A Study on Additive Hazard Model with Additive Frailty for Semi-Competing Risks Data

Youn Nam Kim

The Graduate School
Yonsei University
Department of Biostatistics and Computing

A Study on Additive Hazard Model with Additive Frailty for Semi-Competing Risks Data

A Dissertation
Submitted to the Department of Biostatistics and Computing
and the Graduate School of Yonsei University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Youn Nam Kim

December 2012

This certifies that the dissertation of Youn Nam Kim is approved.

_____

Thesis Supervisor: Dr. Chung Mo Nam

_____

Dr. Jinheum Kim

_____

Dr. Yang-jin Kim

_____

Dr. Inkyung Jung

_____

Dr. Sohee Park

The Graduate School
Yonsei University

December 2012

Contents

List of Tables

List of Figures

Abstracts

In this study, we proposed an illness-death model with Lin and Ying (1994)'s additive hazard and additive frailty for the regression analysis on semi-competing risks problem in a general morbidity/mortality process.

In the competing risks data, the occurrence of any one type of events precludes all other types of events and then prevents observing their occurring times. However, in the general morbidity/mortality process, death can preclude the relapse of colon cancer, but not vice versa. Although relapsed cancer patients are observed in the first event, the time to death for these patients can still be observed after relapse. That is called semi-competing risks data.

Comparing with the Cox-type hazard, the additive hazard function is more natural and properly partitions the effect of the covariate on one transition into the other transition, namely, internal consistency in the illness-death model by Klein (2006). In the proposed model, we adapted the additive frailty to describe the association between the covariates and failure time in terms of the risk difference rather than the risk ratio.

For the inference, we considered a full maximum likelihood on the complete data and incorporated an Estimation-Maximization algorithm to deal with frailty and Gauss-Laguerre quadrature method for calculating the expectations of the functions of frailty. We used the expected observed information matrix for standard errors of the parameter estimates. The piecewise constant baseline hazards are simply assumed to have different constants in the same intervals on three hazards such as relapse, death without relapse and death with relapse.

The frailty was assumed to have a gamma distribution. In the simulation, we found that the regression estimates were robust when the intervals for the baseline hazards was mis-specified or the true frailty distribution deviated from the assumed gamma distribution. The proposed model was applied to the data from a national intergroup trial in the 1980's to study the effectiveness of two adjuvant therapy regimens for the improvement of surgical cure rates in stage III colon cancer. We compared the group treated with levamisole plus fluorouracil with the untreated group using the semi-competing risks model with cancer recurrence and death.

KEY WORDS: Semi-competing risks data, Illness-death model, Additive hazard, Additive frailty, Internal consistency, Gamma distribution, Estimation-Maximization algorithm, Gauss-Laguerre Quadrature, Piecewise constant baseline hazard

I. Introduction

A typical complexity with survival data is that observations may be censored. When you have only one type of event, survival analysis such as Kaplan-Meier curves, log-rank tests, and Cox proportional hazard models are appropriately and widely used. These analyses generally assume that the censoring is non-informative withdrawal. However, if the censoring is informative, the inferences in the general survival analyses may be misleading. For example, in a cancer study, when the risk of relapse is highly related to the risk of death, the time of relapse may informatively be censored by death and vice versa. According to a recent study by Rosenkranz (2010), the analysis of non-fatal events without considering death as a possible outcome can be misleading if the correlation between the events and death is disregarded. In other words, informative censoring can define a competing risk in a competing risks framework. In a competing risks problem, the occurrence of any one type of event precludes all other types of events and then prevents observing their occurring times. That is, all competing events are mutually exclusive. By the way, in a general morbidity/mortality process, death can preclude the occurrence of disease, but not vice versa. Although relapsed cancer patients are observed in a first event, the time to death for these patients can still be observed after relapse. Relapsed cancer patients could survive for some while and die later from any cause. In these patients it is possible to observe both the time to relapse and the time to death with relapse. however, if a cancer patient dies before relapse, we can only observe the time to death

without relapse. This is called a semi-competing risks problem (Fine, Jiang and Chappell, 2001). It consists of one non-terminal event and one terminal event, such as relapse and death in a cancer study. The semi-competing risks is illustrated in Figure 1.

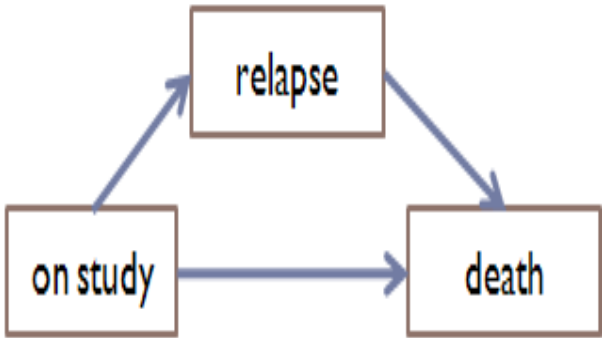

Figure 1. Semi-competing risks

A typical example is a randomized clinical trial of nasopharyngeal cancer. This study was conducted to compare radiation therapy (RT) only with a combination of chemotherapy and radiation therapy (CRT) treatment (Xu, Kalbfleisch and Tai, 2010). Two endpoints of interest were time to tumor recurrence and time to mortality. From the semi-competing risk problems, the tumor recurrence was considered as the non-terminal event while death was considered as the terminal event. Another example is a clinical study with patients undergoing Bone Marrow Transplantation (BMT) for Leukemia. Since patients who relapse are given a wide variety of treatment after relapse, the

three treatment effects on relapse, death with relapse, death without relapse are important measures used in understanding efficacy in BMT studies.

