Published in final edited form as:

Biom J. 2019 January ; 61(1): 187–202. doi:10.1002/bimj.201700326.

A Bayesian Joint Model of Recurrent Events and a Terminal Event

Zheng Li¹, Vernon M. Chinchilli¹, and Ming Wang^{*,1}

¹Department of Public Health Sciences, Penn State College of Medicine, Hershey, PA, U.S.A

Abstract

Recurrent events could be stopped by a terminal event, which commonly occurs in biomedical and clinical studies. In this situation, dependent censoring is encountered because of potential dependence between these two event processes, leading to invalid inference if analyzing recurrent events alone. The joint frailty model is one of the widely used approaches to jointly model these two processes by sharing the same frailty term. One important assumption is that recurrent and terminal event processes are conditionally independent given the subject-level frailty; however, this could be violated when the dependency may also depend on time-varying covariates across recurrences. Furthermore, marginal correlation between two event processes based on traditional frailty modeling has no closed form solution for estimation with vague interpretation. In order to fill these gaps, we propose a novel joint frailty-copula approach to model recurrent events and a terminal event with relaxed assumptions. Metropolis-Hastings within the Gibbs Sampler algorithm is used for parameter estimation. Extensive simulation studies are conducted to evaluate the efficiency, robustness and predictive performance of our proposal. The simulation results show that compared with the joint frailty model, the bias and mean squared error of the proposal is smaller when the conditional independence assumption is violated. Finally, we apply our method into a real example extracted from the MarketScan database to study the association between recurrent strokes and mortality.

Keywords

Bayesian inference; Joint frailty modeling; Recurrent events; Survival copula; Markov chain Monte Carlo

1 Introduction

Time-to-event outcomes are always of primary research interest in clinical trials and biomedical studies, and the Cox proportional hazards (PH) regression model is widely used to investigate the effects of potential risk factors on the time-to-event process. When the event of interest is recurrent, such as successive hospitalized myocardial infarction or heart attack, Cox PH models cannot be directly applied due to the failure to account for the

*Corresponding author: mwang@phs.psu.edu, Phone: +1-717-531-5745.

Conflict of Interest

The authors have declared no conflict of interest.

correlation among recurrent events, and as a result, the estimates of covariate effects might be biased. Another challenge is that the recurrent events might be stopped by a terminal event. Under this context, the censoring mechanism is informative because of highly likely dependence between the recurrent event process and the terminal event process, leading to dependent censoring. Thus, in order to have valid inference, we need to adjust for both the correlation among recurrent events and also the dependent censoring mechanism due to the competing risk of the terminal event.

There exists substantial work on recurrent event analysis in literature. For instance, Lawless (1987) proposed a shared frailty model for times to recurrent events, where a random effect called “frailty” was introduced to account for the within-subject correlation. Conditional on the frailty, the times to recurrent events from the same subject were assumed to be mutually independent. Yue and Chan (1997) proposed a dynamic frailty model, generalizing the joint frailty model by relaxing the joint frailty assumption. In particular, for each subject and each recurrent event, a time-dependent frailty was associated with the corresponding intensity function of the recurrent event. However, those works didn’t consider a terminal event as a semi-competing risk. Among others, relevant research topics on the most recent developments or comprehensive review can be referred to the literature (Lin et al., 1999, 2000; Kelly and Lim, 2000; Cook and Lawless, 2007).

In order to account for dependent censoring due to a terminal event, there are two major categories of joint analysis depending on the research interest, namely, a marginal model approach (Cook and Lawless, 1997; Zeng et al., 2009; Cai et al., 2010) and a frailty model approach (Huang and Wang, 2004; Huang and Liu, 2007; Rondeau et al., 2007). In this work, we focus on the latter which utilizes random effects to account for their correlation. Huang and Wang (2004) proposed jointly modeling the recurrent event process and the terminal event process by sharing a subject-level frailty in the hazard functions of recurrent gap times and the terminal event. Rondeau et al. (2007) further extended the model to analyze the recurrent events in terms of the calendar time. They also proposed to estimate the baseline intensity function by splines instead of the traditional Breslow estimator. A penalized likelihood approach is used to estimate the parameters in the model. Sinha et al. (2008) reviewed current methods from Bayesian perspective. Recently, Mazroui et al. (2012) proposed a general frailty model with extra frailty to adjust for the correlation among recurrent events. Rondeau et al. (2013) proposed a cure frailty model. Yu et al. (2014) proposed a model with time-varying coefficient in the proportional hazard functions. Che and Angus (2016) proposed to use an additive hazard function for the terminal event process. Other references related to this topic can be found in literature (Ghosh and Lin, 2002; Kalbfleisch et al., 2013). All the aforementioned literature assume constant frailty or the latent health status over time for each subject and conditional independence between the terminal event process and the recurrent event process given the frailty; however, their dependency may depend on time-varying covariates such as recurrence-specific ones which are not captured by this subject-level frailty. For example, the recurrence of stroke may depends on the body mass index (BMI). And BMI is a time-varying covariates across the study, the subject-level frailty will fail to capture the change of BMI. Also, we can find other examples in Yue and Chan (1997). Thus, the conditional independence assumption could be violated.

Another concern of the current joint frailty model is that although it can adjust for the dependency between multiple time-to-event processes, the correlation between two event processes is still unclear in terms of the estimation and interpretation. In particular, the majority of the existing works 1) treat the correlation estimate as a nuisance parameter, or 2) have vague interpretation of the dependence with only the association direction but without estimating their correlation in a straightforward manner. In some occasions, researchers may be interested in the correlation between different event outcomes, for example, the association of two types of AIDS events (Shih and Louis, 1995) or the first and second recurrence times to kidney infection after insertion of the catheter on kidney patients (McGilchrist and Aisbett, 1991). The copula plays an important and popular role in such research studies. Clayton (1978) first proposed to use the copula to analyze bivariate time-to-event processes. Zheng and Klein (1995) proposed to model competing risks by a copula approach. Wang and Wells (1997) considered non-parametric estimators for the bivariate survival functions in the copula modeling. Rivest and Wells (2001) incorporated a martingale approach under the copula framework. Joe (2005) proved the asymptotic efficiency of the two-stage estimation procedure for a copula. Cheng and Fine (2012) used a copula model for competing risk data from paired patients. Fu et al. (2013) designed a phase II trial and jointly modeled the progression-free survival (PFS) and overall survival (OS) by a copula, and then, conducted power analysis for a phase III trial based on the estimated model. Most recently, Emura et al. (2015, 2017) proposed a joint frailty-copula model in particular for meta-analysis; however, limited work using a copula approach exists for the joint analysis of recurrent events and a terminal event and substantial interest is drawn from both perspectives of clinical needs and advanced methods development. In addition, Peng et al. (2018) investigated clustered survival data by proposing a flexible semiparametric modeling framework with a copula model to incorporate dependency within clusters and association between two types of event times.

The main purpose of this paper is to develop an novel joint modeling strategy for recurrent gap times and a terminal event under some mild regulations in a full Bayesian framework. This strategy not only accounts for the dependence of recurrent events by a subject-level random frailty, but also for the correlation between recurrent events and a terminal event by applying a copula technique. The efficiency in parameter estimation and statistical inference is gained (i.e., smaller standard deviation estimates) due to potential informative priors from previous studies or literature, reduction of computational load, and easy implementation in statistical software. Importantly, the robustness of our proposal are comprehensively investigated through numerical studies, and also compared to the traditional joint frailty modeling for further evaluation. Also, we conduct dynamic prediction of survival risk based on the history of observed recurrent events for new subjects to further evaluate our proposal's predictive performance.

In the ensuing sections of the paper, we first provide the basic background of survival copula, and present the proposed Bayesian joint frailty-copula approach in section 2. In section 3, we perform extensive simulation studies to evaluate the efficiency and robustness of our proposal. Our simulation results show that the proposed method has lower bias and mean squared error compared with the joint frailty model when the conditional independence assumption is violated, and otherwise, comparable performance still holds. In

section 4, we apply our approach into a real data example to analyze the association between recurrent strokes and death. Finally, we discuss the advantages, limitations of our method and the topics for future study in section 5.

2 Methods

2.1 Copulas

Copulas are widely used in finance and medical research to model the joint distribution of two random variables (Nelsen, 1999). In terms of survival analysis, we are interested in the joint survival probability of two distinct events $\Pr(D \leq d, R \leq r)$ (Georges et al., 2001). Let $S_D(\cdot)$ and $S_R(\cdot)$ denote the marginal survival functions for D and R , respectively. A copula function $C(\cdot, \cdot)$ connecting survival functions $S_D(\cdot) \in [0, 1]$ and $S_R(\cdot) \in [0, 1]$, such that $\Pr(D \leq d, R \leq r) = C(S_D(d), S_R(r))$ is a survival copula (Nelsen, 1999). Next, we will briefly introduce the framework and basic property of survival copulas.

