisson process under model (1). Since m_i denotes the total number of recurrent events for subject i, it follows that given (v_i, X_i, Y_i), m_i has a Poisson distribution with mean $v_i \Lambda_0(Y_i)e^{\gamma_0' X_i}$. Define $F(t) = \Lambda_0(t)/\Lambda_0(\tau)$. Then $F(t)$ can be estimated by

$$\widehat{F}(t) = \prod_{t < s \leq \tau} \left(1 - \frac{\sum_{i=1}^n dN_i^D(s)}{\sum_{i=1}^n \Delta_i(s)N_i^D(s)}\right).$$

Let $X_i^* = (1, X_i')'$, $\alpha_1 = \log \Lambda_0(\tau)$ and $\alpha = (\alpha_1, \gamma_0')'$. Using the generalized estimating equation approach, α can be estimated by solving

$$\frac{1}{n}\sum_{i=1}^n X_i^*(m_i \widehat{F}^{-1}(Y_i) - \exp\{\alpha' X_i^*\}) = 0. \tag{6}$$

Let $\widehat{\alpha} = (\widehat{\alpha}_1, \widehat{\gamma}')'$ denote the solution to the foregoing estimating equation. Then $\Lambda_0(t)$ can be estimated by $\widehat{\Lambda}_0(t) = e^{\widehat{\alpha}_1}\widehat{F}(t)$. Define $\widehat{\Lambda}_i(Y_i) = \widehat{\Lambda}_0(Y_i)e^{\widehat{\gamma}' X_i}$ and $\widehat{V}_i = m_i \widehat{\Lambda}_i(Y_i)^{-1}$.

Let $\mathcal{A}_0(t) = \int_0^t \alpha_0(s)ds$. Then for given η and θ, by replacing V_i^* with $\widehat{V}_i$ in the estimating equation (3), a reasonable estimator for $\mathcal{A}_0(t)$ is the solution to

$$\sum_{i=1}^n dN_i^D(t) - \{\eta' X_i + \theta \widehat{V}_i + \alpha_0(t)\}\Delta_i(t)dt = 0,$$

The solution to the above estimating equation is

$$\widehat{\mathcal{A}}_0(t; \eta, \theta) = \int_0^t \frac{\sum_{i=1}^n dN_i^D(s) - \{\eta' X_i + \theta \widehat{V}_i\}\Delta_i(s)ds}{\sum_{i=1}^n \Delta_i(s)}.$$

On replacement of $\mathcal{A}_0(t)$ with $\widehat{\mathcal{A}}_0(t; \eta, \theta)$ in (4) and (5), and by replacing V_i^* and $\Lambda_i(Y_i)$ with $\widehat{V}_i$ and $\widehat{\Lambda}_i(Y_i)$, respectively, we obtain the following two estimating functions for η and θ:

$$U_1(\eta, \theta) = \sum_{i=1}^{n} \int_0^{\tau} Q(t)\{X_i - \bar{X}(t)\}\Big\{dN_i^D(t) - \{\eta' X_i + \theta \widehat{V}_i\}\Delta_i(t)dt\Big\},$$

$$U_2(\eta, \theta) = \sum_{i=1}^{n} \int_0^{\tau} Q(t)\Big[\{\widehat{V}_i - \bar{V}(t)\}\{dN_i^D(t) - \eta' X_i \Delta_i(t)dt\}$$
$$- \theta\{\widehat{\Omega}_i - \widehat{V}_i \bar{V}(t)\}\Delta_i(t)dt\Big],$$

where

$$\widehat{\Omega}_i = m_i(m_i - 1)\widehat{\Lambda}_i(Y_i)^{-2},$$

and

$$\bar{X}(t) = \frac{\sum_{i=1}^{n} \Delta_i(t) X_i}{\sum_{i=1}^{n} \Delta_i(t)},$$

$$\bar{V}(t) = \frac{\sum_{i=1}^{n} \Delta_i(t) \widehat{V}_i}{\sum_{i=1}^{n} \Delta_i(t)}.$$

Let $\widehat{\eta}$ and $\widehat{\theta}$ denote the solutions to $U_1(\eta, \theta) = 0$ and $U_2(\eta, \theta) = 0$. Then $\widehat{\eta}$ and $\widehat{\theta}$ have the explicit form

$$\begin{pmatrix} \widehat{\eta} \\ \widehat{\theta} \end{pmatrix} = \widehat{A}^{-1}\left\{\frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} Q(t)\begin{pmatrix} X_i - \bar{X}(t) \\ \widehat{V}_i - \bar{V}(t) \end{pmatrix} dN_i^D(t)\right\},$$

where

$$\widehat{A} = \begin{pmatrix} \widehat{A}_{11} & \widehat{A}_{12} \\ \widehat{A}'_{12} & \widehat{A}_{22} \end{pmatrix},$$

and

$$\widehat{A}_{11} = \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} Q(t)\{X_i - \bar{X}(t)\}\{X_i - \bar{X}(t)\}'\Delta_i(t)dt,$$

$$\widehat{A}_{12} = \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} Q(t)\{X_i - \bar{X}(t)\}\{\widehat{V}_i - \bar{V}(t)\}\Delta_i(t)dt,$$

$$\widehat{A}_{22} = \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} Q(t)\{\widehat{\Omega}_i - \widehat{V}_i \bar{V}(t)\}\Delta_i(t)dt.$$

By using the law of large numbers and the consistency of $\widehat{\gamma}$ and $\widehat{\Lambda}_0(t)$ (Wang et al. 2001), one can show that $\widehat{\eta}$ and $\widehat{\theta}$ are consistent. To establish the asymptotic normality

of $\widehat{\eta}$ and $\widehat{\theta}$, let T_{ij} denote the occurrence time of the jth event of the ith subject, and let $P_{1n}(x, y, m)$ and $P_{2n}(x, y, m, \delta)$ denote the joint probability measures of (X_i, Y_i, m_i) and $(X_i, Y_i, m_i, \delta_i)$, respectively. Here we assume that the random vectors (X_i, Y_i, m_i) and $(X_i, Y_i, m_i, \delta_i)$ are defined in the same sample space for all subjects i, and are independent and identically distributed, so the probability measures P_{1n} and P_{2n} do not depend on i. Define

$$\widehat{H}(t) = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{m_i} I(T_{ij} \le t),$$

$$\widehat{R}(t) = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{m_i} I(T_{ij} \le t \le Y_i),$$

$$\widehat{\kappa}_i(t) = \sum_{j=1}^{m_i} \left\{ \int_t^\tau \frac{I(T_{ij} \le u \le Y_i)d\widehat{H}(u)}{\widehat{R}^2(u)} - \frac{I(t < T_{ij} \le \tau)}{\widehat{R}(T_{ij})} \right\},$$

$$\widehat{e}_i = X_i^* \left[\frac{m_i}{\widehat{F}(Y_i)} - \exp\{\widehat{\alpha}' X_i^*\} \right] - \int \frac{x^* m \widehat{\kappa}_i(y)d P_{1n}(x^*, y, m)}{\widehat{F}(y)},$$

and

$$\widehat{D}_1 = \frac{1}{n} \sum_{i=1}^{n} \exp\{\widehat{\alpha}' X_i^*\} X_i^{*\otimes 2},$$

where $v^{\otimes 2} = vv'$ for a column vector v. Furthermore, let $\widehat{\phi}_{1i}$ denote the vector $\widehat{D}_1^{-1}\widehat{e}_i$ without the first entry and $\widehat{\phi}_{2i}$ denote the first entry of $\widehat{D}_1^{-1}\widehat{e}_i$. Set $\widehat{\varphi}_i(t) = \widehat{\kappa}_i(t) + \widehat{\phi}_{2i}$, $\widehat{b}_i(c, w) = \widehat{\varphi}_i(c) + \widehat{\phi}'_{1i}w$ and $\widehat{\xi}_i = (\widehat{\xi}'_{1i}, \widehat{\xi}_{2i})'$, where

