

The Pennsylvania State University
The Graduate School
College of Medicine

JOINT MODELING RECURRENT EVENTS AND A TERMINAL
EVENT WITH FRAILTY-COPULA MODELS

A Dissertation in
Biostatistics
by
Zheng Li

© 2018 Zheng Li

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2018

The dissertation of Zheng Li was reviewed and approved* by the following:

Ming Wang
Assistant Professor of Professor of Public Health Sciences
Co-Chair of Committee
Dissertation Co-Advisor

Vernon M. Chinchilli
Distinguished Professor of Public Health Sciences
Chair of Department of Public Health Sciences
Co-Chair of Committee
Dissertation Co-Advisor

Runze Li
Verne M. Willaman Professor of Statistics

Feng Yue
Assistant Professor of Biochemistry and Molecular Biology
Director of Bioinformatics Division

Nasrollah Ghahramani
Professor of College of Medicine
Interim Chief, Division of Nephrology
Special Member of Committee

Arthur Berg
Associate Professor of Public Health Sciences
Head of Biostatistics Program

*Signatures are on file in the Graduate School.

Abstract

A terminal event can stop a series of recurrent events, which commonly occurs in biomedical and clinical studies. In this situation, the non-informative censoring assumption could fail because of potential dependency between these two event processes, leading to invalid inference if we analyze 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 two processes also depends on time-varying covariates across recurrences. For example, time to death and time to stroke might both depend on the age of the patients. And when we do not include age in the survival models, the subject-level frailty cannot capture the change of the age across recurrences and lead to the violation of the conditional independence assumption. Furthermore, marginal correlation between two event processes based on traditional frailty modeling has no closed form solution for estimation. In order to fill these gaps, we propose a novel joint frailty-copula approach to model recurrent events and a terminal event with modest assumptions under Bayesian framework. 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(MSE) of the propose approach is smaller when the conditional independence assumption is violated. We applied our method into a real example extracted from the MarketScan database to identify potential risk factors and study the association between recurrent strokes and all-cause mortality.

Another important assumption under the joint frailty model is that the correlation between the terminal event and the recurrent events is constant over time.

This is an unrealistic assumption. For example, when we study myocardial infarctions, time to death might be more correlated with the last myocardial infarction compared with the earlier myocardial infarction. We propose a time-varying joint frailty-copula model to further relax this assumption. Under this model, the dynamic correlation between the terminal event and the recurrent event is modeled by a latent Gaussian AR(1) process. The simulation results show that compared with the joint frailty model and the joint frailty-copula model, the bias, SD, MSE and AB of the time-varying frailty copula model are the smallest. Then, we applied our method to analyze the CHS data to identify potential risk factors to myocardial infarction and stroke.

In summary, we propose two methods to jointly model recurrent events and a terminal event. Both methods outperform the traditional joint frailty model when the conditional independence assumption is violated. The time-varying joint frailty-copula model is more flexible, which allows the correlation between the recurrent event process and the terminal event process to change over time. One future topic is to jointly modeling biomarker, recurrent events and death time by the frailty-copula models. Other topics include extending the current model to a cure rate model and modeling multiple types of recurrent events.

Table of Contents

List of Figures	viii
List of Tables	ix
Acknowledgments	xi
Chapter 1	
Introduction	1
1.1 Background	1
1.2 Literature Review	2
1.2.1 Non-recurrent Events	2
1.2.2 Recurrent Events	4
1.2.3 Recurrent Events with Dependent Censoring	5
Chapter 2	
Basic Models and Inference Approach	7
2.1 Introduction	7
2.2 Shared Frailty Models	7
2.3 Joint Frailty Models	10
2.4 Copulas	12
2.5 Parameter Estimation	14
2.5.1 Monte Carlo EM(MCEM) Algorithm	14
2.5.2 Penalized Likelihood Approach	15
Chapter 3	
A Joint Frailty-copula Model for Recurrent Events and a Terminal Event	17
3.1 Introduction	17

3.2	Methods	20
3.2.1	Notation	20
3.2.2	Joint Frailty-copula Model (JFCM)	21
3.2.3	Metropolis-Hastings within the Gibbs Sampler Algorithm	23
3.2.4	Prediction for the Time to the Terminal Event of a New Subject	24
3.3	Simulation Study	26
3.3.1	Simulation Set-up	26
3.3.2	Simulation Results	28
3.4	Real Data Application	29
3.5	Discussion	31
Chapter 4		
A Time-varying Joint Frailty-copula Approach for Recurrent Events and a Terminal Event		37
4.1	Introduction	37
4.2	Methods	38
4.2.1	Notation	38
4.2.2	Time-varying Joint Frailty-copula Model	39
4.2.3	A Clayton Time-varying Joint Frailty-copula Model	40
4.2.4	Bayesian Inference	41
4.3	Simulation	43
4.3.1	Simulation Set-up	43
4.3.2	Simulation Results	44
4.4	Real Data Application	45
4.5	Discussion	48
Chapter 5		
Summary and Future Work		59
5.1	Summary	59
5.2	Future Work	60
Appendix A		
Appendix for chapter 3		62
A.1	Derivation of the likelihood	62
A.2	Metropolis-Hastings within the Gibbs sampler algorithm	63
Appendix B		
Appendix for chapter 4		65

List of Figures

- 3.1 The Brier score (BS) vs the time point t 35
- 3.2 Preliminary analysis on Kaplan-Meier Curves for stroke 35

- 4.1 Forest plots for joint models 47
- 4.2 Kaplan-Meier Curves of MI Between Different Groups 47
- 4.3 Kaplan-Meier Curves of Stroke Events Between Different Groups . . 57
- 4.4 Forest plots for joint models 58

List of Tables

3.1	Summary statistics on joint modeling of recurrent events and a terminal event by a Clayton Copula under the scenario with high failure Rate ($\xi = 60\%$).	33
3.2	Summary statistics on joint modeling of recurrent events and a terminal event by a Clayton Copula under the scenarios with medium failure Rate ($\xi = 50\%$) and low failure rate ($\xi = 40\%$)	34
3.3	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.	36
3.4	Summary of real data analysis on recurrent stroke	36
4.1	Comparison between the time-varying joint Clayton frailty-copula model and the joint frailty model when $\theta_\mu = 1$, censoring rate 40% .	50
4.2	Comparison between the time-varying joint Clayton frailty-copula model and the joint frailty model when $\theta_\mu = 2$, censoring rate 40% .	51
4.3	Comparison between the time-varying joint Clayton frailty-copula model and the joint frailty-copula model when $\theta_\mu = 1$, censoring rate 40%	52
4.4	Comparison between the time-varying joint Clayton frailty-copula model and the joint frailty-copula model when $\theta_\mu = 2$, censoring rate 40%	53
4.5	Comparison between the time-varying joint Clayton frailty-copula model and the joint frailty-copula model when $\theta_\mu = 3$, censoring rate 40%	54
4.6	Joint modeling by the time-varying joint Clayton frailty-copula model when $\theta_\mu = 1$, event rate 20%	55
4.7	Joint modeling by the time-varying joint Clayton frailty-copula model when $\theta_\mu = 1$, event rate 60% and the copula model is misspecified	55

4.8	Baseline Characteristics for Patients with MI Events and No MI Events	56
4.9	Baseline Characteristics for Patients with Stroke Events and No Stroke Events	56
4.10	Joint Modeling MI and Death by TVJFCM	57
4.11	Joint Modeling Stroke and Death by TVJFCM and JFM	58

Acknowledgments

I would like to specially thank Professor Ming Wang and Professor Vernon M. Chinchilli for their help in my research and their support for my summer internship. Also, I would like to thank all the committee members for their time to read my dissertation and attend my defense.

I would like to thank all the professors in my department for teaching lectures, answering my questions, and grading my homework during these years. Because of their hard work, I learned a lot in the statistical field.

Finally, I would like to thank my classmates and the students for the friendship in these four years. Thank you for giving me an incredible student life!

Dedication

I dedicate this dissertation to my parents and my grandparents who support my study for nearly ten years.

Introduction

1.1 Background

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 (Jansen et al., 2018; Ridker et al., 2000), such as successive hospitalized myocardial infarction or heart attack, Cox PH models cannot be directly applied due to the failure to account for the 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 dependency between the recurrent event process and the terminal event process, leading to the violation of the non-informative censoring assumption. Thus, in order to have valid inference, we need to adjust for both the correlation among recurrent events and also the informative censoring mechanism due to the competing risk of the terminal event.

In this dissertation, we include two data examples. One is the MarketScan data set. In this data set, patients with recurrent stroke events are of the main interest. However, death, as a terminal event, will be a semi-competing event which stops the recurrences of stroke events. And some baseline covariates such as gender, age, hypertension, and diabetes may influence both risks of death and stroke. Our

objective is to develop a method which can quantify the association between stroke and death. And also, another goal is to dynamically predict the risk of death based on the observed patient history and covariates.

Another example is a large study for cardiovascular diseases (CHS). In this example, we are interested in two types of recurrent events, myocardial infarction and stroke. But the problem is more complex in that the correlation between the terminal event and the recurrent events possibly changing over time. So in order to have a valid result, we also need to account for the time-varying correlation.

1.2 Literature Review

1.2.1 Non-recurrent Events

When the primary interested event is not recurrent, two survival models are widely used, the Cox proportional hazard (PH) model (Cox, 1992) and the accelerated failure time (AFT) model (Buckley and James, 1979). The interpretation is different for these two models.

Let T_i and C_i respectively denote the event time and the censoring time of the i^{th} subject. $Y_i = \min\{T_i, C_i\}$ is the follow-up time and $D_i = I(T_i < C_i)$ is the failing indicator.

The Cox PH model is a semiparametric model, which is specified via the hazard function

$$\lambda(t) = \lambda_0(t) \exp(\boldsymbol{\beta}' \boldsymbol{x})$$

where \boldsymbol{x} is the associated covariate vector and $\lambda_0(t)$ is the baseline hazard. We do not need to specify the detailed form of $\lambda_0(t)$. Under this model the log hazard is linear. And the hazard functions between different groups are proportional over time (the proportional hazard assumption).

The estimation of the parameters in the model is based on maximum semi-parametric likelihood. Let $\mathcal{R}_i = \{j : Y_j \geq t_i\}$, which is the risk set of individuals who have not experienced the event prior to time t_i . The semiparametric partial

likelihood function is

$$\mathcal{L}(\boldsymbol{\beta}) = \prod_{i=1}^n \left\{ \frac{\exp(\boldsymbol{\beta}' \mathbf{x}_i)}{\sum_{j \in \mathcal{R}_i} \exp(\boldsymbol{\beta}' \mathbf{x}_j)} \right\}^{d_i}, \quad (1.1)$$

where n represents the number of distinct time points during which an event occur.

The baseline cumulative hazard function is estimated by a Breslow estimator,

$$\hat{\Lambda}_{0T}(t) = \sum_{i=1}^n \frac{I(Y_i \leq t_i) D_i}{\sum_{j \in \mathcal{R}_i} \exp(\boldsymbol{\beta}' \mathbf{x}_j)}.$$

Recent articles related to extensions of the Cox PH model are available (Lunn and McNeil, 1995; Sy and Taylor, 2000; Bellera et al., 2010).

The AFT model is in a more direct form to model the event time T_i . It assumes that

$$\log(T_i) = \mu + \boldsymbol{\beta}'_A \mathbf{x}_i + \sigma W_i, \quad (1.2)$$

where σ is the scale parameter and W_i is the residual with density function $f_W(\cdot)$. There are several choices for $f_W(\cdot)$. For example, when W_i follows standard extreme value distribution with density function $f_W(w) = \exp(w - e^w)$, T_i follows a Weibull distribution. Under this special case, we could also prove that $\boldsymbol{\beta}'_A = -\boldsymbol{\beta}'$. The hazard function under this model generally is different than the Cox PH model and can be expressed by

$$h(t) = \exp(-\boldsymbol{\beta}'_A \mathbf{x}_i) h_0 \{(-\boldsymbol{\beta}'_A \mathbf{x}_i) t\}. \quad (1.3)$$

In this model, $\exp(\boldsymbol{\beta}'_A \mathbf{x}_i)$ can be interpreted as an acceleration factor associated with the covariate vector \mathbf{x}_i . The likelihood under this model is

$$\mathcal{L}(\boldsymbol{\beta}) = \prod_{i=1}^n h(y_i)^{\delta_i} S(y_i). \quad (1.4)$$

$\hat{\boldsymbol{\beta}}_A$ is solved by the Newton-Raphson algorithm to maximize the likelihood. Recent works on the AFT model include Wei (1992), Lin and Ying (1995), Kuo and Mallick (1997).

1.2.2 Recurrent Events

Recurrent event times are typically analyzed in two scales, calendar event time or gap time. The calendar event time is calculated from the study onset to the time when recurrent event happens. Suppose the i^{th} patient experienced a total of n_i recurrent events. Let T_{ij} denote the j^{th} calendar recurrent event time. Then, the j^{th} gap time, R_{ij} , is defined by $T_{ij} - T_{i,j-1}$.

When we study T_{ij} , we assume T_{ij} follows a Poisson process and it is appropriate when the recurrent event is caused by external factors. And especially for incidental events, Poisson processes and other counts-based models are useful. Whereas, when we study R_{ij} , R_{ij} is assumed to follow a renewal process. It is assumed that a subject is renewed after each recurrent event occurs. And $R_{ij}, j = 1, \dots, n_i + 1$ are mutually independent. Renewal processes are more suitable for events caused by internal physical cycles. More details can be found in (Cook and Lawless, 2007).

There exists substantial work on recurrent event analysis in the literature to analyze calendar event time and gap event time. Much of the work we refer here can be used to analyze both types of time scales but under calendar event time scale, we are modeling an intensity function and under the recurrent event time scale, we are modeling a hazard function. We first introduce some method to analyze recurrent events. 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 constant frailty assumption. In particular, for each subject and each recurrent event, a time-dependent frailty is associated with the corresponding intensity function of the recurrent event. Among others, relevant research topics on the most recent developments or comprehensive review can be found in the literature (Lin et al., 1999, 2000; Kelly and Lim, 2000; Cook and Lawless, 2007).

Considering nonparametric methods, most of the methods are based on the counting process and unbiased estimating equations. Some important articles include Andersen and Gill (1982), Aalen and Husebye (1991) and Dabrowska et al. (1994). More recently, Lawless and Nadeau (1995) developed robust methods for

estimating rate function and mean function. Martinussen and Scheike (2002) proposed an additive multiplicative hazard model. Andersen et al. (2012) provides detailed reviews of the counting process for survival models.

1.2.3 Recurrent Events with Dependent Censoring

In order to account for the informative 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; Zeng and Lin, 2009) and a frailty model approach (Huang and Wang, 2004; Huang and Liu, 2007; Rondeau et al., 2007).

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 the recurrent event 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. Recently, Rondeau et al. (2013) proposed a cure frailty model. Yu et al. (2014) proposed a model with time-varying coefficient. 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 the articles by Ghosh and Lin (2002); Mazroui et al. (2012); Kalbfleisch et al. (2013).

Cook and Lawless (1997) proposed a marginal approach to study recurrent events and the terminal event process. The main interest of their work is to estimate a marginal rate function and a marginal mean function of the recurrent event process based on methods-of-moment estimator. The correlation between the terminal event time and the recurrent event time is not estimable. Ghosh and Lin (2002) proposed to estimate the rate function of the recurrent event, incorporating the survival probability of the terminal event. Ye et al. (2007) proposed an estimating equation approach to estimate the correlation between the terminal event and the recurrent event under a gamma frailty model. Zeng and Lin (2009) proposed a semiparametric method to estimate the additive hazard for recurrent

events under dependent censoring. Other works can be found in the literature (Zhu et al., 2010; Zeng et al., 2014; Chen et al., 2015).

All the above methods do not consider (1) the violation of the conditional independence assumption and (2) the time-varying correlation between the terminal event and the recurrent event process. In this dissertation, we mainly focus on jointly modeling a recurrent event process and a terminal event model by frailty type models. In Chapter 2, we will give more details about the methods of model fitting for frailty models, including the MCEM algorithm and the penalized likelihood approach. In Chapter 3, we propose a joint frailty-copula model and a dynamically prediction method with application on the MarketScan data set. In Chapter 4, we further extend the joint frailty-copula model by allowing the correlation varying across time-to-event recurrences. We applied this model to solve the problem of the CHS data set. In Chapter 5, we summarize the results and discuss some future topics to study.

Basic Models and Inference

Approach

2.1 Introduction

In this chapter, we mainly introduce some basic models under Cox PH framework to analyze the recurrent events. We first introduce the models proposed by Vaupel et al. (1979) and McGilchrist and Aisbett (1991), which assume non-informative censoring. Then, considering the dependent censoring, we introduce the models proposed by Huang and Wang (2004); Rondeau et al. (2007); Mazroui et al. (2012); Yu et al. (2014); Che and Angus (2016). Copulas, as an alternative way to model survival outcomes, are also introduced. After that, we discuss the MCEM algorithm (Liu et al., 2004) and the penalized likelihood approach (Rondeau et al., 2007) to estimate the parameters in joint frailty models.

2.2 Shared Frailty Models

The first model we introduce here is the gamma shared frailty model. It was first proposed by Vaupel et al. (1979) to account for population heterogeneity. This model is extended to analyze the recurrent events framework in Kelly and Lim (2000). Let \mathcal{F} denote the maximum follow-up time of the study. And let C_i be the censoring time of the patient. Then $Y_i = \min\{\mathcal{F}, C_i\}$. Let T_{ij} be the calendar event

time for the j^{th} recurrent event of the i^{th} patient. And the censoring time for T_{ij} is C_{ij} , which is defined as $Y_i - T_{i,j-1}$. Let $Y_{ij} = \min\{R_{ij}, C_{ij}\}$ and $\Delta_{ij} = I(T_{ij} < C_{ij})$ for the j^{th} recurrent event.

Under this model, the hazard function for the recurrent events is,

$$h_R(r|u_i) = u_i h_{0R}(r) \exp(\boldsymbol{\beta}'_R \mathbf{x}_{i,R}),$$

where $h_{0R}(\cdot)$ is the baseline hazard function. The form of h_{0R} is pre-specified (i.e., exponential baseline hazard or Weibull baseline hazard) and it is associated with a parameter vector $\boldsymbol{\beta}_{0R}$. $\boldsymbol{\beta}_R$ is the coefficient vector associated with the covariate vector $\mathbf{x}_{i,R}$. u_i is a frailty, following a gamma distribution, denoted by $G(k, \lambda)$.

This model could be applied to analyze both the recurrent events in calendar time scale or gap times. The parameter estimates can be obtained by maximizing the semiparametric likelihood.

