

Article

An MM Algorithm for the Frailty-Based Illness Death Model with Semi-Competing Risks Data

Xifen Huang ¹, Jinfeng Xu ^{2,*}, Hao Guo ¹, Jianhua Shi ² and Wenjie Zhao ²¹ School of Mathematics, Yunnan Normal University, Kunming 650092, China² School of Mathematics, Minnan Normal University, Zhangzhou 363000, China

* Correspondence: xjf@mnnu.edu.cn

Abstract: For analyzing multiple events data, the illness death model is often used to investigate the covariate–response association for its easy and direct interpretation as well as the flexibility to accommodate the within-subject dependence. The resulting estimation and inferential procedures often depend on the subjective specification of the parametric frailty distribution. For certain frailty distributions, the computation can be challenging as the estimation involves both the nonparametric component and the parametric component. In this paper, we develop efficient computational methods for analyzing semi-competing risks data in the illness death model with the general frailty, where the Minorization–Maximization (MM) principle is employed for yielding accurate estimation and inferential procedures. Simulation studies are conducted to assess the finite-sample performance of the proposed method. An application to a real data is also provided for illustration.

Keywords: semi-competing risk gamma frailty model; MM algorithm; marginal likelihood; surrogate function; colon cancer data

MSC: 62N01; 62N02

Citation: Huang, X.; Xu, J.; Guo, H.; Shi, J.; Zhao, W. An MM Algorithm for the Frailty-Based Illness Death Model with Semi-Competing Risks Data. *Mathematics* **2022**, *10*, 3702. <https://doi.org/10.3390/math10193702>

Academic Editors: Min Wang, Haijun Gong, Liucang Wu and Songfeng Zheng

Received: 16 August 2022

Accepted: 28 September 2022

Published: 10 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In biomedical studies, one subject may experience multiple types of event times. For example, a subject can experience a terminal event such as death and a non-terminal event such as disease recurrence where the terminal event can censor the non-terminal event but not vice versa. The two event times are usually dependent. This gives rise to semi-competing risks data in which the follow-up of a non-terminal event could be stopped by a terminal event [1]. Such dependency can be modeled by copula; for instance, Ref. [1] investigated the degree of association between the two events under the Clayton copula model, and [2] extended their works to a class of more general copula models and established the asymptotic normality of this estimator. In addition, the Archimedean copula was applied by [3] to model such dependency.

In addition to the copula approach, the semi-competing risk data can be modeled by illness–death model by using the shared frailty to describe the dependence of two events' time. Many researchers ([4–8]) have developed semi-parametric frailty models to analyze the semi-competing risks data, and semi-parametric regression analyses ([9,10]) were also discussed by many researchers. In particular, Ref. [4] proposed a marginal-likelihood approach under the semi-parametric gamma frailty model. Ref. [11] proposed a joint frailty–copula model in the meta-analysis of individual patient data with semi-competing risks. Moreover, Refs. [5,6] proposed semi-parametric Bayesian approaches for the semi-competing risks data. Unlike the classical likelihood, which only involves fixed parameters, the hierarchical likelihood is constructed for both fixed parameters and unobserved frailties at the same time for the semi-parametric shared frailty model by [7]. In addition, Ref. [12] proposed a broad class of semi-parametric transformation models with random effects for

the joint analysis of recurrent events and a terminal event. Ref. [8] developed a novel semi-parametric transformation model for the analysis of semi-competing risks data. In general, the computation involved in the semi-parametric frailty models is usually intensive, and the essential cause comes from the calculation of multiple non-parametric baseline cumulative hazard functions and intractable integrals over the frailty distributions.

Inferences for such semi-competing risks data are generally focused on the covariate treatment effects on the rates of terminal and non-terminal events and the association between the two types of events ([9,13–15]). Some recent research also focused on the parameter estimation procedure of semi-competing risk models.

Ref. [16] provided R packages with many algorithms to deal with independent and cluster-correlated semi-competing risk data. Ref. [17] applied the Bayesian approach to their proposed algorithm to conduct variable selection along with the parameter estimation, but their method is computationally intensive. Similarly, a penalized approach has been applied to this model by [18], and they applied a more efficient proximal gradient method which requires a proper warm start. In addition, Ref. [19] proposed an MCEM scheme to update the parameter estimates. The drawback of this method is that the performance of estimation is highly reliant on the MCMC sample approximates in the E-step, which also requires an accurate starting value. Different from their methods, for the parameter estimation part, we propose applying the MM algorithm in order to achieve better computational efficiency.

The MM algorithm is an important and powerful tool for optimization problems because it can simplify a difficult optimization problem by decomposing a high-dimensional objective function into separable low-dimensional functions. So far, the MM algorithm has a broad range of applications in the field of statistics, such as proportional odds model ([20]), the shared frailty model ([21]), the quantile regression ([22]), and so on. In this paper, we used the MM principle and proposed a profile MM algorithm for the semi-competing risk shared frailty model, which facilitates its pertinent use in high-dimensional situations.

The rest of the paper is organized as follows. In Section 2, we introduce the semi-competing risk shared frailty model. Section 3 presents the estimation procedures based on the MM method. In Section 4, we provide two simulation studies to assess their practical performance. Section 5 illustrates the method by a real data analysis. Some concluding remarks and discussions are given in Section 6.

2. The Semi-Competing Risk Model with Gamma Frailty

Let C_i , T_{i1} , and T_{i2} be the censoring, non-terminal, and terminal event times for the i -th subject, respectively, $i = 1, \dots, n$. Denote by x_i a p -dimensional vector of covariates of the i -th subject. Assume that the censoring time C_i is independent of T_{i1} and T_{i2} given x_i . The observations can be summarized as $Y_{obs} = \{Y_{i1} = T_{i1} \wedge Y_{i2}, Y_{i2} = T_{i2} \wedge C_i, \delta_{i1} = I(T_{i1} \leq Y_{i2}), \delta_{i2} = I(T_{i2} \leq C_i), x_i; i = 1 \dots, n\}$. If the subject fails before the non-terminal event occurs, then we define $T_{i1} = \infty$. The illness–death multi-state model for semi-competing risks data has three states, which are characterized by three intensity or hazard functions:

$$\lambda_1(t_1) = \lim_{\Delta \rightarrow 0} P[T_1 \in [t_1, t_1 + \Delta) | T_1 \geq t_1, T_2 \geq t_1] / \Delta, \quad t_1 > 0, \tag{2.1}$$

$$\lambda_2(t_2) = \lim_{\Delta \rightarrow 0} P[T_2 \in [t_2, t_2 + \Delta) | T_1 \geq t_2, T_2 \geq t_2] / \Delta, \quad t_2 > 0, \tag{2.2}$$

$$\lambda_{12}(t_2|t_1) = \lim_{\Delta \rightarrow 0} P[T_2 \in [t_2, t_2 + \Delta) | T_1 = t_1, T_2 \geq t_2] / \Delta, \quad 0 < t_1 < t_2. \tag{2.3}$$

Equations (2.1) and (2.2) are the usual rough hazard functions for the competing risk portion of the model where a non-terminal or terminal event occurs first. Then, Equation (2.3) defines the hazard rate of the terminal event following the occurrence of the non-terminal event. Usually, $\lambda_{12}(t_2|t_1)$ depend on both t_1 and t_2 . Followed by [4], since the dependence between non-terminal and terminal event times will be modeled by shared frailty later, here we assume a Markov process where the transition probability from the non-terminal event to the terminal event does not depend on the duration of the non-terminal event, i.e., $\lambda_{12}(t_2|t_1) = \lambda_{12}(t_2)$ only depends on t_2 .