$$\widehat{\xi}_{1i} = \int_0^\tau Q(t)\{X_i - \bar{X}(t)\}d\widehat{M}_i^D(t)$$

$$+ \widehat{\theta} \int_0^\tau Q(t) \int \{x - \bar{X}(t)\} \frac{mI(y \ge t)}{\widehat{\Lambda}_0(y)e^{\widehat{\gamma}'x}} \widehat{b}_i(y, x)d P_{1n}(x, y, m)dt,$$

and

$$\widehat{\xi}_{2i} = \int_0^\tau Q(t) \left[\{\widehat{V}_i - \bar{V}(t)\}\{dN_i^D(t) - \Delta_i(t)(\widehat{\eta}' X_i dt + d\widehat{\mathcal{A}}_0(t))\} \right.$$

$$\left. - \widehat{\theta}\{\widehat{\Omega}_i(t) - \widehat{V}_i\bar{V}(t)\}\Delta_i(t)dt \right] - \int Q(y) \frac{m\delta}{\widehat{\Lambda}_0(y)e^{\widehat{\gamma}'x}} \widehat{b}_i(y, x)d P_{2n}(x, y, m, \delta)$$

$$+ \int_0^\tau Q(t) \int \{\widehat{\eta}'x - \widehat{\theta}\bar{V}(t)\} \frac{mI(y \ge t)}{\widehat{\Lambda}_0(y)e^{\widehat{\gamma}'x}} \widehat{b}_i(y, x)d P_{1n}(x, y, m)dt$$

$$+ \int_0^\tau Q(t) \int \frac{mI(y \ge t)}{\widehat{\Lambda}_0(y)e^{\widehat{\gamma}'x}} \widehat{b}_i(y, x)d P_{1n}(x, y, m)d\widehat{\mathcal{A}}_0(t)$$

$$+ \widehat{\theta} \int_0^\tau Q(t) \int \frac{2m(m-1)I(y \ge t)}{\{\widehat{\Lambda}_0(y)e^{\widehat{\gamma}'x}\}^2} \widehat{b}_i(y, x)d P_{1n}(x, y, m)dt,$$

where $d\widehat{M}_i^D(t) = dN_i^D(t) - \Delta_i(t)\{\widehat{\eta}'X_i + \widehat{\theta}\widehat{V}_i\}dt - \Delta_i(t)d\widehat{\mathcal{A}}_0(t)$. The asymptotic joint normality of $\widehat{\eta}$ and $\widehat{\theta}$ is established in the following theorem with the proof given in the Appendix.

Theorem 2.1 *Under the regularity conditions (R1)-(R4) stated in the Appendix,* $n^{1/2}(\widehat{\eta} - \eta_0)$ *and* $n^{1/2}(\widehat{\theta} - \theta_0)$ *have asymptotic joint normal distribution with mean zero and covariance matrix that can be consistently estimated by* $\widehat{A}^{-1}\widehat{\Sigma}\widehat{A}^{-1}$, *where* $\widehat{\Sigma} = n^{-1}\sum_{i=1}^n \widehat{\xi}_i^{\otimes 2}$.

We note finally that the estimates $\widehat{\mathcal{A}}_0(t; \widehat{\eta}, \widehat{\theta})$ and $\widehat{\Lambda}_0(t)$ may not be non decreasing functions. To correct for this, we could replace them above by

$$\widehat{\mathcal{A}}_0^\star(t; \widehat{\eta}, \widehat{\theta}) = \max_{s \leq t} \widehat{\mathcal{A}}_0(s; \widehat{\eta}, \widehat{\theta})$$

$$\widehat{\Lambda}_0^*(t) = \max_{s \leq t} \widehat{\Lambda}_0(s)$$

as suggested in Lin and Ying (1994). Lin and Ying (1994) pointed out however that under regularity conditions, we would have $\widehat{\mathcal{A}}_0^\star(\cdot) - \widehat{\mathcal{A}}_0(\cdot) = o_p(n^{-\frac{1}{2}})$ and $\widehat{\Lambda}_0^*(\cdot) - \widehat{\Lambda}_0(\cdot) = o_p(n^{-\frac{1}{2}})$, so that the asymptotic distribution of $(n^{1/2}(\widehat{\eta} - \eta_0), n^{1/2}(\widehat{\theta} - \theta_0))$ would be unaffected.

2.2 Simulation studies

In this section, we conduct simulation studies to examine the finite sample performance of the estimators proposed for the models in Sect. 2.1. In the simulation studies, for each subject i, we generate the covariate $X_i = (X_{i1}, X_{i2})'$ where X_{i1} is generated from a Bernoulli distribution with success probability 0.5 and X_{i2} is generated from a uniform distribution on (0,1). The latent variable v_i is generated from a gamma distribution with mean 1 and variance 1 if $X_{i1} = 1$ and a uniform distribution on (0.5, 1.5) otherwise. It can be easily verified that v_i is non negative and $E(v_i|X_i) = 1$. Given X_i and v_i, the terminal event time D_i is assumed to follow an additive hazards model:

$$\alpha_i(t) = 1 + \eta_0'X_i + \theta_0 v_i,$$

where $\eta_0 = (-0.5, 0.5)'$ or $(0.5, -0.5)'$, and $\theta_0 = -0.5$ or 0.5. The censoring time C_i is taken as $\min(E + 0.5, \tau)$, where E is exponentially distributed with mean 1 and the end time τ is taken to be $\tau = 2.5$, which yields a censoring percentage ranging between 22 and 39 %.

For the recurrent event process, given X_i, v_i and $Y_i = D_i \wedge C_i$, the recurrent event times are generated from a Poisson process with the intensity function:

$$\lambda_i(t) = v_i \exp\{\gamma_0'X_i\}\lambda_0(t),$$

where $\gamma_0 = (-0.5, 0.5)'$ or $(0.5, -0.5)'$, and $\lambda_0(t) = 2.5(1+t)$. The average number of the recurrent events per subject ranges from 2.31 to 4.23 for different model para-

meters. For each simulation study, we take the weighting function to be $Q(\cdot) \equiv 1$. Sample sizes are chosen to be $n = 200$ and $n = 300$. In order to help assess the accuracy of our simulation estimates, we apply a stratified sampling scheme in the simulation. For each n, we simulate 4000 replications in $m = 20$ groups of 200 replications each. For each replication, we compute the estimator, and average it over the 200 replications within each of the m groups (strata). By the Central Limit Theorem, these 20 averages are nearly iid and normally distributed, and thus a confidence interval for the true parameter using the standard confidence interval based on the t-distribution with $m - 1 = 19$ degrees of freedom can be computed from the tabled values using the standard errors in parentheses next to each simulation estimate. This allows one to assess how many significant digits are likely to be accurate in the simulation estimates.

Tables 1, 2, and 3 present the simulation results for the estimation of γ_0, η_0 and θ_0 respectively. The tables include the bias (BIAS) which is equal to the estimated value minus the true value, the model-based standard deviation estimates (SE), the non model-based standard deviation estimates (ESE), and the empirical coverage probabilities (CP) for the normal-approximation-based nominal 95 % confidence interval based on Theorem 2.1. As mentioned above, the stratified sampling allows us to calculate the simulation standard errors of all the estimates, and these standard errors are reported in parentheses after each estimate in the tables.