For $\forall u \in [0, 1]$ and $\forall v \in [0, 1]$, one important class of copulas called the Archimedean copulas is widely used and includes copulas of the form as

$$C(u, v; \theta) = \varphi^{-1}(\varphi(u; \theta) + \varphi(v; \theta)),$$

where $\varphi(\cdot)$ is a continuous, strictly decreasing convex function with the parameter θ . There are several well-known copulas belonging to this family, for example,

1. Clayton copula:

$$C(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}, \theta \in (0, \infty)$$

1. Frank copula:

$$C(u, v) = -\frac{1}{\theta} \log \left[1 + \frac{(\exp(-\theta u) - 1)(\exp(-\theta v) - 1)}{\exp(-\theta) - 1} \right], \theta \neq 0$$

where θ is a parameter that quantifies the association between U and V . The measures of the correlation such as Kendall's τ can be obtained from the copula function with

$\tau = 4 \int_0^1 \int_0^1 C(u, v) dC(u, v) - 1$. For instance, $\tau = \frac{\theta}{2 + \theta}$ for the Clayton copula.

2.2 Notation

Let D_i and C_i respectively denote the time to the terminal event and the censoring time for the i^{th} subject, $i = 1, \dots, n$. Define $\delta_i = I(D_i < C_i)$ to be the failure status for the terminal event, $Y_i = \min \{D_i, C_i\}$ to be the observed follow-up time, and ξ is the failure rate. Let R_{ij} represent the gap time between the $(j-1)^{th}$ event and the j^{th} event. $T_{ij} = \sum_{j'=1}^j R_{ij'}$ is the calendar time from the study begin to the j^{th} recurrent event. Suppose the i^{th} subject experiences a total of n_i recurrent events. When $j = n_i + 1$, $R_{i, n_i + 1} = Y_i - \sum_{j=1}^{n_i} R_{ij}$, which can be interpreted as the gap time between the n_i^{th} event and the end of the follow up time.

Let $C_{ij} = \max \{ Y_i - T_{i,j-1}, 0 \}$ denote the censoring time for R_{ij} . $\delta_{ij} = 1$ if $R_{ij} < C_{ij}$, otherwise, $\delta_{ij} = 0$.

For the original joint frailty model, the hazard functions for the terminal event time D_i and the recurrent event time R_{ij} of the i^{th} subject are expressed by

$$\begin{aligned}\lambda_{D(d)}(x_{i,T}, w_i) &= \lambda_{0D(d)} \exp \left\{ \beta_D^T x_{i,T} + w_i \right\} \\ \lambda_{R(r)}(x_{i,R}, w_i) &= \lambda_{0R(r)} \exp \left\{ \beta_R^T x_{i,R} + w_i \right\}\end{aligned}$$

where $x_{i,T}$ and $x_{i,R}$ are time-invariant or time-dependent covariate vectors. w_i is the subject-level frailty accounting for the correlation among within-subject recurrences and also captures the dependency between two time-to-event processes. For simplicity, we omit $x_{i,T}$ and $x_{i,R}$ in all the functions from now on, and hazard functions denoted respectively by, $\lambda_D(dw_i)$ and $\lambda_R(dw_i)$. The frailty term is assumed that $w_i \sim N(0, \sigma_w^2)$, but the other parametric distributions such as gamma can still be applicable. The parameters β_D and β_R quantify the effect of the covariates $x_{i,T}$ on D_i and the effects of $x_{i,R}$ on R_{ij} respectively. Denote the observation $\mathcal{Y}_i = \{y_i, \delta_i, \delta_{ij}, r_{ij}, x_{i,T}, x_{i,R}, j = 1, \dots, n_i + 1\}$ for the i^{th} subject, and let $\mathcal{D}_n$ denote the observed data from n subjects, $\mathcal{D}_n = \{\mathcal{Y}_1, \dots, \mathcal{Y}_n\}$.

2.3 Joint Frailty-Copula Model

The regular joint frailty model assumes that $D_i, R_{i1}, \dots, R_{i,n_i+1}$ are mutually independent conditional on the subject-level frailty w_i (Huang and Liu, 2007). However, this conditional independence assumption might be violated when some important time dependent covariates are not included in the hazard functions. Emura et al. (2015) discussed residual dependence in the absence of time-dependent covariates. Similarly, we relax the assumption of conditional independence of w_i by inducing a copula to account for the within-subject correlation. But, we further relax the assumption of the memoryless property of the terminal event time D_i by Emura et al. (2015), i.e., $S_D(d|t_{ij-1}) = S_D(d - t_{ij-1})$ with the observed $(j-1)^{th}$ calendar event time $t_{ij-1} = \sum_{j'=1}^j r_{ij'-1}$, indicating that after the $(j-1)^{th}$ recurrent event, whether the subject will survive the j^{th} recurrent event only depends on the gap time between the $(j-1)^{th}$ and j^{th} recurrent events.

Following the context of Huang and Liu (2007), we first assume a sufficient large constant integer J for every subject independent of data. J can be interpreted as the maximum of recurrent events a subject will experience. This leads to a maximum of $J+1$ gap times for every subject. However, due to the terminal event or subject drop-out, we cannot observe all recurrent events. Assume that the subject experiences a total of n_i recurrent events during the follow-up time. Then, for subject i , we can observe $R_{ij} = r_{ij}$ before the terminal event happens. Assume that the residual dependence is homogeneous across all subjects and unchanged along the time line.

Conditional on w_i , the joint probability of $D_i = t_i$ and $R_{ij} = r_{ij}$ is

$$\Pr(D_i \geq d_i, R_{ij} \geq r_{ij} | w_i) = C(S_D(d_i | w_i), S_R(r_{ij} | w_i)),$$

for any $i, i = 1, \dots, n$ and any $j, j = 1, \dots, J+1$.

Based on the survival copula, we have,

$$\begin{aligned}\Pr(D_i \geq d_i, R_{ij} \geq r_{ij} | w_i) &= C(S_D(d_i | w_i), S_R(r_{ij} | w_i)) \\ \Pr(D_i \geq d_i, R_{ij} = r_{ij} | w_i) &= C_{(01)}(S_D(d_i | w_i), S_R(r_{ij} | w_i))f_R(r_{ij} | w_i) \\ \Pr(D_i = d_i, R_{ij} = r_{ij} | w_i) &= C_{(11)}(S_D(d_i | w_i), S_R(r_{ij} | w_i))f_D(d_i | w_i)f_R(r_{ij} | w_i) \\ \Pr(D_i = d_i, R_{ij} \geq r_{ij} | w_i) &= C_{(10)}(S_D(d_i | w_i), S_R(r_{ij} | w_i))f_D(d_i | w_i)\end{aligned}$$

where $C_{(01)}(u, v) = C(u, v)/v$, and $C_{(10)}(u, v) = C(u, v)/u$, $C_{(11)}(u, v) = C(u, v)/(uv)$.

For example, if $C(\cdot, \cdot)$ is a Clayton copula, then we have

$$\begin{aligned}C_{(01)}(u, v) &= v^{-\theta-1}(u^{-\theta} + v^{-\theta} - 1)^{-\frac{\theta+1}{\theta}} \\ C_{(10)}(u, v) &= u^{-\theta-1}(u^{-\theta} + v^{-\theta} - 1)^{-\frac{\theta+1}{\theta}} \\ C_{(11)}(u, v) &= (\theta+1)(uv)^{-(\theta+1)}(u^{-\theta} + v^{-\theta} - 1)^{-\frac{2\theta+1}{\theta}}\end{aligned}$$

Note that for a subject with $\delta_i = 0$, C_{ij} is $\max(C_i - T_{i,j-1}, 0)$, and C_{ij} is independent of R_{ij} . Given the assumption that $R_{ij}, \dots, R_{i,J+1}$ are mutually independent, conditional on w_i and $D_i = d_i$, the probability of the i^{th} subject who survives through d_i and experiences n_i events given w_i is,

$$\begin{aligned}\Pr(D_i \geq d_i, R_{i1} = r_{i1}, \dots, R_{i,n_i} = r_{i,n_i}, R_{i,n_i+1} \geq c_{i,n_i+1} | w_i) \\ = \Pr(R_{i1} = r_{i1}, \dots, R_{i,n_i} = r_{i,n_i}, R_{i,n_i+1} \geq c_{i,n_i+1} | D_i \geq d_i, w_i) \Pr(D_i \geq d_i | w_i) \\ = \Pr(D_i \geq d_i | w_i) \prod_{j=1}^{n_i} \Pr(R_{ij} = r_{ij} | D_i \geq d_i, w_i) \prod_{j=n_i+1}^J \Pr(R_{ij} \geq c_{ij} | D_i \geq d_i, w_i)\end{aligned}$$

By Bayes' rule,

$$\begin{aligned}\Pr(R_{ij} = r_{ij} | D_i \geq d_i, w_i) &= \frac{C_{(01)}(S_D(d_i | w_i), S_R(r_{ij} | w_i))f_R(r_{ij} | w_i)}{S_D(d_i | w_i)} \\ \Pr(R_{i,n_i+1} \geq c_{i,n_i+1} | D_i \geq d_i, w_i) &= \frac{C(S_D(d_i | w_i), S_R(c_{i,n_i+1} | w_i))}{S_D(d_i | w_i)}\end{aligned}$$

For $j = n_i + 2, \dots, J+1$, C_{ij} is 0. Then, we have,

$$\Pr(R_{ij} \geq c_{ij} | D_i \geq d_i, w_i) = \Pr(R_{ij} \geq 0 | D_i \geq d_i, w_i) = 1.$$