1.1 Analysis for Semi-competing risks data

Even though in a semi-competing risks problem, if you only have a interest of the first event happened, the survival analysis can be applied to a competing risks problem, ignoring the information about the time until the terminal event after the non-terminal event. However, time to death after relapse can give us detailed information for better understanding of the disease process. There is an important utilization on assessing surrogacy in a single trial using the semi-competing risks problem. Ghosh (2009) suggested viewing surrogates and true endpoints as the semi-competing risks data to evaluate the surrogacy. Similarly, Dejardin, Lesaffre and Verbeke (2010) dealt with a mixed model with the relationship of two endpoints with the validation of surrogacy.

There have been several studies about the statistical analysis for the semi-competing risks data. Firstly, Clayton copula model, a representative model in semi-competing risks problem, was proposed by Fine (2001). This model included the joint survivor function of the two events, such as disease progression and death, following the copula distribution with two marginal distributions with dependence parameters. Fine (2001) focused on the estimation of the marginal distribution of the disease progression, considering the dependence of two events. The Clayton copula model has been researched and extended by Fine et al. (2004) developing the estimation methods for the dependence parameters under left truncation. The regression modelling with the

time dependent covariates by Peng and Fine (2007) has also been developed.

Secondly, an illness-death model is another model in the semi-competing risks problem. The illness-death model is a special model of multi-state models. Multi-state models provide a relevant framework for modelling these complex event histories. For example, in multi-state models, a subject is assumed to have capacity to transit from one state of disease to another state of disease according to the transition probability. In the semi-competing risks problem, the non-terminal event during the disease process can be seen as transition from one state to another. The illness-death model which is a typical kind of multi-state models consisting of three states, such as 'health', 'disease', and 'death' with the same structure in the semi-competing risks problem. 'State transition' means the occurrence of the event and the probability from one state to another state can be defined by the hazard function of the occurrence of the event. Frydman and Szarek (2009) illustrated a Cox-type illness-death model to estimate sub-distribution functions corresponding to 'health'→'disease' and 'health'→'death' transitions and the 'disease'→'death' transition intensity using the maximum likelihood methodology. Xu, Kalbfleisch and Tai (2010) proposed the Cox-type illness-death model with shared frailty dealing with common unmeasured associated factors in subjects and used the non-parametric maximum likelihood estimation for the inference. They indicated that Clayton copula models and the approaches involve similar assumptions to those underlying multiple decrement models for latent failure times and the interpretation of the marginal distribution of the non-terminal event is hypothetical.

Most regression models introduced (Fine et al., 2004; Peng and Fine, 2007; Frydman and Szarek, 2009; Dejardin, 2010; Xu, Kalbfleisch and Tai, 2010) dealt with Cox's proportional hazard function. The covariate effects are assumed to

act multiplicatively on some unknown hazard rate. That is, the effect of the covariate is to have constant or time-dependent relative hazard ratio, $\exp(\beta)$ or $\exp(\beta(t))$, over time against the baseline hazard function. Alternately, the additive hazard model can be adapted. In the literature, there are two main types of additive hazard models, namely, Aalen (1989)'s additive hazard model and Lin and Ying (1994)'s additive hazard model. The first, introduced by Aalen (1989), allows the effects of covariates, $\beta(t)$, to act additively and time-dependently on the unknown hazard rate. This model is flexible enough to allow the effects of the covariates to change over time. For this model, estimation is based on least-squares estimation of the cumulative regression functions, $B(t) = \int_0^t \beta(s)ds$. These estimators can be smoothed to obtain estimators of $\beta(t)$. In the second model, proposed by Lin and Ying (1994), the possibly time-dependent covariate effects in the Aalen (1989)'s model are replaced by constants, β. The estimation is based on counting process theory and the technique of the partial-likelihood-based method for the proportional hazards model. Mckeague and Sasieni (1994) mixed up Aalen (1989)'s time-dependent covariates and Lin and Ying (1994)'s time-constant covariates.

To my knowledge, regression models with Clayton copula are all proportional hazard models that have already been formed. By the way, there are several studies about multi-state model or illness-death model with additive hazard function (Aalen et al., 2001; Yin and Gai, 2004; Shu and Klein, 2005; Klein, 2006; Zhang, Akcin and Lim, 2011). The motivation to focus on additive hazard models comes from the notation, by Klein (2006). He said that the additive hazard function is more natural and properly partitions the effect of the covariate on one transition into the other transition. That is, the additive hazard function satisfies internal consistency in multi-state models. Internal

consistency means that the transition probability from one state to one of all possible states is the sum of the transition probabilities from one state to each possible state. For the illness-death model, the transition probability from 'health' to 'disease' or 'death' is the sum of the transition probability from 'health' to 'disease' and the transition probability from 'health' to 'death'. When we model 'disease' and 'death' by proportional hazards in the illness-death model, the model does not always satisfy the internal consistency. By the way, additive hazard always satisfy it.