To make all the definitions clear, and take the estimation of η for example, let $\widehat{\eta}_i$ be the estimate of η for the ith simulation run, where $i = 1, 2, \ldots, 4000$. Also let $\widehat{\sigma}_i$ be the estimated standard deviation of $\widehat{\eta}_i$ according to Theorem 2.1. Define $\bar{\eta}_j$ and $\bar{\sigma}_j$ be the mean of $\widehat{\eta}_i$ and $\widehat{\sigma}_i$ respectively within the jth stratum, where $j = 1, 2, \ldots, 20$. Also let $\tilde{\sigma}_j$ denote the sample standard deviation of $\widehat{\eta}_i$ within the the jth stratum. SE and ESE are the average values of $\{\bar{\sigma}_j\}$ and $\{\tilde{\sigma}_j\}$ respectively over all j, and the sample standard deviations of $\{\bar{\eta}_j\}$, $\{\bar{\sigma}_j\}$ and $\{\tilde{\sigma}_j\}$ over all j and then divided by $\sqrt{m}$ give the simulation standard errors of all three estimates.

Tables 1, 2, and 3 show that our proposed method performs reasonably well. Specifically, the proposed estimators are usually close to the true parameter values, but appear to be slightly biased. We followed up by checking our simulation results based on a larger range of values of n and verified (statistically) that the bias is $O(n^{-1})$ and therefore negligible in terms of the asymptotic normality result of Theorem 2.1, which makes our estimation method at least comparable to the Monte Carlo EM algorithm proposed by Liu et al. (2004). There is also good agreement between the model-based and non model-based standard deviations. The performance of the proposed estimator becomes better when the sample size increases from 200 to 300 as expected. And, the confidence intervals based on Theorem 2.1 all have reasonable estimated coverage probabilities.

3 An application

Now we apply the proposed method to the hospitalization data for heart failure patients from the clinical data repository at University of Virginia Health System. They investigated 1475 heart failure patients in the study, in which about 20 % of the patients died before the censoring time. In the study, they measured three baseline covariates

Table 1 Simulation results for estimation of γ_0

n	γ_0	η_0	θ_0	$\widehat{\gamma}_1$				$\widehat{\gamma}_2$			
				BIAS	SE	ESE	CP	BIAS	SE	ESE	CP
200	$(-0.5, 0.5)'$	$(-0.5, 0.5)'$	-0.5	$-0.006(0.003)$	$0.216(0.002)$	$0.217(0.003)$	0.942	$0.006(0.006)$	$0.3722(0.008)$	$0.376(0.010)$	0.950
			0.5	$0.009(0.005)$	$0.258(0.004)$	$0.261(0.004)$	0.952	$-0.002(0.006)$	$0.398(0.007)$	$0.401(0.007)$	0.938
		$(0.5, -0.5)'$	-0.5	$-0.005(0.003)$	$0.217(0.003)$	$0.217(0.004)$	0.950	$0.003(0.006)$	$0.343(0.006)$	$0.346(0.007)$	0.948
			0.5	$0.013(0.004)$	$0.263(0.001)$	$0.261(0.005)$	0.947	$0.008(0.005)$	$0.395(0.008)$	$0.398(0.009)$	0.956
	$(0.5, -0.5)'$	$(-0.5, 0.5)'$	-0.5	$0.004(0.004)$	$0.226(0.004)$	$0.225(0.004)$	0.953	$0.000(0.006)$	$0.369(0.004)$	$0.373(0.006)$	0.948
			0.5	$-0.005(0.004)$	$0.262(0.003)$	$0.252(0.005)$	0.945	$-0.008(0.006)$	$0.438(0.006)$	$0.446(0.009)$	0.952
		$(0.5, -0.5)'$	-0.5	$0.001(0.004)$	$0.214(0.002)$	$0.213(0.004)$	0.949	$-0.008(0.006)$	$0.392(0.008)$	$0.389(0.010)$	0.947
			0.5	$0.003(0.004)$	$0.257(0.004)$	$0.252(0.003)$	0.953	$-0.001(0.007)$	$0.463(0.007)$	$0.469(0.008)$	0.948
300	$(-0.5, 0.5)'$	$(-0.5, 0.5)'$	-0.5	$-0.004(0.002)$	$0.187(0.002)$	$0.187(0.003)$	0.957	$0.006(0.005)$	$0.328(0.007)$	$0.326(0.008)$	0.954
			0.5	$-0.001(0.004)$	$0.225(0.004)$	$0.221(0.004)$	0.952	$0.006(0.006)$	$0.363(0.005)$	$0.362(0.006)$	0.947
		$(0.5, -0.5)'$	-0.5	$0.005(0.003)$	$0.175(0.004)$	$0.177(0.004)$	0.945	$-0.002(0.004)$	$0.277(0.003)$	$0.278(0.006)$	0.954
			0.5	$-0.008(0.003)$	$0.219(0.003)$	$0.222(0.006)$	0.950	$0.004(0.004)$	$0.336(0.006)$	$0.338(0.011)$	0.946
	$(0.5, -0.5)'$	$(-0.5, 0.5)'$	-0.5	$-0.004(0.003)$	$0.180(0.003)$	$0.182(0.004)$	0.948	$-0.007(0.004)$	$0.306(0.005)$	$0.307(0.006)$	0.953
			0.5	$-0.001(0.004)$	$0.208(0.002)$	$0.207(0.004)$	0.951	$-0.010(0.005)$	$0.369(0.006)$	$0.367(0.006)$	0.965
		$(0.5, -0.5)'$	-0.5	$0.001(0.003)$	$0.183(0.002)$	$0.184(0.003)$	0.952	$0.003(0.005)$	$0.324(0.006)$	$0.326(0.006)$	0.951
			0.5	$0.004(0.004)$	$0.212(0.003)$	$0.213(0.003)$	0.947	$-0.009(0.005)$	$0.383(0.006)$	$0.384(0.009)$	0.945