In practice, Ref. [4] extended the classical semi-competing risk model by incorporating a common frailty (random effect) to model the dependency between non-terminal and terminal event times. That is, given the random effect ω and covariates x , the Cox-type multiplicative models can be expressed in the following form:

$$\lambda_1(t_1|x, \omega) = \omega\lambda_{01}(t_1) \exp\{\beta_1x\}, \quad t_1 > 0, \tag{2.4}$$

$$\lambda_2(t_2|x, \omega) = \omega\lambda_{02}(t_2) \exp\{\beta_2x\}, \quad t_2 > 0, \tag{2.5}$$

$$\lambda_{12}(t_2|t_1, x, \omega) = \omega\lambda_{03}(t_2) \exp\{\beta_3x\}, \quad 0 < t_1 < t_2, \tag{2.6}$$

where $\lambda_{01}(t_1)$, $\lambda_{02}(t_2)$, and $\lambda_{03}(t_2)$ are three baseline hazard functions and ω is a subject-specific random effect or frailty. Usually, the frailties $\omega_i (i = 1, \dots, n)$ are assumed to be independent and identically distributed with a density function with a frailty parameter θ . The common distributions assumed for ω are Gamma($1/\theta, 1/\theta$), Inverse Gaussian (θ, θ^2), and Log-normal($0, \theta$). Since the likelihood of semi-competing risk model with Gamma frailty has the explicit form of expression, we assume that the frailty ω has a Gamma distribution with mean 1, variance θ in the following section, and only show the estimation procedures for the semi-competing risk models (2.4)–(2.6) with $\omega \sim \text{Gamma}(1/\theta, 1/\theta)$, i.e.,

$$g(\omega) = \frac{\omega^{(1/\theta-1)} \exp(-\omega/\theta)}{\Gamma(1/\theta)\theta^{1/\theta}}, \quad \theta > 0.$$

Note that θ measures the dependence between non-terminal and terminal event times and a larger θ indicates a stronger dependence.

For simplicity of expression, we sum up the regression parameters by $\beta = (\beta_1, \beta_2, \beta_3)$ and summarize the three baseline hazard by $\Lambda_0 = (\Lambda_{01}, \Lambda_{02}, \Lambda_{03})$. Then, the parameters of the semi-competing risk model consists of three parts, and we sum up all parameters by $\alpha = (\theta, \beta, \Lambda_0)$.

3. The Estimation via MM Method

3.1. Philosophy of the MM Principle

Assume $\arg \max_{\alpha \in \Theta} \ell(\alpha|Y_{obs})$ is our maximization problem, where $\ell(\alpha|Y_{obs})$ is the objective log-likelihood function, $\alpha = (\alpha_1, \dots, \alpha_q)^T \in \Theta$ are the vector of parameters to be estimated, and Θ is the parameter space.

For such maximization problems, the MM principle provides a general frame for constructing iterative algorithms with monotone convergence which involves two M steps. The first M step aims to construct a surrogate function $Q(\alpha|\alpha^{(k)})$ by a series of algebraic inequalities under the following conditions:

$$Q(\alpha|\alpha^{(k)}) \leq \ell(\alpha|Y_{obs}), \forall \alpha, \alpha^{(k)} \in \Theta, Q(\alpha^{(k)}|\alpha^{(k)}) = \ell(\alpha^{(k)}|Y_{obs}),$$

where $\alpha^{(k)}$ denotes the current estimate of α in the k -th iteration. Once the surrogate function is constructed, the second M step is to maximize the surrogate function $Q(\cdot|\alpha^{(k)})$ instead of the objective log-likelihood function $\ell(\alpha|Y_{obs})$. Then, we update $\alpha^{(k)}$ by $\alpha^{(k+1)}$ as follows:

$$\alpha^{(k+1)} = \arg \max_{\alpha \in \Theta} Q(\alpha|\alpha^{(k)}).$$

3.2. The Estimation Procedure

From Equations (2.4)–(2.6), the observed likelihood function of the semi-competing risk Gamma frailty model can be written as the form of

$$\begin{aligned}
 &L(\theta, \beta, \Lambda_0 | Y_{obs}) \\
 &= \prod_{i=1}^n \left(\lambda_{01}(Y_{i1})^{\delta_{i1}} \lambda_{02}(Y_{i2})^{\delta_{i2}(1-\delta_{i1})} \lambda_{03}(Y_{i2})^{\delta_{i1}\delta_{i2}} (1 + \theta)^{\delta_{i1}\delta_{i2}} \right. \\
 &\quad \times \exp[\delta_{i1}\beta_1 x_i + \delta_{i2}(1 - \delta_{i1})\beta_2 x_i + \delta_{i1}\delta_{i2}\beta_3 x_i] \\
 &\quad \left. \times \{1 + \theta[\Lambda_{01}(Y_{i1})e^{\beta_1 x_i} + \Lambda_{02}(Y_{i2})e^{\beta_2 x_i} + \Lambda_{03}(Y_{i1}, Y_{i2})e^{\beta_3 x_i}]\}^{-\frac{1}{\theta} - \delta_{i1} - \delta_{i2}} \right),
 \end{aligned}$$

then, we have the objective log-likelihood function as follows:

$$\begin{aligned}
 &l(\theta, \beta, \Lambda_0 | Y_{obs}) \\
 &= \sum_{i=1}^n \left(\delta_{i1} \log \lambda_{01}(Y_{i1}) + \delta_{i2}(1 - \delta_{i1}) \log \lambda_{02}(Y_{i2}) + \delta_{i1}\delta_{i2} \log \lambda_{03}(Y_{i2}) \right. \\
 &\quad + \delta_{i1}\beta_1 x_i + \delta_{i2}(1 - \delta_{i1})\beta_2 x_i + \delta_{i1}\delta_{i2}\beta_3 x_i + \delta_{i1}\delta_{i2} \log(1 + \theta) \\
 &\quad \left. - \left(\frac{1}{\theta} + \delta_{i1} + \delta_{i2}\right) \log\{1 + \theta[\Lambda_{01}(Y_{i1})e^{\beta_1 x_i} + \Lambda_{02}(Y_{i2})e^{\beta_2 x_i} + \Lambda_{03}(Y_{i1}, Y_{i2})e^{\beta_3 x_i}]\} \right). \tag{3.1}
 \end{aligned}$$