Then,

$$\begin{aligned} & \Pr(D_i \geq d_i, R_{i1} = r_{i1}, \dots, R_{i, n_i+1} \geq c_{i, n_i+1} | w_i) \\ &= S_D(d_i | w_i) \frac{C(S_D(t_i | w_i), S_R(c_{i, n_i+1} | w_i))}{S_D(d_i | w_i)} \prod_{j=1}^{n_i} \frac{C_{(01)}(S_D(t_i | w_i), S_R(r_{ij} | w_i)) f_R(r_{ij} | w_i)}{S_D(d_i | w_i)} \\ &= C(S_D(t_i | w_i), S_R(c_{i, n_i+1} | w_i)) \prod_{j=1}^{n_i} \frac{C_{(01)}(S_D(t_i | w_i), S_R(r_{ij} | w_i)) f_R(r_{ij} | w_i)}{S_D(d_i | w_i)} \end{aligned}$$

The likelihood conditional on the frailty $\mathbf{w}$ can be expressed below, and the detailed derivation is provided in Appendix A of the Supplementary Materials.

$$\begin{aligned} L(\mathcal{D}_n | \mathbf{w}) &= \prod_{i=1}^n [f_D(y_i | w_i) C_{10}^*(S_D(y_i | w_i), S_R(r_{i, n_i+1} | w_i))]^{\delta_i} \\ &\quad \times \left[S_D^{-n_i}(y_i | w_i) C(S_D(y_i | w_i), S_R(r_{i, n_i+1} | w_i)) \right]^{1-\delta_i} \\ &\quad \times \prod_{j=1}^{n_i} \left[C_{(01)}(S_D(y_i | w_i), S_R(r_{ij} | w_i)) \right]^{1-\delta_i} \left[C_{(11)}(S_D(y_i | w_i), S_R(r_{ij} | w_i)) \right]^{\delta_i} f_R(r_{ij} | w_i) \end{aligned}$$

2.4 Metropolis-Hastings within the Gibbs Sampler Algorithm

We utilize the Bayesian approach for parameter estimation and inference. For simplicity, constant baseline intensity functions $\lambda_{0D}(d) = \exp(\beta_{0D})$ and $\lambda_{0R}(r) = \exp(\beta_{0R})$ are considered; however, this can be extended in a straightforward manner to the setting with non-constant baseline hazards. Let $\Theta = \{\beta_{0R}, \beta_{0D}, \beta_D, \beta_R, \sigma_w^{-2}, \mathbf{w}, \theta\}$ with the dimensionality of M . First, we present the derivation of the posterior distribution of Θ . A Metropolis-Hastings within Gibbs sampler algorithm is used to sample the posterior distribution of $\Pr(\Theta | \mathcal{D}_n)$ (Gilks et al., 1995). A general form of the full conditional distribution of $\Theta_{[m]}$ is

$$\Pr(\Theta_{[m]} | \mathcal{D}_n, \Theta_{[<m]}, \Theta_{[>m]}) \propto L(\Theta) \Pr(\Theta_{[m]}),$$

where $\Theta_{[<m]}$ is the first $(m-1)$ elements of Θ and $\Theta_{[>m]}$, the last $(M-m)$ elements of Θ .

Suppose we want to generate B samples of Θ from $\Pr(\Theta | \mathcal{D}_n)$. The algorithm to get the posterior samples is as follows:

- From $\ell = 1$,
 - From $m = 1$, sample $\Theta_{[m]}^{(\ell)}$ from

$$\Pr(\Theta_{[m]}^{(\ell)} | D_n, \Theta_{[< m]}^{(\ell)}, \Theta_{[> m]}^{(\ell-1)}) \propto L(\Theta_{[< m]}^{(\ell)}, \Theta_{[m]}^{(\ell)}, \Theta_{[> m]}^{(\ell-1)}) \Pr(\Theta_{[m]}^{(\ell)}),$$

– until $m = M$, set $\ell = \ell + 1$.

- until $\ell = B$, end.

In order to sample $\Theta_{[m]}^{(\ell)}$ from $\Pr(\Theta_{[m]}^{(\ell)} | \mathcal{D}_n, \Theta_{[< m]}^{(\ell)}, \Theta_{[> m]}^{(\ell-1)})$, the Metropolis-Hastings algorithm is applied and the procedures are shown below:

1. Generate U from $Unif(0, 1)$
2. Generate $\Theta_{[m]}^N$ by $\Theta_{[m]}^{(\ell-1)} + sN(0, 1)$, step size s
3. Calculate

$$LR = \frac{L(\mathcal{D}_n | \Theta_{[< m]}^{(\ell)}, \Theta_{[m]}^N, \Theta_{[> m]}^{(\ell-1)}) \Pr(\Theta_{[m]}^N)}{L(\mathcal{D}_n | \Theta_{[< m]}^{(\ell)}, \Theta_{[m]}^{(\ell-1)}, \Theta_{[> m]}^{(\ell-1)}) \Pr(\Theta_{[m]}^{(\ell)})}$$

1. If $LR > u$, $\Theta_{[m]}^{(\ell)} = \Theta_{[m]}^N$, else $\Theta_{[m]}^{(\ell)} = \Theta_{[m]}^{(\ell-1)}$

The step size s is chosen so that the acceptance rate is around 0.44. The details of the Metropolis-Hastings within Gibbs algorithm can be found in Appendix B of the Supplementary Materials. Because of efficiency consideration, all the algorithm is coded in C language and called by R.

2.5 Prediction for the Time to the Terminal Event of a New Subject

After the study completion, we can analyze the data based on the posterior distribution $\Theta | \mathcal{D}_n$ from the joint frailty-copula model. Emura et al. (2017) proposed a prediction framework based on the joint frailty-copula model and bivariate paired survival outcomes. By contrary, we are interested in predicting the time to the terminal event considering the history of recurrent events history. We present here subject-level prediction for a new subject N with observed recurrent events history,

$$\mathcal{H}_N(t') = \{r_{N,1}, \dots, r_{N,j}^{s.t.} \sum_{j'=1}^j r_{N,j'} < t' \text{ and } \sum_{j'=1}^{j+1} r_{N,j'} > t'\}.$$

Let w_N denote the frailty of the new subject N . Subject-level prediction for a new subject is $\Pr(D_N > t | D_N > t', \mathcal{H}_N(t'), \mathcal{D}_n, \Theta, w_N)$, denoted by $\pi(t | t', \mathcal{H}_N(t'), \Theta, w_N)$. We can estimate $\pi(t | t', \mathcal{H}_N(t'), \Theta, w_N)$ by plugging in the posterior estimates of $\hat{\Theta}$ and $\hat{w}_N$. The subject-level survival prediction for the new subject can be expressed as,

$$\hat{\pi}(t | t', \mathcal{H}_N(t'), \mathcal{D}_n, \hat{\Theta}, w_N) = \Pr(D_N > t | D_N > t', \mathcal{H}_N(t'), \hat{w}_N, \hat{\Theta}, \mathcal{D}_n),$$

where $\hat{\Theta}$ is the posterior mean of the posterior distribution $\Pr(\Theta | \mathcal{D}_n)$ and $\hat{w}_N$ is the posterior mean of distribution $\Pr(w_N | D_N > t', \mathcal{H}_N(t'), \mathcal{D}_n, \Theta, \mathcal{D}_n)$. We can sample

$\Pr(w_N | D_N > t', \mathcal{H}_N(t'), \Theta, \mathcal{D}_n) \propto \Pr(D_N > t', \mathcal{H}_N(t') | \mathcal{D}_n, \Theta, w_N) \Pr(\Theta | \mathcal{D}_n) \Pr(w_N)$ by

Metropolis Hastings within Gibbs Sampler algorithm. When t' increases,

$\hat{\pi}(t | t', \mathcal{H}_N(t'), \mathcal{D}_n, \hat{\Theta}, \hat{w}_N)$ can be updated dynamically by re-sampling w_N from

$\Pr(w_N | D_N > t', \mathcal{H}_N(t'), \hat{\Theta}, \mathcal{D}_n)$.

Suppose we need to predict the terminal event time for a total of N_t' new subjects. The Brier score (BS) is used to evaluate the bias between the predicted and true risks (Graf et al., 1999), which is defined as

$$E\left\{\left[D(t | t', \mathcal{H}_N(t'), \hat{\Theta}) - \hat{\pi}(t | t', \mathcal{H}_N(t'), \hat{\Theta})\right]^2\right\},$$

where $D(t | t', \mathcal{H}_N(t'), \hat{\Theta})$ is the observed terminal event status which equals 1 if the subject experiences the terminal event in the time interval (t', t) ; otherwise, it is 0. It can be estimated by

$$\widehat{BS}(t', t) = \frac{1}{N_t'} \sum_{i=1}^{N_t'} \hat{G}(t', t) \left\{ \left[D_i(t | t', \mathcal{H}_N(t'), \hat{\Theta}) - \hat{\pi}_i(t | t', \mathcal{H}_N(t'), \hat{\Theta}) \right]^2 \right\},$$

where $\hat{G}(t', t) = I(d_i > t) / \{\hat{S}_0(t) / \hat{S}_0(t')\} + I(t > d_i > t') \delta_i / \{\hat{S}_0(d_i) / \hat{S}_0(t')\}$, accounting for censoring with $\hat{S}_0$, which is estimated based on the Kaplan-Meier curve.