The marginal density function of recurrent events can be derived by

$$\begin{aligned} f_R(r) &= \int_0^\infty u_i h_{0R}(r) \exp(\boldsymbol{\beta}'_R \mathbf{x}_{i,R}) \exp\{-u_i H_{0R}(r) \exp(\boldsymbol{\beta}'_R \mathbf{x}_{i,R})\} f_u(u_i) du_i \\ &= h_{0R}(r) \exp(\boldsymbol{\beta}'_R \mathbf{x}_{i,R}) \\ &\quad \times \int_0^\infty \frac{\lambda^k}{\Gamma(k)} u_i^k \exp\{-u_i(\lambda + H_{0R}(r) \exp(\boldsymbol{\beta}'_R \mathbf{x}_{i,R}))\} du_i \\ &= h_{0R}(r) \exp(\boldsymbol{\beta}'_R \mathbf{x}_{i,R}) k \lambda^k / (\lambda + H_{0R}(r) \exp(\boldsymbol{\beta}'_R \mathbf{x}_{i,R}))^k \end{aligned}$$

Let $\boldsymbol{\beta} = \{\boldsymbol{\beta}_{0R}, \boldsymbol{\beta}_R\}$. Suppose we observe $\mathcal{D}_n = \{y_{ij}, \delta_{ij}, i = 1, \dots, n, j = 1, \dots, n_i + 1\}$. Then the likelihood of the observed data is

$$\mathcal{L}(\boldsymbol{\beta}) = \prod_{i=1}^n \prod_{j=1}^{n_i+1} h_R(y_{ij})^{\delta_{ij}} S_R(y_{ij}).$$

The MLE is solved by a Newton-Raphson algorithm, iteratively updating $\boldsymbol{\beta}^{\text{new}}$ as

$$\boldsymbol{\beta}^{\text{new}} = \boldsymbol{\beta}^{\text{old}} - V(\boldsymbol{\beta}) S(\boldsymbol{\beta}^{\text{old}}),$$

where $S(\boldsymbol{\beta}) = \partial \ell(\boldsymbol{\beta}) / \partial \boldsymbol{\beta}$ and $V(\boldsymbol{\beta}) = (\partial^2 \ell / \partial \boldsymbol{\beta} \partial \boldsymbol{\beta}')^{-1}$.

McGilchrist and Aisbett (1991) proposed a lognormal frailty model as an al-

ternative for the gamma frailty model,

$$h_R(r|u_i) = h_{0R}(r) \exp(\boldsymbol{\beta}'_R \mathbf{x}_{i,R} + \omega_i),$$

where the frailty ω_i follows $\mathcal{N}(0, \sigma_\omega^2)$.

There is no closed form either for the survival function or the density function under this model. McGilchrist and Aisbett (1991) proposed to estimate $\boldsymbol{\beta}_R$ by a restricted likelihood approach (REML).

We treat $\boldsymbol{\omega}$ as the unobserved data. Let $\ell(\boldsymbol{\beta}|\boldsymbol{\omega}, \mathcal{D}_n)$ denote the log likelihood of the complete data set, which can be expressed by,

$$\begin{aligned} \ell(\boldsymbol{\beta}|\boldsymbol{\omega}, \mathcal{D}_n) &= \sum_{i=1}^n \sum_{j=1}^{n_i+1} \log h_R(y_{ij})^{\delta_{ij}} \log S_R(y_{ij}) + \sum_{i=1}^n \log f_\omega(\omega_i) \\ &= \sum_{i=1}^n \sum_{j=1}^{n_i+1} \delta_{ij} \{ \log h_{0R}(y_{ij}) + \boldsymbol{\beta}_R \mathbf{x}_{i,T} + \omega_i \} \\ &\quad - \exp(\boldsymbol{\beta}_R \mathbf{x}_{i,T} + \omega_i) \sum_{i=1}^n \sum_{j=1}^{n_i+1} \int_0^{y_{ij}} h_{0R}(r) dr \\ &\quad + \sum_{i=1}^n \left\{ -\frac{\omega_i^2}{2\sigma_\omega^2} - \log \sqrt{2\pi\sigma_\omega^2} \right\}. \end{aligned}$$

We can decompose the log likelihood into two parts,

$$\ell(\boldsymbol{\beta}|\mathcal{D}_n, \boldsymbol{\omega}) = \ell_1(\boldsymbol{\beta}, \boldsymbol{\omega}) + \ell_2(\boldsymbol{\omega}),$$

where

$$\begin{aligned} \ell_1(\boldsymbol{\beta}, \boldsymbol{\omega}) &= \sum_{i=1}^n \sum_{j=1}^{n_i+1} \log h_R(y_{ij})^{\delta_{ij}} \log S_R(y_{ij}) + \sum_{i=1}^n \log f_\omega(\omega_i) \\ &= \sum_{i=1}^n \sum_{j=1}^{n_i+1} \delta_{ij} \{ \log h_{0R}(y_{ij}) + \boldsymbol{\beta}_R \mathbf{x}_{i,T} + \omega_i \} \\ &\quad - \exp(\boldsymbol{\beta}_R \mathbf{x}_{i,T} + \omega_i) \sum_{i=1}^n \sum_{j=1}^{n_i+1} \int_0^{y_{ij}} h_{0R}(r) dr \end{aligned}$$

and

$$\ell_2(\boldsymbol{\omega}) = \sum_{i=1}^n \left(-\frac{\omega_i^2}{2\sigma_\omega^2} - \log \sqrt{2\pi\sigma_\omega^2} \right)$$

The best linear unbiased predictor (BLUP), which maximize ℓ conditional on σ_ω^2 by the Newton-Raphson algorithm is

$$\begin{bmatrix} \boldsymbol{\beta}_{new} \\ \boldsymbol{\omega}_{new} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\beta}_{old} \\ \boldsymbol{\omega}_{old} \end{bmatrix} - \mathbf{V}^{-1} \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\omega}_{old} \end{bmatrix} + \mathbf{V}^{-1} \begin{bmatrix} \partial\ell_1/\partial\boldsymbol{\beta} \\ \partial\ell_1/\partial\boldsymbol{\omega} \end{bmatrix}_{\boldsymbol{\beta}=\boldsymbol{\beta}_{old}, \boldsymbol{\omega}=\boldsymbol{\omega}_{old}}$$

where

$$\mathbf{V} = \begin{bmatrix} \partial^2\ell_1/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}' & \partial^2\ell_1/\partial\boldsymbol{\beta}\partial\boldsymbol{\omega}' \\ \partial^2\ell_1/\partial\boldsymbol{\omega}\partial\boldsymbol{\beta}' & \partial^2\ell_1/\partial\boldsymbol{\omega}\partial\boldsymbol{\omega}' + \sigma_\omega^{-2}\mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix}.$$

Suppose

$$\mathbf{V}^{-1} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix},$$

where the dimensions of \mathbf{A}_{ij} are the same as \mathbf{V}_{ij} . In each iteration, σ_ω^2 is updated by a consistent estimator $\boldsymbol{\omega}'\boldsymbol{\omega}/n$. The REML estimation of σ_ω^2 is $\boldsymbol{\omega}'\boldsymbol{\omega}/(n-r)$, where $r = \sigma_\omega^{2old} \text{tr}(\mathbf{A}_{22})$.

The advantage of the gamma frailty model is that the likelihood has closed form and the estimation is comparatively easy. However, the lognormal model is more flexible for extensions such as for jointly modeling recurrent events and longitudinal biomarker data. Another extension is the dynamic frailty model. It is easy to specify the joint distribution and the variance covariance matrix of the frailties.

2.3 Joint Frailty Models

The joint frailty model was proposed by (Liu et al., 2004), which is still widely used nowadays. Suppose that T_i is the terminal event time for the i^{th} subject and C_i is the censoring indicator for the terminal event time. Let $D_i = I(T_i < C_i)$ denote the censoring indicator for the terminal event. Under the joint frailty model, the

hazard functions for T_i and T_{ij} are,

$$\begin{aligned} h_T(t_i|u_i) &= u_i h_{0T}(t_i) \exp(\boldsymbol{\beta}'_T \mathbf{x}_{i,T}) \\ h_R(t_{ij}|u_i) &= u_i^\psi h_{0R}(t_{ij}) \exp(\boldsymbol{\beta}'_R \mathbf{x}_{i,R}), \end{aligned} \quad (2.1)$$

where u_i is a frailty, following a gamma distribution $G(\vartheta, 1/\vartheta)$. The mean of the gamma distribution is restricted to be 1 to avoid identifiability problems. ψ is the correlation parameter between the terminal event T_i and the recurrent event time T_{ij} . If $\psi > 0$, T_i and T_{ij} are positively correlated and otherwise, T_i and T_{ij} are negatively correlated.

Suppose there are n subjects in the study and suppose each subject experienced a total of n_i recurrent events. Let $\boldsymbol{\Theta}$ denote the parameter vector which contains all the parameters in the model. The likelihood is

$$\mathcal{L}(\boldsymbol{\Theta}) = \prod_{i=1}^n \int_0^\infty h_T^{d_i}(y_i|u_i) S_T(y_i|u_i) \left\{ \prod_{i=1}^{n_i+1} h_R^{\delta_{ij}}(y_{ij}|u_i) S_R(y_{ij}|u_i) \right\} f_u(u_i) du_i,$$

where $f_u(\cdot)$, $S_T(\cdot)$ and $S_R(\cdot)$ are respectively the density function of u , the conditional survival function of the terminal event time and conditional survival function of the recurrent event time.

Rondeau et al. (2007) extended this model by 1) allowing the recurrent event times to be gap times, and 2) proposing a piecewise spline function for the baseline hazard. Under this model, they induced the recurrent event gap times $R_{ij} = T_{ij} - T_{i,j-1}$. The baseline hazard functions $h_{0T}(t)$ and $h_{0R}(r)$ are estimated by a cubic M-splines (Ramsay, 1988). Suppose there are a total of K knots for the baseline hazard functions $h_{0T}(\cdot)$ and $h_{0R}(\cdot)$, denoted by $\tilde{\mathbf{t}} = \{\tilde{t}_1, \dots, \tilde{t}_K\}$ and $\tilde{\mathbf{r}} = \{\tilde{r}_1, \dots, \tilde{r}_K\}$. The hazard functions are respectively,

$$\begin{aligned} h_{0T}(t) &= \sum_{k=1}^K \beta_{0T,k} M_k(t_k|a, \tilde{\mathbf{t}}) \\ h_{0R}(r) &= \sum_{k=1}^K \beta_{0R,k} M_k(t_k|a, \tilde{\mathbf{r}}), \end{aligned}$$

where $M_k(t_k|a, t)$ is the a -order cubic M-spline basis function, defined by

$$M_k(x|1, t) = \begin{cases} 1/(t_{k-1} - t_k), t_k \leq x < t_{k+1} \\ 0, \text{otherwise} \end{cases}$$

and

$$M_k(x|a, t) = \frac{a \{(x - t_k)M_k(x|a - 1, t) + (t_{k+a} - x)M_{k+1}(x|a - 1, t)\}}{(k - 1)(t_{k+a} - t_k)}.$$

Mazroui et al. (2012) proposed a general frailty model to identify the correlation between the recurrent events and the correlation between the terminal event and the recurrent events.

$$\begin{aligned} h_T(t_i|u_i) &= u_i h_{0T}(t_i) \exp(\boldsymbol{\beta}'_T \mathbf{x}_{i,T}) \\ h_R(t_{ij}|u_i, v_i) &= u_i v_i h_{0R}(t_{ij}) \exp(\boldsymbol{\beta}'_R \mathbf{x}_{i,R}), \end{aligned} \quad (2.2)$$

where u_i accounts for the correlation between the recurrent events and the terminal event. v_i accounts for the correlation between the recurrent events. Yu et al. (2014) include the time-varying coefficients into the model by

$$\begin{aligned} h_T(t_i|u_i) &= h_{0T}(t_i) \exp(Z_i \alpha(t) + \boldsymbol{\beta}'_T \mathbf{x}_{i,T} + \omega_i) \\ h_R(t_{ij}|u_i) &= h_{0R}(t_{ij}) \exp(Z_i \beta(t) + \boldsymbol{\beta}'_R \mathbf{x}_{i,R} + \omega_i), \end{aligned} \quad (2.3)$$

where $\alpha(t)$ and $\beta(t)$ are time-dependent coefficients and ω_i follows a normal distribution, $\mathcal{N}(0, \sigma_\omega^2)$. Che and Angus (2016) proposed a additive hazard functions for the terminal events,

$$h_T(t|u_i) = h_0(t) + \boldsymbol{\beta}'_T \mathbf{x}_{i,T} + \psi u_i,$$

which further relaxed the hazard model assumption.

2.4 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 $\Pr(T \geq t, R \geq r)$

(Georges et al., 2001). Let $S_T(\cdot)$ and $S_R(\cdot)$ denote the marginal survival functions for T and R , respectively. A copula function $C(\cdot, \cdot)$ connects survival functions $S_T(\cdot)$ and $S_R(\cdot)$, such that $\Pr(T \geq t, R \geq r) = C(S_T(t), S_R(r))$ is a survival copula (Nelsen, 1999).

One important class of copulas called the Archimedean copulas is widely used and includes copulas of the form,

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

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:

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

2. 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.

In the survival framework, copulas are often used to study a bivariate time-to-event process, (Shih and Louis, 1995; Fu et al., 2013; Cook et al., 2010). When the relevant events are non-competitive, Shih and Louis (1995) proposed to use a two-stage approach to estimate the parameters in the model. We first estimate marginal survival functions and after that, we plug in the estimated parameters in the model to estimate θ . When two events are competitive or semi-competitive, we need to estimate all the parameters (including θ) in the model by a maximum likelihood approach.

2.5 Parameter Estimation

2.5.1 Monto Carlo EM(MCEM) Algorithm

In order to estimate the parameters in the model, one of the estimation approach is the MCEM algorithm, proposed by Liu et al. (2004). We treat the frailty as the unobserved data. The complete log likelihood can be expressed by

$$\begin{aligned} \ell(\Theta|\mathbf{u}, \mathcal{D}_n) &= \sum_{i=1}^n \left\{ \log h_T^{d_i}(y_i|u_i) S_T(y_i|u_i) + \log f_u(u_i) \right. \\ &\quad \left. + \sum_{j=1}^{n_i} \left[\log h_R^{\delta_{ij}}(y_{ij}|u_i) S_R(y_{ij}|u_i) \right] \right\} \\ &= \sum_{i=1}^n d_i \{ \beta_T \mathbf{x}_{i,T} + \log h_{0T}(y_i) + \psi u_i \} - \sum_{i=1}^n H_{0T}(y_i|u_i) \\ &\quad + \sum_{i=1}^n \sum_{j=1}^{n_i} \{ \delta_{ij} [\beta_R \mathbf{x}_{i,R} + \log h_{0R}(y_{ij}) + u_i] - H_{0R}(y_{ij}|u_i) \} \end{aligned}$$

In the E step, we calculate $Q(\Theta, \Theta^{old}) = E_u \{ \ell(\Theta|\mathbf{u}, \mathcal{D}_n) \}$ with respect to the posterior distribution of $\mathbf{u}|\Theta^{old}, \mathcal{D}_n$, which can be expressed by

$$\begin{aligned} Q(\Theta, \Theta^{old}) &= \sum_{i=1}^n d_i \{ \beta_T \mathbf{x}_{i,T} + \log h_{0T}(y_i) + \psi E_u(u_i) \} - \sum_{i=1}^n E_u \{ H_T(y_i|u_i) \} \\ &\quad + \sum_{i=1}^n \sum_{j=1}^{n_i} \{ \delta_{ij} [\beta_R \mathbf{x}_{i,R} + \log h_{0R}(y_{ij}) + E_u(u_i)] - E_u(H_R(y_{ij}|u_i)) \} \end{aligned}$$

The challenge here is how to calculate the expectation, $E_u(u_i)$, $E_u\{H_R(y_{ij}|u_i)\}$ and $E_u\{H_T(y_i|u_i)\}$. The Metropolis-Hastings algorithm (Hastings, 1970) is used to sample \mathbf{u} from the posterior distribution of $\mathbf{u}|\Theta^{old}, \mathcal{D}_n$, with

$$\begin{aligned} \Pr(\mathbf{u}|\Theta^{old}, \mathcal{D}_n) &\propto \Pr(\mathcal{D}_n|\Theta^{old}, \mathbf{u}) \Pr(\mathbf{u}) \\ &\propto \mathcal{L}(\Theta^{old}|\mathcal{D}_n, \mathbf{u}) \end{aligned}$$

In the M step, we update Θ by maximizing $Q(\Theta, \Theta^{old})$, which can be completed

by a Newton-Raphson algorithm,

$$\Theta^{new} = \Theta^{old} - \mathbf{V}(\Theta) \partial Q(\Theta, \Theta^{old}) / \partial \Theta |_{\Theta = \Theta^{old}}.$$

$\mathbf{V}(\Theta^{old})$ is the inverse of $\partial^2 Q(\Theta, \Theta^{old}) / \partial \Theta \partial \Theta' |_{\Theta = \Theta^{old}}$.

We iteratively update Θ between the E step and M step until the algorithm converges, $\|\Theta^{new} - \Theta^{old}\| < 0.001$, where $\|\cdot\|$ is the Euclidean distance.

Under this method, the observed information matrix of the estimated $\hat{\Theta}$ is derived by Louis's formula (Louis, 1982),

$$\begin{aligned} \mathcal{I}(\hat{\Theta}) = & -\hat{E}_u \left\{ \frac{\partial^2 \ell(\Theta | \mathcal{D}_n, \mathbf{u})}{\partial \Theta \partial \Theta'} \right\}_{\Theta = \hat{\Theta}} - \hat{E}_u \left\{ \frac{\partial \ell(\Theta | \mathcal{D}_n, \mathbf{u})}{\partial \Theta} \frac{\partial \ell(\Theta | \mathcal{D}_n, \mathbf{u})}{\partial \Theta'} \right\}_{\Theta = \hat{\Theta}} \\ & + \hat{E}_u \left\{ \frac{\partial \ell(\Theta | \mathcal{D}_n, \mathbf{u})}{\partial \Theta} \right\}_{\Theta = \hat{\Theta}} \hat{E}_u \left\{ \frac{\partial \ell(\Theta | \mathcal{D}_n, \mathbf{u})}{\partial \Theta'} \right\}_{\Theta = \hat{\Theta}}. \end{aligned}$$

2.5.2 Penalized Likelihood Approach

Although the MCEM algorithm can be applied to estimate Θ , the computation burden is high, especially when n is large because in every E step, we need to sample posterior distribution of \mathbf{u} and calculate the posterior mean. Another approach is to directly maximize the marginal log likelihood. The marginal log likelihood, denoted by $\ell(\Theta)$ is

$$\ell(\Theta) = \log \left[\prod_{i=1}^n \int_0^\infty h_T^{d_i}(y_i | u_i) S_T(y_i | u_i) \left\{ \prod_{i=1}^{n_i+1} h_R^{\delta_{ij}}(y_{ij} | u_i) S_R(y_{ij} | u_i) \right\} f_u(u_i) du_i \right].$$

Since there is no closed form for the integration, Laplace approximation (Laplace, 1986) and Gaussian quadrature (Anderson, 1965) could be used to approximate $\ell(\Theta)$.

Rondeau et al. (2007) proposed a penalized likelihood approach to estimate Θ . The penalized likelihood under this approach can be expressed by,

$$\tilde{\ell}(\Theta) = \ell(\Theta) + \kappa_1 \int_0^\infty h_{0T}''(t)^2 dt + \kappa_2 \int_0^\infty h_{0R}''(r)^2 dt,$$

where $h_{0T}''(t)$ and $h_{0R}''(r)$ are the second derivative of the piece-wise baseline hazard functions; $\kappa_1 \geq 0$ and $\kappa_2 \geq 0$ are the smoothing parameters, chosen to minimize

the cross validation error. We could use an improved Newton-Raphson algorithm, proposed by Marquardt (1963), to optimize $\tilde{\ell}(\Theta)$.

Compared with the MCEM algorithm, the computing burden of the penalized likelihood approach is smaller. However, how to choose appropriate knots for splines is still a problem. If the number of knots is high, the algorithm might not converge or take a long time for convergence.