Based on the MM principle, it is necessary to find a surrogate function for the objective log-likelihood function $l(\theta, \beta, \Lambda_0 | Y_{obs})$ in (3.1). We first denote:

$$A_i^{(k)} = 1 + \theta^k [\Lambda_{01}^{(k)}(Y_{i1})e^{\beta_1^{(k)} x_i} + \Lambda_{02}^{(k)}(Y_{i2})e^{\beta_2^{(k)} x_i} + \Lambda_{03}^{(k)}(Y_{i1}, Y_{i2})e^{\beta_3^{(k)} x_i}]$$

and utilize the supporting hyperplane inequality $-\log(x) \geq -\log(x_0) - \frac{x-x_0}{x_0}$ to deal with the last term of (3.1), then we have the temporary surrogate function as follows:

$$\begin{aligned}
 &Q_1(\theta, \beta, \Lambda_0 | \theta^{(k)}, \beta^{(k)}, \Lambda_0^{(k)}) \\
 &= \sum_{i=1}^n \left\{ \delta_{i1} \log \lambda_{01}(Y_{i1}) + \delta_{i2}(1 - \delta_{i1}) \log \lambda_{02}(Y_{i2}) + \delta_{i1}\delta_{i2} \log \lambda_{03}(Y_{i2}) + \delta_{i1}\beta_1 x_i \right. \\
 &\quad + \delta_{i2}(1 - \delta_{i1})\beta_2 x_i + \delta_{i1}\delta_{i2}\beta_3 x_i + \delta_{i1}\delta_{i2} \log(1 + \theta) - \frac{1}{\theta} \left(\log A_i^{(k)} + \frac{1}{A_i^{(k)}} - 1 \right) \\
 &\quad - \frac{\Lambda_{01}(Y_{i1})e^{\beta_1 x_i} + \Lambda_{02}(Y_{i2})e^{\beta_2 x_i} + \Lambda_{03}(Y_{i1}, Y_{i2})e^{\beta_3 x_i}}{A_i^{(k)}} \\
 &\quad \left. - \frac{(\delta_{i1} + \delta_{i2})\theta[\Lambda_{01}(Y_{i1})e^{\beta_1 x_i} + \Lambda_{02}(Y_{i2})e^{\beta_2 x_i} + \Lambda_{03}(Y_{i1}, Y_{i2})e^{\beta_3 x_i}]}{A_i^{(k)}} + c_1 \right\}.
 \end{aligned}$$

where c_1 is a constant. Following [23,24], we use the profile estimation method and first profile out $\Lambda_{01}, \Lambda_{02}$ and Λ_{03} from $Q_1(\theta, \beta, \Lambda_0 | \theta^{(k)}, \beta^{(k)}, \Lambda_0^{(k)})$ for any given θ and β , this provides the estimate of $\Lambda_{01}, \Lambda_{02}$, and Λ_{03} given θ and β as:

$$d\hat{\Lambda}_{01}(Y_{i1}) = \frac{\delta_{i1}}{\sum_{j=1}^n I(Y_{j1} \geq Y_{i1})\theta(\frac{1}{\theta} + \delta_{i1} + \delta_{i2})e^{\beta_1 x_j} / A_j^{(k)}}, \tag{3.2}$$

$$d\hat{\Lambda}_{02}(Y_{i2}) = \frac{\delta_{i2}(1 - \delta_{i1})}{\sum_{j=1}^n I(Y_{j2} \geq Y_{i2})\theta(\frac{1}{\theta} + \delta_{i1} + \delta_{i2})e^{\beta_2 x_j} / A_j^{(k)}} \tag{3.3}$$

$$d\hat{\Lambda}_{03}(Y_{i2}) = \frac{\delta_{i2}\delta_{i1}}{\sum_{j=1}^n I(Y_{j2} \geq Y_{i2} > Y_{j1})\theta(\frac{1}{\theta} + \delta_{i1} + \delta_{i2})e^{\beta_3 x_j} / A_j^{(k)}} \tag{3.4}$$

Substituting (3.2)–(3.4) into $Q_1(\theta, \beta, \Lambda_0|\theta^{(k)}, \beta^{(k)}, \Lambda_0^{(k)})$, we have:

$$\begin{aligned}
 & Q_2(\theta, \beta|\theta^{(k)}, \beta^{(k)}, \Lambda_0^{(k)}) \\
 &= \sum_{i=1}^n \left\{ -\delta_{i1} \log \left[\sum_{j=1}^n I(Y_{j1} \geq Y_{i1}) \theta \left(\frac{1}{\theta} + \delta_{i1} + \delta_{i2} \right) e^{\beta_1 x_j / A_j^{(k)}} \right] \right. \\
 &\quad - \delta_{i2} (1 - \delta_{i1}) \log \left[\sum_{j=1}^n I(Y_{j2} \geq Y_{i2}) \theta \left(\frac{1}{\theta} + \delta_{i1} + \delta_{i2} \right) e^{\beta_2 x_j / A_j^{(k)}} \right] \\
 &\quad - \delta_{i1} \delta_{i2} \log \left[\sum_{j=1}^n I(Y_{j2} \geq Y_{i2} > Y_{j1}) \theta \left(\frac{1}{\theta} + \delta_{i1} + \delta_{i2} \right) e^{\beta_3 x_j / A_j^{(k)}} \right] \\
 &\quad + \delta_{i1} \beta_1 x_i + \delta_{i2} (1 - \delta_{i1}) \beta_2 x_i + \delta_{i1} \delta_{i2} \beta_3 x_i + \delta_{i1} \delta_{i2} \log(1 + \theta) \\
 &\quad \left. - \frac{1}{\theta} (\log A_i^{(k)} + 1/A_i^{(k)} - 1) + c_2 \right\},
 \end{aligned} \tag{3.5}$$

where c_2 is a constant. For ease of expression, we further denote:

$$\begin{aligned}
 B_{1i}^{(k)} &= \sum_{j=1}^n I(Y_{j1} \geq Y_{i1}) \theta^{(k)} \left(\frac{1}{\theta^{(k)}} + \delta_{i1} + \delta_{i2} \right) e^{\beta_1 x_j / A_j^{(k)}}, \\
 B_{2i}^{(k)} &= \sum_{j=1}^n I(Y_{j2} \geq Y_{i2}) \theta^{(k)} \left(\frac{1}{\theta^{(k)}} + \delta_{i1} + \delta_{i2} \right) e^{\beta_2 x_j / A_j^{(k)}}, \\
 B_{2i}^{(k)} &= \sum_{j=1}^n I(Y_{j2} \geq Y_{i2} > Y_{j1}) \theta^{(k)} \left(\frac{1}{\theta^{(k)}} + \delta_{i1} + \delta_{i2} \right) e^{\beta_3 x_j / A_j^{(k)}}.
 \end{aligned}$$

and also use the supporting hyperplane inequality to deal with the first three terms of $Q_2(\theta, \beta|\theta^{(k)}, \beta^{(k)}, \Lambda_0^{(k)})$ in Equation (3.5), we obtain the surrogate function:

