Statistics and Its Interface Volume 10 (2017) 441–450

Bayesian semiparametric mixed-effects joint models for analysis of longitudinal-competing risks data with skew distribution

Tao Lu

The joint analysis of longitudinal competing risks data has received much attention recently. However, most joint models for this type of data assume parametric functions for both longitudinal and competing risks processes which has its limitation for practical use. Motivated by studying the relationship between two biomarkers modified by time in an AIDS study, we develop the semiparametric mixed-effects joint models for longitudinal-competing risks data analysis. The proposed models differ from existing models in that: i) the commonly used parametric models in the joint models are extended to semiparametric settings to account for irregular data observed in real studies; ii) we employ skew distributions for random errors to account for skewness in data. We propose a Bayesian approach to jointly model two processes which are connected through the share of random effects. An example from a recent AIDS clinical study illustrates the methodology by jointly modeling the viral load and time to death due to AIDS or other reasons to compare potential models with various scenarios and different distribution specifications. The analysis results show a strongly negative relationship between virologic and immunologic biomarkers and CD4 counts reduce risks from both AIDS and other causes. In addition, nonlinear time effects are observed on the viral load at the population level while individual variation is large. These findings may help us to design a better treatment strategy for AIDS patients.

Keywords and phrases: Bayesian inference longitudinal data, Competing risks, Longitudinal data, Competing risks, Partially linear mixed-effects models, Proportional hazard models, Skew distribution, Survival data.

1. INTRODUCTION

In evaluating a patient's success on AIDS treatment, two biomarkers are often used: viral load and CD4 counts. Viral load measures the amount of virus in plasma while CD4 counts is a commonly used measure to assess the strength of a patient's immunological system. A lower viral load and higher CD4 counts may indicate successful treatment of AIDS. After treatment initiation, we expect reduction of viral load and increment of CD4 counts. However, the respond of CD4 counts is usually slower to that of viral load. To model the complex relationship between the virologic and immunologic biomarkers, it is important to take into account the time effect that may modify the relationship. Further, the data on measurements of these biomarkers are collected longitudinally in AIDS studies and trajectories are quite irregular. Thus, it is generally difficult to adopt a parametric model for investigating their relationship in the longitudinal setting. As an example, Figure 1 a) presents the viral load data of 50 subjects from the real data that we will analyze later. It can be seen that there is large inter-subject variation and each subject demonstrates quite irregular pattern.

The situation is further complicated by a terminal event such as death. Because the terminal event may be related to the individual's measurements on biomarkers, the terminal mechanism is not ignorable. Further, there are multiple types of events that may affect the outcome. For example, a patient's death may be caused by HIV infection or other reasons. These risks compete with each other for the survival endpoint. To study the relationship between viral load and CD4 counts which may lead to development of new treatment strategies, it is crucial to evaluate time to death caused by competing risks. Nevertheless, little work has been done for the analysis of the important relationship between two biomarkers in the longitudinal-competing risks data setting.

Joint modeling of longitudinal data and time-to-event data has received much attention recently. Comparing to separate analysis of two processes, the joint model has the advantage of providing more efficient inferences on the time-to-event and longitudinal biomarkers by incorporating all information simultaneously. A common setup for the joint model assumes a mixed-effects model for the longitudinal process and a proportional hazard model for the time-to-event process. The two processes are usually linked through sharing or association of random-effects. The dependency between the longitudinal and time-to-event processes is therefore governed by the latent process characterized by the random-effects. Earlier joint models for longitudinal-survival data analysis deal with univariate time-to-event data [1, 2, 3, 4, 5, 6]. Recently, they are extended to account for more complex cases, such as competing risks data [7, 8, 9, 10, 11], recurrent events [12, 13, 14]. [15] used accelerated failure time model for joint modeling of longitudinal

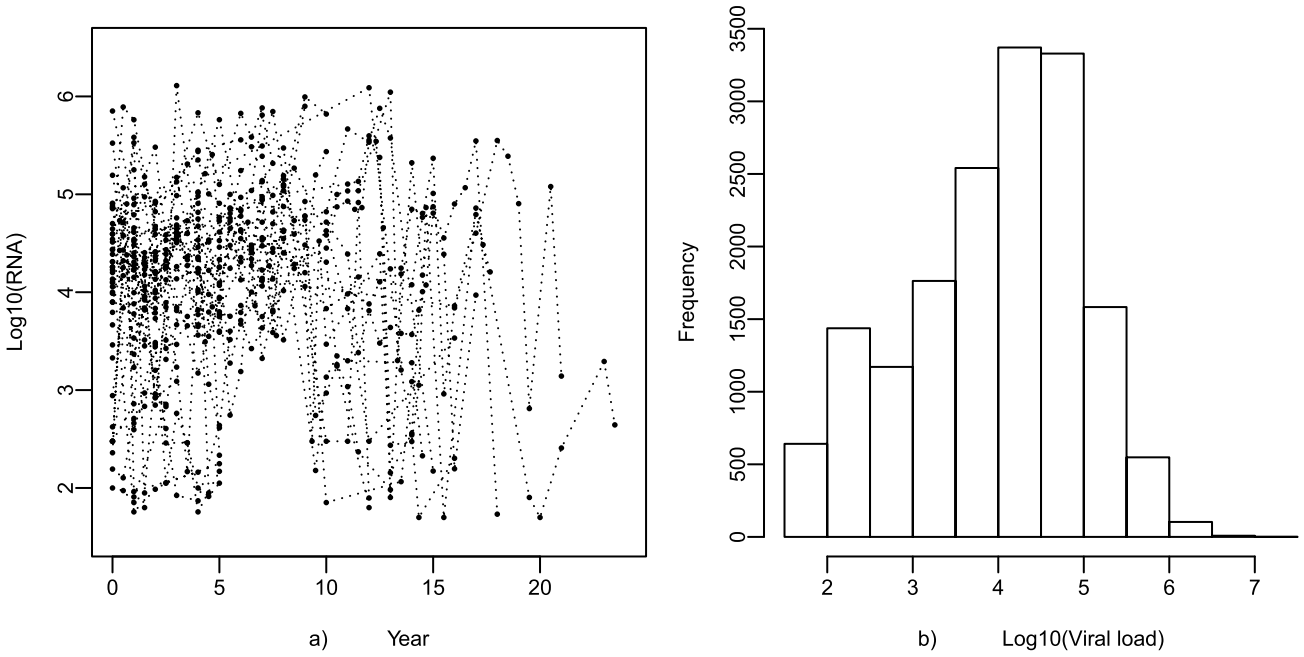

Figure 1. a): viral load trajectories for 50 randomly selected subjects. b): The skewed distribution of $\log_{10}$ viral load of 6,972 patients from MACS study.

and survival data and applied Monte Carlo EM approach to estimate unknown parameters. Also, [16] discussed different approaches to estimate unknown parameters of joint modeling of longitudinal measurements and event time data and then applied a fully parametric approach to modeling Schizophrenic patients data set. [17] used a linear mixed model with skew-normal distribution for longitudinal measurements and a Cox proportional hazard model for time to event variable.

Nevertheless, in all of these studies, at least one of the following two issues stand out: (i) majority of currently existing models assume a linear/genaralized linear mixed effects (LME/GLME) sub-model for the longitudinal process and a proportional hazard sub-model for survival process. The parametric formation of joint models is not sufficient in many cases, i.e. modeling the relationship between viral load and CD4 counts which is modified by irregular time effect. It is thus necessary to extend to more general model frameworks, such as semiparametric models. (ii) symmetric distributions such as normal and t distributions are usually assumed for random errors in these models. The symmetric assumption may lack the robustness against departures from symmetry as well as outliers [18]. Statistical analysis with normal/t assumptions may lead to misleading results. Particularly, skewness often appears in practice. Figure 1 b) shows the histogram of $\log_{10}$ transformed viral load data from an AIDS study. Obviously, the distribution of data is highly skewed even after transformation.