A Joint Frailty-copula Model for Recurrent Events and a Terminal Event

3.1 Introduction

In clinical trials and epidemiology studies, the primary events we are interested in might be recurrent events (Jansen et al., 2018; Ridker et al., 2000). Under the survival framework, the challenge of analyzing this kind of recurrent events is not only to account for the within-subject correlation as we do in a longitudinal study, but also to consider the potential dependent censoring mechanism because of a terminal event. The terminal event such as death, stops the recurrent event process and often is correlated with the recurrent events, leading to the violation of a non-informative censoring assumption. Moreover, the correlation between the terminal event and recurrent events might change over time, which renders it even more difficult when we attempt to recurrent events data.

In terms of joint modeling recurrent events and a terminal event, Huang and Wang (2004) proposed a joint frailty model. They assumed that conditional on a subject-level frailty, the terminal event process and the recurrent event process are independent. Huang and Liu (2007) proposed to extend the joint frailty model to the gap time scale. Rondeau et al. (2007) proposed to use a penalized likelihood

approach to estimate the joint frailty model. More recently, Yu et al. (2014) considered time-varying coefficients in the hazard functions, which make the model even more flexible. Che and Angus (2016), on the other hand, proposed the joint frailty model under an additive hazard function for the terminal event. Other important work can be found in the references (Zeng and Lin, 2009; Cook and Lawless, 1997; Zheng and Klein, 1995).

All the aforementioned articles assumed 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. 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 research works 1) treat the correlation estimate as a nuisance parameter, or 2) have a vague interpretation of the dependence with only the association direction but without estimating their correlation in a straightforward manner. In some situations, 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) 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.

The main purpose of this chapter is to develop a 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, reduction of computational load, and easy implementation in statistical software. Importantly, the robustness of our proposal is 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 chapter, 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.

3.2 Methods

3.2.1 Notation

Let T_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 $d_i = I(T_i < C_i)$ to be the failure status for the terminal event, $Y_i = \min\{T_i, C_i\}$ to be the observed follow-up time, and ξ is the failure rate. Let R_{ij} represent the gap time between the $(j-1)^{\text{th}}$ 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 T_i and the recurrent event time R_{ij} of the i^{th} subject are expressed by

$$\lambda_T(t|\mathbf{x}_{i,T}, \omega_i) = \lambda_{0T}(t) \exp\{\boldsymbol{\beta}_T^T \mathbf{x}_{i,T} + \omega_i\} \quad (3.1)$$

$$\lambda_R(r|\mathbf{x}_{i,R}, \omega_i) = \lambda_{0R}(r) \exp\{\boldsymbol{\beta}_R^T \mathbf{x}_{i,R} + \omega_i\} \quad (3.2)$$

where $\mathbf{x}_{i,T}$ and $\mathbf{x}_{i,R}$ are time-invariant or time-dependent covariate vectors. ω_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 $\mathbf{x}_{i,T}$ and $\mathbf{x}_{i,R}$ in all the functions from now on, and hazard functions denoted respectively by, $\lambda_T(t|\boldsymbol{\omega}_i)$ and $\lambda_R(r|\boldsymbol{\omega}_i)$. The frailty term is assumed that $\omega_i \sim N(0, \sigma_w^2)$, but the other parametric distributions such as gamma can still be applicable. The parameters $\boldsymbol{\beta}_T$ and $\boldsymbol{\beta}_R$ quantify the effect of the covariates $\mathbf{x}_{i,T}$ on T_i and the effects of $\mathbf{x}_{i,R}$ on R_{ij} respectively. Denote the observation $\mathcal{Y}_i = \{y_i, d_i, \delta_{ij}, r_{ij}, \mathbf{x}_{i,T}, \mathbf{x}_{i,R}, j = 1, \dots, n_i + 1\}$ for the i^{th} subject, and let \mathcal{D}_n denote the observed data from n subjects, $\mathcal{D}_n = \{\mathcal{Y}_1, \dots, \mathcal{Y}_n\}$.

3.2.2 Joint Frailty-copula Model (JFCM)

The regular joint frailty model (JFM) assumes that $T_i, R_{i1}, \dots, R_{i, n_i+1}$ are mutually independent conditional on the subject-level frailty ω_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. We relax the assumption of conditional independence of ω_i by inducing a copula to account for the within-subject correlation.

Emura et al. (2015) proposed a joint frailty-copula model. Under this model, the joint survival function of the j^{th} recurrent event time and the terminal event time of the i^{th} patient,

$$\Pr(T_i \geq t_{ij}, R_{ij} \geq r_{ij} | \omega_i, T_i \geq t_{i, j-1}) = C_\theta(S_T(r_{ij} | \omega_i), S_R(r_{ij} | \omega_i)), i = 1, \dots, n,$$

This model implicitly implies $S_T(t_{ij} | t_{i, j-1}) = S_T(r_{ij})$ and the terminal event process is a renewal process. Our methods will relax this renewal assumption.

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 number 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 $R_{ij} = r_{ij}$ for $j = 1, \dots, n_i$ and $R_{ij} \geq c_{ij}$ for $j = n_i + 1, \dots, J + 1$. Assume that the residual dependence is homogeneous across all subjects and unchanged along the time line.

Conditional on ω_i , the joint probability of $T_i \geq t_i$ and $R_{ij} \geq r_{ij}$ is

$$\Pr(T_i \geq t_i, R_{ij} \geq r_{ij} | \omega_i) = C(S_T(t_i | \omega_i), S_R(r_{ij} | \omega_i)),$$

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

Based on the survival copula, we have,

$$\begin{aligned}
\Pr(T_i \geq t_i, R_{ij} \geq r_{ij} | \omega_i) &= C(S_T(t_i | \omega_i), S_R(r_{ij} | \omega_i)) \\
\Pr(T_i \geq t_i, R_{ij} = r_{ij} | \omega_i) &= C_{(01)}^*(S_T(t_i | \omega_i), S_R(r_{ij} | \omega_i)) f_R(r_{ij} | \omega_i) \\
\Pr(T_i = t_i, R_{ij} = r_{ij} | \omega_i) &= C_{(11)}^*(S_T(t_i | \omega_i), S_R(r_{ij} | \omega_i)) f_T(t_i | \omega_i) f_R(r_{ij} | \omega_i) \\
\Pr(T_i = t_i, R_{ij} \geq r_{ij} | \omega_i) &= C_{(10)}^*(S_T(t_i | \omega_i), S_R(r_{ij} | \omega_i)) f_T(t_i | \omega_i)
\end{aligned} \tag{3.3}$$

where $C_{(01)}(u, v) = \partial C(u, v) / \partial v$, and $C_{(10)}(u, v) = \partial C(u, v) / \partial u$, $C_{(11)}(u, v) = \partial^2 C(u, v) / (\partial u \partial v)$. For example, if $C(\cdot, \cdot)$ is a Clayton copula, then we have

$$\begin{aligned}
C_{(01)}(u, v) &= v^{-\theta-1} (u^{-\theta} + v^{-\theta} - 1)^{-\frac{\theta+1}{\theta}} \\
C_{(10)}(u, v) &= u^{-\theta-1} (u^{-\theta} + v^{-\theta} - 1)^{-\frac{\theta+1}{\theta}} \\
C_{(11)}(u, v) &= (\theta + 1) (uv)^{-(\theta+1)} (u^{-\theta} + v^{-\theta} - 1)^{-\frac{2\theta+1}{\theta}}
\end{aligned} \tag{3.4}$$

Note that for a subject with $d_i = 0$, 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 ω_i and $T_i \geq t_i$, the probability of the i^{th} subject who survives through t_i and experiences n_i events given ω_i is,

$$\begin{aligned}
&\Pr(T_i \geq t_i, R_{i1} = r_{i1}, \dots, R_{i,n_i} = r_{i,n_i}, R_{i,n_i+1} \geq c_{i,n_i+1}, \dots, R_{i,J+1} \geq c_{i,J+1} | \omega_i) \\
&= \Pr(R_{i1} = r_{i1}, \dots, R_{i,n_i} = r_{i,n_i}, R_{i,n_i+1} \geq c_{i,n_i+1}, \dots, R_{i,J+1} \geq c_{i,J+1} | T_i \geq t_i, \omega_i) \\
&\quad \times \Pr(T_i \geq t_i | \omega_i) \\
&= \Pr(T_i \geq t_i | \omega_i) \prod_{j=1}^{n_i} \Pr(R_{ij} = r_{ij} | T_i \geq t_i, \omega_i) \prod_{j=n_i+1}^J \Pr(R_{ij} \geq c_{ij} | T_i \geq t_i, \omega_i)
\end{aligned} \tag{3.5}$$

By Bayes' rule,

$$\Pr(R_{ij} = r_{ij} | T_i \geq t_i, \omega_i) = \frac{C_{(01)}^*(S_T(t_i | \omega_i), S_R(r_{ij} | \omega_i)) f_R(r_{ij} | \omega_i)}{S_T(t_i | \omega_i)}$$

$$\Pr(R_{i,n_i+1} \geq c_{i,n_i+1} | T_i \geq t_i, \omega_i) = \frac{C(S_T(t_i | \omega_i), S_R(c_{i,n_i+1} | \omega_i))}{S_T(t_i | \omega_i)}$$

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

$$\Pr(R_{ij} \geq c_{ij} | T_i \geq t_i, \omega_i) = \Pr(R_{ij} \geq 0 | T_i \geq t_i, \omega_i) = 1.$$

Then,

$$\begin{aligned} & \Pr(T_i \geq t_i, R_{i1} = r_{i1}, \dots, R_{i, n_i+1} \geq c_{i, n_i+1}, \dots, R_{i, J+1} \geq c_{i, J+1} | \omega_i) \\ &= S_T(t_i | \omega_i) \frac{C(S_T(t_i | \omega_i), S_R(c_{i, n_i+1} | \omega_i))}{S_T(t_i | \omega_i)} \prod_{j=1}^{n_i} \frac{C_{(01)}^*(S_T(t_i | \omega_i), S_R(r_{ij} | \omega_i)) f_R(r_{ij} | \omega_i)}{S_T(t_i | \omega_i)} \\ &= C(S_T(t_i | \omega_i), S_R(c_{i, n_i+1} | \omega_i)) \prod_{j=1}^{n_i} \frac{C_{(01)}^*(S_T(t_i | \omega_i), S_R(r_{ij} | \omega_i)) f_R(r_{ij} | \omega_i)}{S_T(t_i | \omega_i)} \end{aligned} \quad (3.6)$$

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.

$$\begin{aligned} L(\mathcal{D}_n | \mathbf{w}) &= \prod_{i=1}^n [f_T(y_i | \omega_i) C_{10}^*(S_T(y_i | \omega_i), S_R(r_{i, n_i+1} | \omega_i))]^{d_i} \\ &\quad \times [S_T^{-n_i}(y_i | \omega_i) C(S_T(y_i | \omega_i), S_R(r_{i, n_i+1} | \omega_i))]^{1-d_i} \\ &\quad \times \prod_{j=1}^{n_i} [C_{(01)}(S_T(y_i | \omega_i), S_R(r_{ij} | \omega_i))]^{1-d_i} \\ &\quad \times [C_{(11)}(S_T(y_i | \omega_i), S_R(r_{ij} | \omega_i))]^{d_i} f_R(r_{ij} | \omega_i) \end{aligned} \quad (3.7)$$

3.2.3 Metropolis-Hastings within the Gibbs Sampler Algorithm

We utilize the Bayesian approach for parameter estimation and inference. For simplicity, constant baseline intensity functions $\lambda_{0T}(t) = \exp(\beta_{0T}t)$ and $\lambda_{0R}(t) = \exp(\beta_{0R}t)$ are considered; however, this can be extended in a straightforward manner to the setting with non-constant baseline hazards. Let Θ denote the parameter vector, $\{\beta_{0R}, \beta_{0T}, \boldsymbol{\beta}_T, \boldsymbol{\beta}_R, \sigma_w^{-2}, \mathbf{w}, \theta\}$ 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

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

- 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

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

4. 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.

3.2.4 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. 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 denoted by

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

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,

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

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 $\Pr(w_N | T_N > t', \mathcal{H}_N(t'), \mathcal{D}_n, \Theta, \mathcal{D}_n)$. We can sample $\Pr(w_N | T_N > t', \mathcal{H}_N(t'), \Theta, \mathcal{D}_n) \propto \Pr(T_N > t', \mathcal{H}_N(t') | \mathcal{D}_n, \Theta, w_N) \Pr(\Theta | \mathcal{D}_n) \Pr(w_N)$ by Metropolis Hastings within Gibbs Sampler algorithm. When t' increases, the probability, $\hat{\pi}(t|t', \mathcal{H}_N(t'), \mathcal{D}_n, \hat{\Theta}, \hat{w}_N)$ can be updated dynamically by re-sampling w_N from $\Pr(w_N | T_N > t', \mathcal{H}_N(t'), \hat{\Theta}, \mathcal{D}_n)$.

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 risks 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

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

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

3.3 Simulation Study

3.3.1 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 $f_T(t) = \lambda_{i,T} \exp(-\lambda_{i,T}t)$, 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 $\lambda_{i,T}$ and $\lambda_{ij,R}$ are given respectively,

$$\begin{aligned} \lambda_{i,T} &= \exp(\beta_{0T} + \omega_i + \beta_{T,1}x_{1,i} + \beta_{T,2}x_{2,i}) \\ \lambda_{ij,R} &= \exp(\beta_{0R} + \omega_i + \beta_{R,1}x_{1,i} + \beta_{R,2}x_{2,i}), \end{aligned} \tag{3.8}$$

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, $Bern(0.5)$. $(\beta_{T,1}, \beta_{T,2}, \beta_{R,1}, \beta_{R,2})^T$ is set to be $(1, 1, 2, 2)^T$. $(\beta_{0T}, \beta_{0R})^T$ is set to be $(0.5, 1)^T$.

For each subject i , we first generate U_i from $Unif(0, 1)$. T_i is generated by $-\log(U_i) / \lambda_{i,T}$. Assuming independent censoring for T_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 ($\xi = 60\%$), a medium failure rate ($\xi = 50\%$) and a low failure rate ($\xi = 40\%$). The observed follow-up time $Y_i = \min\{T_i, C_i\}$. Then, considering the j^{th} recurrent event, suppose $S_R(r_{ij} | \omega_i) = V_{ij}$. 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 ($\tau = 0$). Suppose sufficiently large maximum number $J = 1000$ of events occur for each subject. When $T_i \leq C_i$, the conditional distribution of

V_{ij} is $F_V(v|U_i = u_i) = u_i^{-(\theta+1)}(u_i^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$. 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 let $\tilde{W}_{ij} = F_V(v|U_i = u_i)$. 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 $T_i > C_i$, 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} simply 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 the sample size $n = 100, 200$.

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.

In addition, we evaluate the robustness of our method considering the misspecification 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, $\beta_{0T} \propto 1, \beta_{0R} \propto 1, \beta_T \propto \mathbf{1}, \beta_R \propto \mathbf{1}$. Also, σ_w^{-2} are assumed to follow a prior $gamma(\alpha, 1/\alpha)$ 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

posterior mean estimates denoted by EST, standard deviation (SD) of posterior mean estimates, mean squared error (MSE) and the average of absolute bias (AB) of posterior mean estimates.

3.3.2 Simulation Results

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. In Table 3.1, 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 $\beta_{T,1}$ and $\beta_{T,2}$, 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 T_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 performance of our method is still identical in terms of the bias, SD, MSE and AB, compared with the joint frailty model. For example, the average of $\hat{\beta}_{T,1}$ from 250 replicates is 0.981 under our method with SD and MSE respectively 0.234 and 0.055. By contrast, the average $\hat{\beta}_{T,1}$ under joint frailty model is 0.989 with SD and MSE respectively 0.223 and 0.050.

In Table 3.2, 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 of $\hat{\beta}_{R,2}$ is 1.996 under high failure rate, 1.995 under medium failure rate and 1.990 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 3.1, 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 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 Table 3.3, 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 $\beta_{T,1}$, $\beta_{T,2}$, $\beta_{R,1}$ and $\beta_{R,2}$, the parameters associated with $S_T(t|\omega_i)$ and $S_R(r|\omega_i)$ 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 3.1 with the results of Table 3.3, the biases of $\hat{\beta}_{T,1}$, $\hat{\beta}_{T,2}$, $\hat{\beta}_{R,1}$, $\hat{\beta}_{R,2}$ in Table 3.3 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.

3.4 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. The episodes of acute ischemic stroke are diagnosed by the ICD9-CM codes with 434.x and 436.x. 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 in-

cluding diabetes mellitus, cardiovascular disease and heart attack during the past twelve-month. Recently, considerable attention has been given to the association of baseline stroke and co-morbidities with the recurrence of acute ischemic stroke, and substantial literature have shown the higher likelihood of mortality in hospitalized stroke patients (Arabadzhieva 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 death, and also evaluation of the correlation between two event processes of recurrent strokes and death.

A sample of 2122 patients are identified and among them, 597 (28.13%) patients experience at least one stroke event and the death rate is 6.55%. Note that the low mortality rate is a limitation for our study because only hospitalized deaths are recorded for analysis. 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. Preliminary analysis 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 3.2. We fit our proposed model with Clayton copula and compare with the traditional joint frailty model. The results are presented in Table 3.4, including the estimates of log hazard ratios (HR) and 95% credible limits (CL).

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 events (HR=1.46, 95% CI: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 $\tau = 0.32$) 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 results differences between 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.

3.5 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. The algorithm for Bayesian inference is easy to be implemented in statistical software which is accessible upon request, and also informative priors can be incorporated for efficiency improvement if available. 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 identical bias, SD and MSE.

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.