$$\begin{aligned}
 & Q_3(\theta, \beta|\theta^{(k)}, \beta^{(k)}, \Lambda_0^{(k)}) \\
 &= \sum_{i=1}^n \left\{ -\frac{\delta_{i1}}{B_{1i}^{(k)}} \sum_{j=1}^n I(Y_{j1} \geq Y_{i1}) \left[\frac{e^{\beta_1 x_j}}{A_j^{(k)}} + \frac{(\delta_{j1} + \delta_{j2})}{A_j^{(k)}} \theta e^{\beta_1 x_j} \right] \right. \\
 &\quad - \frac{\delta_{i1} (1 - \delta_{i1})}{B_{2i}^{(k)}} \sum_{j=1}^n I(Y_{j2} \geq Y_{i2}) \left[\frac{e^{\beta_2 x_j}}{A_j^{(k)}} + \frac{(\delta_{j1} + \delta_{j2})}{A_j^{(k)}} \theta e^{\beta_2 x_j} \right] \\
 &\quad - \frac{\delta_{i1} \delta_{i2}}{B_{3i}^{(k)}} \sum_{j=1}^n I(Y_{j2} \geq Y_{i2} > Y_{j1}) \left[\frac{e^{\beta_3 x_j}}{A_j^{(k)}} + \frac{(\delta_{j1} + \delta_{j2})}{A_j^{(k)}} \theta e^{\beta_3 x_j} \right] \\
 &\quad + \delta_{i1} \beta_1 x_i + \delta_{i2} (1 - \delta_{i1}) \beta_2 x_i + \delta_{i1} \delta_{i2} \beta_3 x_i \\
 &\quad \left. + \delta_{i1} \delta_{i2} \log(1 + \theta) - \frac{1}{\theta} \left(\log A_i^{(k)} + \frac{1}{A_i^{(k)}} - 1 \right) \right\}.
 \end{aligned} \tag{3.6}$$

To separate the parameters θ with each β_1, β_2 , and β_3 , we further minorize the components $\theta \exp(\beta_1 x_j)$, $\theta \exp(\beta_2 x_j)$ and $\theta \exp(\beta_3 x_j)$ in $Q_3(\theta, \beta|\theta^{(k)}, \beta^{(k)}, \Lambda_0^{(k)})$ by the inequality:

$$-\frac{\theta \exp(\beta_1 x_j)}{\theta^{(k)} \exp(\beta_1^{(k)} x_j)} \geq -\frac{1}{2} \left(\frac{\theta}{\theta^{(k)}} \right)^2 - \frac{1}{2} \frac{\exp(2\beta_1 x_j)}{\exp(2\beta_1^{(k)} x_j)},$$

then we obtain the final surrogate function:

$$\begin{aligned}
 & Q_4(\theta, \beta | \theta^{(k)}, \beta^{(k)}, \Lambda_0^{(k)}) \\
 = & \sum_{i=1}^n \left\{ -\frac{\delta_{i1}}{B_{1i}^{(k)}} \sum_{j=1}^n I(Y_{j1} \geq Y_{i1}) \frac{e^{\beta_1 x_j}}{A_j^{(k)}} - \frac{\delta_{i2}(1 - \delta_{i1})}{B_{2i}^{(k)}} \sum_{j=1}^n I(Y_{j2} \geq Y_{i2}) \frac{e^{\beta_2 x_j}}{A_j^{(k)}} \right. \\
 & - \frac{\delta_{i1} \delta_{i2}}{B_{3i}^{(k)}} \sum_{j=1}^n I(Y_{j2} \geq Y_{i2} > Y_{j1}) \frac{e^{\beta_3 x_j}}{A_j^{(k)}} \\
 & - \left. \frac{\delta_{i1}}{B_{1i}^{(k)}} \sum_{j=1}^n I(Y_{j1} \geq Y_{i1}) \frac{\delta_{j1} + \delta_{j2}}{A_j^{(k)}} \left[\frac{\exp(\beta_1^{(k)} x_j)}{2\theta^{(k)}} \theta^2 + \frac{\theta^{(k)}}{2 \exp(\beta_1^{(k)} x_j)} \exp(2\beta_1 x_j) \right] \right. \\
 & - \left. \frac{\delta_{i2}(1 - \delta_{i1})}{B_{2i}^{(k)}} \sum_{j=1}^n I(Y_{j2} \geq Y_{i2}) \frac{\delta_{j1} + \delta_{j2}}{A_j^{(k)}} \left[\frac{\exp(\beta_2^{(k)} x_j)}{2\theta^{(k)}} \theta^2 + \frac{\theta^{(k)}}{2 \exp(\beta_2^{(k)} x_j)} \exp(2\beta_2 x_j) \right] \right. \\
 & - \left. \frac{\delta_{i1} \delta_{i2}}{B_{3i}^{(k)}} \sum_{j=1}^n I(Y_{j2} \geq Y_{i2} > Y_{j1}) \frac{\delta_{j1} + \delta_{j2}}{A_j^{(k)}} \left[\frac{\exp(\beta_3^{(k)} x_j)}{2\theta^{(k)}} \theta^2 + \frac{\theta^{(k)}}{2 \exp(\beta_3^{(k)} x_j)} \exp(2\beta_3 x_j) \right] \right\} \\
 & + \delta_{i1} \beta_1 x_i + \delta_{i2}(1 - \delta_{i1}) \beta_2 x_i + \delta_{i1} \delta_{i2} \beta_3 x_i + \delta_{i1} \delta_{i2} \log(1 + \theta) - \frac{1}{\theta} \left(\log A_i^{(k)} + \frac{1}{A_i^{(k)}} - 1 \right) \Big\} \\
 \cong & Q_4(\theta | \alpha^{(k)}) + Q_4(\beta_1 | \alpha^{(k)}) + Q_4(\beta_2 | \alpha^{(k)}) + Q_4(\beta_3 | \alpha^{(k)}),
 \end{aligned}$$

where

$$\begin{aligned}
 & Q_4(\theta | \alpha^{(k)}) \\
 = & \sum_{i=1}^n \left\{ \delta_{i1} \delta_{i2} \log(1 + \theta) - \frac{1}{\theta} \left(\log A_i^{(k)} + \frac{1}{A_i^{(k)}} - 1 \right) \right. \\
 & - \frac{\theta^2}{2\theta^{(k)}} \left[\frac{\delta_{i1}}{B_{1i}^{(k)}} \sum_{j=1}^n I(Y_{j1} \geq Y_{i1}) \frac{(\delta_{j1} + \delta_{j2}) \exp(\beta_1^{(k)} x_j)}{A_j^{(k)}} \right. \\
 & + \frac{\delta_{i2}(1 - \delta_{i1})}{B_{2i}^{(k)}} \sum_{j=1}^n I(Y_{j2} \geq Y_{i2}) \frac{(\delta_{j1} + \delta_{j2}) \exp(\beta_2^{(k)} x_j)}{A_j^{(k)}} \\
 & \left. \left. + \frac{\delta_{i1} \delta_{i2}}{B_{3i}^{(k)}} \sum_{j=1}^n I(Y_{j2} \geq Y_{i2} > Y_{j1}) \frac{(\delta_{j1} + \delta_{j2}) \exp(\beta_3^{(k)} x_j)}{A_j^{(k)}} \right] \right\}, \\
 & Q_4(\beta_1 | \alpha^{(k)}) \\
 = & \sum_{i=1}^n \left\{ \delta_{i1} \beta_1 x_i - \frac{\delta_{i1}}{B_{1i}^{(k)}} \sum_{j=1}^n I(Y_{j1} \geq Y_{i1}) \left[\frac{e^{\beta_1 x_j}}{A_j^{(k)}} + \frac{\theta^{(k)} (\delta_{j1} + \delta_{j2}) \exp(2\beta_1 x_j)}{2A_j^{(k)} \exp(\beta_1^{(k)})} \right] \right\}, \\
 & Q_4(\beta_2 | \alpha^{(k)}) \\
 = & \sum_{i=1}^n \left\{ \delta_{i2}(1 - \delta_{i1}) \beta_2 x_i - \frac{\delta_{i2}(1 - \delta_{i1})}{B_{2i}^{(k)}} \sum_{j=1}^n I(Y_{j2} \geq Y_{i2}) \left[\frac{e^{\beta_2 x_j}}{A_j^{(k)}} + \frac{\theta^{(k)} (\delta_{j1} + \delta_{j2}) \exp(2\beta_2 x_j)}{2A_j^{(k)} \exp(\beta_2^{(k)})} \right] \right\}, \\
 & Q_4(\beta_3 | \alpha^{(k)}) \\
 = & \sum_{i=1}^n \left\{ \delta_{i1} \delta_{i2} \beta_3 x_i - \frac{\delta_{i1} \delta_{i2}}{B_{3i}^{(k)}} \sum_{j=1}^n I(Y_{j2} \geq Y_{i2} > Y_{j1}) \left[\frac{e^{\beta_3 x_j}}{A_j^{(k)}} + \frac{\theta^{(k)} (\delta_{j1} + \delta_{j2}) \exp(2\beta_3 x_j)}{2A_j^{(k)} \exp(\beta_3^{(k)})} \right] \right\}.
 \end{aligned}$$