3 Simulation Study

3.1 Simulation Set-up

We evaluate our proposal in terms of efficiency, robustness and predictive accuracy under different scenarios. Here, we assume that the terminal event T_i follows an exponential distribution with the density function $f_D(d) = \lambda_{i,D} \exp(-\lambda_{i,D}d)$, and recurrent gap time R_{ij} follows an exponential distribution with the density function $f_R(r) = \lambda_{ij,R} \exp(-\lambda_{ij,R}r)$. The hazard functions $\lambda_{i,D}$ and $\lambda_{ij,R}$ are given respectively,

$$\begin{aligned} \lambda_{i,D} &= \exp(\beta_{0D} + w_i + \beta_{D,1}x_{1,i} + \beta_{D,2}x_{2,i}) \\ \lambda_{ij,R} &= \exp(\beta_{0R} + w_i + \beta_{R,1}x_{1,i} + \beta_{R,2}x_{2,i}), \end{aligned}$$

where $x_{1,i}$ is a continuous variable generated from $N(0, 1)$ and $x_{2,i}$ is a binary variable generated from a Bernoulli distribution, $Bern(0.5)$. $(\beta_{D,1}, \beta_{D,2}, \beta_{R,1}, \beta_{R,2})^T$ is set to be $(1, 1, 2, 2)^T$. $(\beta_{0D}, \beta_{0R})^T$ is set to be $(0.5, 1)^T$.

For each subject i , we first generate U from $Unif(0, 1)$. D_i is generated by $-\log(U_i)/\lambda_{i,D}$. Assuming independent censoring for D_i , we generate C_i respectively from a uniform distribution $Unif(0, 1)$, $Unif(0, 0.6)$, and $Unif(0, 0.4)$, corresponding to a high failure rate (ξ

= 60%), a medium failure rate ($\xi = 50\%$) and a low failure rate ($\xi = 40\%$). The observed follow-up time $Y_i = \min \{D_i, C_i\}$. Then, considering the j^{th} recurrent event, suppose $S_R(r_{ij}|w_i) = V_{ij}$. Assume that the joint distribution of U_i and V_{ij} is a Clayton copula, $C(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$, where θ is varied by 1 or 2 in different scenarios, respectively corresponding to a low correlation (i.e., $\tau = 0.3$) and a high correlation (i.e., $\tau = 0.5$) between the terminal event process and the recurrent event process. We also simulated a scenario under which the terminal event and the recurrent events are generated independently ($\tau = 0$), where the times of recurrent events and death are generated separately given the subject-level frailty but not based on a joint copula distribution. Suppose sufficiently large maximum number $J = 1000$ of events occur for each subject. When $D_i < C_i$, the conditional distribution of V_{ij} is $F_{V|U}(v | U_i = u_i) = u_i^{-(\theta+1)}(u_i^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$. Since $F_V(v | U_i = u_i)$ follows a uniform distribution $Unif(0, 1)$, we can first generate a $\tilde{W}_{ij} \sim Unif(0, 1)$ and $\tilde{W}_{ij} = F_V(v | U_i = u_i)$. V_{ij} can be generated by

$$V_{ij} = \left((\tilde{W}_{ij}^{-\frac{\theta}{\theta+1}} - 1)u_i^{-\theta} + 1 \right)^{-1/\theta}.$$

When $D_i > C_i$, V_{ij} is generated from $F_V(v | U_i < \exp(-\lambda_i C_i))$ based on the Monte Carlo method. After we have V_{ij} , we can generate the gap time R_{ij} by $-\log(V_{ij})/\lambda_{ij,R}$. We repeated this procedure until T_{ij} is greater than Y_i or $j = J$. For each scenario, we generate 250 Monte Carlo datasets with sample size $n = 100, 200, 300, 400$. The average number of the recurrent events across 250 replicates are 3.491 (SD=7.385) under low censoring rate (40%), 3.230 (SD=7.276) under medium censoring rate (50%), and 2.950 (SD=7.130) under high censoring rate (60%). According to the above methods, the recurrent events are generated for each subject with a minimum of zero events and a maximum of 28 events across all scenarios.

To evaluate the efficiency of our proposal, we consider the regular joint frailty model approach for comparison. Also, in order to evaluate the prediction performance of our proposal, we randomly generate a data set of 200 subjects as a training set to estimate Θ and an independent data set of 200 subjects as a testing set. The training set is generated under a Clayton copula model with $\theta = 2$. Other settings are the same as the setting when we compare the joint frailty model and the joint frailty-copula model. Thus, we predict $\pi(t | t', \mathcal{H}_N(t'))$, where t' is set as 0.03, 0.06 or 0.09 such that the percentage of subjects still at risk is respectively 80%, 70% or 60%, and t takes the values between t' and 0.1 with the increment of 0.01. The BS are thereafter estimated to show the predictive accuracy across various set-ups with the benchmark value calculated based on the Kaplan-Meier estimator.

In addition, we evaluate the robustness of our method considering the mis-specification of copula models. The set up is the same as that we used when we compare the joint frailty model and the joint frailty-copula model. In these settings, the data is generated from a

Frank copula, but we consider a mis-specified copula (i.e., Clayton) for model fitting to show if our method still performs satisfactory.

For all scenarios, non-informative priors are considered, $\Pr(\beta_{0D}) \propto 1$, $\Pr(\beta_{0R}) \propto 1$, $\Pr(\beta_D) \propto 1$, $\Pr(\beta_R) \propto 1$. Also, σ_w^{-2} are assumed to follow a prior $\text{gamma}(\alpha, 1/\alpha)$ with $\alpha = 0.001$ so that the prior is flat enough. It takes 500 iterations for burn-in period and extra 500 iterations for Markov chain Monte Carlo (MCMC) to converge. In order to lower down the dependence of the MCMC sample, the posterior sample is thinned for every 10 iterations. The results are summarized by the average of biases of posterior mean estimates (BIAS), standard deviation (SD) of posterior mean estimates, mean squared error (MSE) and the average of absolute bias (AB) of posterior mean estimates.

3.2 Simulation Results

In Table 1, we compare our method with the regular joint frailty model under the scenarios with a high failure rate ($\xi = 60\%$) and Kendall's τ , 0, 0.3 or 0.5. The true copula is a Clayton copula, which is considered for model fitting. We find out that our method always performs the best in terms of smallest AB and MSE when τ is 0.3 or 0.5. With regards to the regular joint frailty approach, there is no strong bias on the average of posterior mean estimates of $\beta_{D,1}$ and $\beta_{D,2}$, but AB still shows relatively higher bias, and also the estimates of $\beta_{R,1}$ and $\beta_{R,2}$ tend out to be substantially biased. When $\tau = 0.5$ and $n = 200$, the average of $\hat{\beta}_{R,1}$ and $\hat{\beta}_{R,2}$ are respectively 1.997 and 1.996 under the joint frailty-copula approach. By contrast, the mean of $\hat{\beta}_{R,1}$ and $\hat{\beta}_{R,2}$ are respectively 2.132 and 2.147 under the regular joint frailty approach. We also observe a trend that the bias of the frailty model increases when the true association between D_i and R_{ij} increases (i.e., τ increases). Comparatively, there is no trend for our approach that the bias will increase when the association increases. On the other hand, when true $\tau = 0$, the traditional method performs better than our method in terms of smaller SD, MSE and AB even though the outperformance is mild. Our proposed method is still applicable in particular when sample size n tends to be larger. For example, the average of $\hat{\beta}_{D,1}$ from 250 replicates with $n = 200$ is 0.990 under our method with BIAS, SD and MSE as -0.010 , 0.236 and 0.056 respectively. By contrast, the average $\hat{\beta}_{D,1}$ under joint frailty model is 0.989 with BIAS, SD and MSE as -0.011 , 0.223 and 0.050 . When sample size is 300 or 400, similar findings are detected, thus the results are not provided due to space limit.

In Table 2, we evaluate our method under the scenarios with different failure rates. Given the set-ups with medium failure rate and low failure rate, the biases are still small, but they seem greater compared with those under the high failure rate. When $n = 200$ and $\theta = 2$, the average bias of $\hat{\beta}_{R,2}$ is -0.004 under high failure rate, -0.005 under medium failure rate and -0.010 under low failure rate. The bias increases when the censoring rate increases. The MSE and SD still decrease when the sample size increases similar as above. Analogously, when $n = 100$ and $\theta = 2$, the MSE of $\hat{\beta}_{R,2}$ is 0.015 under medium failure rate. If n increases to 200, the MSE is 0.008, nearly half shrinkage.

In Figure 1, the BS increases when t increases. When $t = 0.031$ which is close to $t' = 0.03$, the BS is 0.006. Also, when t increases to 0.1, the BS increases to 0.14. The prediction error decreases when $|t - t'|$ increases. In other words, the prediction is less accurate if we want to predict the terminal event in the farther away from the time point t' , which agrees with our expectation. It is not doubt that the predictive accuracy will be substantially improved with more information accumulated for prediction. In addition, we find that the larger $|t - t'|$ is, the higher improvement our proposal achieves by comparing with the benchmark BS, for instance, the BS based on our proposal decreases 30% when $t = 0.1$ and $t' = 0.03$.