Table 2 Simulation results for estimation of η_0

n	γ_0	η_0	θ_0	$\widehat{\eta}_1$				$\widehat{\eta}_2$			
				BIAS	SE	ESE	CP	BIAS	SE	ESE	CP
200	$(-0.5, 0.5)'$	$(-0.5, 0.5)'$	-0.5	0.014(0.004)	0.198(0.001)	0.200(0.002)	0.947	-0.011(0.005)	0.329(0.001)	0.329(0.005)	0.956
			0.5	0.026(0.007)	0.345(0.005)	0.348(0.007)	0.958	0.007(0.007)	0.451(0.006)	0.452(0.008)	0.950
		$(0.5, -0.5)'$	-0.5	0.022(0.004)	0.214(0.002)	0.223(0.005)	0.946	-0.018(0.004)	0.335(0.002)	0.334(0.005)	0.947
			0.5	0.033(0.007)	0.324(0.009)	0.334(0.013)	0.944	-0.023(0.007)	0.431(0.008)	0.443(0.010)	0.963
	$(0.5, -0.5)'$	$(-0.5, 0.5)'$	-0.5	0.007(0.004)	0.200(0.002)	0.197(0.002)	0.954	-0.013(0.004)	0.332(0.002)	0.334(0.004)	0.951
			0.5	-0.002(0.006)	0.297(0.002)	0.297(0.005)	0.959	0.034(0.009)	0.408(0.005)	0.420(0.007)	0.949
		$(0.5, -0.5)'$	-0.5	0.025(0.005)	0.236(0.005)	0.246(0.006)	0.952	-0.003(0.006)	0.373(0.002)	0.372(0.003)	0.954
			0.5	-0.033(0.009)	0.295(0.007)	0.298(0.008)	0.949	0.035(0.007)	0.497(0.009)	0.505(0.015)	0.945
300	$(-0.5, 0.5)'$	$(-0.5, 0.5)'$	-0.5	0.009(0.002)	0.160(0.001)	0.159(0.003)	0.948	-0.008(0.004)	0.264(0.002)	0.263(0.004)	0.953
			0.5	0.010(0.006)	0.235(0.004)	0.234(0.005)	0.952	0.004(0.005)	0.307(0.005)	0.307(0.008)	0.954
		$(0.5, -0.5)'$	-0.5	0.013(0.003)	0.175(0.003)	0.177(0.004)	0.952	-0.008(0.004)	0.270(0.003)	0.268(0.004)	0.949
			0.5	0.016(0.007)	0.255(0.008)	0.259(0.010)	0.962	-0.015(0.003)	0.337(0.005)	0.341(0.007)	0.952
	$(0.5, -0.5)'$	$(-0.5, 0.5)'$	-0.5	0.002(0.002)	0.158(0.001)	0.158(0.002)	0.948	-0.007(0.004)	0.262(0.001)	0.259(0.003)	0.953
			0.5	-0.001(0.006)	0.210(0.002)	0.213(0.004)	0.959	0.017(0.005)	0.304(0.004)	0.307(0.006)	0.946
		$(0.5, -0.5)'$	-0.5	0.021(0.003)	0.170(0.002)	0.173(0.003)	0.948	-0.002(0.006)	0.278(0.002)	0.278(0.003)	0.953
			0.5	0.011(0.007)	0.238(0.004)	0.240(0.006)	0.953	-0.011(0.006)	0.347(0.005)	0.352(0.008)	0.954

Table 3 Simulation results for estimation of θ_0

γ_0	η_0	θ_0	$n = 200$				$n = 300$			
			BIAS	SE	ESE	CP	BIAS	SE	ESE	CP
$(-0.5, 0.5)'$	$(-0.5, 0.5)'$	-0.5	0.024(0.004)	0.172(0.001)	0.178(0.006)	0.932	0.013(0.003)	0.140(0.002)	0.141(0.005)	0.945
		0.5	0.017(0.003)	0.258(0.006)	0.270(0.009)	0.946	0.008(0.003)	0.174(0.004)	0.172(0.005)	0.949
	$(0.5, -0.5)'$	-0.5	0.015(0.006)	0.209(0.004)	0.216(0.008)	0.947	$-0.007(0.005)$	0.167(0.005)	0.165(0.008)	0.952
		0.5	0.037(0.008)	0.385(0.007)	0.390(0.008)	0.950	0.019(0.007)	0.267(0.006)	0.265(0.008)	0.951
$(0.5, -0.5)'$	$(-0.5, 0.5)'$	-0.5	0.012(0.003)	0.177(0.002)	0.182(0.006)	0.945	$-0.003(0.002)$	0.138(0.001)	0.140(0.004)	0.953
		0.5	$-0.006(0.001)$	0.248(0.004)	0.245(0.007)	0.958	0.001(0.001)	0.154(0.003)	0.156(0.005)	0.957
	$(0.5, -0.5)'$	-0.5	0.032(0.006)	0.241(0.004)	0.253(0.006)	0.947	0.025(0.004)	0.163(0.003)	0.169(0.006)	0.952
		0.5	$-0.021(0.004)$	0.333(0.005)	0.341(0.006)	0.958	0.016(0.004)	0.230(0.004)	0.231(0.007)	0.950

Table 4 Application results for recurrent events

	Est	SE	p value
Race	−0.2225	0.0873	0.0108
Gender	−0.0304	0.0701	0.6647
Age	0.1539	0.0386	0.0001

Table 5 Application results for terminal events

	Our model			Lin and Ying (1994)		
	Est	SE	p value	Est	SE	p value
Race	−0.0080	0.0049	0.1027	−0.2266	0.1277	0.0760
Gender	0.0031	0.0039	0.4266	0.2495	0.1193	0.0365
Age	0.0077	0.0030	0.0103	0.2784	0.0758	0.0002
θ_0	0.0479	0.0061	<0.0001			

for each patient: race, gender and age (ranged from 60 to 90). For individual i, let X_{i1} be a binary indicator of race (white = 1, nonwhite = 0), X_{i2} be a binary indicator of gender (male = 1, female = 0) and X_{i3} be the regulated age (centered at 72 and divided by 10). The dataset also consists of the information of longitudinal medical cost, however, we disregard this part of the data because we are only interested in the analysis of the hospitalization as recurrent event data.

Table 4 shows the application results of model (1) for the recurrent hospitalizations using the procedure proposed in Sect. 2.1. Table 5 shows the application results of model (2) for the terminal event with $Q(\cdot) \equiv 1$ comparing with those in Lin and Ying (1994) without taking into account of the recurrent events.

Est is the estimate of the parameter, and SE is the standard error estimate. The results using our model suggest that white patients tend to be at less risk for hospitalization, and so are younger patients. Age also has a significant effect on the hazard rate of failure, but race is not so significant. The estimate of the parameter θ_0 indicates that, not surprisingly in this application, the correlation between the death rate and the rate of hospitalization is positive, significant at 4.08×10^{-15} level. When comparing our results to those from the additive hazards model alone, we see that covariate effects, especially that of gender, diminish after adjusting for the frailty term, which further strengthens our belief that association between the recurrent event process and the terminal event process is important to be taken into account.

To check the adequacy of the model for the data, following Lin et al. (2000) and Zeng and Cai (2010), we examine the total summation of the residuals for each subject

$$\widehat{M}_i^D = \int_0^{Y_i} \left[dN_i^D(t) - \{\widehat{\eta}' X_i + \widehat{\theta} \widehat{V}_i\} dt - d\widehat{\mathcal{A}}_0(t; \widehat{\eta}, \widehat{\theta}) \right],$$

which has an approximate mean zero and should be approximately independent of X_i under model (2). Thus, a simple graphical procedure for assessing the adequacy
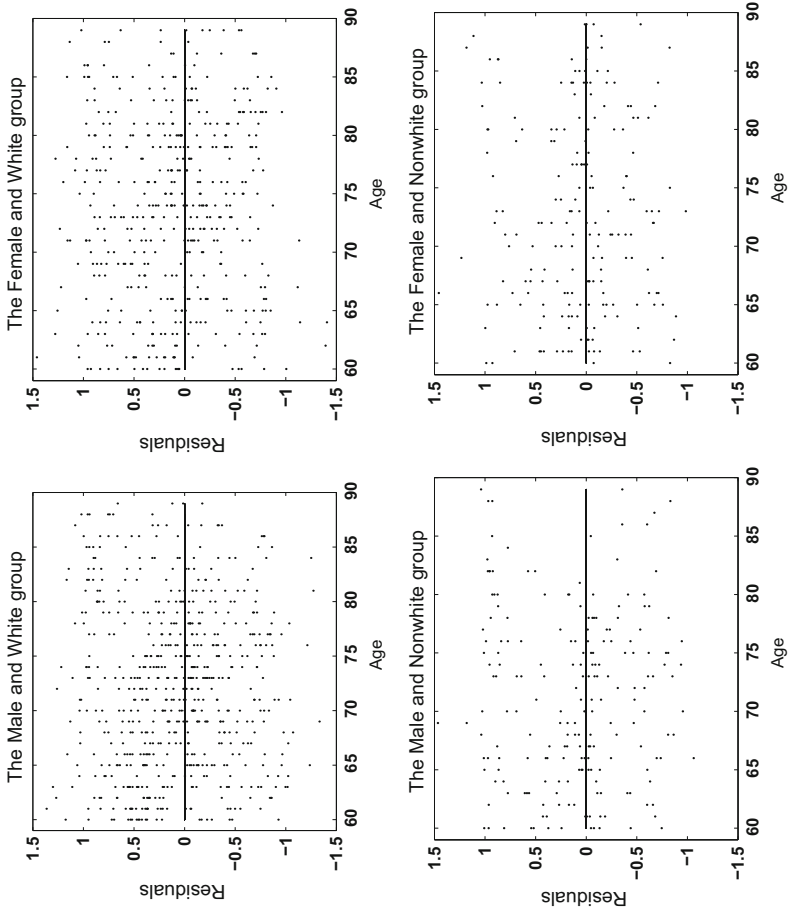
Fig. 1 The residual plot of the failure time process versus age for the subjects in the four groups