From (3.6), it can be seen that the frailty parameter θ and the regression parameters β_1, β_2 , and β_3 are separated from each other. Accordingly, the resulting MM algorithm only

involves a series of separated univariate optimizations in the next maximization step and matrix inversion is not needed. The Algorithm 1 is stated as follows.

Algorithm 1 The estimation procedures via MM method.

Input: $(\theta^{(0)}, \beta^{(0)}, \Lambda_0^{(0)})$
 $k \leftarrow 1$
while $\epsilon_k > 1e - 7$ **do**
 S1. Update $\theta^{(k)}$ and $\beta^{(k)}$ via Equation (3.6) given $(\theta^{(k-1)}, \beta^{(k-1)}, \Lambda_0^{(k-1)})$
 S2. Estimate $\Lambda_0^{(k)}$ using (3.2)–(3.4) given $\theta^{(k)}$ and $\beta^{(k)}$
 S3. $\epsilon_k = l(\theta^{(k)}, \beta^{(k)}, \Lambda_0^{(k)} | Y_{obs}) - l(\theta^{(k-1)}, \beta^{(k-1)}, \Lambda_0^{(k-1)} | Y_{obs})$
 $k \leftarrow k + 1$
end while
Output: $(\theta^{(k)}, \beta^{(k)}, \Lambda_0^{(k)})$

The variance for the estimates in the model involves three non-parametric baseline transition functions, frailty variance, and regression coefficients as shown in [4]. Hence, there is no readily available plug-in formula for estimating it. However, the resampling method such as the bootstrap can be employed to calibrate it and construct confidence interval and the associate inferential procedures.

4. Simulation Study

To evaluate the finite sample performance of the proposed methods, we conducted the following simulation studies. As emphasized in the Section 2, the purpose of incorporating subject specific frailty terms is to account for dependence, which is not taken into account by the measured covariates.

Scenario 1: We first independently simulate n observations from the semi-competing risk models (2.4) and (2.6) with $\omega \sim \text{Gamma}(1/\theta, 1/\theta)$. We consider one covariate X in the model which follows a standard normal distribution. The censoring times were generated from independent uniform distribution to yield censoring proportions of both a terminal and non-terminal event at around 30% or 50%. Let $\lambda_{01}(t_1) = \lambda_{02}(t_2) = 1$, $\lambda_{03}(t_2) = 2$, and $\beta_1 = \beta_2 = \beta_3 = 0.5$. We choose the true value of θ from $\Omega_\theta = \{0.5, 1, 2\}$ and pick the sample size n from $\Omega_n = \{250, 500\}$.

Based on 500 replications, the average values of the biases (Bias) for estimated frailty and regression parameters, their empirical standard deviations (SD), and the average computation times (T) are summarized in Tables 1 and 2. From the results of Tables 1 and 2, we find that the SD of each parameter is relatively small which indicating the effectiveness of MM algorithm for the illness–death model with gamma frailty. It can also be found that most the SDs of θ are larger than that of $\beta_1, \beta_2, \beta_3$ and $\Lambda_{01}, \Lambda_{02}, \Lambda_{03}$. With the increase in the true values of θ , the SD of θ will increase accordingly, while the SD of θ will decrease with the increase in sample size. When the censoring proportion increases from 30% to 50%, the SD of θ is increased accordingly, but the SDs of other parameters, i.e., $\beta_1, \beta_2, \beta_3$ and $\Lambda_{01}, \Lambda_{02}, \Lambda_{03}$, do not show a clear trend of change. A larger value of θ requires longer computation time.

Furthermore, we plot the estimated cumulative baseline hazard function $\Lambda_{01}(t)$ and $\Lambda_{03}(t)$ by the proposed method based on 20 replications (with red color) together with the true cumulative baseline hazard function $\Lambda_{01}(t) = t$ and $\Lambda_{03}(t) = 2t$ (with black color) under the model with gamma frailty in Figures 1 and 2 with sample size 500. Since both cumulative baseline hazard functions $\Lambda_{01}(t)$ and $\Lambda_{02}(t)$ have the same expression, here, we only plot the estimated cumulative baseline hazard function $\Lambda_{01}(t)$.

Table 1. The simulation results of the illness–death model with gamma frailty at 30% censoring proportion.

θ_{true}	Parameter	$n = 250$		Parameter	$n = 500$	
		Bias	SD		Bias	SD
0.5		T = 0.1957			T = 0.6736	
	θ	−0.088	0.128	θ	−0.072	0.091
	β_1	0.045	0.173	β_1	−0.032	0.066
	β_2	0.032	0.196	β_2	−0.021	0.077
	β_3	0.044	0.118	β_3	−0.023	0.087
	$\Lambda_{01}(1)$	−0.036	0.080	$\Lambda_{01}(1)$	−0.024	0.063
	$\Lambda_{02}(1)$	0.023	0.061	$\Lambda_{02}(1)$	−0.013	0.044
	$\Lambda_{03}(1)$	0.045	0.302	$\Lambda_{03}(1)$	−0.038	0.227
1		T = 0.2576			T = 0.8667	
	θ	0.098	0.173	θ	0.088	0.135
	β_1	−0.058	0.096	β_1	−0.032	0.089
	β_2	0.045	0.094	β_2	0.025	0.073
	β_3	−0.043	0.133	β_3	−0.022	0.084
	$\Lambda_{01}(1)$	−0.031	0.067	$\Lambda_{01}(1)$	−0.022	0.045
	$\Lambda_{02}(1)$	0.029	0.058	$\Lambda_{02}(1)$	−0.021	0.034
	$\Lambda_{03}(1)$	0.051	0.336	$\Lambda_{03}(1)$	0.031	0.284
2		T = 0.3157			T = 1.1273	
	θ	0.101	0.184	θ	0.098	0.171
	β_1	0.075	0.097	β_1	−0.045	0.073
	β_2	0.056	0.107	β_2	−0.042	0.069
	β_3	0.061	0.137	β_3	0.051	0.086
	$\Lambda_{01}(1)$	−0.033	0.062	$\Lambda_{01}(1)$	−0.029	0.05
	$\Lambda_{02}(1)$	−0.028	0.055	$\Lambda_{02}(1)$	−0.024	0.04
	$\Lambda_{03}(1)$	−0.054	0.341	$\Lambda_{03}(1)$	−0.045	0.246

Table 2. The simulation results of the illness–death model with gamma frailty at 50% censoring proportion.