Table 3.1. Summary statistics on joint modeling of recurrent events and a terminal event by a Clayton Copula under the scenario with high failure Rate ($\xi = 60\%$).

n	τ	Param	JFCM				JFM			
			EST	SD	MSE	AB	EST	SD	MSE	AB
100	0.0	$\beta_{T,1} = 1$	1.010	0.312	0.097	0.252	1.020	0.294	0.086	0.234
		$\beta_{T,2} = 1$	0.988	0.145	0.021	0.118	0.997	0.138	0.019	0.111
		$\beta_{R,1} = 2$	1.980	0.209	0.044	0.163	1.994	0.197	0.039	0.158
		$\beta_{R,2} = 2$	1.990	0.116	0.013	0.088	1.995	0.113	0.013	0.089
		τ	0.023	0.031	0.002	0.036	-	-	-	-
	0.3	$\beta_{T,1} = 1$	0.999	0.264	0.070	0.207	1.113	0.291	0.097	0.249
		$\beta_{T,2} = 1$	0.994	0.133	0.018	0.106	1.083	0.143	0.027	0.135
		$\beta_{R,1} = 2$	1.991	0.222	0.049	0.170	2.093	0.224	0.059	0.194
		$\beta_{R,2} = 2$	1.994	0.118	0.014	0.090	2.088	0.121	0.022	0.122
		τ	0.329	0.036	0.001	0.029	-	-	-	-
	0.5	$\beta_{T,1} = 1$	0.994	0.224	0.050	0.181	1.152	0.299	0.112	0.269
		$\beta_{T,2} = 1$	0.982	0.120	0.015	0.097	1.116	0.148	0.035	0.154
		$\beta_{R,1} = 2$	1.984	0.213	0.045	0.167	2.143	0.242	0.079	0.226
		$\beta_{R,2} = 2$	1.982	0.116	0.014	0.092	2.140	0.136	0.038	0.158
		τ	0.497	0.023	0.001	0.019	-	-	-	-
200	0.0	$\beta_{T,1} = 1$	0.981	0.234	0.055	0.190	0.989	0.223	0.050	0.181
		$\beta_{T,2} = 1$	1.000	0.114	0.013	0.088	1.008	0.110	0.012	0.087
		$\beta_{R,1} = 2$	1.987	0.148	0.022	0.118	1.987	0.146	0.021	0.117
		$\beta_{R,2} = 2$	2.007	0.084	0.007	0.065	2.002	0.071	0.005	0.056
		τ	0.015	0.020	0.001	0.037	-	-	-	-
	0.3	$\beta_{T,1} = 1$	0.992	0.190	0.036	0.154	1.079	0.227	0.057	0.192
		$\beta_{T,2} = 1$	1.004	0.104	0.011	0.083	1.095	0.112	0.021	0.119
		$\beta_{R,1} = 2$	1.991	0.152	0.023	0.121	2.079	0.163	0.033	0.141
		$\beta_{R,2} = 2$	2.006	0.089	0.008	0.072	2.092	0.088	0.016	0.104
		τ	0.329	0.029	0.001	0.023	-	-	-	-
	0.5	$\beta_{T,1} = 1$	1.002	0.160	0.026	0.127	1.119	0.232	0.068	0.212
		$\beta_{T,2} = 1$	0.998	0.088	0.008	0.069	1.133	0.115	0.031	0.147
		$\beta_{R,1} = 2$	1.997	0.149	0.022	0.119	2.132	0.180	0.050	0.183
		$\beta_{R,2} = 2$	1.996	0.085	0.007	0.068	2.147	0.098	0.031	0.154
		τ	0.500	0.023	0.001	0.019	-	-	-	-

EST, average 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 $\beta_{T,1}$ as an example where $n_{sim} = 250$, $EST = \frac{1}{n_{sim}} \sum \hat{\beta}_{1,i}$, $SD = \sqrt{\frac{1}{n_{sim}-1} \sum (\hat{\beta}_{1,i} - \bar{\hat{\beta}}_1)^2}$, $MSE = \frac{1}{n_{sim}} \sum (\hat{\beta}_{1,i} - \beta_{T,1})^2$, $AB = \frac{1}{n_{sim}} \sum |\hat{\beta}_{1,i} - \beta_{T,1}|$, 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 3.2. Summary statistics on joint modeling of recurrent events and a terminal event by 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 ($\xi = 40\%$)			
			EST	SD	MSE	AB	EST	SD	MSE	AB
100	0.3	$\beta_{T,1} = 1$	1.012	0.289	0.084	0.232	1.002	0.311	0.096	0.241
		$\beta_{T,2} = 1$	0.996	0.143	0.020	0.115	0.994	0.162	0.026	0.129
		$\beta_{R,1} = 2$	1.967	0.228	0.053	0.186	1.975	0.238	0.057	0.182
		$\beta_{R,2} = 2$	1.986	0.129	0.017	0.104	1.979	0.148	0.022	0.118
		τ	0.328	0.039	0.002	0.032	0.332	0.039	0.002	0.031
	0.5	$\beta_{T,1} = 1$	0.999	0.257	0.066	0.203	1.001	0.276	0.076	0.218
		$\beta_{T,2} = 1$	0.995	0.128	0.016	0.104	0.985	0.139	0.019	0.114
		$\beta_{R,1} = 2$	1.980	0.233	0.054	0.183	1.988	0.234	0.055	0.183
		$\beta_{R,2} = 2$	1.991	0.124	0.015	0.100	1.986	0.133	0.018	0.106
		τ	0.495	0.026	0.001	0.021	0.499	0.025	0.001	0.020
200	0.3	$\beta_{T,1} = 1$	0.970	0.195	0.039	0.160	0.954	0.226	0.053	0.186
		$\beta_{T,2} = 1$	0.995	0.107	0.011	0.084	0.992	0.117	0.014	0.090
		$\beta_{R,1} = 2$	1.984	0.162	0.026	0.132	1.968	0.172	0.031	0.139
		$\beta_{R,2} = 2$	1.995	0.090	0.008	0.070	1.987	0.094	0.009	0.075
		τ	0.331	0.032	0.001	0.025	0.337	0.033	0.001	0.027
	0.5	$\beta_{T,1} = 1$	0.992	0.171	0.029	0.135	0.979	0.198	0.039	0.155
		$\beta_{T,2} = 1$	0.996	0.092	0.008	0.073	0.990	0.105	0.011	0.084
		$\beta_{R,1} = 2$	1.991	0.155	0.024	0.125	1.983	0.163	0.027	0.129
		$\beta_{R,2} = 2$	1.995	0.088	0.008	0.071	1.990	0.097	0.009	0.076
		τ	0.498	0.023	0.001	0.018	0.501	0.024	0.001	0.020

EST, average 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 $\beta_{T,1}$ as an example where $n_{sim} = 250$, $EST = \frac{1}{n_{sim}} \sum \hat{\beta}_{1,i}$,

$SD = \sqrt{\frac{1}{n_{sim}-1} \sum (\hat{\beta}_{1,i} - \bar{\hat{\beta}}_1)^2}$, $MSE = \frac{1}{n_{sim}} \sum (\hat{\beta}_{1,i} - \beta_{T,1})^2$, $AB = \frac{1}{n_{sim}} \sum |\hat{\beta}_{1,i} - \beta_{T,1}|$, where n_{sim} is the number of the replicates and $\hat{\beta}_{1,i}$ is the posterior mean estimate of the i^{th} replicate.

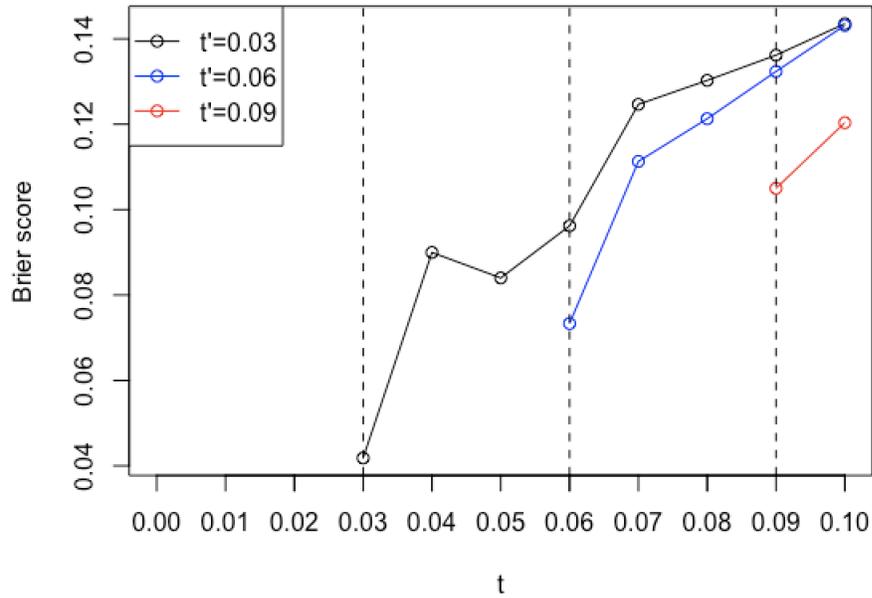


Figure 3.1. The Brier score (BS) vs the time point t

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 $t' = 0.09$.

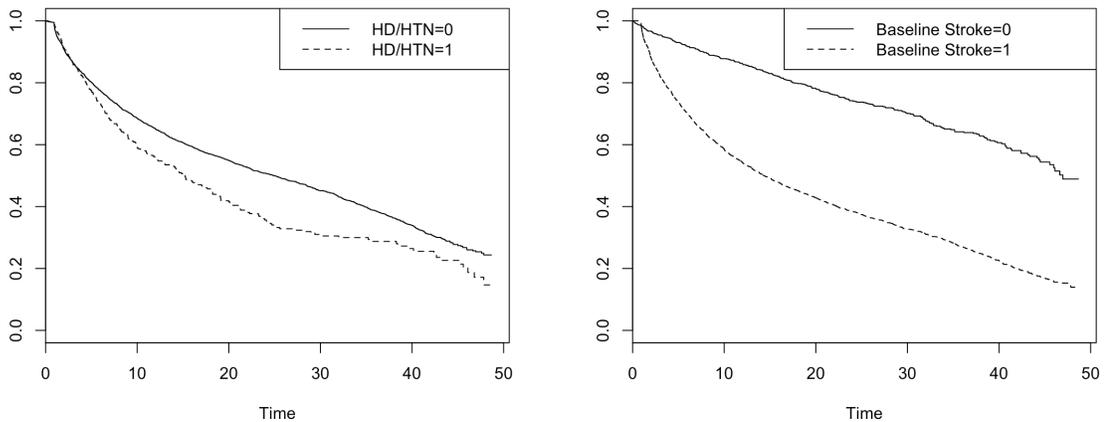


Figure 3.2. Preliminary analysis on Kaplan-Meier Curves for stroke

Table 3.3. 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=200			
		EST	SD	MSE	AB	EST	SD	MSE	AB
0.11	$\beta_{T,1} = 1$	1.004	0.320	0.102	0.255	0.975	0.239	0.057	0.197
	$\beta_{T,2} = 1$	0.979	0.159	0.026	0.126	0.992	0.121	0.015	0.097
	$\beta_{R,1} = 2$	1.995	0.227	0.051	0.182	2.010	0.164	0.027	0.132
	$\beta_{R,2} = 2$	1.992	0.121	0.015	0.094	2.006	0.088	0.008	0.070
	τ	0.083	0.055	0.004	0.052	0.073	0.042	0.003	0.045
0.21	$\beta_{T,1} = 1$	0.971	0.319	0.102	0.253	0.950	0.256	0.068	0.210
	$\beta_{T,2} = 1$	0.966	0.169	0.030	0.137	0.973	0.134	0.019	0.108
	$\beta_{R,1} = 2$	1.984	0.228	0.052	0.176	1.994	0.155	0.024	0.126
	$\beta_{R,2} = 2$	1.988	0.124	0.016	0.098	2.003	0.088	0.008	0.070
	τ	0.163	0.057	0.006	0.061	0.154	0.042	0.005	0.061

EST, average 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 $\beta_{T,1}$ as an example where $n_{sim} = 250$, $EST = \frac{1}{n_{sim}} \sum \hat{\beta}_{1,i}$, $SD = \sqrt{\frac{1}{n_{sim}-1} \sum (\hat{\beta}_{1,i} - \hat{\beta}_1)^2}$, $MSE = \frac{1}{n_{sim}} \sum (\hat{\beta}_{1,i} - \beta_{T,1})^2$, $AB = \frac{1}{n_{sim}} \sum |\hat{\beta}_{1,i} - \beta_{T,1}|$, 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 3.4. Summary of real data analysis on recurrent stroke

	Covariates	JFCM			JFM		
		EST	2.50%	97.50%	EST	2.50%	97.50%
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
	Female	-0.603	-0.852	-0.249	-0.152	-0.547	0.242
	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
	Female	-0.359	-0.619	-0.258	-0.245	-0.470	-0.021
	Stroke	1.647	1.462	1.862	2.318	2.065	2.572
Correlation	τ	0.320	0.049	0.547			

A Time-varying Joint Frailty-copula Approach for Recurrent Events and a Terminal Event

4.1 Introduction

The current joint frailty models assume that the association between the terminal event time and the recurrent event time is constant. The assumption of a constant frailty over time under the gamma frailty model yields a closed-form expression for Kendall's τ (Oakes, 1989). In reality, the terminal event might be more correlated with the most recent recurrent event compared with the earlier recurrent event. For example, the death might be more associated with the previous myocardial infarction event compared with the first myocardial infarction event.

Copulas often are used to model time-varying correlation in the financial market (Wang et al., 2011; Almeida and Czado, 2012; Aloui et al., 2013; Li and Zeng, 2018). In these models, the correlation is modeled via time series, e.g., AR(1) process. In the survival framework, copulas also have been adopted to study recurrent events (Cook et al., 2010; Meyer and Romeo, 2015). Typically, a multi-dimensional copula is used to connect marginal survival functions of the recurrent events. Emura et al. (2015) proposed a joint frailty-copula model for meta-analysis and further as a special case, extended it to jointly model recurrent events and the terminal event.

But it only would be valid when the terminal event process is a renewal process and the correlation is still a constant.

In this chapter, we propose a time-varying joint frailty-copula model, which 1) relaxes the conditional independence assumption under the joint frailty model, 2) relaxes the assumption that the terminal event should be a renewal process, 3) allows the correlation between the terminal event and the recurrent events to change over time, and 4) models the correlation on subject level instead of the marginal level. We perform extensive simulation studies, comparing our method with the joint frailty model and the joint frailty-copula model in terms of the bias, standard deviation, mean squared error, and absolute bias. Also, we evaluate our model when the copula model is misspecified.

In the remainder of this chapter, we introduce the details of the time-varying joint frailty-copula models in section 4.2. In section 4.3, we apply the simulation plan and results. In section 4.4, we applied our method to a real data set (CHS data set) to analyze recurrent stroke, recurrent MI and death.

4.2 Methods

4.2.1 Notation

Let i denote the i^{th} subject and let j denote the j^{th} patient. Suppose that each subject experiences a maximum of J recurrent events without the terminal event. The terminal event will stop the recurrent event. Let T_i denote the terminal event time of the i^{th} subject and let R_{ij} denote the recurrent event gap time between the j^{th} and $(j-1)^{th}$ recurrent event. Let C_i denote the censoring time for T_i and $Y_i = \min\{T_i, C_i\}$ be the follow up time for the i^{th} subject. $D_i = I(T_i < C_i)$ is the failing indicator of the subject i . We assume T_i and C_i are independent. Further, suppose T_{ij} is the calendar recurrent event time calculate from the beginning of the study. Then, $C_{ij} = \max\{Y_i - T_{i,j-1}, 0\}$ naturally is the censoring time for R_{ij} . Let $\Delta_{ij} = \min\{R_{ij}, C_{ij}\}$. C_{ij} contains the information about T_i , and in many cases T_i is correlated with R_{ij} . Because the non-informative censoring mechanism is violated, we should not model the R_{ij} 's alone. We need to jointly model T_i and R_{ij} .

4.2.2 Time-varying Joint Frailty-copula Model

Next, we consider relaxing the assumption of constant correlation via a time-varying joint frailty-copula model.

Suppose the hazard functions for the terminal event time T_i and the recurrent event time R_{ij} are

$$\begin{aligned}\log h_T(t_i) &= \log h_{0T}(t_i) + \boldsymbol{\beta}'_T \mathbf{x}_{i,T} + \omega_i \\ \log h_R(r_{ij}) &= \log h_{0R}(r_{ij}) + \boldsymbol{\beta}'_R \mathbf{x}_{i,R} + \omega_i,\end{aligned}\tag{4.1}$$

where $\omega_i \sim \mathcal{N}(0, \sigma_w^2)$. $h_{0T}(\cdot)$ and $h_{0R}(\cdot)$ are baseline hazard functions and can be modeled by piecewise splines with parameter $\boldsymbol{\beta}_{0T}$ and $\boldsymbol{\beta}_{0R}$. The survival functions of T_i and R_{ij} are respectively, $S_T(t_i|\omega_i) = \exp\left\{-\int_0^{t_i} h_T(t|\omega_i)dt\right\}$ and $S_R(r_{ij}|\omega_i) = \exp\left\{-\int_0^{r_{ij}} h_R(r|\omega_i)dr\right\}$.

For the j^{th} recurrent event time and the terminal event time of the i^{th} patient we assume the correlation parameter in the copula function is θ_{ij} . So the joint survival function is

$$\Pr(T_i \geq t_i, R_{ij} \geq r_{ij}|\omega_i) = C_{\theta_{ij}}(S_T(t_i|\omega_i), S_R(r_{ij}|\omega_i)), i = 1, \dots, n, j = 1, \dots, J,$$

where $C_{\theta_{ij}}(\cdot, \cdot)$ is a survival Archimedean copula function, θ_{ij} is the parameter in the copula function to quantify the association between the terminal event and the j^{th} recurrent event time. We relax the assumption in the previous joint frailty-copula model, which assumes $\theta_{11} = \dots = \theta_{nJ} = \theta$. Also, compared with the traditional time-varying copulas, the correlation parameter is modeled on subject level, which means that the correlation between the terminal event and the recurrent event could be different for each subject and each time point.

We assume that the overall correlation (marginal level) between T_i and R_{ij} across time is quantified by θ_μ . And $\theta_{ij} = g(\theta_\mu, \gamma_{ij})$. γ_{ij} is a latent effect which leads to the change of θ_{ij} and $g(\cdot, \cdot)$ is a function chosen to transform θ_μ, γ_{ij} to θ_{ij} . We assume that $\gamma_{i1} \sim \mathcal{N}(0, \sigma_\epsilon^2/(1 - \rho^2))$ and

$$\gamma_{ij} = \rho\gamma_{i,j-1} + \epsilon_{ij}, j \geq 2$$

where $|\rho| < 1$ and $\epsilon_{i1}, \dots, \epsilon_{iJ} \sim_{i.i.d} \mathcal{N}(0, \sigma_\epsilon^2)$. So, marginally, we have

$$\gamma_{i1}, \dots, \gamma_{iJ} \sim \mathcal{N}(0, \sigma_\epsilon^2 / (1 - \rho^2)).$$

Suppose a subject experienced n_i recurrent events prior to death or study withdrawal, and let $\boldsymbol{\gamma}_i = \{\gamma_{i1}, \dots, \gamma_{i, n_i+1}\}$. $\boldsymbol{\gamma}_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_i)$ where

$$\boldsymbol{\Sigma}_i = \frac{\sigma_\epsilon^2}{1 - \rho^2} \begin{bmatrix} 1 & \rho & \dots & \rho^{n_i} \\ \rho & 1 & \dots & \rho^{n_i-1} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{n_i} & \rho^{n_i-1} & \dots & 1 \end{bmatrix}.$$

The overall model structure is plotted in Figure 4.2.2.