The non-terminal event may substantially change the risk of terminal event to occur. That means that the time to non-terminal event and the time to terminal event is no longer independent. The hazards for two events in a subject are not completely independent on his/her biological or environmental characteristics. Therefore the appropriate analysis on the semi-competing risks problem requires to deal with the dependency. Frailty, introduced by Vaupel, Manton and Stallard (1979), is defined as the impact of heterogeneity in individual frailty on the dynamics of mortality. They said, an individual with higher value of frailty will have a higher probability of dying, thus more frailty individuals are more likely to die first. Duchateu et al. (2002) researched Cox's proportional models with shared frailty for the common impact of center's heterogeneity in a multi-center clinical study. Multiplicative shared frailty in the hazard function mostly assumed to follow gamma distribution with mean 1 and variance. The variance of shared frailty is interpreted as the degree of the center's or subject's dynamics, that is the higher variance the higher heterogeneity. The Clayton copula model is jointed two marginal distributions with dependence parameters (Fine, 2001). Xu, Kalbfleisch and Tai (2010) proposed the illness-death model with Cox's proportional hazard and a shared frailty on the semi-competing risks problem.

Recently, Chong et al. (2011) presented a frailty model with additive hazard for the semi-competing risk; International Society for Computational Biology (ISCB). They dealt with additive covariate effects and multiplicative frailty effects on baseline hazard.

1.2 Objectives

In this study, we propose the illness-death model with Lin and Ying (1994)'s additive hazard and additive frailty for the regression analysis on the semi-competing risks data. Additive frailty was applied to assess the within-group association for the case of paired observations by Lin and Ying (1997) and to the cluster failure data by Silva and Turkman (2004). Frailty was additively inserted into the hazard function that maintained an additive structure in turn. Lin and Ying (1997) explained that the additive frailty described the association between the covariates and failure time in terms of the risk difference rather than the risk ratio. There are some studies about multi-state models or illness-death models with the additive hazard function with multiplicative frailty (Pipper and Martinussen, 2004; Chong et al., 2011; Martinussen, Scheike and Zucker, 2011). This study could be the first clinical research showing the proposed model on the semi-competing risks problem.

1.3 Outlines

The illness-death model with Cox's proportional hazard and a shared frailty

(Xu, Kalbfleisch and Tai, 2010) is introduced in Section II. The notation and model are described in Section II. The proposed models are presented in Section III, the simulation to assess the performance and the robustness of the proposed models in Section IV and application to colon cancer data in Section V, finally, the conclusion and the discussion in Section VI.

II. Semi-competing risks models

2.1 Notation and framework

In this study, we considered a clinical study of cancer patients to measure the treatment effect on cancer relapse and death in the present of semi-competing risks. In Figure 1, this situation consists of three compartments, including 'on study', 'relapse' and 'death'. The 'on study' is the state occupied at time $t=0$. For convenience, let us call 'relapse' as a non-terminal event and 'death' as a terminal event.

Let T_{i1}, T_{i2}, and C_i be the time to relapse, death and censoring times for the ith patient, $i=1,...,n$. The observed event times are $Y_{i1}=\min(T_{i1},T_{i2},C_i)$ and $Y_{i2}=\min(T_{i2},C_i)$. The censoring indicator is δ_{ij}, where $j=1,2$ for the ith patient, $i=1,...,n$. Y_{i1} is the observed time for relapse, Y_{i2} is for death and their observable time space is $0<Y_{i1}\leqq Y_{i2}$. For example, that if $\delta_{i1}=0$ and $\delta_{i2}=1$, then $Y_{i1}=Y_{i2}=T_{i2}$. We consider common covariates, z_i with different effects on three hazards of relapse, death without relapse and death with relapse. The observation data can be described as $O_i=\{Y_{i1},Y_{i2},\delta_{i1},\delta_{i2},z_i\}$, $i=1,...,n$.

2.2 Illness-death models

The illness-death process is fully characterized through the transition probabilities between one state and another state or through transition intensities representing the instantaneous hazard of progression to another state conditionally on occupying a state, defined as follows:

$$\lambda_1(t_1) = \lim_{\Delta \to 0} \frac{P\{T_1 \in [t_1, t_1+\Delta) | T_1 \geqq t_1, T_2 \geqq t_1\}}{\Delta}, \quad (1)$$

$$\lambda_2(t_2) = \lim_{\Delta \to 0} \frac{P\{T_2 \in [t_2, t_2+\Delta) | T_1 \geqq t_2, T_2 \geqq t_2\}}{\Delta}, \quad (2)$$

$$\lambda_{12}(t_2|t_1) = \lim_{\Delta \to 0} \frac{P\{T_2 \in [t_2, t_2+\Delta) | T_1 = t_1, T_2 \geqq t_2\}}{\Delta}, \quad 0 < t_1 \leqq t_2. \quad (3)$$

The hazards $\lambda_1(t_1)$ and $\lambda_2(t_2)$ are the usual cause specific or crude hazard functions for the part of competing risks in which either the terminal or nonterminal event may occur first. The hazard $\lambda_{12}(t_2|t_1)$ defines the rate of death after the occurrence of relapse at time $T_1 = t_1$.