θ_{true}	Parameter	$n = 250$		Parameter	$n = 500$	
		Bias	SD		Bias	SD
0.5		T = 0.2250			T = 0.7070	
	θ	−0.148	0.132	θ	−0.114	0.104
	β_1	−0.029	0.107	β_1	−0.024	0.078
	β_2	−0.036	0.116	β_2	−0.031	0.083
	β_3	−0.016	0.175	β_3	−0.014	0.113
	$\Lambda_{01}(1)$	−0.031	0.08	$\Lambda_{01}(1)$	0.029	0.055
	$\Lambda_{02}(1)$	0.021	0.055	$\Lambda_{02}(1)$	0.022	0.039
	$\Lambda_{03}(1)$	0.053	0.232	$\Lambda_{03}(1)$	−0.048	0.173
1		T = 0.2966			T = 1.0153	
	θ	−0.161	0.176	θ	−0.143	0.147
	β_1	−0.032	0.120	β_1	−0.028	0.083
	β_2	0.033	0.129	β_2	−0.027	0.088
	β_3	0.02	0.187	β_3	0.018	0.126
	$\Lambda_{01}(1)$	−0.037	0.066	$\Lambda_{01}(1)$	−0.034	0.048
	$\Lambda_{02}(1)$	0.028	0.051	$\Lambda_{02}(1)$	−0.025	0.035
	$\Lambda_{03}(1)$	0.061	0.297	$\Lambda_{03}(1)$	−0.055	0.201
2		T = 0.3883			T = 1.2907	
	θ	0.173	0.194	θ	−0.161	0.172
	β_1	0.041	0.119	β_1	−0.034	0.083
	β_2	−0.036	0.119	β_2	−0.031	0.086
	β_3	0.025	0.184	β_3	0.022	0.138
	$\Lambda_{01}(1)$	−0.041	0.063	$\Lambda_{01}(1)$	−0.038	0.044
	$\Lambda_{02}(1)$	−0.032	0.047	$\Lambda_{02}(1)$	−0.029	0.038
	$\Lambda_{03}(1)$	−0.068	0.293	$\Lambda_{03}(1)$	−0.060	0.232

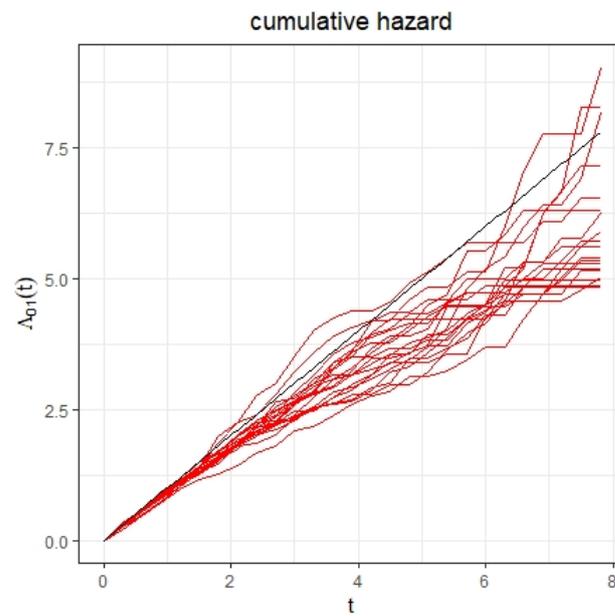


Figure 1. Simulation results for estimated cumulative baseline hazard function $\Lambda_{01}(t)$ by the proposed method based on 20 replications (with red color) together with the true cumulative baseline hazard function $\Lambda_{01}(t) = t$ (with black color) under the model with gamma frailty in Scenario 1 with sample size 500.

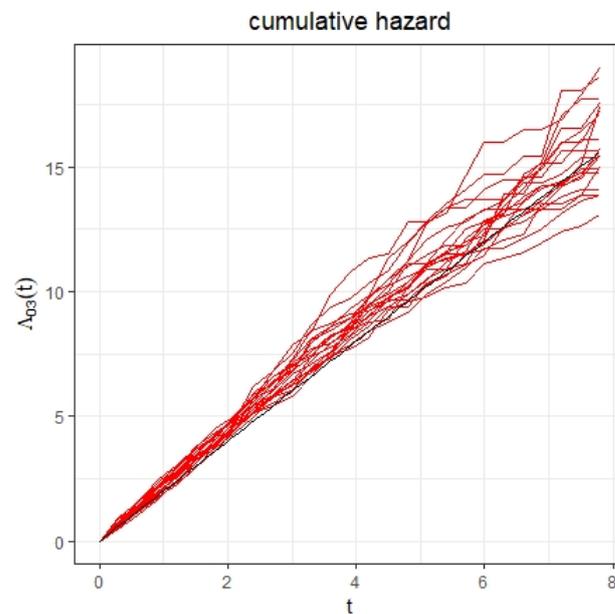


Figure 2. Simulation results for estimated cumulative baseline hazard function $\Lambda_{03}(t)$ by the proposed method based on 20 replications (with red color) together with the true cumulative baseline hazard function $\Lambda_{03}(t) = 2t$ (with black color) under the model with gamma frailty in Scenario 1 with sample size 500.

Scenario 2: We then independently simulate n observations from the semi-competing risk models (2.4)–(2.6) with $\omega \sim \text{Log-normal}(0, \theta)$. Similarly, we consider one covariate X with standard normal distribution in the model. The censoring times were generated from independent uniform distribution to yield censoring proportions of non-terminal event at around 65% and terminal event at around 30%. Let $\lambda_{01}(t_1) = 0.5, \lambda_{02}(t_2) = 1, \lambda_{03}(t_2) = 3$

and $\beta_1 = \beta_2 = 0.5, \beta_3 = 1$. We choose the true value of θ from $\Omega_\theta = \{0.1, 0.5, 1\}$ and pick the sample size n from $\Omega_n = \{250, 500\}$.

Based on 500 replications, the average values of the biases (Bias) of estimated frailty and regression parameters, their empirical standard deviations (SD), and the average computation times (T) are summarized in Table 3. From the results of Table 3, we find that the SDs of parameters $\beta_1, \beta_2, \beta_3$ and $\Lambda_{01}, \Lambda_{02}, \Lambda_{03}$ are relatively small, especially when the sample size is large. It can also be found that the SDs of θ are larger than that of other parameters. With the increase in the true values of θ , the SD of θ will increase accordingly, while the SD of θ will decrease with the increase in sample size. Compared to the model in Scenario 1, some parameters such as the frailty parameter in the semi-competing risk model with log-normal frailty are more difficult to estimate. Similar with Scenario 1, a larger value of θ requires longer computation time.

Table 3. The simulation results of the illness–death model with log-normal frailty.