Let $\boldsymbol{\gamma} = \{\boldsymbol{\gamma}_1, \dots, \boldsymbol{\gamma}_n\}$, and let $\mathcal{D}_n = \{y_i, r_{ij}, d_i, \delta_{ij}, i = 1, \dots, n, j = 1, \dots, n_i\}$ denote the observed data from n subjects. Suppose $\boldsymbol{\Theta}$ denotes the vector of all parameters in the model, $\boldsymbol{\Theta} = \{\boldsymbol{\beta}_{0T}, \boldsymbol{\beta}_{0R}, \boldsymbol{\beta}_T, \boldsymbol{\beta}_R, \sigma_\epsilon^2, \sigma_\omega^2, \theta_\mu\}$. Let $C_{01, \theta_{ij}}(u, v) = \partial C_{\theta_{ij}}(u, v) / \partial v$, $C_{10, \theta_{ij}}(u, v) = \partial C_{\theta_{ij}}(u, v) / \partial u$, and $C_{11, \theta_{ij}}(u, v) = \partial^2 C_{\theta_{ij}}(u, v) / \partial u \partial v$. The likelihood of the observed data given $\boldsymbol{\omega}$ and $\boldsymbol{\gamma}$ is,

$$\begin{aligned} \mathcal{L}(\boldsymbol{\Theta} | \boldsymbol{\omega}, \boldsymbol{\gamma}) &\propto \prod_{i=1}^n h_T^{d_i}(y_i | \omega_i) S_T(y_i | \omega_i) \\ &\times \prod_{j=1}^{n_i+1} \left\{ \frac{C_{01, \theta_{ij}}(S_T(y_i | \omega_i), S_R(r_{ij} | \omega_i)) f_R(r_{ij} | \omega_i)}{S_T(y_i | \omega_i)} \right\}^{(1-d_i)\delta_{ij}} \\ &\times \{C_{11, \theta_{ij}}(S_T(y_i | \omega_i), S_R(r_{ij} | \omega_i)) f_R(r_{ij} | \omega_i)\}^{d_i \delta_{ij}} \\ &\times \{C_{10, \theta_{ij}}(S_T(y_i | \omega_i), S_R(r_{ij} | \omega_i))\}^{d_i(1-\delta_{ij})} \\ &\times \left\{ \frac{C_{\theta_{ij}}(S_T(y_i | \omega_i), S_R(r_{ij} | \omega_i))}{S_T(y_i | \omega_i)} \right\}^{(1-d_i)(1-\delta_{ij})}, \end{aligned}$$

4.2.3 A Clayton Time-varying Joint Frailty-copula Model

The Clayton copula (Oakes, 1989) is often used to joint model survival functions. We use it as an example to present our method. A time-varying joint frailty-copula

model can be expressed as

$$C_{\theta_{ij}}(S_T(t_i|\omega_i), \omega_i) = (S_T(t_i|\omega_i)^{-\theta_{ij}} + S_R(r_{ij}|\omega_i)^{-\theta_{ij}} - 1)^{-1/\theta_{ij}}, \theta_{ij} \in (0, \infty),$$

where the correlation between the i^{th} event and the j^{th} recurrent event is Kendall's $\tau_{ij} = \theta_{ij}/(\theta_{ij} + 2)$. So the correlation is changing after each recurrent event because of the health status change. Under the Clayton copula, we have

$$\begin{aligned} C_{01, \theta_{ij}}^*(S_T(t_i|\omega_i), S_R(r_{ij}|\omega_i)) &= S_T(t_i|\omega_i)^{-\theta_{ij}-1} (S_T(t_i|\omega_i)^{-\theta_{ij}} \\ &\quad + S_R(r_{ij}|\omega_i)^{-\theta_{ij}} - 1)^{-\frac{\theta_{ij}+1}{\theta_{ij}}} \\ C_{10, \theta_{ij}}^*(S_T(t_i|\omega_i), S_R(r_{ij}|\omega_i)) &= S_T(t_i|\omega_i)^{-\theta_{ij}-1} (S_T(t_i|\omega_i)^{-\theta_{ij}} \\ &\quad + S_R(r_{ij}|\omega_i)^{-\theta_{ij}} - 1)^{-\frac{\theta_{ij}+1}{\theta_{ij}}} \\ C_{11, \theta_{ij}}^*(S_T(t_i|\omega_i), S_R(r_{ij}|\omega_i)) &= (\theta_{ij} + 1) \{S_T(t_i|\omega_i) S_R(r_{ij}|\omega_i)\}^{-(\theta_{ij}+1)} (S_T(t_i|\omega_i)^{-\theta_{ij}} \\ &\quad + S_R(r_{ij}|\omega_i)^{-\theta_{ij}} - 1)^{-\frac{2\theta_{ij}+1}{\theta_{ij}}} \end{aligned} \tag{4.2}$$

Because $\theta_{ij} > 0$, we choose $g(\cdot, \cdot)$ such that $\theta_{ij} = \theta_\mu \exp(\gamma_{ij})$, where θ_μ can be interpreted as the overall correlation between the terminal event and the recurrent events in the study.

4.2.4 Bayesian Inference

We choose non-informative priors for the parameters in the model. Let $G^{-1}(\cdot)$ denote the inverse gamma distribution and $U(\cdot)$ denote the uniform distribution. We have,

$$\Pr(\beta_{0T}) \propto 1, \Pr(\beta_T) \propto 1, \Pr(\beta_{0R}) \propto 1, \Pr(\beta_R) \propto 1,$$

$$\theta_\mu \sim G(\alpha, \beta), \sigma_\omega^2 \sim G^{-1}(\alpha, \beta), \sigma_\epsilon^2 \sim G^{-1}(\alpha, \beta), \rho \sim U(-1, 1),$$

where α and β are chosen such that the density curve is flat enough, typically $\alpha = 0.001$ and $\beta = 1000$.

We treat ω and γ as unobserved quantities. We apply the Bayesian data augmentation (BA) algorithm (Tanner and Wong, 1987) to handle the missing data problem. If we want to generate a total of M samples of Θ from its posterior

distribution, then in the ℓ^{th} iteration,

1. (I Step) draw missing value $\boldsymbol{\gamma}^{(\ell)}, \boldsymbol{\omega}^{(\ell)}$ from the posterior predictive distribution, $\boldsymbol{\gamma}|\boldsymbol{\Theta}^{(\ell-1)}, \mathcal{D}_n, \boldsymbol{\omega}$ and $\boldsymbol{\omega}|\boldsymbol{\Theta}^{(\ell-1)}, \mathcal{D}_n, \boldsymbol{\omega}^{(\ell)}$
2. (P Step) draw $\boldsymbol{\Theta}^{(\ell)}$ from the posterior distribution $\boldsymbol{\Theta}|\mathcal{D}_n, \boldsymbol{\omega}^{(\ell)}, \boldsymbol{\gamma}^{(\ell)}$.

Note that step 1 (I step) is imputing the missing value of $(\boldsymbol{\gamma}, \boldsymbol{\omega})$ based on the current value of $\boldsymbol{\Theta}$ and step 2 (P step) is updating the current value of $\boldsymbol{\Theta}$ based on the imputed data set. We iteratively sample between step 1 and step 2, until the MCMC chain converges and we have a total of M posterior samples of $\boldsymbol{\Theta}$.

The challenge in the BA algorithm is how to sample the parameter and the missing data from their conditional distribution. In the I step, we sample γ_{ij} from

$$\Pr(\gamma_{ij}|\boldsymbol{\gamma}_{-ij}, \boldsymbol{\omega}, \boldsymbol{\Theta}, \mathcal{D}_n) \propto \mathcal{L}(\boldsymbol{\Theta}|\boldsymbol{\omega}, \boldsymbol{\gamma}) \Pr(\gamma_{ij}|\boldsymbol{\gamma}_{-ij}, \rho, \sigma_\epsilon^2)$$

and ω_i from

$$\Pr(\omega_i|\boldsymbol{\omega}_{-i}, \boldsymbol{\gamma}, \boldsymbol{\Theta}, \mathcal{D}_n) \propto \mathcal{L}(\boldsymbol{\Theta}|\boldsymbol{\omega}, \boldsymbol{\gamma}) \Pr(\omega_i|\boldsymbol{\omega}_{-i}, \sigma_\omega^2)$$

by the Metropolis-Hastings (MH) algorithm. Note that $\gamma_{ij}|\boldsymbol{\gamma}_{-ij}, \rho, \sigma_\epsilon^2$ still will still be normal,

$$\gamma_{ij}|\boldsymbol{\gamma}_{-ij}, \rho, \sigma_\epsilon^2 \sim \mathcal{N}\left(\frac{\rho}{1+\rho^2}(\gamma_{i,j+1} + \gamma_{i,j-1}), \frac{\sigma_\epsilon^2}{1+\rho^2}\right).$$

In the P step, similarly we can draw $\boldsymbol{\beta}_{0T}, \boldsymbol{\beta}_{0R}, \boldsymbol{\beta}_T, \boldsymbol{\beta}_R$ and θ_μ by the MH algorithm. For $\sigma_\epsilon^2, \sigma_\omega^2$ and ρ , there are closed forms for the posterior distribution, where

$$\sigma_\epsilon^2|\boldsymbol{\gamma}, \boldsymbol{\omega}, \boldsymbol{\Theta}_{-\sigma_\epsilon^2}, \mathcal{D}_n \sim G^{-1}\left(\frac{N+2\alpha}{2}, \left\{\frac{1}{\beta} + \frac{\sum_{i=1}^n \sum_{j=2}^{n_i+1} (\gamma_{ij} - \rho\gamma_{i,j-1})^2 + (1-\rho^2) \sum_{i=1}^n \gamma_{i1}^2}{2}\right\}^{-1}\right),$$

$$\rho|\boldsymbol{\gamma}, \boldsymbol{\omega}, \boldsymbol{\Theta}_{-\rho}, \mathcal{D}_n \sim \mathcal{N}_{(-1,1)}\left(\frac{\sum_{i=1}^n \sum_{j=2}^{n_i+1} \gamma_{ij}\gamma_{i,j-1}}{\sum_{i=1}^n \sum_{j=2}^{n_i+1} \gamma_{i,j-1}^2}, \frac{\sigma_\epsilon^2}{\sum_{i=1}^n \sum_{j=2}^{n_i+1} \gamma_{i,j-1}^2}\right),$$

and

$$\sigma_{\omega}^2 | \boldsymbol{\omega}, \mathcal{D}_n \sim G^{-1} \left(\alpha + n/2, \left\{ \beta^{-1} + 0.5 \sum_{j=1}^n \omega_j^2 \right\}^{-1} \right).$$

The derivation of the posterior distribution is shown in the Appendix. After we have the posterior distribution of Θ we can estimate Θ by the posterior mean.

4.3 Simulation

4.3.1 Simulation Set-up

We performed extensive simulation study to evaluate our method. We first compare our method with the joint frailty model and the joint frailty-copula model when model is correctly specified. And next, we evaluate our model under model misspecification.

When the copula model is correctly specified, the data is generated by a Clayton copula and when we fit the time-varying copula model we also chose the Clayton copula model. Consider the i^{th} patient ($i = 1, \dots, n$), where the hazard functions of the recurrent events and the terminal event are respectfully,

$$\begin{aligned} \log h_T(t_i) &= \beta_{0T} + x_{i1}\beta_{T,1} + x_{i2}\beta_{T,2} + \omega_i \\ \log h_R(r_{ij}) &= \beta_{0R} + x_{i1}\beta_{T,1} + x_{i2}\beta_{T,2} + \omega_i, \end{aligned} \tag{4.3}$$

where x_{i1} is generated from standard normal distribution and x_{i2} is generated by a Bernoulli distribution, $Bern(0.5)$. $(\beta_{T,1}, \beta_{T,2}, \beta_{R,1}, \beta_{R,2})'$ is set to be $(1, 1, 2, 2)'$. The baseline hazard $(\beta_{0T}, \beta_{0R})'$ is $(0.5, 1)$. σ_{ω}^2 is 0.5. After we have the hazards, we first generate T_i by inverting the exponential distribution, $exp(h_T(t_i))$. When $T_i \leq C_i$, the conditional distribution of V_{ij} is $F_V(v|U_i = u_i) = u_i^{-(\theta_{ij}+1)}(u_i^{-\theta} + v^{-\theta_{ij}} - 1)^{-1/\theta_{ij}}$. 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 let $\tilde{W}_{ij} = F_V(v|U_i = u_i)$. 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 $T_i > C_i$, 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} simply by $-\log(V_{ij})/\lambda_{ij,R}$. We repeat this procedure until T_{ij} is greater than Y_i or $j = J$.

The sample size n is varied from 100 to 200. The correlation between γ_{ij} and $\gamma_{i,j-1}$, ρ , is varied from 0.1 to 0.3. The overall correlation parameter, θ_μ , is changed from 1 to 2. The censoring rate is varied from 80% to 40%. Suppose sufficiently large maximum number $J = 1000$ of events occur for each subject. For each scenario, we generate 250 Monte Carlo datasets with the sample size $n = 100, 200$. We summarize the average of the estimates, standard deviation (SD), mean squared error (MSE) and absolute bias (AB) under these replicates.

When the model is misspecified, we generate the data by a Frank copula and when we fit the model, we still choose a Clayton copula. Other settings are similar to the previous settings in terms of n, n_i and ρ .

4.3.2 Simulation Results

We first compare our method with the joint frailty model. The result is presented in Table 4.1 and Table 4.2, which corresponds to $\theta_\mu = 1$ and $\theta_\mu = 2$. The censoring rate is 40%. For the TVJFCM, we do not observe strong bias both in Table 4.1 and Table 4.2. The AB, MSE, and SD decreases as the sample size n increases. For example the AB for $\beta_{T,1}$ is 0.206 when the sample size is 100. And it drops to 0.111 when the sample size increases to 200.

Comparing TVJFCM and JFM, the bias, MSE and AB of TVJFCM is smaller than JFM in both Table 4.1 and Table 4.2. Taking the case when $n = 100$ and $\rho = 0.1$, the bias of $\beta_{T,1}$ is 0.053 for TVJFCM and 0.138 for JFM. When the correlation ρ increase from 0.1 to 0.3 or when θ_μ increases from 1 to 2, the bias of JFM increases but TVJFCM doesn't. For example, still considering the sample size at 100, the bias of $\beta_{T,1}$ is 0.138 when $\rho = 0.1$ and it increases to 0.141 when $\rho = 0.1$ for JFM. Whereas, the bias of $\beta_{T,1}$ is 0.053 when $\rho = 0.1$ and it increases to 0.033 when $\rho = 0.3$ for TVJFCM.

In Table 4.3, Table 4.4 and Table 4.5, we compared our method with the joint frailty-copula model. The bias, SD, MSE and AB of the TVJFCM is generally smaller than the JFCM. Especially, when θ_μ increases from 1 to 3, the discrepancy between these two methods further increases. When $\theta_\mu = 3$, the AB of θ_μ is 0.158

under TVJFCM and it is 0.176 under JFCM. However, the performance of JFCM is improved when the sample size increases and as a result, the discrepancy between TVJFCM and JFCM decreases. Comparing TVJFCM, JFCM and JFM, the bias, MSE and AB of the TVJFCM is the smallest. JFCM is still better than JFM.

In Table 4.6, we evaluate the TVJFCM, controlling θ_μ to be 1, and increasing the censoring rate from 40% to 80%. Compare the result in 4.1, the bias, SD, MSE and AB is smaller when the censoring rate is low. The bias of $\beta_{T,1}$ is 0.161 when the censoring rate is 80% and sample size is 100. By contrary, the bias is 0.053 when the censoring rate is 40%. Increasing sample size from 100 to 200, the bias, SD and MSE still decreases a lot. In Table 4.7, the model misspecification problem is considered. The data is generated by a Frank copula, where $\theta_\mu = 1$ but we fit a Clayton copula. The bias, SD, AB and MSE are all greater compared with the result in Table 4.1. For example, the AB of $\beta_{T,1}$ is 0.305 when the model is misspecified and it is 0.206 when model is correctly specified.

In summary, the bias, MSE, SD and AB of TVJFCM is the smallest compared with JFCM and JFM. And when the censoring rate increases, although bias of the TVJFCM increases, they are still small and can be improved by increasing the sample size. When the copula model is misspecified, the bias of JFCM increases so model selection is preferred.

4.4 Real Data Application

We applied our model to analyze the data from the Cardiovascular Health Study (CHS), which is funded by National Heart, Lung and Blood Institute (NHLBI) to study potential risk factors for cardiovascular disease (CVD). The study started recruiting patients in 1989 and ended in 1999. At the beginning, 5201 men and women were enrolled (1990) from Forsyth County, North Carolina; Sacramento County, California; Washington County, Maryland; and Pittsburgh, Pennsylvania. 687 African Americans were enrolled after the initial baseline study. So, a total of 5888 patients were enrolled. 58% of the patients are women and 42% of the patients are men. The patients enrolled in this study are all above 65 years old at baseline. Patients have been followed up for 18 years in our data set. Risk factors includes blood pressure, lipids and subclinical disease such as echocardiography of the heart,

carotid ultrasound, and cranial magnetic-resonance imaging (MRI), are measured during patients visits. The primary outcomes in this study includes coronary heart disease (CHD), angina, heart failure (HF), stroke, transient ischemic attack (TIA), claudication, and mortality. The incidence of CVD events, the number of recurrent events were also recorded during the patient visits. The data set can be requested online at, <http://chs-nhlbi.org>. Some baseline characteristics of the patients are examined in Table 4.8. Race is 1 for the white and Race is 0 for non-white. The Gender is set equal to 1 for males and otherwise, it is 0.

In this study, we are particular interested in the correlation between recurrent myocardial infarction (MI) and all-cause death as a terminal event. MI is defined as part of the myocardium death due to an occlusion of a coronary artery from causes like embolus, spasm, thrombosis, or the rupture of a plaque. The covariates we are interested in include the gender, hypertension, race, systolic blood pressure (SBP) and baseline MI status(MIB). The death rate is 81.78% up to the year 2006. The number of recurrent events range from 0 to 5. Around 18% of the patients have at least one MI event. 803 patients have only one MI event. As an exploratory analysis, we summarize the distribution of baseline characteristics between patients with MI recurrences and without recurrences in Table 4.8. Chi-squared test is used to test whether these factors are associated with the recurrence of MI. The p-value of gender and baseline MI status is very small. We further plot the Kaplan-Meier curves across different gender groups and MIB group in Figure 4.2. There is a big difference in survival curves between different groups. Finally, we fit a joint frailty model and a Clayton time-varying frailty-copula model to analyze the data. Since hypertension is highly insignificant in the exploratory analysis, we do not include it into the model. The result is shown in Table 4.10. Under our model, MIB, gender, SBP and race are all associated with both the hazard of recurrent MI events and the hazard of death. Comparatively, the effect of MIB and gender is more strongly associated with the MI and death. For example, compared with patients with no MI event previously, the death hazard increases by 38% (HR=1.38, 95%CL=(1.28,1.4)) and the MI hazard increases by 89% (HR=1.89, 95%CL=(1.63,2.20)). Under the joint frailty model, most results are identical the same except for race. Under the joint frailty model, race is not a risk factor for death.

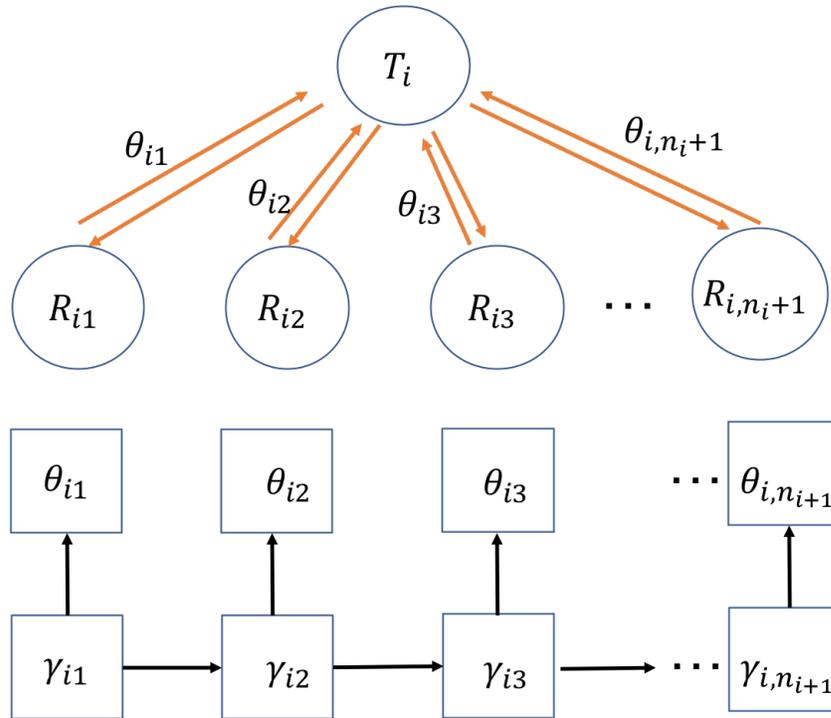


Figure 4.1. Forest plots for joint models

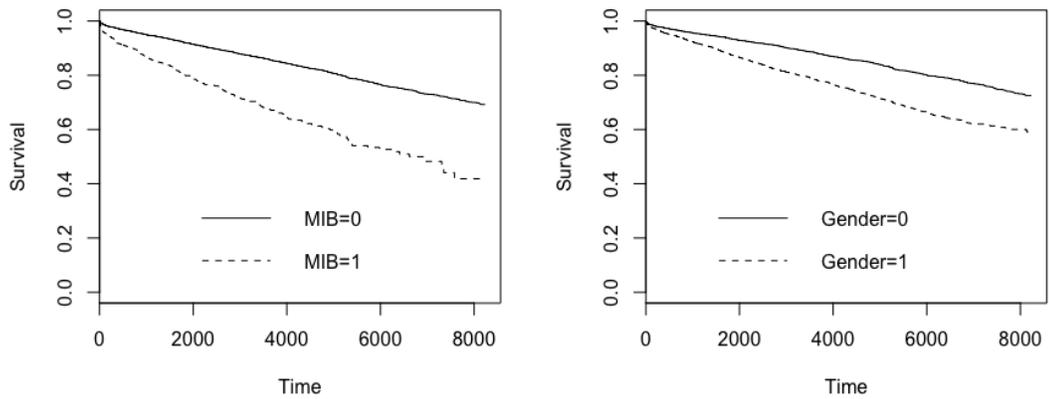


Figure 4.2. Kaplan-Meier Curves of MI Between Different Groups

*Under log rank test, p-values for both graphs are < 0.001

Another potential interest is the correlation between recurrent stroke and death. The covariates we are interested in includes the gender, race, SBP and baseline stroke status(STB). The stroke event is defined as the rapid onset of neurologic deficit or subarachnoid hemorrhage and 1) Greater than 24 hours unless death; 2) CT/MRI Lesion CT/MRI form; 3) not secondary to brain trauma, tumor or infection. The exploratory analysis of the survival curves is shown in Table 4.9 and Figure 4.4. As shown in Figure 4.4, the baseline stroke status and gender have a strong effect on the recurrence of the stroke (log rank test p-value<0.001). The result in Table 4.11 show that gender, STB, SBP and race are associated with both death process and recurrent stroke process. Note that if we fit the model by joint frailty model, SBP is not significant.