In Table 3, we evaluate our method when the copula model is mis-specified. The copula for data generation is a Frank copula, while we use a Clayton copula to fit the data. the estimates of $\beta_{D,1}$, $\beta_{D,2}$, $\beta_{R,1}$ and $\beta_{R,2}$, the parameters associated with $S_D(d|w_i)$ and $S_R(t|w_i)$ still perform satisfactory without strong bias. For the estimates of θ , we observe a strong bias because of model mis-specification. When $\theta = 1$ under a Frank copula, Kendal's τ is 0.11. However, the mean of the estimates of τ under a Clayton copula is 0.083, where the bias is about 0.028. Compared with the results of Table 1 with the results of Table 3, the biases of $\hat{\beta}_{D,1}$, $\hat{\beta}_{D,2}$, $\hat{\beta}_{R,1}$, $\hat{\beta}_{R,2}$ in Table 3 are almost identical. When the copula model is mis-specified, the bias of the covariate effects estimates are still negligible if the copula margins can be correctly specified.

Moreover, note that in majority of real data application and clinical practice, the positive correlation recurrent and terminal event processes is expected and assumed (Lu and Liu, 2014; Lin et al., 2015); however, negative correlation could also be possible. To explore the performance of our proposal in such situations, we conduct additional simulation studies by considering a more general Clayton copula model (Berkes et al., 2009), $C(u, v) = \max\{(u^{-\theta} + v^{-\theta} - 1), 0\}^{-1/\theta}$, $\theta \in [-1, 0) \cup (0, \infty)$, with the prior of θ defined by $\theta + 1 \sim \text{Gamma}(0.001, 1000)$. We simulated datasets with true $\theta = -0.1, -0.2$, sample size $n = 100, 200$, and the results (not shown due to space limit) indicate that our approach is still applicable with satisfactory performance.

4 Real Data Application

We apply the proposed method to a real data application on recurrent acute ischemic strokes. Our real data example is obtained from the MarketScan database between January, 2011 and December 2014, including the subjects who are aged 45-54 with surgical and medical admission for inpatient acute care hospitalization. A recurrent stroke is defined as any recurrent stroke occurring more than 28 days after the incident stroke (Coull and Rothwell, 2004). The baseline characteristics for enrolled subjects were tracked, for instance, gender, baseline stroke status, in-hospitalized mortality and pre-existing co-morbidity conditions including diabetes, cardiovascular disease and heart attack during the past twelve-month. Recently, literature have shown the higher likelihood of mortality in hospitalized stroke patients (Arabadzchieva et al., 2015; Feng et al., 2010). The goal of this study is to utilize our proposal for rigorous investigation of the effects of baseline factors on the risk of stroke occurrence or all-cause death, and further evaluation of the correlation between two event processes of recurrent strokes and all-cause death.

A sample of 2,122 patients are identified and among them, 597 (28.13%) patients experience at least one stroke event and 139 patients are dead with the death rate as 6.55%. Note that the low mortality rate is a limitation for our study because only hospitalized deaths are recorded for analysis. The censoring time is defined as the time when patients dropped out of the study. The average of the number of recurrent stroke events is 4 with $SD = 4.42$, and 37.7% of the patients experience only one stroke among those who have stroke recurrence. Based upon the independence assumption for all recurrent events, preliminary analyses on Kaplan-Meier curves show the evidence of higher risk of stroke for the patients with heart disease or hypertension (HD/HTN) or baseline stroke compared to those without HD/HTN or baseline stroke, which can be referred to Figure 2. We fit our proposed model with Clayton copula and compare with the traditional joint frailty model. The results are presented in Table 4, including the estimates of log hazard ratios (EST) and 95% credible limit (CL) represented by 2.5% CL and 97.5% CL. Note that the same prior set-ups as simulation studies are utilized for the hazard function and the variance of the frailty σ_w^2 when we fit the Clayton frailty-copula model. However, the iterations are increased to 1,000 for burn-in period and ensuring MCMC convergence. Similarly, the posterior sample is thinned for every 10 iterations to lower down the dependence of the MCMC sample.

Based on our proposal, there is a strong evidence that patients with HD/HTN are at increased risk for death ($HR=2.858$, 95% CL: 2.190-3.677) and stroke ($HR=1.46$, 95% CL: 1.165-1.613). Baseline stroke status and gender also has a significant association with both death and stroke. The hazards of death and stroke are respectively increased by 2.52 and 4.19 times if a patient have stroke at baseline. Also, females are detected to have significantly less risk of death and stroke compared to males, where the hazard for death is decreased by about 45% and that for stroke by 30%. Diabetes has an effect with trend towards significance on death and stroke; however, compared with baseline stroke and HD/HTN, the effect is not strong with the hazards for death and stroke increased by 19% and 16% respectively. We observe mild correlation (Kendall's $\tau = 0.32$, 95% CL: 0.049-0.547) between death and recurrent stroke events, indicating the necessity to adjust for the residual dependence by including potential time-dependent covariates in the model due to the different results from two models with regards to diabetes and gender effects. In particular, larger gender disparities on risk of death and stroke are identified, and for diabetes, the effects tend to be smaller. More importantly, all have narrower credible limits indicating improved efficiency after adjusting for the correlation between death and stroke.

We performed five-fold cross-validation to evaluate our prediction performance, where $K = 5$ is denoted as the total number of partitions. We randomly split data in five equal size partitions stratified by death status. At each time, we fit the model by excluding one partition, and then we calculate the crossvalidation BS (CVBS) upon the work by Emura et al. (2017). For instance, we first consider the time point, $t' = 2.33$ month, when half of the recurrent events occur in the database. If we want to predict the death risk of patients at the time point $t = 5.00$ month, the CVBS = 0.005 in comparison with the benchmark value of CVBS as 0.008. Furthermore, if we change t' to 3.33 and 4.33, the CVBS will decrease to 0.004 and less than 0.001 correspondingly, whilst the benchmark CVBS values are 0.006 and 0.003 respectively.

5 Discussion

We propose a joint frailty-copula method under the Bayesian framework to jointly model recurrent events and a terminal event. This method can be utilized when the conditional independence assumption is violated for the terminal and recurrent event processes, and also can provide direct estimate of their association. Based on numerical studies via simulation and real data application, our proposal achieves satisfactory performances in terms of smaller MSE and AB compared with regular joint frailty models, even though under the scenarios of independence between two event processes, our method still has comparable performance with regards to bias, SD and MSE. However, the advantage of our method, compared with the traditional frailty model, depends on the copula assumption. However, the performance could deteriorate if the assumption is violated. Of note is that the existing R package “joint.Cox” built up upon the work by Emura et al. (2015), and can be used to fit the joint frailty-copula model mainly for meta-analysis by utilizing the penalized likelihood approach. There are several advantages of our proposal: 1) The assumption of the memoryless property for the terminal event time D_i (i.e., $S_D(d|d) = S_D(d - d')$) in Emura et al. (2015) is relaxed, and more flexibility and applicability is provided by considering the time to the terminal event since the study initiation; 2) Our method is proposed under the Bayesian framework, which is easy to be implemented in statistical software and also informative priors can be incorporated for efficiency improvement if available. 3) The computation burden of the algorithm based on “joint.Cox” could be intensive, and the convergence issue for likelihood-based approaches may be encountered. Our program is coded by C language for the Gibbs sampler algorithm and called by R software, which can substantially improve the computing efficiency. The program is accessible to public upon request.

In addition, our approach is not sensitive to the magnitude of the dependence between the terminal and recurrent event processes and also the terminal event failure rate. Also, with regards to the copula mis-specification problem, the bias of the covariate effect estimation is still small when the copula margins are correctly specified. Therefore, in order to overcome the model mis-specification problem, more flexible copula margins are preferred (i.e., spline). After we choose the flexible margins, we could conduct copula model selection based on deviance information criteria. Another method to solve this problem could be the Bayesian model averaging technique, where we can update the weight of the candidate model based on the data, and the covariate effect estimation could be a weighted average from different candidate models.

Currently, we assume the number of the potential recurrent events is sufficiently large enough. However, this will not always be appropriate in practice. Some subjects might never have recurrent events because their hazards are extremely low. A cure model or a pattern mixture model can be considered in joint modeling to solve this problem, which is worthwhile for further study. On the other hand, we assume constant dependency, however, in some occasions, their dependence could vary over time. For instance, stronger correlation may be detected when death occur immediately after one subsequent event compared to the prior event. In such situation, dynamic relationship between recurrent events and death using

a time-varying copula can be incorporated into the model by replacing θ by $\theta(t)$, which could be another direction for future study.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank the existing work by Emura et al. (2015, 2017) for idea initiation and research motivation. The project described was supported in part by research grant U01 DK082183 from the National Institute of Digestive, Diabetes and Kidney Diseases of the National Institutes of Health, U.S. Department of Health and Human Services and by the National Center for Advancing Transnational Sciences, Grant KL2 TR000126 and TR002015. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Appendix

A. The likelihood of the joint frailty-copula model

For a subject with $\delta_i = 1$, $C_{ij} = \max\{D_i - T_{i,j-1}, 0\}$, and C_{ij} is independent of R_{ij} given $D_i = d_i$ and w_i . Similarly we can express the probability of the i_{th} subject who is terminated at d_i and experienced n_i events given w_i ,

$$\begin{aligned} \Pr(D_i = d_i, R_{i1} = r_{i1}, \dots, R_{i, n_i+1} \geq c_{i, n_i+1}, \dots, R_{i, J+1} \geq c_{i, J+1} | w_i) \\ = C_{(10)}^*(S_D(t_i | w_i), S_R(c_{i, n_i+1} | w_i)) \prod_{j=1}^{n_i} C_{(11)}^*(S_D(t_i | w_i), S_R(r_{ij} | w_i)) f_R(r_{ij} | w_i) \end{aligned}$$