θ_{true}	Parameter	$n = 250$		Parameter	$n = 500$	
		Bias	SD		Bias	SD
0.1		T = 0.4923			T = 1.3003	
	θ	−0.139	0.106	θ	−0.131	0.098
	β_1	0.030	0.114	β_1	0.030	0.088
	β_2	0.035	0.084	β_2	0.024	0.064
	β_3	0.075	0.143	β_3	0.075	0.120
	$\Lambda_{01}(1)$	0.004	0.067	$\Lambda_{01}(1)$	0.015	0.054
	$\Lambda_{02}(1)$	−0.104	0.084	$\Lambda_{02}(1)$	−0.112	0.066
	$\Lambda_{03}(1)$	−0.050	0.489	$\Lambda_{03}(1)$	−0.043	0.350
0.5		T = 0.9377			T = 2.6653	
	θ	−0.163	0.147	θ	−0.159	0.144
	β_1	−0.050	0.110	β_1	−0.038	0.091
	β_2	−0.049	0.076	β_2	−0.026	0.062
	β_3	−0.157	0.142	β_3	−0.101	0.107
	$\Lambda_{01}(1)$	−0.028	0.058	$\Lambda_{01}(1)$	−0.024	0.050
	$\Lambda_{02}(1)$	−0.155	0.071	$\Lambda_{02}(1)$	−0.151	0.054
	$\Lambda_{03}(1)$	0.262	0.569	$\Lambda_{03}(1)$	0.262	0.382
1		T = 1.2817			T = 3.6046	
	θ	−0.178	0.145	θ	−0.160	0.145
	β_1	−0.072	0.109	β_1	−0.064	0.088
	β_2	−0.065	0.081	β_2	−0.063	0.064
	β_3	−0.225	0.136	β_3	−0.179	0.105
	$\Lambda_{01}(1)$	−0.072	0.060	$\Lambda_{01}(1)$	−0.050	0.046
	$\Lambda_{02}(1)$	−0.227	0.072	$\Lambda_{02}(1)$	−0.166	0.056
	$\Lambda_{03}(1)$	0.619	0.694	$\Lambda_{03}(1)$	0.484	0.514

5. Real Data Analysis

Colon cancer is a serious type of cancer and has a high mortality rate. This type of cancer is not easy to cure since even if all the apparent diseased tissue can be surgically removed, there still exists other residual tumor parts which are not observable. Therefore, the cancer-recurrence event is commonly observed in clinical trials. For those patients who experience a recurrence, they will be subject to a higher risk of mortality. Thus, it is crucial to find the therapy which can significantly reduce the cancer-recurrence rate. In the following illustration, the colon cancer data provided by [25] are applied to our model. A total of 929 patients with Stage C disease are included for the modeling. Among them, 304 patients received levamisole plus fluorouracil, which is the covariate where the effect will be tested. A total of 468 patients developed recurrence, and 414 of them died. There are three stages recorded: State 1 stands for the stage that the patient is alive and disease-free. Stage 2 represents the condition that the patient is alive but with a recurrent diagnosis of cancer, while stage 3 denotes death. A general therapy for this type of cancer is the combination of

levamisole and fluorouracil, which is denoted by the treatment Lev+5-FU [26]. Previous research has proven the effectiveness of this therapy on reducing the cancer-recurrence rate. We will show the significance of this covariate on increasing the state transition hazard rate under our model framework.

Similar to the simulation study, we report the estimate of $\beta_1, \beta_2,$ and β_3 along with the corresponding baseline hazard. The observations (X_i, C_i, T_i) from the original dataset are re-sampled with replacement in non-parametric way, where $\hat{\alpha}^{(k)} = (\hat{\theta}^{(k)}, \hat{\beta}_1^{(k)}, \hat{\beta}_2^{(k)}, \hat{\beta}_3^{(k)})$ is the parameter estimation result under k^{th} bootstrap. Let $\bar{\alpha}$ be the mean of estimated parameters with the number of resamples equals to r . The estimated standard error is:

$$SE = \left(\frac{1}{r-1} \sum_{k=1}^r (\hat{\alpha}^{(k)} - \bar{\alpha})^2 \right)^{\frac{1}{2}}$$

In addition, the 95% bootstrap CI from Table 4 is constructed using 2.5% and 97.5% quantiles of $\{\hat{\alpha}^{(k)}\}_{k=1}^r$ with $r = 1000$. As shown in Table 4, the treatment with therapy Lev+5-FU can significantly reduce the probability of occurrence of recurrent event. However, we can observe that this therapy has little effect to the decrease in mortality. Both $\hat{\beta}_2$ and $\hat{\beta}_3$ have a p -value around 0.4, which indicates the insignificance of these two parameters. The cumulative hazard rate for different transitions are presented by Figure 3 and it indicates that a patient who experiences the recurrence of this disease will have very high mortality rate in the following years. In addition, we also observe the high hazard rate for the recurrence of colon cancer, which shows the importance of therapies such as Lev+5-FU, which can significantly reduce the occurrence of such event.

Table 4. The fitting results of colon cancer data.

Parameter	Est.	SE	p-Value	95% Bootstrap CI
θ	71.56	3.50	<0.001	[66.97, 75.62]
β_1	−0.735	0.180	<0.001	[−1.038, −0.473]
β_2	−0.036	0.373	0.397	[−0.696, 0.532]
β_3	0.054	0.167	0.378	[−0.220, 0.313]

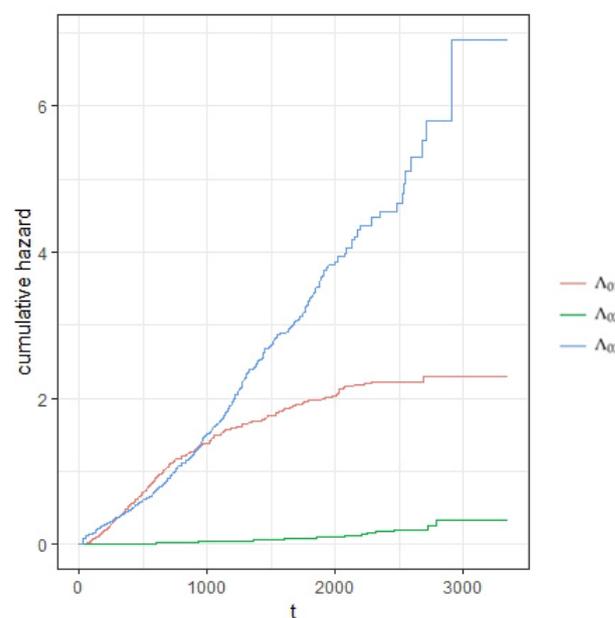


Figure 3. Cumulative hazard for different transitions in terms of days.

6. Discussion

In this paper, we proposed an efficient MM algorithm for the non-parametric maximum likelihood estimation of the illness–death model with a gamma frailty. The paper is motivated by the feature of colon cancer data while the existing method proposed by [4] has a high computational cost on the parameter estimation process. The proposed MM method can decompose the high-dimensional optimization problem into a sum of univariate optimization problems by the construction of a simple surrogate function which provides accurate and efficient simulation results. This MM approach avoids matrix inversion and can provide a toolkit for developing more efficient algorithms in a broad range of statistical optimization problems. Therefore, the proposed MM algorithm can help to estimate the parameters from the semi-competing risk model more efficiently.