In this paper, we develop a jointly modeling approach to investigate the relationship between HIV viral load and CD4 counts in presence of skewed data. Particularly, we employ a Bayesian approach to examine jointly a partially linear mixed-effects model with skew distribution for HIV viral load and a cause-specific proportional hazard model [19] for competing risks survival data. From methodological perspective, we relax the parametric assumptions for the analysis of longitudinal-competing risks data under such a general framework that handles skew distribution in longitudinal endpoints. The two components of the joint models are associated through sharing of random-effects. [20] developed a new model for analyzing longitudinal-competing risks survival data in which a *mixed-effects varying-coefficient model* with skewed distribution was employed for the longitudinal process and a cause-specific *varying-coefficient hazard model with random-effects* was adopted for the survival process. The linkage between longitudinal process and competing risk time-to-event process is established by the *association* of random-effects between two models. In this paper, we proposed a *partially linear mixed-effects model with skew distribution* for the longitudinal process and a cause-specific proportional hazard model for the competing risks process and the connection between two processes is through the *sharing* of random effects which is computationally more convenient. To the best of our knowledge, the incorporation of partially linear mixed-effects model in joint modeling of longitudinal-competing risk data has never been investigated.

We employ a multivariate skew-t (ST) distribution [21, 22, 23, 24, 25, 26] to develop joint models for longitudinal-competing risk survival data with skew distribution. We de-

velop associated statistical methodologies to compare performance with those based on symmetric distributions. It is noted that ST distribution is approximate to skew-normal (SN) distribution when its degrees of freedom approach infinity and the SN distribution reduces to normal distribution if the skewness parameter is zero. Therefore, the joint models built on SN and normal distributions are just special cases of that built on the ST distribution. For completeness, Appendix A (Web supplementary material, http://intlpress.com/site/pub/pages/journals/items/sii/content/vols/0010/0003/s002) briefly discusses the multivariate skew distributions.

The rest of article is organized as follows. In Section 2, we describe the data set that motivated this research and the joint statistical models that account for asymmetric distribution in response are introduced. The Bayesian approach that estimates the parameters in the joint model is presented in Section 3. In Section 4, the proposed joint models and inferential method are applied to AIDS data and analysis results are presented. Simulation studies are conducted to assess the performance of the proposed model in Section 5. The article is concluded with discussion in Section 6.

2. MOTIVATING DATA AND JOINT MODELS

2.1 Motivating data

The data that motivates this study is from a Multicenter AIDS Cohort Study (MACS) [27], an ongoing prospective study of the natural and treated histories of HIV-1 infection in homosexual and bisexual men conducted by multiple sites (Baltimore, Chicago, Pittsburgh, and Los Angeles). A total of 6,972 men at risk of HIV infection have been enrolled in three cohorts (first cohort $n = 4{,}954$, from 1984 to 1985; second cohort $n = 668$, from 1987 to 1991; third cohort $n = 1{,}350$, from 2001 to 2003). The study participants had baseline and semiannual follow-up visits. There are 1598 subjects with more than three measurements on HIV viral load and CD4 counts. Among all patients, 7.3% died from AIDS related diseases (referred to as AIDS death) and 23% died from other causes (referred to as other death). A $\log_{10}$ transformation of viral load and standardized CD4 counts were used in the analysis for the purpose of stabilizing variation of measurement errors.

Figure 1 a) shows that the trajectories of individual viral load are quite irregular. Thus, the relationship between viral load and study time can not be modeled by a parametrical function. We adopt a semiparametrical model in which a linear relationship between viral load and CD4 counts is assumed and a nonparametric function is employed to model the modification effect of time. Further, Figure 1 b) shows that viral load is highly skewed even after transformation. Therefore, a normality assumption is not quite realistic for this dataset. Alternatively, we adopt an asymmetric skew-t (ST) distribution [21, 22, 23, 24, 25, 26] rather than a symmetric normal distribution.

2.2 ST partially linear mixed-effects model

To model the relationship between viral load and CD4 counts accounting for irregular time effects, we adopt a partially linear mixed-effects model for the longitudinal process:

$$\begin{aligned} &\boldsymbol{y}_i = \beta_i \boldsymbol{z}_i + \boldsymbol{g}_i(t_i) + \boldsymbol{\varepsilon}_i, \\ &\beta_i = \beta + \varpi_i, \\ &\boldsymbol{g}_i(t_i) = \boldsymbol{g}(t_i) + \boldsymbol{\Delta}_i(t_i) \end{aligned} \tag{1}$$

where $\boldsymbol{y}_i = (y_{i1}, \cdots, y_{in_i})^T$ and $\boldsymbol{z}_i = (z_{i1}, \cdots, z_{in_i})^T$. y_{ij} and z_{ij} are the viral load and CD4 counts for the ith subject at time t_{ij}, respectively; $\boldsymbol{t}_i = (t_{i1}, \cdots, t_{in_i})^T$; $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \cdots, \varepsilon_{in_i})^T$. β_i is the individual coefficient that quantifies the relationship between viral load and CD4 counts for individual i; β is the population coefficient (fixed-effects) and ϖ_i is the random-effects variable that quantifies the departure from the population for individual i. We assume ϖ_i is normally distributed with mean 0 and variance σ^2_ϖ. Both $\boldsymbol{g}(.)$ and $\boldsymbol{\Delta}_i(.)$ are unknown smoothing functions. $\boldsymbol{g}(.)$ stands for the population smoothing curve while $\boldsymbol{\Delta}_i(.)$ represents the random-effects. Altogether, $\boldsymbol{g}_i(.)$ is the smoothing curve for individual i. The error term $\boldsymbol{\varepsilon}_i$ and random smoothing function $\boldsymbol{g}_i(.)$ are zero mean stochastic processes and are independent from each other. ϖ_i is independent of both $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{g}_i(.)$.

We employ a regression spline method to approximate $\boldsymbol{g}(.)$ and $\boldsymbol{\Delta}_i(.)$ with a linear combination of spline basis functions $\boldsymbol{\Theta}_p(t) = (\theta_0(t), \cdots, \theta_p(t))^T$ and $\boldsymbol{\Phi}_r(t) = (\phi_0(t), \cdots, \phi_r(t))^T$ respectively:

$$\begin{aligned} &\boldsymbol{g}_p(t) \approx \textstyle\sum_{k=1}^{p} \xi_k \theta_k(t) = \boldsymbol{\Theta}_p(t)^T \boldsymbol{\xi}_p \\ &\boldsymbol{\Delta}_{i,r}(t) \approx \textstyle\sum_{k=1}^{r} \chi_{ki} \phi_k(t) = \boldsymbol{\Phi}_r(t)^T \boldsymbol{\chi}_{ri} \end{aligned} \tag{2}$$

where $\boldsymbol{\xi}_p = (\xi_1, \cdots, \xi_p)^T$ is a p-dimensional vector of fixed-effects, $\boldsymbol{\chi}_{ri} = (\chi_{1i}, \cdots, \chi_{ri})^T$ is a r-dimensional vector of random-effects; the dimensional numbers p and r are determined by Deviance Information Criterion (DIC). Note that $\boldsymbol{g}_p(.)$ is an approximation of true population smoothing curve $\boldsymbol{g}(.)$ based on a series of spline basis of order p. Based on the assumption of $\boldsymbol{\Delta}_i(.)$, we treat $\boldsymbol{\chi}_{ri}$ as i.i.d. realizations of a zero-mean random vector. Denote $\boldsymbol{\Theta}_{pi} = (\boldsymbol{\Theta}_p(t_{i1}), \ldots, \boldsymbol{\Theta}_p(t_{in_i}))^T$ and $\boldsymbol{\Phi}_{ri} = (\boldsymbol{\Phi}_r(t_{i1}), \ldots, \boldsymbol{\Phi}_r(t_{in_i}))^T$. Plug (2) into (1):

$$\boldsymbol{y}_i = \beta \boldsymbol{z}_i + \varpi_i \boldsymbol{z}_i + \boldsymbol{\Theta}_{pi} \boldsymbol{\xi}_p + \boldsymbol{\Phi}_{ri} \boldsymbol{\chi}_{ri} + \boldsymbol{\varepsilon}_i \tag{3}$$