In general, the hazard $\lambda_{12}(t_2|t_1)$ can depend on both t_1 and t_2. In a Markov model, however, $\lambda_{12}(t_2|t_1)$ depends only on t_2. That is, in the Markov illness-death model, the three hazards are independent of one another. A semi-Markov process allows that $\lambda_{12}(t_2|t_1)$ depends on the sojourn time from $T_1 = t_1$ to $T_2 = t_2$. However, frailty can be regarded as the heterogeneity among three hazards in a patient. Xu, Kalbfleisch and Tai (2010) considered multiplicative frailty in the Markov illness-death model. The hazards are

assumed to be independent of one another, given a frailty. Denoting the frailty by a random variable $u > 0$ with $E(u) = 1$, conditional transition functions for the ith patient, analogous to equations (1)-(3), are as follows:

$$\lambda_1(t_1|u) = u\lambda_{01}(t_1), \tag{4}$$

$$\lambda_2(t_2|u) = u\lambda_{02}(t_2), \tag{5}$$

$$\lambda_{12}(t_2|t_1,u) = u\lambda_{12}(t_2), \qquad 0 < t_1 \leqq t_2. \tag{6}$$

Conditional on frailty u, this is just a standard Markov illness-death model. The frailty u was assumed to have a gamma distribution with mean 1, variance θ, and density $\theta^{-1/\theta}u^{1/\theta-1}\exp^{-u/\theta}/\Gamma(1/\theta)$.

The general model was described with incorporating covariates z. Xu, Kalbfleisch and Tai (2010) dealt with cox-type models and adapted the maximum likelihood methods for the inference. Incorporating a vector of covariates $z = (z_1, z_2, \ldots, z_p)'$ in the conditional hazards, analogous to equations (4)-(6), we have

$$\lambda_1(t_1|u,z) = u\lambda_{01}(t_1)\exp^{\beta_1' z}, \tag{7}$$

$$\lambda_2(t_2|u,z) = u\lambda_{02}(t_2)\exp^{\beta_2' z}, \tag{8}$$

and

$$\lambda_{12}(t_2|t_1,u,z) = u\lambda_{12}(t_2)\exp^{\beta_3' z}, \qquad 0 < t_1 \leqq t_2, \tag{9}$$

where $\beta_1 = (\beta_{011}, \beta_{012}, \ldots, \beta_{01p})'$, $\beta_2 = (\beta_{021}, \beta_{022}, \ldots, \beta_{02p})'$ and

$\beta_{12} = (\beta_{121}, \beta_{122}, ..., \beta_{12p})'$.

The frailty u has the gamma distribution and is distributed independently of covariates z. An alternative approach to regression analysis would incorporate the covariates in the marginal hazards, analogous to equations (7)-(9) as follows:

$$\lambda_1(t_1|u,z) = [1+\theta\Lambda_{01}(t_1)]^{-1}\lambda_{01}(t_1)\exp^{\beta_1'z}, \quad (10)$$

$$\lambda_2(t_2|u,z) = [1+\theta\Lambda_{02}(t_1)]^{-1}\lambda_{02}(t_2)\exp^{\beta_2'z}, \quad (11)$$

$$\lambda_{12}(t_2|t_1,u,z) = (1+\theta)[1+\theta\{\Lambda_{12}(t_2)-\Lambda_{12}(t_1)\}]^{-1}\lambda_{12}(t_2)\exp^{\beta_3'z}, \quad 0 < t_1 \leqq t_2, \quad (12)$$

where $\Lambda_{01}(t) = \int_0^t \lambda_{01}(s)ds$, $\Lambda_{02}(t) = \int_0^t \lambda_{02}(s)ds$, and $\Lambda_{12}(t) = \int_0^t \lambda_{12}(s)ds$.

The marginal model is non-Markovian unless frailty u is constant ($\theta = 0$).

2.3 Estimations for illness-death models

There are two different modeling strategies using conditional intensities or marginal intensities. It is important to note that the regression parameters have different interpretations, corresponding to the regression effects on conditional intensities and marginal intensities, respectively. Xu, Kalbfleisch and Tai (2010) focused on the conditional approach. Under the models (7)-(9),

the conditional likelihood, given a frailty u_i, based on the data $O_i = \{Y_{i1}, Y_{i2}, \delta_{i1}, \delta_{i2}, z_i\}$ for the i^{th} patient, $i = 1,...,n$, is