In summary, gender, race, SBP, and baseline event history are associated with MI, stroke and death event times. Patients with events history and male has the higher of MI and stroke. Our model find two more associated covariates compared with the joint frailty model. We plot a forest plot to show the effects of the risk factors (Figure 4.4). Our result is more similar compared with a previous publication by Yanez et al. (2009), Psaty et al. (2001) and Bansal et al. (2017).

4.5 Discussion

We propose a time-varying joint frailty-copula approach to jointly model recurrent event gap times and a terminal event time which 1) relaxes the conditional independence assumption, 2) captures time dependent correlation between the terminal event process and the recurrent event process and 3) models the correlation between the terminal event and the recurrent event process on the marginal level. In the simulation study, we don't observe strong bias in all scenarios when the copula model is correctly specified. The absolute bias, standard error and mean squared error continuously drop when the sample size increases. The proposed model successfully addressed the problem of the dependent censoring when we analyze recurrent events data.

Compared with the traditional joint frailty model or the joint frailty-copula model, our model is more flexible which relaxes the assumption of conditional independence and constant correlation over time. So in the simulation study, we

observe smaller absolute bias and smaller MSE compared with the traditional methods.

We extended the time-varying copula model to analyze recurrent event gap times. Furthermore, we combined the frailty and the copula to allow extra correlation between the terminal event and the recurrent event given the within-subject frailty. Last, we apply the Bayesian augmentation algorithm and avoid integrating the likelihood with respect to the frailty, which is less tractable.

The limitation of our method is that our modeling approach depends on correctly choosing the copula. This is a common model selection problem. Typically, we can use information criteria like AIC, BIC, DIC or deviance to overcome it. Another approach to select the correct model is the cross-validation error. We could choose a copula model which minimizes the cross-validation error.

Another potential problem is related to the Gibbs sampler algorithm. In the latent Gaussian process, it will take a long time for the Gibbs sampler algorithm to converge if the correlation between γ_{ij} is high (i.e., $\rho = 0.9$). We could invoke an ordered relaxation for the Gibbs sampler algorithm to solve the problem (Neal, 1998).

Some future works are considered. The biomarker is important in clinical trial study or genetic study. And it is typically modeled by a mixed effect model. So one interesting topic is to jointly model the longitudinal biomarker process, the recurrent event process and the terminal event process.

Also, in clinical trials, some of the patients might be cured and will not have recurrent events after the treatment. If we ignore this part of the patients, the estimation of the hazard might be biased. So another work is to consider extending our model to the cure rate time-varying joint frailty-copula model which adjusts the bias resulted by the cured patients.

Table 4.1. Comparison between the time-varying joint Clayton frailty-copula model and the joint frailty model when $\theta_\mu = 1$, censoring rate 40%

n	ρ	Param	TVFCJM				JFM			
			EST	SD	MSE	AB	EST	SD	MSE	AB
100	0.1	$\beta_{T,1} = 1$	1.053	0.249	0.064	0.206	1.138	0.292	0.104	0.255
		$\beta_{T,2} = 1$	1.004	0.099	0.010	0.078	1.128	0.160	0.042	0.161
		$\beta_{R,1} = 2$	2.027	0.165	0.028	0.133	2.060	0.178	0.035	0.150
		$\beta_{R,2} = 2$	1.996	0.086	0.007	0.068	2.070	0.108	0.016	0.101
		$\rho = 0.1$	0.098	0.057	0.003	0.046	-	-	-	-
	$\theta_\mu = 1$	1.015	0.157	0.025	0.129	-	-	-	-	
	0.3	$\beta_{T,1} = 1$	1.033	0.216	0.048	0.173	1.141	0.292	0.105	0.256
		$\beta_{T,2} = 1$	1.015	0.102	0.011	0.081	1.127	0.161	0.042	0.162
		$\beta_{R,1} = 2$	2.018	0.159	0.026	0.126	2.064	0.183	0.037	0.154
		$\beta_{R,2} = 2$	2.008	0.086	0.007	0.067	2.072	0.105	0.016	0.103
$\rho = 0.3$		0.289	0.061	0.004	0.049	-	-	-	-	
$\theta_\mu = 1$	0.996	0.146	0.021	0.116	-	-	-	-		
200	0.1	$\beta_{T,1} = 1$	1.016	0.138	0.019	0.111	1.140	0.202	0.060	0.195
		$\beta_{T,2} = 1$	1.008	0.070	0.005	0.055	1.122	0.107	0.026	0.134
		$\beta_{R,1} = 2$	2.010	0.108	0.012	0.088	2.082	0.126	0.023	0.119
		$\beta_{R,2} = 2$	2.000	0.059	0.003	0.046	2.067	0.066	0.009	0.078
		$\rho = 0.1$	0.097	0.043	0.002	0.035	-	-	-	-
	$\theta_\mu = 1$	1.004	0.126	0.016	0.099	-	-	-	-	
	0.3	$\beta_{T,1} = 1$	1.002	0.137	0.019	0.108	1.139	0.203	0.060	0.194
		$\beta_{T,2} = 1$	1.003	0.071	0.005	0.056	1.121	0.107	0.026	0.133
		$\beta_{R,1} = 2$	2.001	0.111	0.012	0.086	2.083	0.136	0.025	0.125
		$\beta_{R,2} = 2$	1.998	0.058	0.003	0.046	2.070	0.070	0.010	0.082
$\rho = 0.3$		0.289	0.042	0.002	0.033	-	-	-	-	
$\theta_\mu = 1$	1.006	0.126	0.016	0.099	-	-	-	-		

Table 4.2. Comparison between the time-varying joint Clayton frailty-copula model and the joint frailty model when $\theta_\mu = 2$, censoring rate 40%

n	ρ	Param	TJFCM				JFM			
			EST	SD	MSE	AB	EST	SD	MSE	AB
100	0.1	$\beta_{T,1} = 1$	1.000	0.186	0.034	0.143	1.182	0.301	0.124	0.279
		$\beta_{T,2} = 1$	1.012	0.085	0.007	0.067	1.164	0.167	0.054	0.187
		$\beta_{R,1} = 2$	1.997	0.156	0.024	0.124	2.135	0.207	0.061	0.196
		$\beta_{R,2} = 2$	2.004	0.077	0.006	0.061	2.119	0.116	0.028	0.136
		$\rho = 0.1$	0.094	0.067	0.005	0.052	-	-	-	-
		$\theta_\mu = 2$	2.022	0.187	0.035	0.149	-	-	-	-
	0.3	$\beta_{T,1} = 1$	1.004	0.169	0.029	0.131	1.181	0.300	0.122	0.277
		$\beta_{T,2} = 1$	1.018	0.086	0.008	0.070	1.164	0.166	0.055	0.188
		$\beta_{R,1} = 2$	2.002	0.143	0.020	0.116	2.135	0.195	0.056	0.186
		$\beta_{R,2} = 2$	2.005	0.078	0.006	0.063	2.122	0.121	0.029	0.140
200	0.1	$\beta_{T,1} = 1$	1.003	0.115	0.013	0.091	1.176	0.211	0.075	0.219
		$\beta_{T,2} = 1$	1.006	0.058	0.003	0.046	1.157	0.110	0.037	0.163
		$\beta_{R,1} = 2$	2.007	0.106	0.011	0.084	2.134	0.147	0.040	0.163
		$\beta_{R,2} = 2$	2.004	0.052	0.003	0.041	2.126	0.083	0.023	0.129
		$\rho = 0.1$	0.096	0.051	0.003	0.042	-	-	-	-
		$\theta_\mu = 2$	1.985	0.174	0.030	0.138	-	-	-	-
	0.3	$\beta_{T,1} = 1$	1.017	0.120	0.015	0.096	1.175	0.210	0.075	0.218
		$\beta_{T,2} = 1$	1.008	0.054	0.003	0.043	1.157	0.111	0.037	0.164
		$\beta_{R,1} = 2$	2.021	0.105	0.011	0.082	2.136	0.142	0.039	0.162
		$\beta_{R,2} = 2$	2.007	0.051	0.003	0.041	2.125	0.084	0.023	0.128
	$\rho = 0.3$	0.287	0.043	0.002	0.037	-	-	-	-	
	$\theta_\mu = 2$	1.987	0.162	0.026	0.130	-	-	-	-	

Table 4.3. Comparison between the time-varying joint Clayton frailty-copula model and the joint frailty-copula model when $\theta_\mu = 1$, censoring rate 40%

n	ρ	Parameter	TVJFCM				JFCM			
			EST	SD	MSE	AB	EST	SD	MSE	AB
100	0.1	$\beta_{T,1} = 1$	1.053	0.249	0.064	0.206	1.031	0.241	0.059	0.195
		$\beta_{T,2} = 1$	1.004	0.099	0.010	0.078	1.024	0.115	0.014	0.092
		$\beta_{R,1} = 2$	2.027	0.165	0.028	0.133	2.015	0.174	0.030	0.141
		$\beta_{R,2} = 2$	1.996	0.086	0.007	0.068	2.014	0.095	0.009	0.073
		$\rho = 0.1$	0.098	0.057	0.003	0.046	-	-	-	-
		$\theta_\mu = 1$	1.015	0.157	0.025	0.129	0.974	0.158	0.025	0.125
	0.3	$\beta_{T,1} = 1$	1.033	0.216	0.048	0.173	1.036	0.236	0.057	0.190
		$\beta_{T,2} = 1$	1.015	0.102	0.011	0.081	1.013	0.105	0.011	0.082
		$\beta_{R,1} = 2$	2.018	0.159	0.026	0.126	2.015	0.163	0.027	0.132
		$\beta_{R,2} = 2$	2.008	0.086	0.007	0.067	2.005	0.079	0.006	0.062
		$\rho = 0.3$	0.289	0.061	0.004	0.049	-	-	-	-
		$\theta_\mu = 1$	0.996	0.146	0.021	0.116	0.992	0.162	0.026	0.131
200	0.1	$\beta_{T,1} = 1$	1.016	0.138	0.019	0.111	1.015	0.139	0.019	0.112
		$\beta_{T,2} = 1$	1.008	0.070	0.005	0.055	1.008	0.072	0.005	0.055
		$\beta_{R,1} = 2$	2.010	0.108	0.012	0.088	2.011	0.110	0.012	0.092
		$\beta_{R,2} = 2$	2.000	0.059	0.003	0.046	2.000	0.058	0.003	0.045
		$\rho = 0.1$	0.097	0.043	0.002	0.035	-	-	-	-
		$\theta_\mu = 1$	1.004	0.126	0.016	0.099	1.010	0.120	0.014	0.096
	0.3	$\beta_{T,1} = 1$	1.002	0.137	0.019	0.108	1.004	0.152	0.023	0.123
		$\beta_{T,2} = 1$	1.003	0.071	0.005	0.056	1.005	0.073	0.005	0.057
		$\beta_{R,1} = 2$	2.001	0.111	0.012	0.086	2.002	0.115	0.013	0.091
		$\beta_{R,2} = 2$	1.998	0.058	0.003	0.046	1.998	0.055	0.003	0.044
		$\rho = 0.3$	0.289	0.042	0.002	0.033	-	-	-	-
		$\theta_\mu = 1$	1.006	0.126	0.016	0.099	1.008	0.112	0.013	0.088

Table 4.4. Comparison between the time-varying joint Clayton frailty-copula model and the joint frailty-copula model when $\theta_\mu = 2$, censoring rate 40%

n	ρ	Parameter	TVJFCM				JFCM			
			EST	SD	MSE	AB	EST	SD	MSE	AB
100	0.1	$\beta_{T,1} = 1$	1.000	0.186	0.034	0.143	1.007	0.185	0.034	0.141
		$\beta_{T,2} = 1$	1.012	0.085	0.007	0.067	1.012	0.093	0.009	0.074
		$\beta_{R,1} = 2$	1.997	0.156	0.024	0.124	2.000	0.158	0.025	0.126
		$\beta_{R,2} = 2$	2.004	0.077	0.006	0.061	2.005	0.081	0.007	0.065
		$\rho = 0.1$	0.094	0.067	0.005	0.052	-	-	-	-
		$\theta_\mu = 2$	2.022	0.187	0.035	0.149	1.974	0.176	0.031	0.136
	0.3	$\beta_{T,1} = 1$	1.004	0.169	0.029	0.131	1.014	0.171	0.029	0.133
		$\beta_{T,2} = 1$	1.018	0.086	0.008	0.070	1.017	0.095	0.009	0.076
		$\beta_{R,1} = 2$	2.002	0.143	0.020	0.116	2.010	0.150	0.023	0.119
		$\beta_{R,2} = 2$	2.005	0.078	0.006	0.063	2.005	0.084	0.007	0.066
		$\rho = 0.3$	0.288	0.066	0.005	0.054	-	-	-	-
		$\theta_\mu = 2$	2.003	0.177	0.031	0.139	1.983	0.199	0.040	0.156
200	0.1	$\beta_{T,1} = 1$	1.003	0.115	0.013	0.091	1.010	0.121	0.015	0.096
		$\beta_{T,2} = 1$	1.006	0.058	0.003	0.046	1.003	0.058	0.003	0.047
		$\beta_{R,1} = 2$	2.007	0.106	0.011	0.084	2.014	0.112	0.013	0.089
		$\beta_{R,2} = 2$	2.004	0.052	0.003	0.041	2.003	0.053	0.003	0.042
		$\rho = 0.1$	0.096	0.051	0.003	0.042	-	-	-	-
		$\theta_\mu = 2$	1.985	0.174	0.030	0.138	1.990	0.165	0.027	0.130
	0.3	$\beta_{T,1} = 1$	1.017	0.120	0.015	0.096	1.006	0.126	0.016	0.100
		$\beta_{T,2} = 1$	1.008	0.054	0.003	0.043	1.005	0.058	0.003	0.046
		$\beta_{R,1} = 2$	2.021	0.105	0.011	0.082	2.010	0.108	0.012	0.088
		$\beta_{R,2} = 2$	2.007	0.051	0.003	0.041	2.004	0.053	0.003	0.042
		$\rho = 0.3$	0.287	0.043	0.002	0.037	-	-	-	-
		$\theta_\mu = 2$	1.987	0.162	0.026	0.130	1.999	0.171	0.029	0.136

Table 4.5. Comparison between the time-varying joint Clayton frailty-copula model and the joint frailty-copula model when $\theta_\mu = 3$, censoring rate 40%

n	ρ	Parameter	TVJFCM				JFCM			
			EST	SD	MSE	AB	EST	SD	MSE	AB
100	0.1	$\beta_{T,1} = 1$	1.014	0.153	0.024	0.123	0.999	0.165	0.027	0.129
		$\beta_{T,2} = 1$	1.007	0.067	0.004	0.053	0.996	0.075	0.006	0.058
		$\beta_{R,1} = 2$	2.007	0.137	0.019	0.107	1.992	0.153	0.023	0.122
		$\beta_{R,2} = 2$	2.007	0.065	0.004	0.051	1.996	0.072	0.005	0.056
		$\rho = 0.1$	0.094	0.066	0.004	0.053	-	-	-	-
		$\theta_\mu = 3$	2.975	0.180	0.033	0.146	2.961	0.202	0.042	0.163
	0.3	$\beta_{T,1} = 1$	1.015	0.156	0.024	0.123	1.007	0.155	0.024	0.122
		$\beta_{T,2} = 1$	1.005	0.071	0.005	0.057	0.999	0.080	0.006	0.061
		$\beta_{R,1} = 2$	2.011	0.140	0.020	0.107	2.004	0.147	0.021	0.115
		$\beta_{R,2} = 2$	2.007	0.067	0.004	0.053	2.001	0.075	0.006	0.058
		$\rho = 0.3$	0.287	0.066	0.004	0.054	-	-	-	-
		$\theta_\mu = 3$	3.000	0.198	0.039	0.158	2.968	0.218	0.048	0.176
200	0.1	$\beta_{T,1} = 1$	0.998	0.096	0.009	0.076	0.998	0.108	0.012	0.086
		$\beta_{T,2} = 1$	1.002	0.049	0.002	0.038	0.999	0.050	0.002	0.038
		$\beta_{R,1} = 2$	1.999	0.083	0.007	0.066	1.998	0.102	0.010	0.081
		$\beta_{R,2} = 2$	2.000	0.044	0.002	0.034	1.997	0.046	0.002	0.035
		$\rho = 0.1$	0.096	0.040	0.002	0.033	-	-	-	-
		$\theta_\mu = 3$	2.982	0.171	0.030	0.133	2.984	0.184	0.034	0.148
	0.3	$\beta_{T,1} = 1$	0.995	0.096	0.009	0.075	0.997	0.111	0.012	0.087
		$\beta_{T,2} = 1$	1.000	0.046	0.002	0.036	0.998	0.052	0.003	0.041
		$\beta_{R,1} = 2$	1.997	0.091	0.008	0.072	1.999	0.103	0.011	0.083
		$\beta_{R,2} = 2$	1.999	0.044	0.002	0.034	1.996	0.048	0.002	0.039
		ρ	0.294	0.047	0.002	0.037	-	-	-	-
		$\theta_\mu = 3$	2.999	0.180	0.032	0.140	2.987	0.186	0.035	0.148

Table 4.6. Joint modeling by the time-varying joint Clayton frailty-copula model when $\theta_\mu = 1$, event rate 20%