Let $\boldsymbol{X}_i = (\boldsymbol{z}_i, \boldsymbol{\Theta}_{pi})$, $\boldsymbol{Z}_i = (\boldsymbol{z}_i, \boldsymbol{\Phi}_{ri})$, $\boldsymbol{\zeta} = (\beta, \boldsymbol{\xi}_p^T)^T$ and $\boldsymbol{b}_i = (\varpi_i, \boldsymbol{\chi}_{ri}^T)^T$. We can write (3) as

$$\boldsymbol{y}_i = \boldsymbol{X}_i \boldsymbol{\zeta} + \boldsymbol{Z}_i \boldsymbol{b}_i + \boldsymbol{\varepsilon}_i \tag{4}$$

which is a standard LME model if we assume $\boldsymbol{X}_i$ and $\boldsymbol{Z}_i$ are the fixed-effects and random-effects design matrix, respectively; $\boldsymbol{\zeta}$ and $\boldsymbol{b}_i$ are the fixed-effects and random-effects parameter vectors, respectively; $\boldsymbol{b}_i \sim N(0, \boldsymbol{\Sigma})$ and $\boldsymbol{\varepsilon}_i \sim N(0, \boldsymbol{R}_i)$. However, in practice, the viral load $\boldsymbol{y}_i$ are

most likely not normally distributed. In this case, we may assume $\boldsymbol{\varepsilon}_i \overset{\text{iid}}{\sim} ST_{n_i,\nu}\left(-J(\nu)\delta_e \mathbf{1}_{n_i}, \sigma^2 \boldsymbol{I}_{n_i}, \delta_e \boldsymbol{I}_{n_i}\right)$, which follows a multivariate ST distribution with degrees of freedom ν, unknown scale parameter σ^2 and skewness parameter δ_e, where $J(\nu) = (\nu/\pi)^{1/2}[\Gamma((\nu-1)/2)/\Gamma(\nu/2)]$, and $\mathbf{1}_{n_i} = (1,\ldots,1)^T$.

2.3 Cause-specific proportional hazard model

In the follow-up of a study, a subject may experience multiple types of failure or could be right censored. We assume there are K types of failure and use indicator d_i to represent each type. For example, $d_i = 0$ indicates an independent censoring event and $d_i = k$ indicates the kth type of failure, where $k = 1,\ldots,K$. Denote D_i as the failure time, C_i as the censoring time and $T_i = \min(D_i, C_i)$ as the observed time for subject i. To model the competing risks survival process, we consider a cause-specific hazard model.

$$\lambda_{ik}(t) = \lambda_{0k}(t)\exp(\psi_k z_i + \boldsymbol{\chi}_{ri}^T \boldsymbol{c}_k), \quad k = 1,\ldots,K \tag{5}$$

where the function $\lambda_{ik}(t)$ is the instantaneous hazard rate from cause k at time t for subject i, the coefficient ψ_k represents the effect of the covariate z_i to the kth type of failure, $\boldsymbol{c}_k = (c_{k1},\cdots,c_{kr})^T$ is the parameter vector that links the longitudinal and competing risks survival processes. In particular, $\boldsymbol{c}_k$ measures the strength of association between viral load and each type of failure. Note that $\boldsymbol{\chi}_{ri}$ is the random effects coefficients for the spline basis in (2) and $\chi_{1i},\cdots,\chi_{ri}$ collectively describe the nonlinear time effect on the change of viral load in subject i relative to the population. This makes the interpretation of the parameters $\boldsymbol{c}_k$ less straightforward as $\boldsymbol{c}_k$ is the coefficients associated with $\boldsymbol{\chi}_{ri}$. Intuitively, the interpretation for any c_{ki} $(i = 1,\cdots,r)$ is that for each unit increase of subject-specific χ_{ki}, the hazard ratio for the kth event is $\exp(c_{ki})$. The baseline hazard $\lambda_{0k}(t)$ is approximated by a piecewise step function, i.e. $\lambda_{0k}(t) = \lambda_{0k,j}$ for t in the interval between $t_{k,j-1}$ and $t_{k,j}$ where $0 < t_{k,1} < \cdots < t_{k,J} < \infty$ is the partition of time. For future reference, we denote $\lambda_{0;k} = (\lambda_{0k,1},\cdots,\lambda_{0k,J})$. To select the appropriate number of knots, we suggest to first fit the data with a large number of pieces, and then combine the adjacent intervals with similar baseline estimates. For the real data application, we used a piecewise constant baseline hazard function with five knots for the event of HIV related death and other death, respectively, and the time points defining the steps are taken to be 5 equally split percentiles of the observed dropout times for the two event types, respectively.

The corresponding survival function for the kth competing terminal event is $S_k(t_i) = \exp\big[-\int_0^{t_i}\lambda_{0k}(t)\exp(\psi_k z_i + \boldsymbol{\chi}_{ri}^T\boldsymbol{c}_k)dt\big]$. The likelihood of the competing terminal events is

$$L(\boldsymbol{\Upsilon}) = \prod_{k=1}^{K}\Big\{[\lambda_{0k}(t_i)]^{I(d_i=k)}S_k(t_i)\Big\}, \tag{6}$$

where the parameter vector $\boldsymbol{\Upsilon} = (\psi_1,\cdots,\psi_K,\boldsymbol{c}_1^T,\cdots,\boldsymbol{c}_K^T)^T$.

3. BAYESIAN INFERENCES FOR THE JOINT MODEL

The previously discussed longitudinal process and time-to-event process are inherently connected. The computational cost associated with the joint likelihood for the longitudinal data with skewness as well as the competing risks survival data is usually high. Often times, the Frequentist approach fails to converge [28]. We propose a fully Bayesian approach for the joint models (1) and (5) which handle longitudinal-competing risk survival data with skew distribution. The Markov chain Monte Carlo (MCMC) technique is employed to estimate all parameters simultaneously based on the joint likelihood.

Denote $\boldsymbol{\theta} = \{\boldsymbol{\zeta}, \boldsymbol{\Upsilon}, \lambda_{0;k}, \sigma^2, \boldsymbol{\Sigma}, \nu, \delta_e\}$ as the collection of unknown population parameters in models (4) and (5). The prior distributions for $\boldsymbol{\theta}$ are as follows:

$$\begin{aligned}&\boldsymbol{\zeta} \sim N(\boldsymbol{\tau}_1, \boldsymbol{\Lambda}_1),\ 1/\sigma^2 \sim \Gamma(\omega_1,\omega_2),\ \boldsymbol{\Sigma} \sim IW(\boldsymbol{\Omega},\rho),\\ &\delta_e \sim N(0,\gamma), \nu \sim Exp(\nu_0)I(\nu > 3),\ \boldsymbol{\Upsilon} \sim N(\boldsymbol{\tau}_2, \boldsymbol{\Lambda}_2),\\ &\lambda_{0;k} \sim \Gamma(\gamma_{k1},\gamma_{k2})\end{aligned} \tag{7}$$

where the mutually independent Normal (N), Inverse Wishart (IW), Gamma (Γ) and Exponential (Exp) prior distributions are chosen to facilitate computations. The hyperparameter matrices $\boldsymbol{\Lambda}_1$, $\boldsymbol{\Lambda}_2$ and $\boldsymbol{\Omega}$ are assumed to be diagonal for implementation convenience. The exponential prior for ν is truncated to lie above 3 to make variance of ST distribution well-defined.