$$
\begin{aligned}
L_{c,i} &= \lambda_1(Y_{i1})^{\delta_{i1}} \lambda_2(Y_{i1})^{(1-\delta_{i1})\delta_{i2}} \lambda_{12}(Y_{i2})^{\delta_{i1}\delta_{i2}} \qquad (13)\\
&\times \exp\left[\delta_{i1}\beta_1' z_i + (1-\delta_{i1})\delta_{i2}\beta_2' z_i + \delta_{i1}\delta_{i2}\beta_3' z_i\right]\\
&\times \exp\left[-u_i\left\{\Lambda_{i1}(Y_{i1})\exp^{\beta_1' z_i} + \Lambda_{i2}(Y_{i1})\exp^{\beta_2' z_i} + \delta_{i1}\Lambda_{i12}(Y_{i1}, Y_{i2})\exp^{\beta_3' z_i}\right\}\right],
\end{aligned}
$$

where $\Lambda_{12}(s,t) = \Lambda_{12}(t) - \Lambda_{12}(s)$. The full likelihood is $L_{f,i} = L_{c,i} \times g(u_i|\theta)$. Averaging over the distribution of u_i gives the observable likelihood function for the i^{th} patient, $i = 1,...,n$, as

$$
\begin{aligned}
L_{o,i} &= \lambda_1(Y_{i1})^{\delta_{i1}} \lambda_2(Y_{i1})^{(1-\delta_{i1})\delta_{i2}} \lambda_{12}(Y_{i2})^{\delta_{i1}\delta_{i2}} \qquad (14)\\
&\times \exp\left[\delta_{i1}\beta_1' z_i + (1-\delta_{i1})\delta_{i2}\beta_2' z_i + \delta_{i1}\delta_{i2}\beta_3' z_i\right]\\
&\times (1+\theta)^{\delta_{i1}\delta_2}\left[1 + \theta\left\{\Lambda_{i1}(Y_{i1})\exp^{\beta_1 z_i} + \Lambda_{i2}(Y_{i1})\exp^{\beta_2 z_i} + \delta_{i1}\Lambda_{i12}(Y_{i1}, Y_{i2})\exp^{\beta_3 z_i}\right\}\right]^{-1/\theta - \delta_{i1} - \delta_{i2}}.
\end{aligned}
$$

Let $t_{r1}, t_{r2}, ..., t_{rm}$ be the ordered distinct relapse times, $t_{d1}, t_{d2}, ..., t_{df}$ be the ordered distinct death times without relapse and $t_{rd1}, t_{rd2}, ..., t_{rdg}$ be the ordered distinct death times with relapse. We let $\lambda_{01} = (\lambda_{011}, \lambda_{012}, ..., \lambda_{01m})'$, $\lambda_{02} = (\lambda_{021}, \lambda_{022}, ..., \lambda_{02f})'$ and $\lambda_{12} = (\lambda_{121}, \lambda_{122}, ..., \lambda_{12g})'$ where $\lambda_{01j} = d\Lambda_{01}(t_{rj})$, $j = 1,...,m$, $\lambda_{02j} = d\Lambda_{02}(t_{dj})$, $j = 1,...,f$ and $\lambda_{12j} = d\Lambda_{12}(t_{rdj})$, $j = 1,...,g$. Let $\eta = (\theta, \beta_1', \beta_2', \beta_3', \lambda_{01}', \lambda_{02}', \lambda_{12}')'$ be the $(m+f+g+4)$-dimensional vector of all

parameters. The maximum likelihood estimates are obtained by taking the derivative of the log likelihood with respect to η and solving the corresponding equations.

A algorithm is available to compute the estimates by using an Newton-Raphson method. Xu, Kalbfleisch and Tai (2010) proved that these estimators, under relatively mild conditions, are asymptotically normal with a covariance matrix estimated in outline in Murphy (1995).

The estimate of the asymptotic variance of $(\hat{\theta},\widehat{\beta_1},\widehat{\beta_2},\widehat{\beta_3})$ was estimated using the relevant submatrix from the inverse of the observed information matrix, $\hat{I}^{-1}$ for η. Similarly, the asymptotic covariance function of $\widehat{\Lambda_{01}},\widehat{\Lambda_{02}}$ and $\widehat{\Lambda_{12}}$ was obtained from the submatrices of $\hat{I}^{-1}$ which was associated with $\lambda_{01},\lambda_{02}$ and λ_{12}, respectively. The profile likelihood approach could provide an alternative approach for estimating asymptotic variances for large sample sizes.

III. Proposed models

3.1 Proposed models

We proposed an additive type model on the semi-competing risk. Frailty can be considered as an additive risk from the heterogeneity among three hazards in a patient. The hazards are assumed to be independent of one another, given the frailty. Denoting the frailty by a random variable $u > 0$ with $E(u) = 1$, conditional transition functions for the ith patient, analogous to equations (4)-(6), are as follows:

$$\lambda_1(t_1|u_i) = \lambda_{01}(t_1) + u, \tag{15}$$

$$\lambda_2(t_2|u) = \lambda_{02}(t_2) + u, \tag{16}$$

$$\lambda_{12}(t_2|t_1,u) = \lambda_{12}(t_2) + u, \qquad 0 < t_1 \leqq t_2. \tag{17}$$