of the assumed model is to plot the residual against the covariate X_i. In Fig. 1, we divide the patients into four different groups based on their race and gender (Male and White, Male and Nonwhite, Female and White, Female and Nonwhite), and plot the cumulative residuals $\widehat{M}_i^D$ versus age for each group. It can be seen from the figure that the residuals have no trend, which indicates that our model fits the data well.

More formally, we apply some model checking techniques to check the goodness-of-fit of the proposed model by doing randomness tests on the distribution of the Median-Absolute-Deviation (MAD) normalized residuals, where the MAD of the residuals is defined as

$$MAD = median_i(|\widehat{M}_i^D - median_j(\widehat{M}_j^D)|).$$

First we define the measure of discrepancy between two cumulative distributions functions (CDF) to be $D(\mathcal{F}_1, \mathcal{F}_2) = \sup_{-\infty \leq t \leq \infty} |\mathcal{F}_1(t) - \mathcal{F}_2(t)|$. Let $\mathcal{G}$ be the empirical CDF of the MAD-divided residuals from the fit to the hospitalization data. Then we generate $n = 20$ independent sets of residuals by fitting the model using the estimated parameters from the fit to the hospitalization data, which gives us n empirical CDFs that are independent and identically distributed estimates of the null CDF of the normalized residuals. Let $\mathcal{F}_n$ be the average of these n empirical CDFs and $D(\mathcal{G}, \mathcal{F}_n)$ the discrepancy between the real data fit from the simulated null distribution.

To estimate the p-value associated with $D(\mathcal{G}, \mathcal{F}_n)$, we apply bootstrap simulations of $N = 1000$ replications to generate N independent and identically distributed observations $\{D_1, D_2, \ldots, D_N\}$. For each replication i ($i = 1, 2, \ldots, N$), we generate $n = 20$ new independent runs of the simulation using real-data fitted parameters, and again this gives us n estimates of the null CDFs of the normalized residuals. Let $\mathcal{H}_n$ be the average of these n empirical CDFs. We then generate one additional empirical CDF $\mathcal{T}$ using the real-data fitted parameters, and $D_i = D(\mathcal{T}, \mathcal{H}_n)$. The bootstrap estimate of the p-value associated with $D(\mathcal{G}, \mathcal{F}_n)$ is the fraction of cases in which $D_i < D(\mathcal{G}, \mathcal{F}_n)$, $i = 1, 2, \ldots, N$.

The algorithm yields the statistic $D(\mathcal{G}, \mathcal{F}_n) = 0.021$ with a p-value of 0.430 based on 1000 realizations of $D(\mathcal{T}, \mathcal{H}_n)$. This result indicates that there is no significant evidence to reject the goodness-of-fit of our model.

4 Conclusions

Frailty models are common for modeling jointly repeated measures and survival time data (Henderson et al. 2000; Lin et al. 2002). Here, we have proposed a joint model of the recurrent event process and the terminal event through a common subject-specific latent variable, in which the proportional intensity model is used for modeling the recurrent event process and the additive hazards model is used for modeling the terminal event time. The latent variable (frailty) is assumed to act as a multiplicative factor in the intensity function and an additive factor in the hazard function, and hence induces correlation between the recurrent event process and the terminal event time. A specific feature of our proposed model is that the frailty distribution is treated as a nuisance para-

meter and no parametric assumptions are imposed. Our estimation procedures can be easily implemented, and simulation results show that the proposed methods work well.

Our analysis is applicable under settings where recurrent event data are available for a large number of processes exhibiting a relatively small number of recurrent events. These types of processes arise frequently in medical studies, where information is often available on many individuals, each of whom may experience transient clinical events repeatedly over a period of observation. If the study ends not only with the censoring scheme but also can end with a terminal event such as death, this is where our proposed model would be most useful. Therefore our work can significantly influence applications in the biomedical fields. More precisely, by performing inferences based on the estimated parameters, we are able to analyze different effects of all kinds of possible determinants (such as age, gender, race, environmental condition, genetic information, or medical treatment) on the rates of the repeated clinical events and the risk of the death, and also to infer the correlation between the recurrent event rate and the risk of the terminal event . Note that a terminal event does not necessarily have to be death. For example, a major injury that causes the retirement of a race horse can be treated as a terminal event, while the injuries during its career are considered to be the recurrent event process. All these types of data sets can be analyzed through our model.

In applying our model to the heart failure data, the frailty term coefficient in (2) allows us to determine that the recurrent event process (the rate of hospitalization) and the terminal event process (death rate) are significantly correlated (positively, in this case). To our knowledge, ours is the first model of this type to allow explicit inference on this correlation, making it particularly important in biomedical applications.

5 Appendix: Proofs of asymptotic results

In this section, we will use the same notation defined above in Sect. 2.1, and all limits are taken as $n \to \infty$. Let $V_i = m_i \{\Lambda_0(Y_i) e^{\gamma_0' X_i}\}^{-1}$ and $\Omega_i = m_i(m_i-1)\{\Lambda_0(Y_i) e^{\gamma_0' X_i}\}^{-2}$. Recall that $(N_i(\cdot), Y_i, \delta_i, X_i, D_i, C_i, v_i)$, $i = 1, 2, \ldots$ are assumed to be independent and identically distributed copies of $(N(\cdot), Y, \delta, X, D, C, v)$, each i corresponding to a subject sampled at random.

In order to study the asymptotic distributions of $\widehat{\eta}$ and $\widehat{\theta}$, we need the following regularity conditions:

(R1) $P(Y \geq \tau, v > 0) > 0$, $P(Y > \tau_\varepsilon) = 1$, where $\tau_\varepsilon = \inf\{t : \Lambda_0(t) > \varepsilon\}$ for some $\varepsilon > 0$, and $E\{N(\tau)^2\} < \infty$.

(R2) $G(t) = E\{vI(Y \geq t) \exp(\gamma_0' X)\}$ is a continuous function for $t \in [0, \tau]$.

(R3) The weight function $Q(t)$ has bounded variation and converges to a deterministic function $q(t)$ in probability uniformly in $t \in [0, \tau]$;

(R4) A is non singular and $\Sigma = E(\xi_i \xi_i')$ exists.

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{12}' & A_{22} \end{pmatrix},$$

$$A_{11} = E\left\{ \int_0^\tau q(t)\{X - \bar{x}(t)\}^{\otimes 2} \Delta(t) dN(t) \right\},$$

$$A_{22} = E\left\{ \int_0^\tau q(t)\{V - \bar{v}(t)\}^2 \Delta(t) dN(t) \right\},$$

$$A_{12} = E\left\{ \int_0^\tau q(t)\{X - \bar{x}(t)\}\{V - \bar{v}(t)\} \Delta(t) dN(t) \right\},$$

where $\Delta(t) = I(t \leq Y)$, ξ_i is defined in Lemma 5.2 below, and $\bar{x}(t)$ and $\bar{v}(t)$ are the limits (in probability) respectively, of $\bar{X}(t)$ and $\bar{V}(t)$, which are well defined by regularity condition R1.