Through introduction of a $n_i \times 1$ random vectors $\boldsymbol{w}_{e_i}$ based on the stochastic representation for ST distributions (see Appendix A in detail), we formulate the longitudinal response model (4) in association with cause-specific competing risk hazard model (5) hierarchically:

$$\begin{aligned}&\boldsymbol{y}_i|\boldsymbol{z}_i,\boldsymbol{b}_i,\boldsymbol{w}_{e_i};\boldsymbol{\zeta},\sigma^2,\delta_e,\nu \sim t_{n_i,\nu+n_i}(\boldsymbol{X}_i\boldsymbol{\zeta} + \boldsymbol{Z}_i\boldsymbol{b}_i +\\ &\delta_e[\boldsymbol{w}_{e_i} - J(\nu+n_i)\mathbf{1}_{n_i}], \omega_i\sigma^2\boldsymbol{I}_{n_i}), \boldsymbol{w}_{e_i} \sim t_{n_i,\nu}(\mathbf{0},\boldsymbol{I}_{n_i})\\ &I(\boldsymbol{w}_{e_i} > \mathbf{0}),\ \ \boldsymbol{b}_i \sim N(0,\boldsymbol{\Sigma}),\ \ t_i \sim F_c(t_i|\boldsymbol{\Upsilon},\lambda_{0;k}).\end{aligned} \tag{8}$$

where $\omega_i = (\nu + \boldsymbol{w}_{e_i}^T\boldsymbol{w}_{e_i})/(\nu+n_i)$, $t_{n_i,\nu}(\boldsymbol{\mu},\boldsymbol{A})$ denote the n_i-variate t distribution with parameters $\boldsymbol{\mu}$, $\boldsymbol{A}$ and degrees of freedom ν, $I(\boldsymbol{w} > 0)$ is an indicator function and $\boldsymbol{w} = |\boldsymbol{\varsigma}|$ with $\boldsymbol{\varsigma} \sim t_{n_i,\nu}(0,\boldsymbol{I}_{n_i})$. $F_c(\cdot)$ is the distribution function related to the competing risks survival model (5).

In the Bayesian framework, we specify the models for the observed data and the prior distributions for the unknown model parameters. Subsequently, based on the posterior distributions, we make statistical inference for the unknown parameters. Denote $f(\cdot)$, $f(\cdot|\cdot)$ and $\pi(\cdot)$ as a generic density function, a conditional density function and a prior density function, respectively. If we assume that $\boldsymbol{\zeta}, \boldsymbol{\Upsilon}, \lambda_{0;k}, \sigma^2, \boldsymbol{\Sigma}, \nu, \delta_e$ are independent of each other, then

$\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\zeta})\pi(\boldsymbol{\Upsilon})\pi(\lambda_{0;k})\pi(\sigma^2)\pi(\boldsymbol{\Sigma})\pi(\nu)\pi(\delta_e)$. As a result, the joint posterior density of $\boldsymbol{\theta}$ is

(9)
$$f(\boldsymbol{\theta}|\mathcal{D}) \propto \{\prod_i^n \int f(\boldsymbol{y}_i|\boldsymbol{z}_i, \boldsymbol{b}_i, \boldsymbol{w}_{e_i}; \boldsymbol{\zeta}, \sigma^2, \delta_e, \nu) f(\boldsymbol{w}_{e_i}|\boldsymbol{w}_{e_i} > \mathbf{0}) f(\boldsymbol{b}_i) F_c(t_i|\boldsymbol{\Upsilon}, \lambda_{0;k}) d\boldsymbol{b}_i\} \pi(\boldsymbol{\theta})$$

Generally, the integrals in (9) are of high dimension and do not have closed form. It is often not accurate to approximate the integrals numerically. Therefore, it is prohibitive to directly calculate the posterior distribution of $\boldsymbol{\theta}$ based on the observed data. Alternatively, the MCMC procedure is powerful for drawing samples from posterior distributions, based on (9), by employing Metropolis-Hastings (M-H) algorithm along with the the Gibbs sampler.

4. APPLICATION TO MACS DATA

4.1 Model specification

We have briefly described the dataset that motivated this research in Section (2.1). As discussed earlier, the relationship between viral load and CD4 counts is modified by time effects (Figure 1, a). In addition, the viral load in $\log_{10}$ scale is highly skewed (Figure 1, b). Therefore, It is critical to consider the skew distribution, such as ST, in the semiparametric mixed-effects model. From biological point of view, the longitudinal process and competing risk time-to-event process are inherently connected. Ignoring either one may lead to severe bias on parameter estimation. The proposed joint models take into consideration all of these factors. Thus, we believe it will outperform other potential models. Toward this end, we firstly compare three statistical models with different specification of random errors for the longitudinal response model.

- **Model N:** A joint model with the independent multivariate normal distribution of random errors for response model (4).
- **Model ST:** A joint model with the independent multivariate ST distribution of random errors for response model (4).
- **Model SN:** A joint model with the independent multivariate SN distribution of random errors for response model (4)

By comparing Models N through SN, we test if the skewed distribution assumed for error terms works better than that of symmetric distribution. Furthermore, we examine following scenarios: 1) to investigate how association between longitudinal and competing risk processes contributes to modeling results, we fit a model (**Model S**) by setting the coefficients $\boldsymbol{c}_k$ in the cause-specific hazard model (5) to 0 which indicates no association between two processes; 2) we examine a model (**Model R**) in which the competing risk mechanism is not taken into account. That is, we fit a joint model in which the cause-specific hazard model (5) is only applied to the first survival outcome while treating other types of risk as independent censoring. By comparing this reduced model with the complete model that adopts competing risk hazard model, we want to find out how other types of risk treated as independent censoring influences modeling results.

In terms of MACS data, we are interested in two cause-specific time-to-events: HIV death and other death. Thus, the competing risk hazard models for the two risks are:

(10)
$$\begin{aligned} \lambda_{i1}(t) &= \lambda_{01}(t)\exp(\psi_1 z_i + \boldsymbol{\chi}_{ri}^T \boldsymbol{c}_1), \\ \lambda_{i2}(t) &= \lambda_{02}(t)\exp(\psi_2 z_i + \boldsymbol{\chi}_{ri}^T \boldsymbol{c}_2) \end{aligned}$$

To approximate the nonlinear time effects in model (4), we use the cubic B-spline basis functions for $g(t)$ and $\Delta_i(t)$. The smoothing parameters p and r are determined by DIC and the location of knots is selected at the quantiles of the data [29]. For MACS dataset, the optimal smoothing parameters are set at 3 for p and r. To stabilize the variance and reach quick convergence, we use the $\log_{10}$ transformed viral load and standardized CD4 counts.