The dependence of T_2 on T_1 is described by the conditional (given u) explanatory hazard difference $\lambda_{12}(t_2) - \lambda_2(t_2)$ as well as by the common frailty u. We assume the frailty u has a gamma distribution with mean 1 and variance θ. Incorporating covariates z in the conditional hazards, the general functions, analogous to equations (15)-(17), are as follows:

$$\lambda_1(t_1|u,z) = \lambda_{01}(t_1) + u + \beta_1'z, \tag{18}$$

$$\lambda_2(t_2|u,z) = \lambda_{02}(t_2) + u + \beta_2'z, \tag{19}$$

$$\lambda_{12}(t_2|t_1,u,z) = \lambda_{12}(t_2) + u + \beta_3'z, \qquad 0 < t_1 \leqq t_2. \tag{20}$$

We simply assume constant additive risks of covariates on the hazards. This is the excess risk of the event occurred due to given covariates, suggested by Lin and Ying (1994). The conditional hazard given u in equations (18), (19), and (20) together with the gamma frailty distribution can be seen to correspond to a marginal model with hazards as defined in equations (18)-(20). The marginal functions are

$$\lambda_1(t_1 \mid z) = \lambda_{01}(t_1) + \beta_1'z + (1+2\theta t_1)^{-1}, \tag{21}$$

$$\lambda_2(t_2 \mid z) = \lambda_{02}(t_2) + \beta_2'z + (1+2\theta t_2)^{-1}, \tag{22}$$

and

$$\lambda_{12}(t_2 \mid t_1, z) = \lambda_{12}(t_2) + \beta_3'z + \{1+\theta(t_1+t_2)\}^{-1}\left[\lambda_{01}(t_1) + \beta_1'z + \{1+\theta(t_1+t_2)\}^{-1}\right]^{-1}$$
$$\times \left[\lambda_{01}(t_1) + \beta_1'z + (1+\theta)\{1+\theta(t_1+t_2)\}^{-1}\right], \qquad 0 < t_1 \leqq t_2. \tag{23}$$

The proof of the marginal hazard is described in Appendix A.

3.2 Estimations for proposed models

Estimations for the model is based on the log likelihood function. Under the model (18)-(20), the conditional likelihood, given a frailty u_i, based on the data $O_i = \{Y_{i1}, Y_{i2}, \delta_{i1}, \delta_{i2}, z_i\}$ for the i^{th} patient, $i = 1,...,n$, is

$$\begin{aligned} L_{c,i}' = & \{\lambda_{01}(Y_{i1}) + u_i + \beta_1' z_i\}^{\delta_{i1}} \{\lambda_{02}(Y_{i1}) + u_i + \beta_2' z_i\}^{(1-\delta_{i1})\delta_{i2}} \{\lambda_{12}(Y_{i2}) + u_i + \beta_3' z_i\}^{\delta_{i1}\delta_{i2}} \\ & \times \exp[-\{\Lambda_{i1}(Y_{i1}) + \Lambda_{i2}(Y_{i1}) + \delta_{i1}\Lambda_{i12}(Y_{i1}, Y_{i2})\}] \\ & \times \exp[-\{\beta_1' z Y_{i1_i} + \beta_2' z_i Y_{i1} + \delta_{i1}\beta_3' z_i (Y_{i2} - Y_{i1}) + u_i (Y_{i1} + Y_{i2})\}]. \end{aligned} \tag{24}$$

Averaging over the distribution of u_i and taking a product over i gives the observable likelihood function (Appendix B). Because the observable likelihood function is too complicated to estimate the parameters, we consider the full likelihood that we may have if the frailties are observed.

The full likelihood for the i^{th} patient, $i = 1,...,n$, is given by

$$L_{f,i}' = L_{c,i}' \times g(u_i|\theta). \tag{25}$$

The full likelihood gives the likelihood for the complete data with a unknown frailty. The Estimation-Maximization (EM) algorithm is a natural choice for parameter estimations. The EM algorithm usually provides a mean of

maximizing complex likelihoods. In the expectation (E) step of the algorithm, the expected value of $L_{f,i}{}'$ is computed with the current estimates of the parameters and the observable data. In the maximization (M) step of the algorithm, estimates of the parameters which maximize the expected value of $L_{f,i}{}'$ from the E step are obtained. The algorithm iterates between these two steps until the estimates converge.

The baseline hazard is simply assumed to be piecewise constant in this study. The piecewise constant approach allows the hazard rate to change in non-parametric ways. Implementing this model is straightforward. The hazard rate is constant within a interval but can be different between intervals. The researcher splits into some intervals of the time periods and estimates an exponential hazard with each time interval.