Define $R(t) = G(t)\Lambda_0(t)$,

$$H(t) = \int_0^t G(u) d\Lambda_0(u),$$

$$D_1 = E\{\exp\{\alpha_0' X_i^*\} X_i^{*\otimes 2}\},$$

$$\kappa_i(t) = \sum_{j=1}^{m_i} \left\{ \int_t^\tau \frac{I(T_{ij} \leq u \leq Y_i) dH(u)}{R^2(u)} - \frac{I(t < T_{ij} \leq \tau)}{R(T_{ij})} \right\},$$

and

$$e_i = X_i^* \left[\frac{m_i}{F(Y_i)} - \exp\{\alpha_0' X_i^*\} \right] - \int \frac{x^* m \kappa_i(y) dP_1(x^*, y, m)}{F(y)},$$

where T_{ij} denotes the occurrence time of the jth event of the ith subject, and $P_1(x^*, y, m)$ is the joint probability measure of (X_i^*, Y_i, m_i). Note that X_i and X_i^* share the same random variables, hereafter we use $P_1(x, y, m)$ to denote the joint probability measure of (X_i, Y_i, m_i). Let ϕ_{1i} denote the vector $D_1^{-1} e_i$ without the first entry and ϕ_{2i} denote the first entry of $D_1^{-1} e_i$. Set $\varphi_i(t) = \kappa_i(t) + \phi_{2i}$, and $b_i(y, x) = \varphi_i(y) + \phi_{1i}' x$.

Under conditions (R1) and (R2), it follows from Wang et al. (2001) that

$$n^{1/2}\{\widehat{\Lambda}_0(t) - \Lambda_0(t)\} = n^{-1/2} \Lambda_0(t) \sum_{i=1}^n \varphi_i(t) + o_p(1) \tag{7}$$

and

$$n^{1/2}\{\widehat{\gamma} - \gamma_0\} = n^{-1/2} \sum_{i=1}^n \phi_{1i} + o_p(1) \tag{8}$$

Lemma 5.1 (Lemma A.1 of Lin and Ying 2001) *Let $H_n(t)$ and $M_n(t)$ be two sequences of bounded processes. Suppose that $H_n(t)$ is monotone and converges uniformly to $H(t)$ in probability (i.e., $\sup_t |H_n(t) - H(t)|$ converges to 0 in probability) and that $M_n(t)$ converges weakly to a zero-mean process with continuous sample paths. Then $\int_0^t \{H_n(s) - H(s)\} dM_n(s)$ converges in probability to 0 uniformly in t.*

Lemma 5.2 *Under the joint models specified in* Sect. 2 *and condition (R4),* $n^{-1/2} \sum_{i=1}^{n} \xi_i$ *has an asymptotically normal distribution with mean zero and covariance matrix* $\Sigma = E(\xi_i \xi_i')$, *where* $\xi_i = (\xi_{1i}', \xi_{2i}')'$ *with*

$$
\xi_{1i} = \int_0^\tau q(t)\{X_i - \bar{x}(t)\}dM_i^D(t)
$$
$$
+ \theta_0 \int_0^\tau q(t) \int \{x - \bar{x}(t)\} \frac{mI(y \geq t)}{\Lambda_0(y)e^{\gamma_0'x}} b_i(y, x)dP_1(x, y, m)dt,
$$

and

$$
\xi_{2i} = \int_0^\tau q(t)\Big[\{V_i - \bar{v}(t)\}\{dN_i^D(t) - (\eta_0'X_i + \alpha_0(t))\Delta_i(t)dt\}
$$
$$
- \theta_0\{\Omega_i(t) - V_i\bar{v}(t)\}\Delta_i(t)dt\Big] - \int q(y) \frac{m\delta}{\Lambda_0(y)e^{\gamma_0'x}} b_i(y, x)dP_2(x, y, m, \delta)
$$
$$
+ \int_0^\tau q(t) \int \{\eta_0'x + \alpha_0(t) - \theta_0\bar{v}(t)\} \frac{mI(y \geq t)}{\Lambda_0(y)e^{\gamma_0'x}} b_i(y, x)dP_1(x, y, m)dt
$$
$$
+ \theta_0 \int_0^\tau q(t) \int \frac{2m(m-1)I(y \geq t)}{\{\Lambda_0(y)e^{\gamma_0'x}\}^2} b_i(y, x)dP_1(x, y, m)dt,
$$

where $dM_i^D(t) = dN_i^D(t) - \Delta_i(t)\{\eta_0'X_i + \theta_0 V_i + \alpha_0(t)\}dt$, *and* $P_1(x, y, m)$ *and* $P_2(x, y, m, \delta)$ *denote the joint probability measures of* (X_i, Y_i, m_i) *and* $(X_i, Y_i, m_i, \delta_i)$, *respectively.*

Proof Note that under the joint models specified in Sect. 2.1, $\xi_i, i = 1, \ldots, n$ are independent and identically distributed. Recalling the Central Limit Theorem and condition (R4), it suffices to verify that $E(\xi_i) = 0$. By the definition of $dM_i^D(t)$, under model (2) and the Poisson model (1), it has mean zero given the covariates. Hence, the first term of ξ_{1i} has mean zero. In addition, for fixed x and y, under model (1), we have

$$
E(b_i(y, x)) = E(\kappa_i(y)) + E(\phi_{2i}) + x'E(\phi_{1i}) = 0.
$$

By the Fubini theorem, we exchange the order of expectation, and can verify that the second term of ξ_{1i} has mean zero. So, we have verified that $E(\xi_{1i}) = 0$. Similarly, we can verify $E(\xi_{2i}) = 0$, which completes the proof.

Lemma 5.3 *Under the regularity conditions (R1)-(R3), we have*

$$
\widehat{V}_i - V_i = -m_i\{\Lambda_0(Y_i)e^{\gamma_0'X_i}\}^{-1}n^{-1}\left\{\sum_{j=1}^{n}\varphi_j(t) + \sum_{j=1}^{n}\phi_{1j}'X_i\right\} + o_p(n^{-1/2}) \quad (9)
$$

and

$$\widehat{\Omega}_i - \Omega_i = -2m_i(m_i-1)\{\Lambda_0(Y_i)e^{\gamma_0'X_i}\}^{-2}n^{-1}\Big\{\sum_{j=1}^n \varphi_j(t) + \sum_{j=1}^n \phi_{1j}'X_i\Big\} + o_p(n^{-1/2})$$
(10)

Proof By using the Taylor expansion theorem and (7), (8), we have for (9)

$$
\begin{aligned}
\widehat{V}_i - V_i &= m_i\{\widehat{\Lambda}_0(Y_i)e^{\widehat{\gamma}'X_i}\}^{-1} - m_i\{\Lambda_0(Y_i)e^{\gamma_0'X_i}\}^{-1} \\
&= -m_i\{\Lambda_0(Y_i)e^{\gamma_0'X_i}\}^{-2}\{e^{\gamma_0'X_i}\{\widehat{\Lambda}_0(Y_i)-\Lambda_0(Y_i)\}+\Lambda_0(Y_i)e^{\gamma_0'X_i}(\widehat{\gamma}-\gamma_0)'X_i\} \\
&\quad + o(\|\widehat{\gamma}-\gamma_0\| + |\widehat{\Lambda}_0(Y_i) - \Lambda_0(Y_i)|) \\
&= -m_i\{\Lambda_0(Y_i)e^{\gamma_0'X_i}\}^{-1}n^{-1}\Big\{\sum_{j=1}^n \varphi_j(t) + \sum_{j=1}^n \phi_{1j}'X_i\Big\} + o_p(n^{-1/2}),
\end{aligned}
$$

where $o_p(.)$ is independent of i since (7) holds uniformly in t (Huang and Wang 2004). Similarly, we can show that (10) holds.