ρ	Parameter	n=100				n=200			
		EST	SD	MSE	AB	EST	SD	MSE	AB
0.1	$\beta_{T,1} = 1$	1.161	0.600	0.385	0.474	1.046	0.362	0.133	0.282
	$\beta_{T,2} = 1$	1.074	0.283	0.085	0.22	1.021	0.16	0.026	0.123
	$\beta_{R,1} = 2$	2.029	0.296	0.088	0.235	2.005	0.171	0.029	0.135
	$\beta_{R,2} = 2$	2.014	0.139	0.019	0.111	2.002	0.082	0.007	0.064
	$\rho = 0.1$	0.096	0.101	0.010	0.079	0.095	0.063	0.004	0.050
	$\theta_\mu = 1$	0.982	0.187	0.035	0.143	0.975	0.159	0.026	0.128
0.3	$\beta_{T,1} = 1$	1.144	0.582	0.358	0.453	1.058	0.377	0.145	0.281
	$\beta_{T,2} = 1$	1.072	0.277	0.081	0.21	1.028	0.162	0.027	0.123
	$\beta_{R,1} = 2$	2.028	0.283	0.081	0.227	2.011	0.182	0.033	0.145
	$\beta_{R,2} = 2$	2.016	0.134	0.018	0.105	2.007	0.089	0.008	0.072
	$\rho = 0.3$	0.291	0.095	0.009	0.075	0.293	0.065	0.004	0.053
	$\theta_\mu = 1$	0.979	0.182	0.033	0.142	0.978	0.154	0.024	0.125

Table 4.7. Joint modeling by the time-varying joint Clayton frailty-copula model when $\theta_\mu = 1$, event rate 60% and the copula model is misspecified

ρ	Parameter	n=100				n=200			
		EST	SD	MSE	AB	EST	SD	MSE	AB
0.1	$\beta_{T,1} = 1$	0.841	0.335	0.137	0.305	0.878	0.218	0.062	0.203
	$\beta_{T,2} = 1$	0.865	0.163	0.045	0.171	0.884	0.129	0.030	0.140
	$\beta_{R,1} = 2$	2.005	0.172	0.029	0.131	2.009	0.113	0.013	0.089
	$\beta_{R,2} = 2$	1.985	0.087	0.008	0.069	2.005	0.062	0.004	0.049
	$\rho = 0.1$	0.099	0.065	0.004	0.053	0.099	0.047	0.002	0.037
	$\tau_\mu = 0.11$	0.206	0.060	0.013	0.099	0.158	0.043	0.004	0.054
0.3	$\beta_{T,1} = 1$	0.838	0.339	0.141	0.306	0.878	0.205	0.057	0.196
	$\beta_{T,2} = 1$	0.872	0.160	0.042	0.167	0.884	0.131	0.031	0.140
	$\beta_{R,1} = 2$	1.986	0.173	0.030	0.136	2.003	0.110	0.012	0.090
	$\beta_{R,2} = 2$	1.993	0.087	0.008	0.067	2.006	0.060	0.004	0.046
	$\rho = 0.3$	0.292	0.063	0.004	0.051	0.293	0.045	0.002	0.036
	$\tau_\mu = 0.11$	0.204	0.057	0.012	0.096	0.155	0.042	0.004	0.051

Table 4.8. Baseline Characteristics for Patients with MI Events and No MI Events

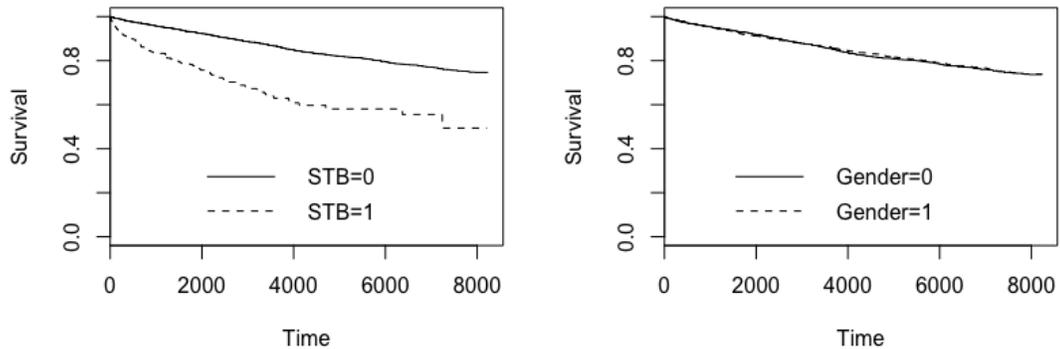
	No MI	MI	p-value
Total	2499	3097	
Hypertension			
No	354(44.58%)	440(55.42%)	
Yes	2142(44.63%)	2657(55.37%)	1.000
Gender			
Female	1538(47.51%)	1699(52.49%)	
Male	961(40.69%)	1401(59.31%)	<0.001
Race			
White	2077(44.37%)	2604(55.63%)	
Black	403(45.74%)	478(54.26%)	
Other	19(51.35%)	18(48.65%)	0.547
MIB			
No	2387(46.84%)	2709(53.16%)	
Yes	112(22.27%)	391(77.73%)	<0.001
SBP			
Mean(SD)	133.6(11.65)	138.89(11.08)	0.503

Table 4.9. Baseline Characteristics for Patients with Stroke Events and No Stroke Events

	No ST	ST	p-value
Total	2437	3041	
Hypertension			
No	352(44.73%)	435(55.27%)	
Yes	2082(44.45%)	2602(55.55%)	0.904
Gender			
Female	1490(47.53%)	1645(52.47%)	
Male	947(40.42%)	1396(59.58%)	<0.001
Race			
White	2027(44.21%)	2558(55.79%)	
Black	389(45.55%)	465(54.45%)	
Other	21(53.85%)	18(46.15%)	0.400
STB			
No	2379(45.21%)	2883(54.79%)	
Yes	58(26.85%)	158(73.15%)	<0.001
SBP			
Mean(SD)	133.40(11.64)	138.73(11.08)	0.620

Table 4.10. Joint Modeling MI and Death by TVJFCM

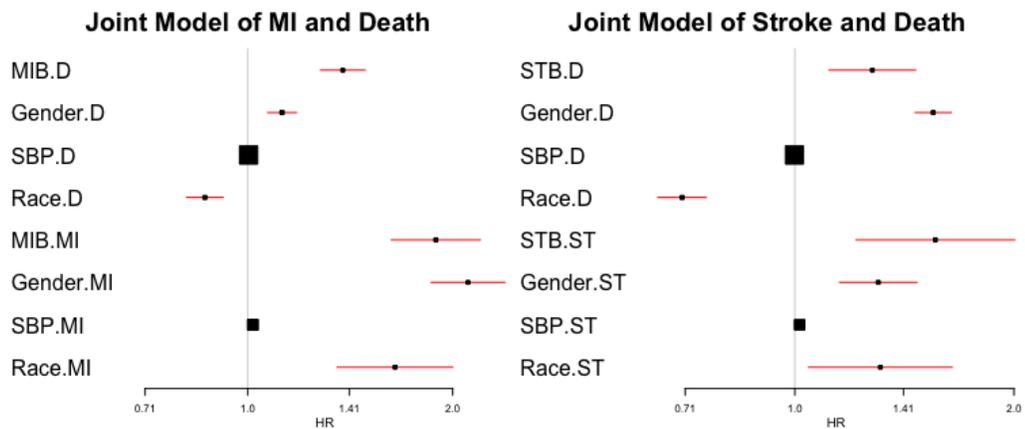
	Covariates	JFM			TVJFCM		
		EST	2.5%	97.5%	EST	2.5%	97.5%
Time to Death	MIB=1	0.513	0.435	0.591	0.322	0.246	0.397
	Gender=1	0.265	0.214	0.315	0.116	0.068	0.165
	SBP	0.006	0.005	0.008	0.003	0.002	0.004
	Race=1	-0.001	-0.068	0.067	-0.145	-0.206	-0.083
Time to MI	MIB=1	0.635	0.476	0.795	0.636	0.486	0.787
	Gender=1	0.775	0.658	0.891	0.745	0.620	0.869
	SBP	0.013	0.011	0.016	0.018	0.015	0.020
	Race=1	0.491	0.307	0.674	0.498	0.303	0.694
Correlation	θ	-	-	-	0.537	0.030	1.044
	ρ	-	-	-	0.092	0.059	0.125

**Figure 4.3.** Kaplan-Meier Curves of Stroke Events Between Different Groups

*Under log rank test, p-values for both graphs are < 0.001

Table 4.11. Joint Modeling Stroke and Death by TVJFCM and JFM

	Covariates	JFM			TVJFCM		
		EST	2.5%	97.5%	EST	2.5%	97.5%
Time to Death	STB=1	0.263	0.120	0.405	0.245	0.108	0.382
	Gender=1	0.286	0.230	0.342	0.438	0.381	0.494
	SBP	0.000	-0.002	0.001	-0.001	-0.002	0.000
	Race=1	-0.410	-0.486	-0.334	-0.357	-0.433	-0.280
Time to Stroke	STB=1	0.382	0.057	0.708	0.444	0.194	0.695
	Gender=1	0.405	0.274	0.536	0.264	0.142	0.386
	SBP	0.019	0.016	0.022	0.015	0.012	0.018
	Race=1	0.045	-0.144	0.235	0.271	0.043	0.498
Correlation	θ	-	-	-	3.366	2.948	3.784
	ρ	-	-	-	0.471	0.420	0.522

**Figure 4.4.** Forest plots for joint models

Summary and Future Work

5.1 Summary

Most of the literature current on joint models are built on the work of Liu et al. (2004). As a result, when the conditional independence assumption or the constant correlation assumption is violated, the inference would be biased. Yue and Chan (1997) and McGilchrist and Aisbett (1991) although considered dynamic frailty models to relax the conditional independence assumption, they do not consider the terminal event process. Also with more frailties, the integral of the likelihood is less tractable.

Emura et al. (2015) proposed the joint frailty-copula model for meta-analysis. In Chapter 3, we extend the frailty-copula model to analyze the recurrent gap times with dependent censoring. The frailty-copula models do not specify the distributions of the time-dependent frailties and how the frailties are correlated. Instead, we use a survival copula to connect the conditional survival function. So it is more robust compared to the dynamic frailty approach, and the integral of the likelihood is more tractable. In Chapter 4, we further incorporate a latent AR(1) process to model the correlation between the terminal event and the recurrent events, which relaxes the assumption of the constant correlation over time by proposing a time-varying joint frailty-copula model. In order to avoid the problem of non-tractable likelihood integration, we propose to use a Bayesian augmentation algorithm to estimate the parameters in the model. The Bayesian augmentation algorithm first imputes missing values and then draws posterior samples based on

imputed data set.

In the simulation study, both models outperform the traditional joint frailty model when the conditional independence assumption is violated. We also see when the assumption is not violated that the bias of the proposed model is identical but the standard error and mean squared error is comparatively greater. This is expected, but the proposed models still are more appealing because the assumptions for the traditional models are too simplistic. When the correlation is not a constant, the bias, standard error and mean squared error are all improved by the time-varying frailty copula model.

5.2 Future Work

As mentioned in Chapter 3 and Chapter 4, there still are some interesting topics, which could be studied based on our model. One interesting topic to consider is jointly modeling biomarker data, recurrent event gap times, and the terminal event. Under the new model, the hazard functions are

$$y(t|\omega_i) = m_i(t) + \epsilon(t) \quad (5.1)$$

$$\lambda_R(r|\omega_i) = \lambda_{0R}(r) \exp \{ \boldsymbol{\beta}'_R \mathbf{x}_{i,R} + \alpha_R m_i(r) \} \quad (5.2)$$

$$\lambda_T(t|\omega_i) = \lambda_{0T}(t) \exp \{ \boldsymbol{\beta}'_T \mathbf{x}_{i,T} + \alpha_T m_i(t) \} \quad (5.3)$$

where $m_i(t)$ is the mean expression level of a biomarker at time t and it is modeled by a mixed-effects model, $m_i(t) = \beta_0 + \boldsymbol{\beta}_1 x + \omega_i$. $\epsilon(t_{ij})$ s are mutually independent and following a normal distribution, $\mathcal{N}(0, \sigma_\epsilon^2)$. The likelihood function of this model needs to further incorporate the biomarker likelihood. The EM algorithm and the BA algorithm can both be applied to estimate the parameters.

Another potential situation to consider is multiple types of the recurrent events with dependent censoring. In the CHS study, there are multiple types of the recurrent event, like stroke or MI. There is a correlation between MI and stroke and they are both stopped by death. Lin et al. (2017) considered a Bayesian approach to jointly model these three time-to-event processes,

$$\lambda_{R,k}(r|\omega_i) = \lambda_{0R,k}(r) \exp \{ \boldsymbol{\beta}'_R \mathbf{x}_{i,R} + \omega_{ik} \} \quad (5.4)$$

$$\lambda_T(t|\omega_i) = \lambda_{0T}(t) \exp \left\{ \boldsymbol{\beta}'_T \mathbf{x}_{i,T} + \sum_{k=1}^K \psi_k \omega_{ik} \right\}, \quad (5.5)$$

where $\lambda_{R,k}$ is the hazard function for the k^{th} ($k = 1, \dots, K$) type of the recurrent event. ω_{ik} is the frailty associated with the i^{th} patient and the k^{th} type of the recurrent event. The effect of the frailties on the terminal event is denoted by ψ_k . This model still assumes conditional independence given ω_{ik} . Still, no literature consider relaxing this assumption. A potential method is to use a multivariate copula model to connect these conditional survival functions.

In clinical trials, a portion of the patients might be cured after the treatment, and the hazard may become 0 (Liu et al., 2016). A cure rate model under the joint frailty-copula model might be possible. The hazard functions under this model can be specified by

$$\lambda_R(r|\omega_i) = \begin{cases} \lambda_{0R}(r) \exp(\boldsymbol{\beta}'_R \mathbf{x}_{i,R} + \omega_i) & \text{if not cured} \\ 0 & \text{if cured} \end{cases} \quad (5.6)$$

$$\lambda_T(t|\omega_i) = \lambda_{0T}(t) \exp \{ \boldsymbol{\beta}'_T \mathbf{x}_{i,T} + \omega_i \}, \quad (5.7)$$

The probability of the i^{th} patient to be cured is $p_i = \exp(\boldsymbol{\beta}'_c \mathbf{x}_{c,i}) / (1 + \exp(\boldsymbol{\beta}'_c \mathbf{x}_{c,i}))$, where $\mathbf{x}_{c,i}$ is the covariate vector associated with the cure rate.

Appendix for chapter 3

A.1 Derivation of the likelihood

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

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

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

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

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

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

A.2 Metropolis-Hastings within the Gibbs sampler algorithm

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

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

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

- (d) If $LR > U$, $\beta_{0T}^{(l)} = \beta_{0T}^N$. Otherwise, $\beta_{0T}^{(l)} = \beta_{0T}^{(l-1)}$.
2. Sample $\beta_{0R}^{(l)}$ by $\Pr(\beta_{0R}^{(l)}|\beta_{0T}^{(l)}, \beta_T^{(l-1)}, \beta_R^{(l-1)}, \mathbf{w}^{(l-1)}, \theta^{(l-1)}, \sigma_w^{-2(l-1)}, D_n)$, identically.
3. Sample $\beta_T^{(l)}$ by $\Pr(\beta_T^{(l)}|\beta_{0T}^{(l)}, \beta_{0R}^{(l)}, \beta_R^{(l-1)}, \mathbf{w}^{(l-1)}, \theta^{(l-1)}, \sigma_w^{-2(l-1)}, D_n)$, identically.
4. Sample $\beta_R^{(l)}$ by $\Pr(\beta_R^{(l)}|\beta_{0T}^{(l)}, \beta_{0R}^{(l)}, \beta_T^{(l)}, \mathbf{w}^{(l-1)}, \theta^{(l-1)}, \sigma_w^{-2(l-1)}, D_n)$, similarly as in step 1.
5. For $i = 1, \dots, n$, sample $w_i^{(l)}$ from

$$\Pr(w_i^{(l)}|\beta_{0T}^{(l)}, \beta_{0R}^{(l)}, \beta_T^{(l)}, \beta_R^{(l)}, \mathbf{w}_{[-i]}^{(l-1)}, \theta^{(l-1)}, \sigma_w^{-2(l-1)}, D_n),$$

- (a) Generate $w_i^N = w_i^{(l-1)} + sN(0, 1)$, where s is the step size of the random walk
- (b) Generate $U \sim Unif(0, 1)$
- (c) Let $L_i(\cdot)$ denote the likelihood of the i^{th} subject. Calculate

$$LR = \frac{L_i(\boldsymbol{\beta}_{0T}^{(l)}, \boldsymbol{\beta}_{0R}^{(l)}, \boldsymbol{\beta}_T^{(l)}, \boldsymbol{\beta}_R^{(l)}, w_i^N, \theta^{(l-1)}, \sigma_w^{-2(l-1)}, \mathcal{D}_n) \Pr(w_i^N)}{L_i(\boldsymbol{\beta}_{0T}^{(l)}, \boldsymbol{\beta}_{0R}^{(l)}, \boldsymbol{\beta}_T^{(l)}, \boldsymbol{\beta}_R^{(l)}, w_i^{(l-1)}, \theta^{(l-1)}, \sigma_w^{-2(l-1)}, \mathcal{D}_n) \Pr(w_i^{(l-1)})}$$

- (d) If $LR > U$, $w_i^{(l)} = w_i^N$. Otherwise, $w_i^{(l)} = w_i^{(l-1)}$
6. Sample $\theta^{(l)}$ by $\Pr(\theta | \boldsymbol{\beta}_{0T}^{(l)}, \boldsymbol{\beta}_{0R}^{(l)}, \boldsymbol{\beta}_T^{(l)}, \boldsymbol{\beta}_R^{(l)}, \mathbf{w}^{(l)}, \sigma_w^{-2(l-1)}, \mathcal{D}_n)$ similarly as in step 1.
7. Sample $\sigma_w^{-2(l)}$ by $\Pr(\sigma_w^{-2} | \boldsymbol{\beta}_{0T}^{(l)}, \boldsymbol{\beta}_{0R}^{(l)}, \boldsymbol{\beta}_T^{(l)}, \boldsymbol{\beta}_R^{(l)}, \mathbf{w}^{(l)}, \mathcal{D}_n)$, which is a gamma distribution, i.e.,

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

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.

Appendix for chapter 4

We first derive the conditional distribution of $\gamma_{ij}|\gamma_{-ij}, \boldsymbol{\omega}, \boldsymbol{\Theta}$. Since $\boldsymbol{\omega}$ is independent of γ_{ij} , the distribution of $\gamma_{ij}|\gamma_{-ij}, \boldsymbol{\omega}, \boldsymbol{\Theta}$ is equivalent to $\gamma_{ij}|\gamma_{-ij}, \rho, \sigma_\epsilon^2$.

For $j = 1$, it only depends on γ_{i2}, ρ and σ_ϵ^2 . We have the joint distribution,

$$(\gamma_{i1}, \gamma_{i2})|\sigma_\epsilon^2, \rho \sim \mathcal{N}\left(\mathbf{0}, \frac{\sigma_\epsilon^2}{1 - \rho^2} \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right).$$

So $\gamma_{i1}|\gamma_{-i1}, \sigma_\epsilon^2, \rho \sim \mathcal{N}(\rho\gamma_{i2}, \sigma_\epsilon^2)$.