Let $S_1=(s_0, s_1]$, $S_2=(s_1, s_2]$, ..., and $S_K=(s_{K-1}, s_K)$, where $s_0=0$ and $s_K=\infty$, be the intervals in piecewise constant baseline hazards $\lambda_{01}(t)$, $\lambda_{02}(t)$, and $\lambda_{12}(t)$. We let $\lambda_{01}=(\lambda_{011},\lambda_{012},...,\lambda_{01K})'$, $\lambda_{02}=(\lambda_{021},\lambda_{022},...,\lambda_{02K})'$ and $\lambda_{12}=(\lambda_{121},\lambda_{122},...,\lambda_{12K})'$ where $\lambda_{01k}=d\Lambda_{01}(t_{rk})$, $\lambda_{02k}=d\Lambda_{02}(t_{dk})$ and $\lambda_{12k}=d\Lambda_{12}(t_{rdk})$, $k=1,...,K$, be the constant baseline parameters for intervals. Let $\eta=(\theta,\beta_1{}',\beta_2{}',\beta_3{}',\lambda_{01}{}',\lambda_{02}{}',\lambda_{12}{}')'$ be the $(3\times K+4)$-dimensional vector of parameters.

For the convenience, let $d_{1ik}=I(Y_{i1}\in S_k)$ and $d_{2ik}=I(Y_{i2}\in S_k)$ be the indicators, whether the observations Y_{i1} and Y_{i2} are included in the interval S_k ,$k=1,...,K$ or not. The baseline hazard functions can be expressed by

$$\lambda_{01}(Y_{i1})=\lambda_{011}d_{1i1}+\lambda_{012}d_{1i2}+...+\lambda_{01K}d_{1iK}=\sum_{k=1}^{K}\lambda_{01k}d_{1ik},$$

$$\lambda_{02}(Y_{i1}) = \lambda_{021}d_{1i1} + \lambda_{022}d_{1i2} + \ldots + \lambda_{02K}d_{1iK} = \sum_{k=1}^{K}\lambda_{02k}d_{1ik},$$

and

$$\lambda_{12}(Y_{i2}) = \lambda_{121}d_{2i1} + \lambda_{122}d_{2i2} + \ldots + \lambda_{12K}d_{2iK} = \sum_{k=1}^{K}\lambda_{12k}d_{2ik}.$$

The cumulative baseline hazard functions are

$$\Lambda_{01}(Y_{i1}) = \sum_{k=1}^{K}\left\{(Y_{i1} - s_{k-1})d_{1ik} + (s_k - s_{k-1})\left(1 - \sum_{l=1}^{k}d_{1il}\right)\right\}\lambda_{01k},$$

$$\Lambda_{02}(Y_{i1}) = \sum_{k=1}^{K}\left\{(Y_{i1} - s_{k-1})d_{1ik} + (s_k - s_{k-1})\left(1 - \sum_{l=1}^{k}d_{1il}\right)\right\}\lambda_{02k},$$

and

$$\begin{aligned}\Lambda_{12}(Y_{i1}, Y_{i2}) = \Lambda_{12}(Y_{i2}) - \Lambda_{12}(Y_{i1}) = &\{(Y_{i2} - Y_{i1})d_{1i1}d_{2i1} + (s_1 - Y_{i1})d_{1i1}(1 - d_{2i1})\}\lambda_{121} \\ &+ \sum_{k=2}^{K}\left\{(Y_{i2} - Y_{i1})d_{1ik}d_{2ik} + (s_k - Y_{i1})d_{1ik}\left(1 - \sum_{l=1}^{k}d_{2il}\right)\right. \\ &\left. + (s_k - s_{k-1})\left(\sum_{l=1}^{k-1}d_{1il}\right)\left(1 - \sum_{l=1}^{k}d_{2il}\right) + (Y_{i2} - s_{k-1})\left(\sum_{l=1}^{k-1}d_{1il}\right)d_{2ik}\right\}\lambda_{12k},\end{aligned}$$

where $\Lambda_{12}(s,t) = \Lambda_{12}(t) - \Lambda_{12}(s)$.

Let us add some notations, $a_{i1}(Y_{i1})$, $a_{i2}(Y_{i1})$, $a_{i12}(Y_{i2})$, $A_{i1}(Y_{i1})$, $A_{i2}(Y_{i1})$ and $A_{i12}(Y_{i1}, Y_{i2})$, for convenience, as follows:

$$a_{i1}(Y_{i1}) = \lambda_{01}(Y_{i1}) + \beta_1'z_i,$$

$$a_{i2}(Y_{i1}) = \lambda_{02}(Y_{i1}) + \beta_2' z_i,$$
$$a_{i12}(Y_{i2}) = \lambda_{12}(Y_{i2}) + \beta_3' z_i,$$

$$A_{i1}(Y_{i1}) = \Lambda_{01}(Y_{i1}) + \beta_1' z_i Y_{i1},$$
$$A_{i2}(Y_{i1}) = \Lambda_{02}(Y_{i1}) + \beta_2' z_i Y_{i1},$$
$$A_{i12}(Y_{i1}, Y_{i2}) = \Lambda_{12}(Y_{i1}, Y_{i2}) + \beta_3' z_i (Y_{i2} - Y_{i1}).$$