Then, the likelihood of observing $\mathcal{D}_n$, given random frailty w is,

$$\begin{aligned} \mathcal{L}(\mathcal{D}_n | w) &= \prod_{i=1}^n \left(S_D(y_i | w_i) \prod_{j=1}^{n_i} \frac{C_{(01)}^*(S_D(t_i | w_i), S_R(r_{ij} | w_i)) f_R(r_{ij} | w_i)}{S_D(y_i | w_i)} \right)^{1-\delta_i} \\ &\quad \times \left(f_D(y_i | w_i) \prod_{j=1}^{n_i} \frac{C_{(11)}^*(S_D(y_i | w_i), S_R(r_{ij} | w_i)) f_D(y_i | w_i) f_R(r_{ij} | w_i)}{f_D(y_i | w_i)} \right)^{\delta_i} \\ &\quad \times \left(\frac{C_{(10)}^*(S_D(y_i | w_i), S_R(c_{i, n_i+1} | w_i)) f_D(y_i | w_i)}{f_D(y_i | w_i)} \right)^{\delta_i} \\ &\quad \times \left(\frac{C(S_D(y_i | w_i), S_R(c_{i, n_i+1} | w_i))}{S_D(t_{i, n_i+1} | w_i)} \right)^{1-\delta_i} \end{aligned}$$

After simplifying the above equation, we can express the likelihood by

$$\begin{aligned}\mathcal{L}(\mathcal{D}_n | \mathbf{w}) &= \prod_{i=1}^n [f_D(y_i | w_i) C_{10}^*(S_D(y_i | w_i), S_R(c_{i, n_i} + 1 | w_i))]^{\delta_i} \\ &\quad \times \left[S_T^{-n_i}(y_i | w_i) C(S_D(y_i | w_i), S_R(c_{i, n_i} + 1 | w_i)) \right]^{1 - \delta_i} \\ &\quad \times \prod_{j=1}^{n_i} \left[C_{(01)}(S_D(y_i | w_i), S_R(r_{ij} | w_i)) \right]^{1 - \delta_i} \left[C_{(11)}(S_D(y_i | w_i), S_R(r_{ij} | w_i)) \right]^{\delta_i} f_R(r_{ij} | w_i)\end{aligned}$$

B. Metropolis-Hastings within the Gibbs sampler algorithm

For $l = 1, \dots, M$, in the l^{th} iteration,

1. Sample $\beta_{0D}^{(l)}$ by $\Pr(\beta_{0D} | \beta_{0R}^{(l-1)}, \beta_T^{(l-1)}, \beta_R^{(l-1)}, \mathbf{w}^{(l-1)}, \theta^{(l-1)}, \sigma_w^{-2(l-1)}, D_n)$
 - a. Generate $\beta_{0D}^N = \beta_{0D}^{(l-1)} + sN(0, 1)$, where s is the step size of the random walk
 - b. Generate $U \sim \text{Unif}(0, 1)$
 - c. Calculate

$$LR = \frac{L(\beta_{0D}^N, \beta_{0R}^{(l-1)}, \beta_T^{(l-1)}, \beta_R^{(l-1)}, \mathbf{w}^{(l-1)}, \theta^{(l-1)}, \sigma_w^{-2(l-1)}, D_n)}{L(\beta_{0D}^{(l-1)}, \beta_{0R}^{(l-1)}, \beta_T^{(l-1)}, \beta_R^{(l-1)}, \mathbf{w}^{(l-1)}, \theta^{(l-1)}, \sigma_w^{-2(l-1)}, D_n)}.$$

- d. If $LR > U$, $\beta_{0D}^{(l)} = \beta_{0D}^N$. Otherwise, $\beta_{0D}^{(l)} = \beta_{0D}^{(l-1)}$.
2. Sample $\beta_{0R}^{(l)}$ by $\Pr(\beta_{0R} | \beta_{0D}^{(l)}, \beta_T^{(l-1)}, \beta_R^{(l-1)}, \mathbf{w}^{(l-1)}, \theta^{(l-1)}, \sigma_w^{-2(l-1)}, D_n)$, similarly as in step 1.
3. Sample $\beta_T^{(l)}$ by $\Pr(\beta_T | \beta_{0D}^{(l)}, \beta_{0R}^{(l)}, \beta_R^{(l-1)}, \mathbf{w}^{(l-1)}, \theta^{(l-1)}, \sigma_w^{-2(l-1)}, D_n)$, similarly as in step 1.
4. Sample $\beta_R^{(d)}$ by $\Pr(\beta_R | \beta_{0D}^{(l)}, \beta_{0R}^{(l)}, \beta_D^{(l)}, \mathbf{w}^{(d-1)}, \theta^{(l-1)}, \sigma_w^{-2(l-1)}, D_n)$, similarly as in step 1.
5. For $i = 1, \dots, n$, sample $w_i^{(l)}$ from $\Pr(w_i^{(l)} | \beta_{0D}^{(l)}, \beta_{0R}^{(l)}, \beta_D^{(l)}, \beta_R^{(l)}, \mathbf{w}_{[-i]}^{(l-1)}, \theta^{(l-1)}, \sigma_w^{-2(l-1)}, D_n)$
 - a. Generate $w_i^N = w_i^{(l-1)} + sN(0, 1)$, where s is the step size of the random walk
 - b. Generate $U \sim \text{Unif}(0, 1)$
 - c. Let $L_i(\cdot)$ denote the likelihood of the i^{th} subject. Calculate

$$LR = \frac{L_i(\beta_{0D}^{(l)}, \beta_{0R}^{(l)}, \beta_T^{(l)}, \beta_R^{(l)}, w_i^N, \theta^{(d-1)}, \sigma_w^{-2(l-1)}, D_n) \Pr(w_i^N)}{L_i(\beta_{0D}^{(l)}, \beta_{0R}^{(l)}, \beta_T^{(l)}, \beta_R^{(l)}, w_i^{(l-1)}, \theta^{(l-1)}, \sigma_w^{-2(l-1)}, D_n) \Pr(w_i^{(l-1)})}$$

d. If $LR > U$, $w_i^{(l)} = w_i^N$. Otherwise, $w_i^{(l)} = w_i^{(l-1)}$

6. Sample $\theta^{(l)}$ by $\Pr(\theta | \beta_{0D}^{(l)}, \beta_{0R}^{(l)}, \beta_D^{(l)}, \beta_R^{(l)}, w^{(l)}, \sigma_w^{-2(l-1)}, D_n)$ similarly as in step 1.
7. Sample $\sigma_w^{-2(l)}$ by $\Pr(\sigma_w^{-2} | \beta_{0D}^{(l)}, \beta_{0R}^{(l)}, \beta_D^{(l)}, \beta_R^{(l)}, w^{(l)}, D_n)$, which is a gamma distribution, i.e.,

$$\text{gamma}(\alpha + n / 2, (\alpha + w^{T(l)} w^{(l)} / 2)^{-1})$$

The step size in the algorithm is chosen so that the acceptance rate is around 0.44. The algorithm is programmed in C language for efficiency. After we get the posterior sample, we estimate Θ via the posterior mean to minimize the expected squared error loss function.

References

- Arabadzheva D, Kaprelyan A, Tzoukova A, and Kanazirev B (2015). Cardiovascular comorbidity in acute ischemic stroke. Scientific cooperations medical workshops.
- Berkes P, Wood F, and Pillow JW (2009). Characterizing neural dependencies with copula models. In Advances in neural information processing systems, pages 129–136.
- Cai J, Zeng D, and Pan W (2010). Semiparametric proportional means model for marker data contingent on recurrent event. Lifetime data analysis, 16(2):250–270. [PubMed: 20012357]
- Che X and Angus J (2016). A new joint model of recurrent event data with the additive hazards model for the terminal event time. Metrika, 79(7):763–787.
- Cheng Y and Fine JP (2012). Cumulative incidence association models for bivariate competing risks data. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 74(2):183–202. [PubMed: 22505835]
- Clayton DG (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. Biometrika, 65(1):141–151.
- Cook RJ and Lawless J (2007). The statistical analysis of recurrent events. Springer Science & Business Media.
- Cook RJ and Lawless JF (1997). Marginal analysis of recurrent events and a terminating event. Statistics in medicine, 16(8):911–924. [PubMed: 9160488]
- Coull AJ and Rothwell PM (2004). Underestimation of the early risk of recurrent stroke: evidence of the need for a standard definition. Stroke, 35:1925–1929. [PubMed: 15192248]
- Emura T, Nakatochi M, Matsui S, Michimae H, and Rondeau V (2017). Personalized dynamic prediction of death according to tumour progression and high-dimensional genetic factors: meta-analysis with a joint model. Statistical methods in medical research, page 0962280216688032.
- Emura T, Nakatochi M, Murotani K, and Rondeau V (2015). A joint frailty-copula model between tumour progression and death for meta-analysis. Statistical methods in medical research, page 0962280215604510.
- Feng W, Hendry RM, and Adams RJ (2010). Risk of recurrent stroke, myocardial infarction, or death in hospitalized stroke patients. Neurology, 74(7):588–593. [PubMed: 20157161]