We take weakly informative prior distribution for the parameters in the joint models: (i) fixed-effects are taken to be independent normal distributions $N(0, 100)$ for each component of the population parameter vectors $\boldsymbol{\zeta}$ and $\boldsymbol{\Upsilon}$. (ii) for inverse of the scale parameters σ^2, we assume a limiting non-informative gamma prior distribution, $\Gamma(0.01, 0.01)$ so that the distribution has mean 1 and variance 100. (iii) the degree of freedom parameters ν follows truncated exponential distribution with $\nu_0 = 0.5$. (iv) for the skewness parameters δ_e, we choose independent normal distribution $N(0, 100)$. (v) inverse Wishart distributions $IW(\boldsymbol{\Omega}, \rho)$ is taken as priors for the variance-covariance matrice of the random-effects $\boldsymbol{\Sigma}$, with covariance matrices $\boldsymbol{\Omega} = diag(0.01, 0.01, 0.01, 0.01)$ and $\rho = 4$, respectively. (vi) for each of piecewise baseline hazard $\lambda_{0k,j}$, gamma distributions $\Gamma(0.1, 0.1)$ are employed.

The proposed method is implemented with WinBUGS software (Lunn et al., 2000). The MCMC scheme iterates between the Gibbs sampler and the Metropolis-Hastings algorithm for drawing samples from posterior distributions of parameters in the joint models. Based on MCMC samples, we are able to draw statistical inference for interested parameters. Specifically, we are interested in the posterior means and quantiles. Convergence of MCMC algorithm is assessed by standard tools in WinBUGS software such as trace plots and Gelman-Rubin diagnostics [30]. We run multiple chains. After convergency, we keep sampling for certain period and extract samples every few iterations. The collected samples are then used to make inferences for the posterior distribution of interested parameters.

When convergence is achieved, for each of three chains, after an initial number of 50,000 burn-in iterations, every 50th MCMC sample is retained from the next 50,000. Thus we obtain 3,000 samples of targeted posterior distributions of the unknown parameters for statistical inference.

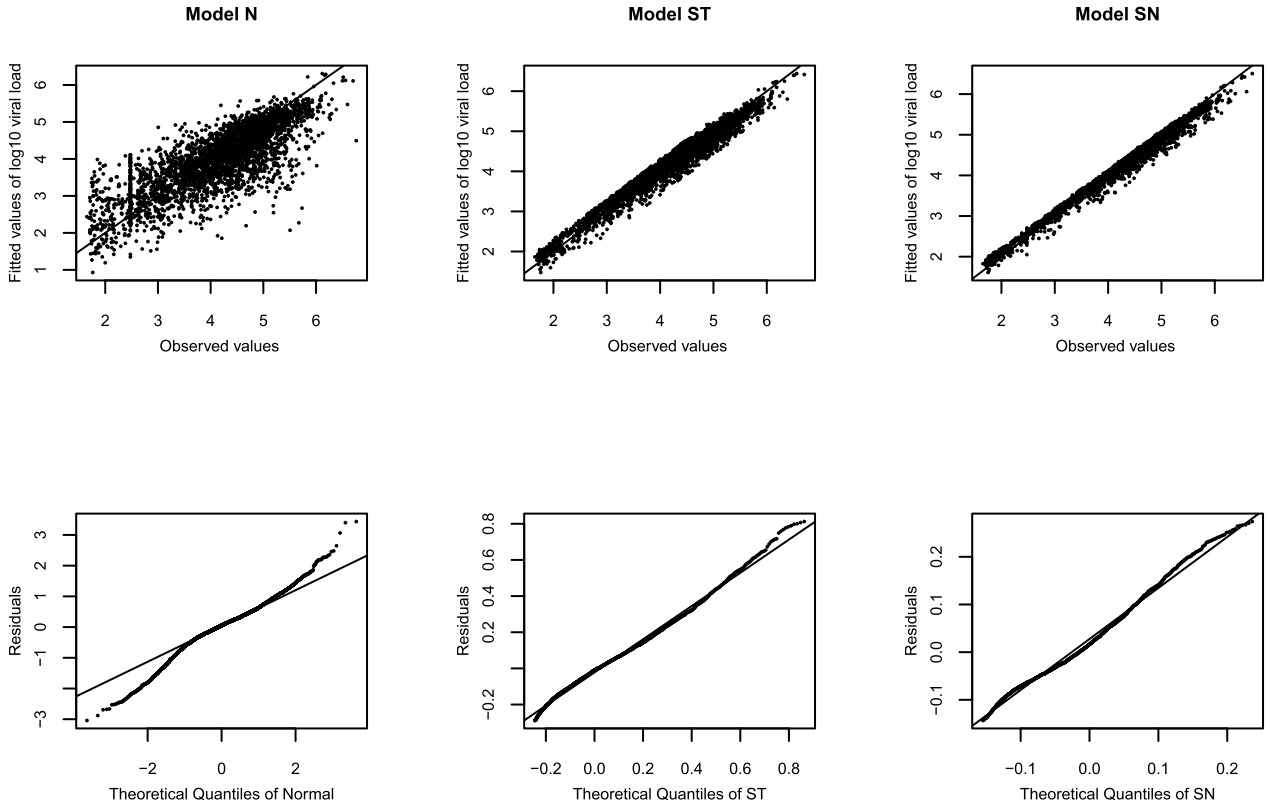

Figure 2. Diagnostics of model fitting for the three models. Top panel: fitted versus observed values of $\log_{10}$ viral load; bottom panel: Q-Q plots for residuals in three models.

In Winbugs, the Metropolis-Hastings algorithm is based on a symmetric normal proposal distribution, whose standard deviation is tuned over the first 4000 iterations in order to get an acceptance rate of between 20% and 40%. All summary statistics for the model will ignore information from this adapting phase.

4.2 Model comparison

Figure 2 shows the diagnostic plots to assess the goodness-of-fit of the three models. It can be seen that Models ST and SN which assume asymmetric distribution for random errors provide a better fit to observed data, compared with the normal model. The examination of Q-Q plots of residuals also confirms that there are fewer outliers from Models ST/SN than normal model. Thus, Models ST/SN fits the data better than Model N.

Table 1 presents the population posterior means, standard deviation and 95% credible interval for the parameters in three models. For parameter estimates of the response model, β has a significantly negative posterior mean in all three models, implying that there is negative relationship between HIV viral load and CD4 counts at the population level. This finding is consistent with biological mechanism since CD4 cells are one of the primary targets for HIV infection. In addition, we find that the absolute values of estimates for β based on Model N is larger than that from Model ST/SN, suggesting an overestimated association between the two biomarkers by the model assuming normal distribution for random errors. The estimates of two parameters that quantify the effect of CD4 cells on survival of two competing risks (ψ_1 for HIV-related death and ψ_2 for other death) are all negative for three models, indicating a reduced risk of experiencing two kinds of death due to more CD4 cells. It is noted that there is more beneficial effect of CD4 cells based on Models ST/SN as the parameter estimates of ψ_1 and ψ_2 are more negative than those in Model N. The estimates of many parameters that associate the longitudinal and survival processes $(c_{11}, c_{12}, c_{13}, c_{21}, c_{22}, c_{23})$ are significantly different from zero in all models, indicating a strong correlation between the two processes. Table 2 presents additional parameter estimates of the models. For the scale parameter σ^2, the posterior mean value (0.63) in Model N is much larger than that of any other corresponding posterior means in Models SN (0.11) and ST (0.14). This result is not surprising because Model N does not account for skewness in data. The estimates of skewness parameters (δ_e) are significantly negative in both Models ST and SN, which confirms the left-skewed viral load data. Based on above discussion, we can see that accounting for significant skewness, when the data exhibit skewness, provides a better model fit to the

Table 1. The estimated posterior mean (PM) for parameters in the joint model, and the corresponding standard deviation (SD) and lower limit (L_{CI}) and upper limit (U_{CI}) of 95% equal-tail credible interval (CI)