In this paper, the explicit form of the marginal likelihood for the illness–death model can be derived. Therefore, the iteration steps do not involve integration calculation. However, in general, the marginal likelihood of the illness–death model with shared frailty is usually hard to obtain as an explicit form of the marginal likelihood is not available. For example, if the illness–death model involves a log-normal frailty or with correlated frailties, the intractable integral of the shared frailty will lead to a more complicated marginal likelihood. Even in the case of gamma frailty, the complicated model setup with three covariates and three non-parametric baseline hazard causes difficulties in parameter estimation. The inclusion of intractable integrals given general shared frailty will lead to a much higher computational cost and lower estimation accuracy. Therefore, even though our method can be extended to handling the illness–death model with general frailty, further modifications should be made to the algorithm to improve the estimation efficiency.

Author Contributions: Data curation, X.H., J.X., H.G. and W.Z.; Formal analysis, X.H., J.X. and H.G.; Investigation, X.H. and J.S.; Methodology, J.X. and J.S.; Project administration, J.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fine, J.P.; Jiang, H.; Chappell, R. On semi-competing risks data. *Biometrika* **2001**, *88*, 907–919. [[CrossRef](#)]
2. Wang, W. Estimating the association parameter for copula models under dependent censoring. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **2003**, *65*, 257–273. [[CrossRef](#)]
3. Lakhal, L.; Rivest, L.P.; Abdous, B. Estimating survival and association in a semicompeting risks model. *Biometrics* **2008**, *64*, 180–188. [[CrossRef](#)] [[PubMed](#)]
4. Xu, J.; Kalbfleisch, J.D.; Tai, B. Statistical analysis of illness–death processes and semicompeting risks data. *Biometrics* **2010**, *66*, 716–725. [[CrossRef](#)]
5. Zhang, Y.; Chen, M.H.; Ibrahim, J.G.; Zeng, D.; Chen, Q.; Pan, Z.; Xue, X. Bayesian gamma frailty models for survival data with semi-competing risks and treatment switching. *Lifetime Data Anal.* **2014**, *20*, 76–105. [[CrossRef](#)] [[PubMed](#)]
6. Lee, K.H.; Haneuse, S.; Schrag, D.; Dominici, F. Bayesian semiparametric analysis of semicompeting risks data: Investigating hospital readmission after a pancreatic cancer diagnosis. *J. R. Stat. Soc. Ser. C (Appl. Stat.)* **2015**, *64*, 253–273. [[CrossRef](#)]
7. Ha, I.D.; Xiang, L.; Peng, M.; Jeong, J.H.; Lee, Y. Frailty modelling approaches for semi-competing risks data. *Lifetime Data Anal.* **2020**, *26*, 109–133. [[CrossRef](#)]
8. Jiang, F.; Haneuse, S. A semi-parametric transformation frailty model for semi-competing risks survival data. *Scand. J. Stat.* **2017**, *44*, 112–129. [[CrossRef](#)]
9. Chen, Y.H. Maximum likelihood analysis of semicompeting risks data with semiparametric regression models. *Lifetime Data Anal.* **2012**, *18*, 36–57. [[CrossRef](#)]
10. Becker, M.P.; Yang, I.; Lange, K. EM algorithms without missing data. *Stat. Methods Med Res.* **1997**, *6*, 38–54. [[CrossRef](#)]
11. Wu, B.H.; Michimae, H.; Emura, T. Meta-analysis of individual patient data with semi-competing risks under the Weibull joint frailty–copula model. *Comput. Stat.* **2020**, *35*, 1525–1552. [[CrossRef](#)]
12. Zeng, D.; Lin, D. Semiparametric transformation models with random effects for joint analysis of recurrent and terminal events. *Biometrics* **2009**, *65*, 746–752. [[CrossRef](#)] [[PubMed](#)]

13. Zhou, R.; Zhu, H.; Bondy, M.; Ning, J. Semiparametric model for semi-competing risks data with application to breast cancer study. *Lifetime Data Anal.* **2016**, *22*, 456–471. [[CrossRef](#)] [[PubMed](#)]
14. Hsieh, J.J.; Huang, Y.T. Regression analysis based on conditional likelihood approach under semi-competing risks data. *Lifetime Data Anal.* **2012**, *18*, 302–320. [[CrossRef](#)]
15. Peng, L.; Fine, J.P. Regression modeling of semicompeting risks data. *Biometrics* **2007**, *63*, 96–108. [[CrossRef](#)]
16. Alvares, D.; Haneuse, S.; Lee, C.; Lee, K.H. SemiCompRisks: An R package for the analysis of independent and cluster-correlated semi-competing risks data. *R J.* **2019**, *11*, 376. [[CrossRef](#)]
17. Chapple, A.G.; Vannucci, M.; Thall, P.F.; Lin, S. Bayesian variable selection for a semi-competing risks model with three hazard functions. *Comput. Stat. Data Anal.* **2017**, *112*, 170–185. [[CrossRef](#)]
18. Reeder, H.T.; Lu, J.; Haneuse, S. Penalized estimation of frailty-based illness-death models for semi-competing risks. *arXiv* **2022**, arXiv:2202.00618.
19. Peng, M.; Xiang, L.; Wang, S. Semiparametric regression analysis of clustered survival data with semi-competing risks. *Comput. Stat. Data Anal.* **2018**, *124*, 53–70. [[CrossRef](#)]
20. Hunter, D.R.; Lange, K. Computing estimates in the proportional odds model. *Ann. Inst. Stat. Math.* **2002**, *54*, 155–168. [[CrossRef](#)]
21. Huang, X.; Xu, J.; Tian, G. On profile MM algorithms for gamma frailty survival models. *Stat. Sin.* **2019**, *29*, 895–916. [[CrossRef](#)]
22. Hunter, D.R.; Lange, K. Quantile regression via an MM algorithm. *J. Comput. Graph. Stat.* **2000**, *9*, 60–77.
23. Johansen, S. An extension of Cox's regression model. *Int. Stat. Rev. Int. Stat.* **1983**, *51*, 165–174. [[CrossRef](#)]
24. Klein, J.P. Semiparametric estimation of random effects using the Cox model based on the EM algorithm. *Biometrics* **1992**, *48*, 795–806. [[CrossRef](#)] [[PubMed](#)]
25. Moertel, C.G.; Fleming, T.R.; Macdonald, J.S.; Haller, D.G.; Laurie, J.A.; Goodman, P.J.; Ungerleider, J.S.; Emerson, W.A.; Tormey, D.C.; Glick, J.H.; et al. Levamisole and fluorouracil for adjuvant therapy of resected colon carcinoma. *N. Engl. J. Med.* **1990**, *322*, 352–358. [[CrossRef](#)]
26. Laurie, J.A.; Moertel, C.G.; Fleming, T.R.; Wieand, H.S.; Leigh, J.E.; Rubin, J.; McCormack, G.W.; Gerstner, J.B.; Krook, J.E.; Malliard, J. Surgical adjuvant therapy of large-bowel carcinoma: An evaluation of levamisole and the combination of levamisole and fluorouracil. The North Central Cancer Treatment Group and the Mayo Clinic. *J. Clin. Oncol.* **1989**, *7*, 1447–1456. [[CrossRef](#)] [[PubMed](#)]