- Fu H, Wang Y, Liu J, Kulkarni PM, and Melemed AS (2013). Joint modeling of progression-free survival and overall survival by a bayesian normal induced copula estimation model. *Statistics in medicine*, 32(2):240–254. [PubMed: 22806764]
- Georges P, Lamy A-G, Nicolas E, Quibel G, and Roncalli T (2001). Multivariate survival modelling: a unified approach with copulas.
- Ghosh D and Lin D (2002). Marginal regression models for recurrent and terminal events. *Statistica Sinica*, pages 663–688.
- Gilks WR, Best N, and Tan K (1995). Adaptive rejection metropolis sampling within gibbs sampling. *Applied Statistics*, pages 455–472.
- Graf E, Schmoor C, Sauerbrei W, and Schumacher M (1999). Assessment and comparison of prognostic classification schemes for survival data. *Statistics in medicine*, 18(17-18):2529–2545. [PubMed: 10474158]
- Huang C-Y and Wang M-C (2004). Joint modeling and estimation for recurrent event processes and failure time data. *Journal of the American Statistical Association*, 99(468):1153–1165. [PubMed: 24068850]
- Huang X and Liu L (2007). A joint frailty model for survival and gap times between recurrent events. *Biometrics*, 63(2):389–397. [PubMed: 17688491]
- Joe H (2005). Asymptotic efficiency of the two-stage estimation method for copula-based models. *Journal of Multivariate Analysis*, 94(2):401–419.
- Kalbfleisch JD, Schaabel DE, Ye Y, and Gong Q (2013). An estimating function approach to the analysis of recurrent and terminal events. *Biometrics*, 69(2):366–374. [PubMed: 23651362]
- Kelly PJ and Lim LL-Y (2000). Survival analysis for recurrent event data: an application to childhood infectious diseases. *Statistics in medicine*, 19(1):13–33. [PubMed: 10623910]
- Lawless JF (1987). Regression methods for poisson process data. *Journal of the American Statistical Association*, 82(399):808–815.
- Lin D, Sun W, and Ying Z (1999). Nonparametric estimation of the gap time distribution for serial events with censored data. *Biometrika*, 86(1):59–70.
- Lin D, Wei L, Yang I, and Ying Z (2000). Semiparametric regression for the mean and rate functions of recurrent events. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62(4):711–730.
- Lin L-A, Luo S, Chen BE, and Davis BR (2015). Bayesian analysis of multi-type recurrent events and dependent termination with nonparametric covariate functions. *Statistical methods in medical research*, page 0962280215613378.
- Lu L and Liu C (2014). Analysis of correlated recurrent and terminal events data in sas. In *Proceedings of the Northeast SAS Users Group Conference*. <http://www.lexjansen.com/nesug/nesug08/sa/sa16.pdf>, accessed May, volume 14.
- Mazroui Y, Mathoulin-Pelissier S, Soubeyran P, and Rondeau V (2012). General joint frailty model for recurrent event data with a dependent terminal event: application to follicular lymphoma data. *Statistics in medicine*, 31(11-12):1162–1176. [PubMed: 22307954]
- McGilchrist C and Aisbett C (1991). Regression with frailty in survival analysis. *Biometrics*, pages 461–466. [PubMed: 1912255]
- Nelsen RB (1999). Introduction. In *An Introduction to Copulas*, pages 1–4. Springer.
- Peng M, Xiang L, and Wang S (2018). Semiparametric regression analysis of clustered survival data with semi-competing risks. *Computational Statistics Data Analysis*, 124:53–70.
- Rivest L-P and Wells MT (2001). A martingale approach to the copula-graphic estimator for the survival function under dependent censoring. *Journal of Multivariate Analysis*, 79(1):138–155.
- Rondeau V, Mathoulin-Pelissier S, Jacqmin-Gadda H, Brouste V, and Soubeyran P (2007). Joint frailty models for recurring events and death using maximum penalized likelihood estimation: application on cancer events. *Biostatistics*, 8(4):708–721. [PubMed: 17267392]
- Rondeau V, Schaffner E, Corbière F, Gonzalez JR, and Mathoulin-Pélissier S (2013). Cure frailty models for survival data: Application to recurrences for breast cancer and to hospital readmissions for colorectal cancer. *Statistical methods in medical research*, 22(3):243–260. [PubMed: 21632696]

- Shih JH and Louis TA (1995). Inferences on the association parameter in copula models for bivariate survival data. *Biometrics*, pages 1384–1399. [PubMed: 8589230]
- Sinha D, Maiti T, Ibrahim JG, and Ouyang B (2008). Current methods for recurrent events data with dependent termination: a bayesian perspective. *Journal of the American Statistical Association*, 103(482):866–878. [PubMed: 19169433]
- Wang W and Wells MT (1997). Nonparametric estimators of the bivariate survival function under simplified censoring conditions. *Biometrika*, 84(4):863–880.
- Yu Z, Liu L, Bravata DM, and Williams LS (2014). Joint model of recurrent events and a terminal event with time-varying coefficients. *Biometrical Journal*, 56(2):183–197. [PubMed: 24285466]
- Yue H and Chan K (1997). A dynamic frailty model for multivariate survival data. *Biometrics*, pages 785–793. [PubMed: 9333346]
- Zeng D, Chen Q, and Ibrahim JG (2009). Gamma frailty transformation models for multivariate survival times. *Biometrika*, 96(2):277–291. [PubMed: 24948838]
- Zheng M and Klein JP (1995). Estimates of marginal survival for dependent competing risks based on an assumed copula. *Biometrika*, 82(1):127–138.

Figure 1.

Brier score (BS) vs the time point t

The black line is the BS curve when t increases from 0.03 to 0.10 with 0.01 increment, controlling that $t' = 0.03$. The blue line is the BS curve when t increases from 0.06 to 0.10 with 0.01 increment, controlling that $t' = 0.06$. The red line is the BS curve when t increases from 0.09 to 0.10 with 0.01 increment, controlling that $t' = 0.09$.

Figure 2.

Preliminary analyses on time-to-stroke based on Kaplan-Meier curves stratified by baseline heart disease or hypertension (HD/HTN) or baseline stroke with both p-values < 0.001 based on log-rank tests

Table 1

Summary statistics on joint modeling of recurrent events and a terminal event with a Clayton copula under the scenario with high failure rate ($\xi = 60\%$).

n	τ	Parameter	Frailty-copula				Frailty			
			BIAS	SD	MSE	AB	BIAS	SD	MSE	AB
100	0	$\beta_{D,1} = 1$	0.028	0.319	0.102	0.251	0.020	0.294	0.086	0.234
		$\beta_{D,2} = 1$	-0.007	0.146	0.021	0.116	0.003	0.138	0.019	0.111
		$\beta_{R,1} = 2$	-0.001	0.218	0.047	0.177	0.006	0.197	0.039	0.158
		$\beta_{R,2} = 2$	0.003	0.121	0.015	0.096	0.005	0.113	0.013	0.089
		τ	0.011	0.018	0.000	0.011	-	-	-	-
	0.3	$\beta_{D,1} = 1$	-0.001	0.264	0.070	0.207	0.113	0.291	0.097	0.249
		$\beta_{D,2} = 1$	-0.006	0.133	0.018	0.106	0.083	0.143	0.027	0.135
		$\beta_{R,1} = 2$	-0.009	0.222	0.049	0.170	0.093	0.224	0.059	0.194
		$\beta_{R,2} = 2$	-0.006	0.118	0.014	0.090	0.088	0.121	0.022	0.122
		τ	-0.004	0.036	0.001	0.029	-	-	-	-
	0.5	$\beta_{T,1} = 1$	-0.006	0.224	0.050	0.181	0.152	0.299	0.112	0.269
		$\beta_{D,2} = 1$	-0.018	0.120	0.015	0.097	0.116	0.148	0.035	0.154
		$\beta_{R,1} = 2$	-0.016	0.213	0.045	0.167	0.143	0.242	0.079	0.226
		$\beta_{R,2} = 2$	-0.018	0.116	0.014	0.092	0.140	0.136	0.038	0.158
		τ	-0.003	0.023	0.001	0.019	-	-	-	-
200	0	$\beta_{D,1} = 1$	-0.010	0.236	0.056	0.190	-0.011	0.223	0.050	0.181
		$\beta_{D,2} = 1$	0.003	0.112	0.013	0.089	0.008	0.110	0.012	0.087
		$\beta_{R,1} = 2$	-0.007	0.156	0.024	0.122	-0.013	0.146	0.021	0.117
		$\beta_{R,2} = 2$	0.003	0.077	0.006	0.059	0.002	0.071	0.005	0.056
		τ	0.007	0.011	0.000	0.007	-	-	-	-
	0.3	$\beta_{D,1} = 1$	-0.008	0.190	0.036	0.154	0.079	0.227	0.057	0.192
		$\beta_{D,2} = 1$	0.004	0.104	0.011	0.083	0.095	0.112	0.021	0.119
		$\beta_{R,1} = 2$	-0.009	0.152	0.023	0.121	0.079	0.163	0.033	0.141
		$\beta_{R,2} = 2$	0.006	0.089	0.008	0.072	0.092	0.088	0.016	0.104
		τ	-0.004	0.029	0.001	0.023	-	-	-	-
	0.5	$\beta_{D,1} = 1$	0.002	0.160	0.026	0.127	0.119	0.232	0.068	0.212
		$\beta_{D,2} = 1$	-0.002	0.088	0.008	0.069	0.133	0.115	0.031	0.147
		$\beta_{R,1} = 2$	-0.003	0.149	0.022	0.119	0.132	0.180	0.050	0.183
		$\beta_{R,2} = 2$	-0.004	0.085	0.007	0.068	0.147	0.098	0.031	0.154
		τ	0.000	0.023	0.001	0.019	-	-	-	-