Lemma 5.4 *Under the regularity conditions (R1)-(R3),we have*

$$
\begin{aligned}
n^{-1/2}U_1(\eta_0, \theta_0) &= n^{-1/2}\sum_{i=1}^n \int_0^\tau Q(t)\{X_i - \bar{x}(t)\}dM_i^D(t) \\
&\quad - \theta_0 n^{-1/2}\sum_{i=1}^n \int_0^\tau Q(t)\{X_i - \bar{x}(t)\}\{\widehat{V}_i - V_i\}\Delta_i(t)dt + o_p(1),
\end{aligned}
$$
(11)

and

$$
\begin{aligned}
n^{-1/2}U_2(\eta_0, \theta_0) &= n^{-1/2}\sum_{i=1}^n \int_0^\tau Q(t)\Big[\{V_i - \bar{v}(t)\}\{dN_i^D(t) - (\eta_0'X_i + \alpha_0(t))\Delta_i(t)dt\} \\
&\quad - \theta_0\{\Omega_i(t) - V_i\bar{v}(t)\}\Delta_i(t)dt\Big] \\
&\quad + n^{-1/2}\sum_{i=1}^n \int_0^\tau Q(t)\{\widehat{V}_i - V_i\}\{dN_i^D(t) - (\eta_0'X_i + \alpha_0(t))\Delta_i(t)dt\} \\
&\quad - \theta_0 n^{-1/2}\sum_{i=1}^n \int_0^\tau Q(t)\{\widehat{\Omega}_i - \Omega_i - \{\widehat{V}_i - V_i\}\bar{v}(t)\}\Delta_i(t)dt + o_p(1).
\end{aligned}
$$
(12)

Proof Straightforward calculations give

$$n^{-1/2}U_1(\eta_0, \theta_0) = n^{-1/2}\sum_{i=1}^n \int_0^\tau Q(t)\{X_i - \bar{x}(t)\}dM_i^D(t)$$

$$- \theta_0 n^{-1/2} \sum_{i=1}^{n} \int_0^{\tau} Q(t)\{X_i - \bar{x}(t)\}\{\widehat{V}_i - V_i\}\Delta_i(t)dt$$

$$- \int_0^{\tau} Q(t)\{\bar{X}(t) - \bar{x}(t)\}n^{-1/2} \sum_{i=1}^{n} \{dM_i^D(t) - \theta_0\{\widehat{V}_i - V_i\}\Delta_i(t)dt\}.$$

To obtain (11), it suffices to show that the third term is $o_p(1)$. We first show that

$$\int_0^{\tau} Q(t)\{\bar{X}(t) - \bar{x}(t)\}n^{-1/2} \sum_{i=1}^{n} dM_i^D(t) = o_p(1) \tag{13}$$

To see this, let $M^n(t) = \int_0^t -1/2 \sum_{i=1}^{n} dM_i^D(t)$. Then, by the functional Central Limit Theorem(Vaart and Wellner 1996), under conditions(R1)-(R3), $M^n(t)$ converges weakly to a Gaussian process G with continuous paths. By the Skorohod embedding theorem, there exists a sequence $\tilde{M}^n(t)$ which is equal to $M^n(t)$ in law and converges uniformly to a Gaussian process $\tilde{G}$ with continuous paths. In addition, by the uniform law of large numbers, $Q(t)\{\bar{X}(t) - \bar{x}(t)\}$ converges uniformly to 0 almost surely. Then, integrating by parts and applying Lemma 5.1, we have that

$$\int_0^{\tau} Q(t)\{\bar{X}(t) - \bar{x}(t)\}d\tilde{M}^n(t) = o_p(1),$$

which implies that

$$\int_0^{\tau} Q(t)\{\bar{X}(t) - \bar{x}(t)\}dM^n(t) = o_p(1).$$

Hence, (13) is obtained. Applying the expansion (9), we have

$$\int_0^{\tau} Q(t)\{\bar{X}(t) - \bar{x}(t)\}n^{-1/2} \sum_{i=1}^{n} \{\widehat{V}_i - V_i\}\Delta_i(t)dt$$

$$= -\int_0^{\tau} Q(t)\{\bar{X}(t) - \bar{x}(t)\}\Big\{\{n^{-1} \sum_{i=1}^{n} m_i\{\Lambda_0(Y_i)e^{\gamma_0' X_i}\}^{-1}\Delta_i(t)\}\{n^{-1/2} \sum_{j=1}^{n} \varphi_j(t)\}$$

$$+ \{n^{-1} \sum_{i=1}^{n} m_i\{\Lambda_0(Y_i)e^{\gamma_0' X_i}\}^{-1}X_i\Delta_i(t)\}'\{n^{-1/2} \sum_{j=1}^{n} \phi_{1j}\}\Big\}dt + o_p(1). \tag{14}$$

By the uniform law of large numbers, we have that $\bar{X}(t) - \bar{x}(t) = o_p(1)$, $n^{-1} \sum_{i=1}^{n} m_i\{\Lambda_0(Y_i)e^{\gamma_0' X_i}\}^{-1}\Delta_i(t) = O_p(1)$ and $n^{-1} \sum_{i=1}^{n} m_i\{\Lambda_0(Y_i)e^{\gamma_0' X_i}\}^{-1}X_i$ $\Delta_i(t) = O_p(1)$ uniformly in t. By the functional form of the Central Limit Theorem, we have $n^{-1/2} \sum_{j=1}^{n} \varphi_j(t) = O_p(1)$ uniformly in t. $n^{-1/2} \sum_{j=1}^{n} \phi_{1j} = O_p(1)$ follows by the Central Limit Theorem. Note that $Q(t)$ converges uniformly to $q(t)$ by

assumption (R3), so applying Lemma 5.1 again we obtain

$$\int_0^\tau Q(t)\{\bar{X}(t) - \bar{x}(t)\}n^{-1/2}\sum_{i=1}^n\{\widehat{V}_i - V_i\}\Delta_i(t)dt = o_p(1),$$

which completes the proof of (11). In a similar way, we can show that (12) also holds.

Lemma 5.5 *Under the regularity conditions (R1)-(R3), $n^{-1/2}U(\eta_0, \theta_0)$ has asymptotically a normal distribution with mean zero and covariance matrix $\Sigma = E(\xi_i\xi_i')$, where $U(\eta, \theta) = (U_1(\eta, \theta)', U_2(\eta, \theta)')'$ and ξ_i is defined in* Lemma 5.2.

Proof Combing the results of Lemmas 5.3 and 5.4, with justifications similar to those in the proof of (11), we obtain that

$$n^{-1/2}U(\eta_0, \theta_0) = n^{-1/2}\sum_{i=1}^n\xi_i + o_p(1).$$

Hence, the proof can be completed by applying Lemma 5.2.