Model		β	ξ_1	ξ_2	ξ_3	ψ_1	ψ_2	c_{11}	c_{12}	c_{13}	c_{21}	c_{22}	c_{23}
N	PM	-0.51	4.19	-1.36	4.34	-1.52	-1.9	-1.81	-1.54	0.11	-2.48	1.95	-0.18
	L_{CI}	-0.67	4.05	-1.79	2.89	-2.41	-2.67	-2.43	-2.71	-1.47	-2.94	1.08	-0.43
	U_{CI}	-0.34	4.44	-0.86	5.81	-0.65	-1.13	-1.17	-0.38	1.69	-2.02	2.91	-0.01
	SD	0.07	0.07	0.24	0.75	0.46	0.39	0.33	0.6	0.79	0.25	0.49	0.11
ST	PM	-0.33	4.02	-0.79	3.23	-1.83	-2.39	-1.97	-1.25	0.13	-2.72	1.84	0.12
	L_{CI}	-0.56	3.88	-1.17	1.81	-2.79	-3.49	-2.55	-2.2	-0.77	-3.29	1.07	-0.04
	U_{CI}	-0.11	4.1	-0.37	4.68	-0.86	-1.56	-1.31	-0.72	1.02	-2.11	2.69	0.3
	SD	0.1	0.06	0.2	0.74	0.48	0.5	0.3	0.49	0.44	0.29	0.38	0.08
SN	PM	-0.31	3.99	-0.5	3.5	-1.97	-2.47	-1.23	-1.37	0.2	-2.33	1.35	0.2
	L_{CI}	-0.45	3.94	-1.03	1.81	-2.77	-3.43	-1.95	-2.52	-0.15	-2.94	0.44	0.02
	U_{CI}	-0.22	4.06	-0.02	5.73	-1.21	-1.72	-0.56	-0.19	0.58	-1.78	2.28	0.39
	SD	0.06	0.03	0.25	1.0	0.39	0.44	0.36	0.57	0.17	0.28	0.42	0.1
SN2	PM	-0.35	4.02	-0.63	3.62	-1.93	-2.58	-1.34	-0.94	0.26	-2.3	1.82	0.3
	L_{CI}	-0.51	3.89	-1.05	1.81	-2.8	-3.82	-1.75	-1.74	0.04	-2.79	0.99	0.06
	U_{CI}	-0.2	4.14	-0.11	5.43	-0.97	-1.39	-0.91	-0.12	0.48	-1.81	2.65	0.55
	SD	0.08	0.06	0.21	0.89	0.48	0.62	0.22	0.39	0.11	0.25	0.41	0.12

Table 2. The estimated posterior mean (PM) for parameters of scale, skewness, degree of freedom and the corresponding standard deviation (SD), lower limit (L_{CI}) and upper limit (U_{CI}) of 95% equal-tail credible interval (CI) as well as DIC and EPD values for MACS study

Model		σ^2	δ_e	ν	DIC	EPD
N	PM	0.63	–	–	6124	1.44
	U_{CI}	0.67	–	–		
	SD	0.02	–	–		
ST	PM	0.14	-0.69	3.11	4302	0.53
	U_{CI}	0.19	-0.48	3.18		
	SD	0.03	0.12	0.03		
SN	PM	0.11	-1.25	–	3727	0.36
	U_{CI}	0.15	-1.18	–		
	SD	0.02	0.03	–		
SN2	PM	0.12	-1.14	–	3711	0.35
	U_{CI}	0.15	-1.03	–		
	SD	0.02	0.07	–		

data and gives more accurate estimates to the parameters.

To further investigate whether Model SN provides better fit to the data than either Model ST or Model N, Deviance Information Criterion (DIC) by [31] are obtained and found to be 6124, 4302 and 3727 for Models N, ST and SN, respectively. Model SN has the smallest DIC, confirming that Model SN is superior to Models ST and N in fitting the data. Further, we adopt an alternative model comparison technique: expected predictive deviance (EPD) which is defined as $EPD = E\left\{\sum_{i,j}(y_{rep,ij} - y_{obs,ij})^2\right\}$ where the predictive value $y_{rep,ij}$ is a replicate of the observed $y_{obs,ij}$ and the expectation is taken over the posterior distribution of the model parameters $\boldsymbol{\theta}$ [32]. The best model comes with the least EPD. We find that EPDs are 1.44, 0.53 and 0.36 for Models N, ST and SN, respectively. Thus, based on these values, Model SN is relatively better than Models ST and N. These findings are consistent with those displayed in the goodness-of-fit in Figure 2, indicating that Model SN outperforms both Models ST and N. In summary, our results suggest that it is substantial to assume an SN distribution for the response model in order to achieve reliable results, in particular if the data exhibit skewness.

4.3 Results based on Model SN

We report the analysis results and interpretation based on Model SN. Using SN distribution for random errors, we investigate a few alternative models that adopt different structures for the joint models. Firstly, we test whether the competing risks sub-model has any additive effect on improving model fitting to the longitudinal response data. To do so, the coefficients $\boldsymbol{c}_k$ in the cause-specific hazard model (5) are set to 0 by which we separate the two processes. We find that the DIC for the separate model (Model S) increased to 3,920 from 3,727 in Model SN and EPD is also larger in Model S (0.47). Secondly, to investigate whether accounting for competing risks is superior to the univariate time-to-event data analysis in terms of model fitting, we substitute the cause-specific hazard model (5) by a univariate hazard model in

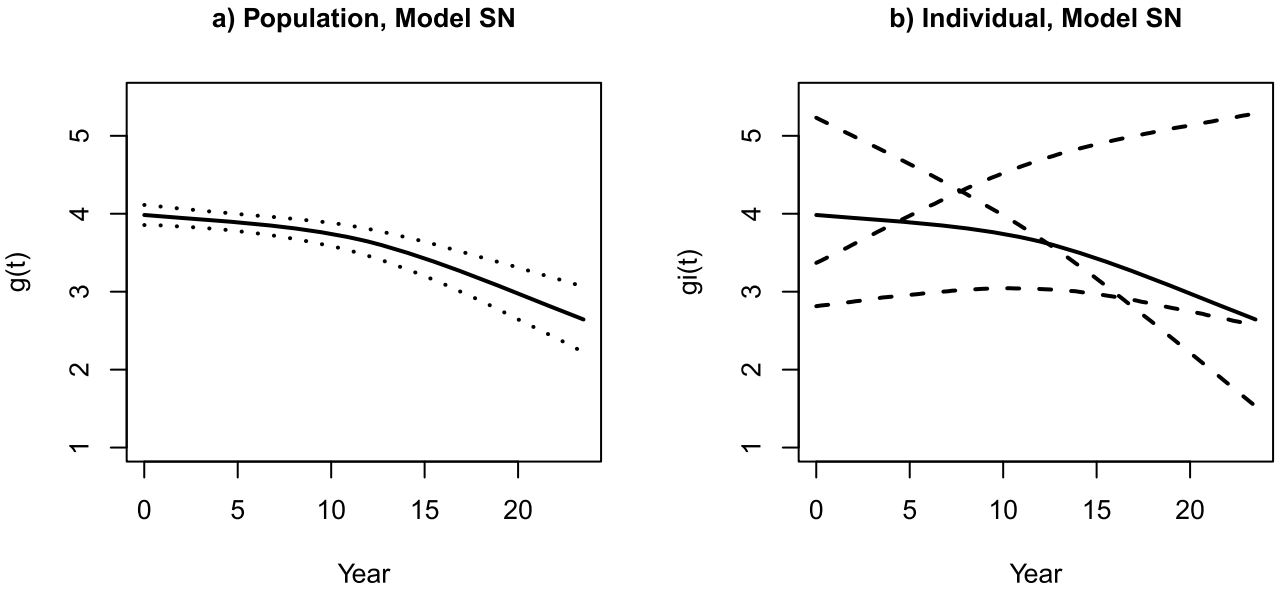

Figure 3. a) The population estimating curve of $g(t)$ for longitudinal model (1) based on the best selected Model SN. The estimates (solid line) along with the 95% pointwise credible interval (dotted line) are presented. b) Illustration of three represented individual estimating curve (dashed lines) of $g_i(t)$ based on Model SN. Population curves (solid lines) are shown for comparison.

which HIV-related death is treated as the only type of failure and other death is treated as independent censoring. We find that the reduce model (Model R) has a larger DIC (3,782) as well as EPD (0.41) than Model SN. As a result, Model SN that accounts for competing risks mechanism by adopting the cause-specific hazard model (5) is superior to alternative models that do not consider this fact.