BIAS, bias of posterior mean estimates from Monte Carlo datasets. SD, standard deviation of posterior mean estimates from Monte Carlo datasets. MSE, means squared error of posterior mean estimates. AB, average of absolute bias of posterior mean estimates from Monte Carlo datasets. Take

$$\beta_{D,1} \text{ as an example where } n_{sim} = 250, \text{ BIAS} = \frac{1}{n_{sim}} \sum (\hat{\beta}_{1,i} - \beta_{D,1}), \text{ SD} = \sqrt{\frac{1}{n_{sim}-1} \sum (\hat{\beta}_{1,i} - \bar{\hat{\beta}}_1)^2},$$

$\text{MSE} = \frac{1}{n_{sim}} \sum (\hat{\beta}_{1,i} - \beta_{1,i})^2$, $\text{AB} = \frac{1}{n_{sim}} \sum | \hat{\beta}_{1,i} - \beta_{1,i} |$, where n_{sim} is the number of the replicates and $\hat{\beta}_{1,i}$ is the posterior mean estimate of the i^{th} replicate.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Joint modeling of recurrent events and a terminal event with a Clayton copula under the scenarios with medium failure rate ($\xi = 50\%$) and low failure rate ($\xi = 40\%$)

n	τ	Parameter	Medium Failure ($\xi = 50\%$)				Low Failure ($\xi = 40\%$)			
			BIAS	SD	MSE	AB	BIAS	SD	MSE	AB
100	0.3	$\beta_{D,1} = 1$	0.012	0.289	0.084	0.232	0.002	0.311	0.096	0.241
		$\beta_{D,2} = 1$	-0.004	0.143	0.020	0.115	-0.006	0.162	0.026	0.129
		$\beta_{R,1} = 2$	-0.033	0.228	0.053	0.186	-0.025	0.238	0.057	0.182
		$\beta_{R,2} = 2$	-0.014	0.129	0.017	0.104	-0.021	0.148	0.022	0.118
		τ	-0.006	0.039	0.002	0.032	0.032	0.039	0.002	0.031
	0.5	$\beta_{D,1} = 1$	-0.001	0.257	0.066	0.203	0.001	0.276	0.076	0.218
		$\beta_{D,2} = 1$	-0.005	0.128	0.016	0.104	-0.015	0.139	0.019	0.114
		$\beta_{R,1} = 2$	-0.020	0.233	0.054	0.183	-0.012	0.234	0.055	0.183
		$\beta_{R,2} = 2$	-0.009	0.124	0.015	0.100	-0.014	0.133	0.018	0.106
		τ	-0.005	0.026	0.001	0.021	-0.001	0.025	0.001	0.020
200	0.3	$\beta_{D,1} = 1$	-0.030	0.195	0.039	0.160	-0.046	0.226	0.053	0.186
		$\beta_{D,2} = 1$	-0.005	0.107	0.011	0.084	-0.008	0.117	0.014	0.090
		$\beta_{R,1} = 2$	-0.016	0.162	0.026	0.132	-0.032	0.172	0.031	0.139
		$\beta_{R,2} = 2$	-0.005	0.090	0.008	0.070	-0.013	0.094	0.009	0.075
		τ	0.031	0.032	0.001	0.025	0.037	0.033	0.001	0.027
	0.5	$\beta_{D,1} = 1$	-0.008	0.171	0.029	0.135	-0.021	0.198	0.039	0.155
		$\beta_{D,2} = 1$	-0.004	0.092	0.008	0.073	-0.010	0.105	0.011	0.084
		$\beta_{R,1} = 2$	-0.009	0.155	0.024	0.125	-0.017	0.163	0.027	0.129
		$\beta_{R,2} = 2$	-0.005	0.088	0.008	0.071	-0.010	0.097	0.009	0.076
		τ	-0.002	0.023	0.001	0.018	0.001	0.024	0.001	0.020

BIAS, bias of posterior mean estimates from Monte Carlo datasets. SD, standard deviation of posterior mean estimates from Monte Carlo datasets. MSE, means squared error of posterior mean estimates. AB, average of absolute bias of posterior mean estimates from Monte Carlo datasets. Take

$$\beta_{D,1} \text{ as an example where } n_{sim} = 250, \text{ BIAS} = \frac{1}{n_{sim}} \sum (\hat{\beta}_{1,i} - \beta_{D,1}), \text{ SD} = \sqrt{\frac{1}{n_{sim}-1} \sum (\hat{\beta}_{1,i} - \bar{\hat{\beta}}_1)^2},$$

$$\text{MSE} = \frac{1}{n_{sim}} \sum (\hat{\beta}_{1,i} - \beta_{D,1})^2, \text{ AB} = \frac{1}{n_{sim}} \sum |\hat{\beta}_{1,i} - \beta_{D,1}|, \text{ where } n_{sim} \text{ is the number of the replicates and } \hat{\beta}_{1,i} \text{ is the posterior mean estimate of the } i^{th} \text{ replicate.}$$

Table 3

Summary statistics on joint modeling of recurrent events and a terminal event under the scenarios of high failure rate ($\xi = 60\%$). The true copula is a Frank copula, and the model fitting utilizes a Clayton copula.

τ	Parameter	n=100				n=100			
		BIAS	SD	MSE	AB	BIAS	SD	MSE	AB
0.11	$\beta_{D,1} = 1$	0.013	0.310	0.096	0.247	-0.012	0.231	0.053	0.181
	$\beta_{D,2} = 1$	-0.001	0.150	0.022	0.118	0.009	0.120	0.014	0.097
	$\beta_{R,1} = 2$	0.005	0.221	0.049	0.179	-0.001	0.159	0.025	0.126
	$\beta_{R,2} = 2$	0.009	0.107	0.012	0.087	0.020	0.080	0.007	0.065
	τ	-0.048	0.052	0.005	0.061	-0.053	0.041	0.005	0.058
0.21	$\beta_{D,1} = 1$	-0.009	0.325	0.105	0.259	-0.034	0.261	0.069	0.215
	$\beta_{D,2} = 1$	-0.032	0.157	0.026	0.129	-0.022	0.122	0.015	0.098
	$\beta_{R,1} = 2$	0.006	0.230	0.053	0.183	0.008	0.177	0.031	0.139
	$\beta_{R,2} = 2$	-0.001	0.111	0.012	0.089	0.014	0.088	0.008	0.071
	τ	-0.047	0.050	0.005	0.055	-0.057	0.041	0.005	0.060

BIAS, bias of posterior mean estimates from Monte Carlo datasets. SD, standard deviation of posterior mean estimates from Monte Carlo datasets. MSE, means squared error of posterior mean estimates. AB, average of absolute bias of posterior mean estimates from Monte Carlo datasets. Take

$$\beta_{D,1} \text{ as an example where } n_{sim} = 250, \text{ BIAS} = \frac{1}{n_{sim}} \sum (\hat{\beta}_{1,i} - \beta_{D,1}), \text{ SD} = \sqrt{\frac{1}{n_{sim}-1} \sum (\hat{\beta}_{1,i} - \bar{\hat{\beta}}_1)^2},$$

$\text{MSE} = \frac{1}{n_{sim}} \sum (\hat{\beta}_{1,i} - \beta_{D,1})^2, \text{ AB} = \frac{1}{n_{sim}} \sum |\hat{\beta}_{1,i} - \beta_{D,1}|$, where n_{sim} is the number of the replicates and $\hat{\beta}_{1,i}$ is the posterior mean estimate of the i^{th} replicate.

Table 4

Summary of real data analysis on recurrent stroke.

		Frailty-copula			Frailty		
	Covariates	EST	2.5% CL	97.5% CL	EST	2.5% CL	97.5% CL
Time to Death	HD/HTN	1.050	0.784	1.302	1.092	0.542	1.641
	Diabetes	0.175	0.051	0.550	0.236	-0.173	0.646
	Gender (=Female)	-0.603	-0.852	-0.249	-0.152	-0.547	0.242
	Baseline Stroke	1.258	0.764	2.175	0.908	0.501	1.314
Time to Stroke	HD/HTN	0.390	0.153	0.487	0.536	0.182	0.891
	Diabetes	0.154	-0.049	0.260	0.402	0.170	0.634
	Gender (=Female)	-0.359	-0.619	-0.258	-0.245	-0.470	-0.021
	Baseline Stroke	1.647	1.462	1.862	2.318	2.065	2.572
Correlation	τ	0.320	0.049	0.547			

EST: the estimates of log hazard ratios except for the correlation coefficient τ ; 2.5% CL: 2.5% credible limit; 97.5% CL: 97.5% credible limit