Consider $j = 2, \dots, n_i$,

$$\begin{aligned} \Pr(\gamma_{ij}|\gamma_{-ij}, \sigma_\epsilon^2, \rho) &\propto \Pr(\gamma_{ij}, \gamma_{-ij}|\sigma_\epsilon^2, \rho) \\ &= \Pr(\gamma_{i, n_i}|\gamma_{i, n_i-1}, \sigma_\epsilon^2, \rho) \dots \Pr(\gamma_{i2}|\gamma_{i1}, \sigma_\epsilon^2, \rho) \Pr(\gamma_{i1}|\sigma_\epsilon^2, \rho) \\ &\propto \Pr(\gamma_{i, j+1}|\gamma_{i, j}, \sigma_\epsilon^2, \rho) \Pr(\gamma_{ij}|\gamma_{i, j-1}, \sigma_\epsilon^2, \rho) \\ &\propto \exp\left\{-\frac{(\gamma_{i, j+1} - \rho\gamma_j)^2 + (\gamma_j - \rho\gamma_{i, j-1})^2}{2\sigma_\epsilon^2}\right\} \\ &\propto \exp\left[-\frac{\{\gamma_{ij} - \rho/(1 + \rho^2)(\gamma_{i, j+1} - \gamma_{i, j-1})\}^2}{2\sigma_\epsilon^2/(1 + \rho^2)}\right] \end{aligned}$$

So, $\gamma_{ij}|\gamma_{-ij}, \sigma_\epsilon^2, \rho \sim \mathcal{N}(\rho/(1 + \rho^2)(\gamma_{i, j+1} + \gamma_{i, j-1}), \sigma_\epsilon^2/(1 + \rho^2))$.

For $j = n_i + 1$, from $\gamma_{i, n_i+1} = \rho\gamma_{i, n_i} + \epsilon_{i, n_i+1}$, we have $\gamma_{i, n_i+1}|\gamma_{-i, n_i+1}, \sigma_\epsilon^2, \rho \sim \mathcal{N}(\rho\gamma_{i, n_i}, \sigma_\epsilon^2)$.

Next, we derive the conditional distribution for σ_ϵ^2 . The density function of σ_ϵ^2

is $f(\sigma_\epsilon^2) = \sigma_\epsilon^{-2\alpha-2} \exp(-1/\beta\sigma_\epsilon^2)$. The posterior is,

$$\begin{aligned}
\Pr(\sigma_\epsilon^2 | \boldsymbol{\gamma}, \boldsymbol{\omega}, \boldsymbol{\Theta}_{-\sigma_\epsilon^2}, \mathcal{D}_n) &\propto \Pr(\mathcal{D}_n | \boldsymbol{\Theta}, \boldsymbol{\omega}, \boldsymbol{\gamma}) \Pr(\boldsymbol{\omega}, \boldsymbol{\gamma} | \boldsymbol{\Theta}) \\
&\propto \Pr(\boldsymbol{\omega} | \sigma_\omega^2) \Pr(\boldsymbol{\gamma} | \rho, \sigma_\epsilon^2) \Pr(\sigma_\epsilon^2) \\
&\propto \Pr(\boldsymbol{\gamma} | \sigma_\epsilon^2, \rho) \Pr(\sigma_\epsilon^2) \\
&\propto \sigma_\epsilon^{-2\alpha-2} \exp\left(-\frac{1}{\beta\sigma_\epsilon^2}\right) \prod_{i=1}^n \sqrt{\frac{1-\rho^2}{2\pi\sigma_\epsilon^2}} \exp\left\{\frac{\gamma_{i1}^2(1-\rho^2)}{2\sigma_\epsilon^2}\right\} \\
&\quad \times \prod_{i=1}^n \prod_{j=2}^{n_i+1} \frac{1}{\sqrt{2\pi\sigma_\epsilon^2}} \exp\left\{-\frac{(\gamma_{ij} - \rho\gamma_{i,j-1})^2}{2\sigma_\epsilon^2}\right\} \\
&= \sigma_\epsilon^{-2\left(\frac{2\alpha+N}{2}-1\right)} \exp\left\{-\frac{\frac{1}{\beta} + 0.5 \sum_{i=1}^n \sum_{j=2}^{n_i+1} (\gamma_{ij} - \rho\gamma_{i,j-1})^2}{\sigma_\epsilon^2}\right\} \\
&\quad \times \exp\left\{-\frac{0.5 \sum_{i=1}^n (1-\rho^2)\gamma_{i1}^2}{\sigma_\epsilon^2}\right\},
\end{aligned}$$

which is the form of inverse gamma distribution. So,

$$\sigma_\epsilon^2 | \boldsymbol{\gamma}, \boldsymbol{\omega}, \boldsymbol{\Theta}_{-\sigma_\epsilon^2}, \mathcal{D}_n \sim G^{-1}\left(\frac{N+2\alpha}{2}, \left\{\frac{1}{\beta} + \frac{\sum_{i=1}^n \sum_{j=2}^{n_i+1} (\gamma_{ij} - \rho\gamma_{i,j-1})^2 + (1-\rho^2) \sum_{i=1}^n \gamma_{i1}^2}{2}\right\}^{-1}\right).$$

Finally, we show that the posterior of ρ is,

$$\begin{aligned}
\Pr(\rho | \boldsymbol{\gamma}, \boldsymbol{\omega}, \boldsymbol{\Theta}_{-\rho}, \mathcal{D}_n) &\propto \Pr(\boldsymbol{\gamma} | \sigma_\epsilon^2, \rho) \Pr(\rho) \\
&\propto \prod_{i=1}^n \prod_{j=2}^{n_i+1} \frac{1}{\sqrt{2\pi\sigma_\epsilon^2}} \exp\left\{-\frac{(\gamma_{ij} - \rho\gamma_{i,j-1})^2}{2\sigma_\epsilon^2}\right\} \\
&\propto \exp\left\{-\frac{\sum_{i=1}^n \sum_{j=2}^{n_i+1} (\gamma_{ij} - \rho\gamma_{i,j-1})^2}{2\sigma_\epsilon^2}\right\},
\end{aligned}$$

where we use $\Pr(\boldsymbol{\gamma}_{-i1} | \gamma_{i1}, \sigma_\epsilon^2, \rho)$ to approximate $\Pr(\boldsymbol{\gamma}_i | \sigma_\epsilon^2, \rho)$. So we have,

$$\rho | \boldsymbol{\Theta}_{-\rho}, \boldsymbol{\gamma}, \boldsymbol{\omega}, \mathcal{D}_n \sim \mathcal{N}_{(-1,1)}\left(\frac{\sum_{i=1}^n \sum_{j=2}^{n_i+1} \gamma_{ij}\gamma_{i,j-1}}{\sum_{i=1}^n \sum_{j=2}^{n_i+1} \gamma_{i,j-1}^2}, \frac{\sigma_\epsilon^2}{\sum_{i=1}^n \sum_{j=2}^{n_i+1} \gamma_{i,j-1}^2}\right).$$

Bibliography

- Aalen, O. O. and Husebye, E. (1991). Statistical analysis of repeated events forming renewal processes. *Statistics in Medicine*, 10(8):1227–1240.
- Almeida, C. and Czado, C. (2012). Efficient bayesian inference for stochastic time-varying copula models. *Computational Statistics and Data Analysis*, 56(6):1511–1527.
- Aloui, R., Hammoudeh, S., and Nguyen, D. K. (2013). A time-varying copula approach to oil and stock market dependence: The case of transition economies. *Energy Economics*, 39:208–221.
- Andersen, P. K., Borgan, O., Gill, R. D., and Keiding, N. (2012). *Statistical Models Based on Counting Processes*. Springer Science & Business Media.
- Andersen, P. K. and Gill, R. D. (1982). Cox’s regression model for counting processes: a large sample study. *The Annals of Statistics*, pages 1100–1120.
- Anderson, D. G. (1965). Gaussian quadrature formulae for $\int_0^1 \ln(x) f(x) dx$. *Mathematics of Computation*, pages 477–481.
- Arabadzhieva, D., Kaprelyan, A., Tzoukeva, A., and Kanazirev, B. (2015). Cardiovascular comorbidity in acute ischemic stroke. *Scientific cooperations medical workshops*.
- Bansal, N., Katz, R., Robinson-Cohen, C., Odden, M. C., Dalrymple, L., Shlipak, M. G., Sarnak, M. J., Siscovick, D. S., Zelnick, L., Psaty, B. M., et al. (2017). Absolute rates of heart failure, coronary heart disease, and stroke in chronic kidney disease: an analysis of 3 community-based cohort studies. *JAMA cardiology*, 2(3):314–318.
- Bellera, C. A., MacGrogan, G., Debled, M., de Lara, C. T., Brouste, V., and Mathoulin-Pélissier, S. (2010). Variables with time-varying effects and the cox

- model: some statistical concepts illustrated with a prognostic factor study in breast cancer. *BMC Medical Research Methodology*, 10(1):20.
- Buckley, J. and James, I. (1979). Linear regression with censored data. *Biometrika*, 66(3):429–436.
- 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.
- Chen, C.-M., Chuang, Y.-W., and Shen, P.-S. (2015). Two-stage estimation for multivariate recurrent event data with a dependent terminal event. *Biometrical Journal*, 57(2):215–233.
- Cheng, Y. and Fine, J. P. (2012). Cumulative incidence association models for bivariate competing risks data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 74(2):183–202.
- Clayton, D. G. (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, R. J. and Lawless, J. (2007). *The statistical analysis of recurrent events*. Springer Science & Business Media.
- Cook, R. J. and Lawless, J. F. (1997). Marginal analysis of recurrent events and a terminating event. *Statistics in Medicine*, 16(8):911–924.
- Cook, R. J., Lawless, J. F., and Lee, K.-A. (2010). A copula-based mixed poisson model for bivariate recurrent events under event-dependent censoring. *Statistics in Medicine*, 29(6):694–707.
- Coull, A. J. and Rothwell, P. M. (2004). Underestimation of the early risk of recurrent stroke: evidence of the need for a standard definition. *Stroke*, 35:1925–1929.
- Cox, D. R. (1992). Regression models and life-tables. In *Breakthroughs in Statistics*, pages 527–541. Springer.
- Dabrowska, D. M., Sun, G.-W., and Horowitz, M. M. (1994). Cox regression in a markov renewal model: an application to the analysis of bone marrow transplant data. *Journal of the American Statistical Association*, 89(427):867–877.
- 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, R. M., and Adams, R. J. (2010). Risk of recurrent stroke, myocardial infarction, or death in hospitalized stroke patients. *Neurology*, 74(7):588–593.
- Fu, H., Wang, Y., Liu, J., Kulkarni, P. M., and Melemed, A. S. (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.
- 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, W. R., 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.
- Hastings, W. K. (1970). Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109.
- 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.
- Huang, X. and Liu, L. (2007). A joint frailty model for survival and gap times between recurrent events. *Biometrics*, 63(2):389–397.
- Jansen, H., Jansch, A., Breitling, L. P., Hoppe, L., Dallmeier, D., Schmucker, R., Brenner, H., Koenig, W., and Rothenbacher, D. (2018). Hs-troponins for the prediction of recurrent cardiovascular events in patients with established chd—a comparative analysis from the karola study. *International journal of cardiology*, 250:247–252.
- Joe, H. (2005). Asymptotic efficiency of the two-stage estimation method for copula-based models. *Journal of Multivariate Analysis*, 94(2):401–419.
- Kalbfleisch, J. D., Schaubel, D. E., Ye, Y., and Gong, Q. (2013). An estimating function approach to the analysis of recurrent and terminal events. *Biometrics*, 69(2):366–374.
- Kelly, P. J. and Lim, L. L.-Y. (2000). Survival analysis for recurrent event data: an application to childhood infectious diseases. *Statistics in Medicine*, 19(1):13–33.

- Kuo, L. and Mallick, B. (1997). Bayesian semiparametric inference for the accelerated failure-time model. *Canadian Journal of Statistics*, 25(4):457–472.
- Laplace, P. S. (1986). Memoir on the probability of the causes of events. *Statistical Science*, 1(3):364–378.
- Lawless, J. F. (1987). Regression methods for poisson process data. *Journal of the American Statistical Association*, 82(399):808–815.
- Lawless, J. F. and Nadeau, C. (1995). Some simple robust methods for the analysis of recurrent events. *Technometrics*, 37(2):158–168.
- Li, B. and Zeng, Z. (2018). Time-varying dependence structures of equity markets of china, asean and the usa. *Applied Economics Letters*, 25(2):87–91.
- 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, D. and Ying, Z. (1995). Semiparametric inference for the accelerated life model with time-dependent covariates. *Journal of statistical planning and inference*, 44(1):47–63.
- Lin, L.-A., Luo, S., Chen, B. E., and Davis, B. R. (2017). Bayesian analysis of multi-type recurrent events and dependent termination with nonparametric covariate functions. *Statistical methods in medical research*, 26(6):2869–2884.
- Liu, L., Huang, X., Yaroshinsky, A., and Cormier, J. N. (2016). Joint frailty models for zero-inflated recurrent events in the presence of a terminal event. *Biometrics*, 72(1):204–214.
- Liu, L., Wolfe, R. A., and Huang, X. (2004). Shared frailty models for recurrent events and a terminal event. *Biometrics*, 60(3):747–756.
- Louis, T. A. (1982). Finding the observed information matrix when using the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 226–233.
- Lunn, M. and McNeil, D. (1995). Applying cox regression to competing risks. *Biometrics*, pages 524–532.
- Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics*, 11(2):431–441.

- Martinussen, T. and Scheike, T. H. (2002). A flexible additive multiplicative hazard model. *Biometrika*, 89(2):283–298.
- 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.
- McGilchrist, C. and Aisbett, C. (1991). Regression with frailty in survival analysis. *Biometrics*, pages 461–466.
- Meyer, R. and Romeo, J. S. (2015). Bayesian semiparametric analysis of recurrent failure time data using copulas. *Biometrical Journal*, 57(6):982–1001.
- Neal, R. M. (1998). Suppressing random walks in markov chain monte carlo using ordered overrelaxation. In *Learning in graphical models*, pages 205–228. Springer.
- Nelsen, R. B. (1999). Introduction. In *An Introduction to Copulas*, pages 1–4. Springer.
- Oakes, D. (1989). Bivariate survival models induced by frailties. *Journal of the American Statistical Association*, 84(406):487–493.
- Psaty, B. M., Furberg, C. D., Kuller, L. H., Cushman, M., Savage, P. J., Levine, D., O’leary, D. H., Bryan, R. N., Anderson, M., and Lumley, T. (2001). Association between blood pressure level and the risk of myocardial infarction, stroke, and total mortality: the cardiovascular health study. *Archives of internal medicine*, 161(9):1183–1192.
- Ramsay, J. O. (1988). Monotone regression splines in action. *Statistical Science*, pages 425–441.
- Ridker, P. M., Rifai, N., Pfeffer, M., Sacks, F., Lepage, S., Braunwald, E., et al. (2000). Elevation of tumor necrosis factor- α and increased risk of recurrent coronary events after myocardial infarction. *Circulation*, 101(18):2149–2153.
- Rivest, L.-P. and Wells, M. T. (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.

- Rondeau, V., Schaffner, E., Corbière, F., Gonzalez, J. R., 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.
- Shih, J. H. and Louis, T. A. (1995). Inferences on the association parameter in copula models for bivariate survival data. *Biometrics*, pages 1384–1399.
- Sy, J. P. and Taylor, J. M. (2000). Estimation in a cox proportional hazards cure model. *Biometrics*, 56(1):227–236.
- Tanner, M. A. and Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American statistical Association*, 82(398):528–540.
- Vaupel, J. W., Manton, K. G., and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 16(3):439–454.
- Wang, K., Chen, Y., and Huang, S. (2011). The dynamic dependence between the chinese market and other international stock markets: A time-varying copula approach. *International Review of Economics & Finance*, 20(4):654–664.
- Wang, W. and Wells, M. T. (1997). Nonparametric estimators of the bivariate survival function under simplified censoring conditions. *Biometrika*, 84(4):863–880.
- Wei, L.-J. (1992). The accelerated failure time model: a useful alternative to the cox regression model in survival analysis. *Statistics in medicine*, 11(14-15):1871–1879.
- Yanez, N. D., Burke, G. L., Manolio, T., Gardin, J. M., and Polak, J. (2009). Sibling history of myocardial infarction or stroke and risk of cardiovascular disease in the elderly: the cardiovascular health study. *Annals of epidemiology*, 19(12):858–866.
- Ye, Y., Kalbfleisch, J. D., and Schaubel, D. E. (2007). Semiparametric analysis of correlated recurrent and terminal events. *Biometrics*, 63(1):78–87.
- Yu, Z., Liu, L., Bravata, D. M., and Williams, L. S. (2014). Joint model of recurrent events and a terminal event with time-varying coefficients. *Biometrical Journal*, 56(2):183–197.
- Yue, H. and Chan, K. (1997). A dynamic frailty model for multivariate survival data. *Biometrics*, pages 785–793.
- Zeng, D., Chen, Q., and Ibrahim, J. G. (2009). Gamma frailty transformation models for multivariate survival times. *Biometrika*, 96(2):277–291.

- Zeng, D., Ibrahim, J. G., Chen, M.-H., Hu, K., and Jia, C. (2014). Multivariate recurrent events in the presence of multivariate informative censoring with applications to bleeding and transfusion events in myelodysplastic syndrome. *Journal of Biopharmaceutical Statistics*, 24(2):429–442.
- Zeng, D. and Lin, D. (2009). Semiparametric transformation models with random effects for joint analysis of recurrent and terminal events. *Biometrics*, 65(3):746–752.
- Zheng, M. and Klein, J. P. (1995). Estimates of marginal survival for dependent competing risks based on an assumed copula. *Biometrika*, 82(1):127–138.
- Zhu, L., Sun, J., Tong, X., and Srivastava, D. K. (2010). Regression analysis of multivariate recurrent event data with a dependent terminal event. *Lifetime data analysis*, 16(4):478–490.

Vita
Zheng Li

Education

- Ph.D., Biostatistics** Expected May, 2018
The Pennsylvania State University, Hershey, PA
- Master of Science in Public Health, Biotatistics** May, 2014
Emory University, Atlanta, GA
- Bachelor of Science, Statistics** July, 2011
Shanghai University of Finance and Economics, Shanghai, China

Recent Publications

1. **Li, Z.**, Chinchilli, V.M., Wang, M. (2017). A Bayesian Joint Frailty-copula Model for Recurrent Events and a Terminal Event. Manuscript submitted for publication.
2. Wang, M., Matthews, S. A., Iskandarani, K., Li, Y., **Li, Z.**, Chinchilli, V. M., and Zhang, L. (2017). Spatial-temporal analysis of prostate cancer incidence from the Pennsylvania Cancer Registry, 2000-2011. *Geospatial Health*, 12(2).
3. Wang, M., **Li, Z.**, Lee, E. Y., Lewis, M. M., Zhang, L., Sterling, N. W., ..., Huang, X. (2017). Predicting the multi-domain progression of Parkinsons disease: a Bayesian multivariate generalized linear mixed-effect model. *BMC Medical Research Methodology*, 17(1), 147.
4. Chen, Z.,**Li, Z.**, Zhuang, R., Yuan, Y., Kutner, M., Owonikoko, T., ..., Kowalski, J. (2017). Adaptive Estimation of Personalized Maximum Tolerated Dose in Cancer Phase I Clinical Trials Based on All Toxicities and Individual Genomic Profile. *PloS one*, 12(1), e0170187.
5. Wang, M., **Li, Z.**, Reeves, B. (2016). Re:Receiver operating characteristic curve inference from a sample with a limit of detection. *American Journal of Epidemiology*, 184(7), 552-553.
6. Wang, M., Kong, L., **Li, Z.**, Zhang, L. (2015). Covariance estimators for generalized estimating equations (GEE) in longitudinal analysis with small samples. *Statistics in Medicine*.