We have seen the negative relationship (estimates of β) between HIV viral load and CD4 counts in Table 1. However, the relationship is modified by the time effect for which we assume a nonparametric function $\boldsymbol{g}(t)$ in model (1). The population estimating curve of $\boldsymbol{g}(t)$ against time along with 95% credible interval is presented in Figure 3 (a). We observe that the time effect decreases monotonically at a faster pace after year 10. At the individual level, we observe quite different features than the population. Figure 3 b) presents a few patterns for the estimating curves from three representative patients. The population estimating curves are also shown for comparison. In contrast to the population curve of $\boldsymbol{g}(t)$, the individual estimates of $\boldsymbol{g}_i(t)$ demonstrate large variation. For example, subject 140 has a steeper descending curve than the population curve. On the contrary, subject 298 shows an increasing trend instead of the decreasing curve observed at the population level. It seems that for this subject, HIV viral load will accumulate more as the treatment continues, holding the CD4 counts at the same level. There are also subjects, such as 63, who show relatively constant curves, implying minimal time effect on HIV viral load. In terms of the effect of CD4 counts on survival of each risk, we find it is more effective on reducing the risk of other death than that of HIV-related death based on comparison of the estimates for ψ_1 and ψ_2 in Model SN of Table 1.

5. SIMULATION

We carry out a simulation study to evaluate the performance of the proposed joint models. The design of simulated data is similar to the joint models used for the real data. Specifically, we adopt the mixed-effects model (4) to generate longitudinal outcomes. The measurement time t_{ij} are same to those in the real data. The CD4 counts z_{ij} are taken from real data for each individual. Similar to the real data, we simulate two competing risks according to the cause-specific proportional hazard model (10). To generate event time data, we use constant baseline hazard of 0.1 and 0.2 for two risks, respectively. An exponential distribution with mean equal to 0.1 is used to generate censoring time. We set the true values of model parameters as those obtained from real data analysis for Model SN which are listed in Tables 1 and 2. To save computing time, we simulate data for 300 subjects in the simulation study. One advantage of the proposed joint model is the capability of handling skewness exhibited in data. To simulate data with skewness, we generate the random errors for model (4) as follows: firstly, data are sampled from a Gamma distribution $\Gamma(3,1)$; secondly, the samples are subtracted by 3 which yields skewed data with mean 0 and variance 3. Under the settings described above, we simulate 200 data sets.

For each simulated data set, we adopted similar models and MCMC sampling schema that were used for the real

Table 3. Simulation results based on 200 simulated data set. The true values, bias (%), coverage rate (%) of 95% credible interval and standard deviation (SD) are presented for each parameter

		Model N			Model ST			Model SN		
Parameter	True value	Bias	Coverage	SD	Bias	Coverage	SD	Bias	Coverage	SD
β	-0.31	7.59	90.5	0.24	2.34	93.5	0.05	1.36	95.0	0.02
ξ_1	3.99	-5.99	89.5	1.37	-1.10	94.0	0.42	-0.88	94.5	0.41
ξ_2	-0.5	-5.10	92.0	0.36	-2.38	98.0	0.22	-1.93	98.0	0.16
ξ_3	3.5	-3.80	91.5	1.02	-2.03	96.5	0.64	-1.67	99.0	0.53
ψ_1	-1.97	-8.12	91.0	0.96	-3.09	94.5	0.56	-1.72	96.0	0.29
ψ_2	-2.47	-3.72	90.5	1.32	-0.46	99.0	0.36	-0.35	98.5	0.67
c_{11}	-1.23	-4.66	92.0	0.53	-1.59	99.0	0.28	-0.96	96.0	0.18
c_{12}	-1.37	12.78	93.0	0.61	8.54	94.5	0.19	5.81	95.5	0.13
c_{13}	0.2	14.62	96.0	0.92	6.51	93.5	0.09	4.08	96.5	0.03
c_{21}	-2.33	17.10	89.0	1.42	7.36	98.0	1.10	4.82	96.0	0.78
c_{22}	1.35	14.08	94.5	1.19	5.25	98.5	0.56	3.97	97.5	0.37
c_{23}	0.2	13.29	92.5	0.15	6.28	93.5	0.02	4.30	95.5	0.07

data analysis. Vague priors with large variances are used in the Bayesian inference. We fit each of Models ST, SN and N to each simulated data set and compare their performance. Table 3 presents the simulation results. As can be seen, both Models ST (bias ranges from -3.09% to 8.54%) and SN (bias ranges from -1.93% to 5.81%) are performing well with estimates close to true values. However, Model N produces larger bias (ranges from -8.12% to 17.1%). In summary, the simulation results show that it is important to account for skewness present in the data.

6. DISCUSSION

Motivated by the study of relationship between two biomarkers in an AIDS study, we developed a Bayesian approach for joint models of longitudinal-competing risks data with skewness. A partially linear mixed-effects model was employed to model the longitudinal process that characterizes the relationship between two biomarkers which is modified by irregular time effects. A cause-specific semiparametric proportional hazard model with random-effects was proposed to examine the association of time-to-event competing risks. The two components of joint models were connected through the sharing of random-effects between longitudinal and survival processes. We applied the proposed Bayesian modeling approach to simultaneously estimate all parameters in the joint models. The modeling results for the MACS data show that there is strongly negative relationship between virologic and immunologic biomarkers at the population level. The time effects on the longitudinal endpoint show a curvature at the population level while the large variation is seen at the individual level. When skewness is present in data, it is critical to adopt an appropriate statistical method for drawing robust conclusions. To the best of our knowledge, there are very few studies on longitudinal-competing risks data with skewness. Our results show that skew distributions are more suitable than symmetric distributions.

In Bayesian analysis, it is critical to perform sensitivity analysis to see if the posterior estimates change significantly when priors are different. Toward this end, we carried out sensitivity analysis by employing a few sets of different values for the hyper-parameters in (7) and re-run the MCMC sampling scheme. We observe that the conclusions are similar to those presented in the article. Thus, we are confident that the obtained results are robust against hyper-parameter values. In model (1), we adopted the regression spline basis to represent the unknown smoothing function. There are a lot of alternative ways for approximating the unknown function, such as local polynomial kernel and smoothing splines. It is interesting to compare the modeling results based on various nonparametric methods.

In the MACS study, data are collected from four centers. As one referee pointed out, there might exist center variation. Toward this end, we model the center heterogeneity via the addition of a frailty term, a center-level random effect with a multiplicative effect on the hazard function. We find that the parameter estimating results are essentially unchanged, regardless of small numerical fluctuation. Therefore, the center effect does not affect the conclusions.