Lemma 5.6 *Under the regularity conditions (R1)–(R4),*

$$n^{1/2}\begin{pmatrix}\widehat{\eta} - \eta_0 \\ \widehat{\theta} - \theta_0\end{pmatrix} = A^{-1}n^{-1/2}U(\eta_0, \theta_0) + o_p(1).$$

Proof Note that $-n^{-1}\partial U(\eta_0, \theta_0)/\partial(\eta', \theta') = \widehat{A}$ is independent of η_0 and θ_0 and converges in probability to A, which is defined in condition (R4). Straightforward calculation gives

$$n^{-1}U(\widehat{\eta}, \widehat{\theta}) - n^{-1}U(\eta_0, \theta_0) = -\widehat{A}\begin{pmatrix}\widehat{\eta} - \eta_0 \\ \widehat{\theta} - \theta_0\end{pmatrix}.$$

Therefore, based on the consistency of $\widehat{\eta}$ and $\widehat{\theta}$ proved in Sect. 2.1, we complete the proof.

Proof of Theorem 2.1

Proof Combining Lemmas 5.5 and 5.6, we have, $\widehat{\eta}$ and $\widehat{\theta}$ converge in probability to η_0 and θ_0, respectively. In addition $n^{1/2}(\widehat{\eta} - \eta_0)$ and $n^{1/2}(\widehat{\theta} - \theta_0)$ have asymptotically a joint normal distribution with mean zero and covariance matrix $A^{-1}\Sigma A^{-1}$.

References

Andersen PK, Gill RD (1982) Cox's regression model for counting processes: a large sample study. Ann Stat 10:1100–1120

Andersen PK, Borgan O, Gill RD, Keiding N (1993) Statistical models based on counting process. Springer, New York

Breslow NE, Day NE (1987) Statistical methods in cancer research. Volume II: the design and analysis of cohort studies. IARC Sci Publ 82:1–406

Byar DP (1980) The Veterans administration study of chemoprophylaxis for recurrent stage I bladder tumors: comparisons of placebo, pyridoxine and topical thiotepa. In: Pavone-Macaluso M, Smith PH, Edsmyr F (eds) Bladder tumors and other topics in urological oncology. Plenum, New York, pp 363–370

Chang SH, Wang MC (1999) Conditional regression analysis for recurrence time data. J Am Stat Assoc 94:1221–1230

Chiang CL (1968) Introduction to stochastic processes in biostatistics. Wiley, New York

Cook RJ, Lawless JF (1997) Marginal analysis of recurrent events and a terminating event. Stat Med 16:911–924

Cook RJ, Lawless JF (2007) The statistical analysis of recurrent events. Springer, New York

Dunson DB, Herring AH (2005) Bayesian model selection and averaging in additive and proportional hazards models. Lifetime Data Anal 11:213–232

Gail MH, Santner TJ, Brown CC (1980) An analysis of comparative carcinogenesis experiments based on multiple times to tumor. Biometrics 36:255–266

Ghosh D (2004) Accelerated rates regression models for recurrent failure time data. Lifetime Data Anal 10:247–261

Ghosh D, Lin DY (2000) Nonparametric analysis of recurrent events and death. Biometrics 56:554–562

Ghosh D, Lin DY (2002) Marginal regression models for recurrent and terminal events. Stat Sin 12:663–688

Henderson R, Diggle P, Dobson A (2000) Joint modelling of longitudinal measurements and event time data. Biostatistics 1:465–480

Huang CY, Wang MC (2004) Joint modeling and estimation of recurrent event processes and failure time data. J Am Stat Assoc 99:1153–1165

Lawless JF, Nadeau C (1995) Some simple robust methods for the analysis of recurrent events. Technometrics 37:158–168

Lee EW, Wei LJ, Amato DA (1992) Cox-type regression analysis for large numbers of small groups of correlated failure time observations. In: Klein JP, Goel PK (eds) Survival analysis: state of the art. Kluwer Acad. Publ, Dordrecht, pp 237–247

Liang KY, Zeger SL (1986) Longitudinal data analysis using generalized linear models. Biometrika 73:13–22

Lin DY, Ying Z (1994) Semiparametric analysis of the additive risk model. Biometrika 81:61–71

Lin DY, Ying Z (2001) Semiparametric and nonparametric regression analysis of longitudinal data. J Am Stat Assoc 96:103–126

Lin DY, Wei LJ, Yang I, Ying Z (2000) Semiparametric regression for the mean and rate function of recurrent events. J R Stat Soc B 69:711–730

Lin DY, Wei LJ, Ying Z (2001) Semiparametric transformation models for point processes. J Am Stat Assoc 96:620–628

Lin H, Turnbull BW, McCulloch CE, Slate EH (2002) Latent class models for joint analysis of longitudinal biomarker and event process data: Application to longitudinal prostate-specific antigen readings and prostate cancer. J Am Stat Assoc 97:53–65

Liu L, Wolfe RA, Huang X (2004) Shared frailty models for recurrent events and a terminal event. Biometrics 60:747–756

Liu YY, Wu YS, Cai JW, Zhou HB (2010) Additive-multiplicative rates model for recurrent events. Lifetime Data Anal 16:353–373

Mazroui Y, Mathoulin-Pelissier S, Soubeyran P, Rondeau V (2010) General joint frailty model for recurrent event data with a dependent terminal event: application to follicular lymphoma data. Stat Med 31:1162–1176

Miloslavsky M, Keles S, Van der Laan MJ, Butler S (2004) Recurrent events analysis in the presence of time-dependent covariates and dependent censoring. J R Stat Soc B 66:239–257

Murphy SA (1994) Consistency in a proportional hazards model incorporating a random effect. Ann Stat 22:712–731

Murphy SA (1995) Asymptotic theory for the frailty model. Ann Stat 23:182–198

Nielsen GG, Gill RD, Andersen PK, Sorensen TIA (1992) A counting process approach to maximum likelihood estimation in frailty models. Scand J Stat 19:25–43

Pan Q, Schaubel DE (2009) Flexible estimation of differences in treatment-specific recurrent event means in the presence of a terminating event. Biometrics 65:753–761

Pepe MS, Cai J (1993) Some graphical displays and marginal regression analysis for recurrent failure times and time-dependent covariates. J Am Stat Assoc 88:811–820

Prentice RL, Williams BJ, Peterson AV (1981) On the regression analysis of multivariate failure time data. Biometrika 68:373–379

Ross SM (1989) Introduction to probability models. Academic Press, New York

Schaubel DE, Zeng D, Cai J (2006) A semiparametric additive rates model for recurrent event data. Lifetime Data Anal 12:389–406

Sun L, Su B (2008) A class of accelerated means regression models for recurrent event data. Lifetime Data Anal 14:357–375

Sun L, Tong X, Zhou X (2011) A class of Box–Cox transformation models for recurrent event data. Lifetime Data Anal 17:280–301

Sun L, Zhao X, Zhou J (2012) A class of mixed models for recurrent event data. Can J Stat (in press)

van der Vaart AW, Wellner JA (1996) Weak convergence and empirical processes. Springer, New York

Wang MC, Qin J, Chiang CT (2001) Analyzing recurrent event data with informative censoring. J Am Stat Assoc 96:1057–1065

Wei LJ, Lin DY, Weissfeld L (1989) Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. J Am Stat Assoc 84:1065–1073

Ye Y, Kalbfleisch JD, Schaubel DE (2007) Semiparametric analysis of correlated recurrent and terminal events. Biometrics 63:78–87

Zeng D, Cai J (2010) Semiparametric additive rate model for recurrent events with informative terminal event. Biometrika 97:699–712

Zeng D, Lin DY (2006) Efficient estimation of semiparametric transformation models for counting processes. Biometrika 93:627–640

Zeng D, Lin DY (2007) Semiparametric transformation models with random effects for recurrent events. J Am Stat Assoc 102:167–180

Zeng D, Lin DY (2009) Semiparametric transformation models with random effects for joint analysis of recurrent and terminal events. Biometrics 65:746–752