As another referee pointed out, the CD4 counts usually contain substantial measurement error in practice. To account for measurement error in covariate, we consider an LME model for the CD4 covariate process:

$$(11)\quad z_{ij} = (\alpha_1 + a_{1i}) + (\alpha_2 + a_{2i})t_{ij} + (\alpha_3 + a_{3i})t_{ij}^2 + \epsilon_{ij}$$

where $z_{ij}^* = (\alpha_1 + a_{1i}) + (\alpha_2 + a_{2i})t_{ij} + (\alpha_3 + a_{3i})t_{ij}^2$ is the true CD4 counts, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)^T$ is the population (fixed-effects) parameter vector, $\boldsymbol{a} = (a_{1i}, a_{2i}, a_{3i})^T$ is an individual-specific random-effects parameter vector which is assumed to follow normal distribution centered at zero with $\boldsymbol{\Sigma}_a$ being the the unrestricted covariance matrix for random-effects. We then substitute the true CD4 counts into Models (4) and (10). The parameter estimation results are listed in Tables 1 and 2 under Model SN2. We find that the rela-

tionship between viral load and CD4 counts were slightly underestimated (β: -0.35 vs. -0.31) by Model SN in which CD4 measurement errors were not accounted for.

It might be meaningful to extend the normal distribution assumption to skewed-t (ST) or skewed normal (SN) distribution for the random-effects. In fact, [33] have studied NLME models where both model errors and random-effects were assumed to have skew normal distribution. It was found, in comparison of random-effects with normal and random-effects with SN, that the modeling results were very similar and not significantly different. Along with this finding, the random-effects are assumed to have a normal distribution in the joint models considered in this paper.

Alternative to the cause-specific model, we could adopt a mixture sub-model [34], in which the probability for each failure type is modeled with the logistic model.

Received 10 June 2015

REFERENCES

[1] Faucett C. L., Thomas D. C. Simultaneously modelling censored survival data and repeatedly measured covariates: a Gibbs sampling approach. Statistics in Medicine. 1996;15(15):1663–1685.

[2] Wulfsohn M. S., Tsiatis A. A. A joint model for survival and longitudinal data measured with error. Biometrics. 1997;53:330–339. MR1450186

[3] Henderson R., Diggle P., Dobson A. Joint modelling of longitudinal measurements and event time data. Biostatistics. 2000;1(4):465–480.

[4] Brown E. R., Ibrahim J. G. Bayesian approaches to joint cure-rate and longitudinal models with applications to cancer vaccine trials. Biometrics. 2003;59(3):686–693.

[5] Albert P. S., Shih J. H. On estimating the relationship between longitudinal measurements and time-to-event data using a simple two-stage procedure. Biometrics. 2010;66:983–987. MR2758235

[6] Rizopoulos D. Dynamic Predictions and Prospective Accuracy in Joint Models for Longitudinal and Time-to-Event Data. Biometrics. 2011;67(3):819–829.

[7] Elashoff R. M., Li G., Li N. A joint model for longitudinal measurements and survival data in the presence of multiple failure types. Biometrics. 2008 December;64:762–771. MR2526626

[8] Hu W., Li G., Li N. A Bayesian approach to joint analysis of longitudinal measurements and competing risks failure time data. Statistics in Medicine. 2009;28:1601–19.

[9] Yu B., Ghosh P. Joint modeling for cognitive trajectory and risk of dementia in the presence of death. Biometrics. 2010;66(1):294–300.

[10] Huang X., Li G., Elashoff R. M., Pan J. A general joint model for longitudinal measurements and competing risks survival data with heterogeneous random effects. Lifetime Data Analysis. 2011;17:80–100.

[11] Li N., Elashoff R. M., Li G., Tseng C. H. Joint analysis of bivariate longitudinal ordinal outcomes and competing risks survival times with nonparametric distributions for random effects. Statistics in Medicine. 2012;31:1707–1721.

[12] Huang C. Y., Wang M. C. Joint modeling and estimation for recurrent event processes and failure time data. Journal of the American Statistical Association. 2004;99(468):1153–1165.

[13] Han J., Slate E. H., Pea E. A. Parametric latent class joint model for a longitudinal biomarker and recurrent events. Statistics in Medicine. 2007;26:5285–5302. MR2415667

[14] Kim S., Zeng D., Chambless L., Li Y. Joint models of longitudinal data and recurrent events with informative terminal event. Statistics in Biosciences. 2012;4:262–281.

[15] Tseng Y. K., Hsieh F., Wang J. L. Joint modelling of accelerated failure time and longitudinal data. Biometrika. 2005;92(3):587–603. MR2202648

[16] Diggle P. J., Sousa I., Chetwynd A. G. Joint modelling of repeated measurements and time-to-event outcomes: the fourth Armitage lecture. Statistics in Medicine. 2008;27:2981–2998.

[17] Baghfalaki T., Ganjali M. A Bayesian approach for joint modeling of skew-normal lonigtudianl measurement and time to event data. Revstat. 2015;13(2):169–191.

[18] Verbeke G., Lesaffre E. A linear mixed-effects model with heterogeneity in random-effects population. Journal of the American Statistical Association. 1996;91:217–221.

[19] Prentice R., Kalbfleisch J., Peterson A., Flournoy N., Farewell V., Breslow N. The analysis of failure times in the presence of competing risks. Biometrics. 1978;34:541–554.

[20] Lu T., Wang M., Liu G., Dong G. H., Qian F. Mixed-effects varying-coefficient model with skewed distribution coupled with cause-specific varying-coefficient hazard model with random-effects for longitudinal-competing risks data analysis. Journal of Biopharmaceutical Statistics. 2016;26:519–533.

[21] Sahu S., Dey D., Branco M. A new class of multivariate skew distributions with applications to Bayesian regression models. The Canadian Journal of Statistics. 2003;31:129–150.

[22] Azzalini A., Capitanio A. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. Journal of Royal Statistical Society, Series B. 2003;65:367–389. MR1983753

[23] Arellano-Valle R., Genton M. On fundamental skew distributions. Journal of Multivariate Analysis. 2005;96:93–116.

[24] Azzalini A., Genton M. Robust likelihood methods based on the skew-t and related distributions. International Statistical Review. 2008;76:106–129.

[25] Jara A., Quintana F., Martin E. Linear mixed models with skew-elliptical distributions: a Bayesian approach. Computational Statistics and Data Analysis. 2008;52:5033–5045.

[26] Ho H., Lin T. Robust linear mixed models using the skew-t distribution with application to schizophrenia data. Biometrical Journal. 2010;52:449–469.

[27] Kaslow R., Ostrow D., Detels R., Phair J., Polk B., Rinaldo C. The Multicenter AIDS Cohort Study: rationale, organization, and selected characteristics of the participants. American Journal of Epidemiology. 1987;126:310–318.

[28] Wu L. A joint model for nonlinear mixed-effects models with censoring and covariates measured with error. Journal of the American Statistical Association. 2002;97:955–964.

[29] Eubank R. Nonparametric Regression and Spline Smoothing. New York: Dekker; 1999. MR1680784

[30] Gelman A., Rubin D. Inference from iterative simulation using multiple sequences. Statistical Science. 1992;7:457–511.

[31] Spiegelhalter D., Best N., Carlin B., Van der Linde A. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society, Series B. 2002;64:583–639.

[32] Gelman A., Carlin J., Stern H., Rubin D. Bayesian Data Analysis. London: Chapman and Hall; 2003.

[33] Huang Y., Dagne G. A Bayesian approach to joint mixed-effects models with a skew-normal distribution and measurement errors in covariates. Biometrics. 2011;67:260–269.

[34] Ng S., McLachlan G. An EM-based semi-parametric mixture model approach to the regression analysis of competing-risk data. Statistics in Medicine. 2003;22:1011–1097.

Tao Lu
Department of Mathematics and Statistics
University of Nevada
Reno, NV
USA
E-mail address: stat.lu11@gmail.com