The full likelihood (25) for the i^{th} patient, $i = 1,...,n$, can be expressed by

$$\begin{aligned} L_{f,i}' &= \exp[-\{\Lambda_{01}(Y_{i1}) + (u_i + \beta_1' z_i) Y_{i1}\} - \{\Lambda_{02}(Y_{i1}) + (u_i + \beta_2' z_i) Y_{i1}\} \\ &\quad - \delta_{i1}\{\Lambda_{12}(Y_{i1}, Y_{i2}) + (u_i + \beta_3' z_i)(Y_{i2} - Y_{i1})\}] \\ &\times (\lambda_{01}(Y_{i1}) + u_i + \beta_1' z_i)^{\delta_{i1}} (\lambda_{02}(Y_{i1}) + u_i + \beta_2' z_i)^{(1-\delta_{i1})\delta_{i2}} (\lambda_{12}(Y_{i2}) + u_i + \beta_3' z_i)^{\delta_{i1}\delta_{i2}} \\ &\times f(u_i|\theta) \\ &= \exp[-\{A_{i1}(Y_{i1}) + u_i Y_{i1}\} - \{A_{i2}(Y_{i1}) + u_i Y_{i1}\} \\ &\quad - \delta_{i1}\{A_{i12}(Y_{i1}, Y_{i2}) + u_i (Y_{i2} - Y_{i1})\}] \\ &\times (a_{i1} + u_i)^{\delta_{i1}} (a_{i2} + u_i)^{(1-\delta_{i1})\delta_{i2}} (a_{i12} + u_i)^{\delta_{i1}\delta_{i2}} \times g(u_i|\theta). \end{aligned} \tag{26}$$

The log full likelihood for the i^{th} patient, $i = 1,...,n$, is

$$\begin{aligned} l_i = &-\{A_{i1}(Y_{i1}) + A_{i2}(Y_{i1}) + A_{i12}(Y_{i1}, Y_{i2}) + (Y_{i1} + Y_{i2}) u_i\} \\ &+ \delta_{i1} \log(a_{i1} + u_i) + (1 - \delta_{i1}) \delta_{i2} \log(a_{i2} + u_i) + \delta_{i1}\delta_{i2} \log(a_{i12} + u_i) \end{aligned}$$

$$+(\theta^{-1}-1)\log(u_i)-\theta^{-1}u_i-\log(\Gamma(1/\theta))-\theta^{-1}\log(\theta). \qquad (27)$$

The log full likelihood function for all patients is $l=\sum_{i=1}^{n} l_i$.

Numerical integration techniques, Gauss-Laguerre quadrature, were applied to calculate the expectation of the function of frailty such as $E(u_i|O_i)$, $E(\log(u_i)|O_i)$ and $E((const.+u_i)^{-1}|O_i)$, $i=1,...,n$, from the density of $f(u_i|O_i)$ in E-step of EM algorithm. A brief introduction to the Gauss-Laguerre quadrature is given in Appendix C.

In M-step, parameter estimates are obtained by maximizing the full log likelihood as if the frailty, an estimation in E-step is known. The partial derivatives for $\eta=(\theta,\beta_1',\beta_2',\beta_3',\lambda_{01}',\lambda_{02}',\lambda_{12}')'$ are

$$U_{\beta_1}=\sum_{i=1}^{n}\left\{-Y_{i1}+\delta_{i1}(a_{i1}+u_i)^{-1}\right\}z_i,$$

$$U_{\beta_2}=\sum_{i=1}^{n}\left\{-Y_{i1}+(1-\delta_{i1})\delta_{i2}(a_{i2}+u_i)^{-1}\right\}z_i,$$

$$U_{\beta_3}=\sum_{i=1}^{n}\left\{-(Y_{i2}-Y_{i1})+\delta_{i1}\delta_{i2}(a_{i12}+u_i)^{-1}\right\}z_i,$$

$$U_{\lambda_{01k}}=-\sum_{i=1}^{n}\left[\sum_{k=1}^{K}\left\{(Y_{i1}-s_{k-1})d_{1ik}+(s_k-s_{k-1})\left(1-\sum_{l=1}^{k}d_{1il}\right)\right\}\right]+\sum_{i=1}^{n}\delta_{i1}(a_{i1}+u_i)^{-1},$$
$$k=1,2,...,K,$$

$$U_{\lambda_{02k}}=-\sum_{i=1}^{n}\left[\sum_{k=1}^{K}\left\{(Y_{i1}-s_{k-1})d_{1ik}+(s_k-s_{k-1})\left(1-\sum_{l=1}^{k}d_{1il}\right)\right\}\right]+\sum_{i=1}^{n}(1-\delta_{i1})\delta_{i2}(a_{i2}+u_i)^{-1},$$
$$k=1,2,...,K,$$

$$U_{\lambda_{121}} = -\sum_{i=1}^{n}\{(Y_{i2}-Y_{i1})d_{1i1}d_{2i1}+(s_1-Y_{i1})d_{1i1}(1-d_{2i1})\}+\sum_{i=1}^{n}\delta_{i1}\delta_{i2}(a_{i12}+u_i)^{-1},$$

$$U_{\lambda_{12k}} = -\sum_{i=1}^{n}\left[\sum_{k=2}^{K}\left\{(Y_{i2}-Y_{i1})d_{1ik}\,d_{2ik}+(s_k-Y_{i1})d_{1ik}\left(1-\sum_{l=1}^{k}d_{2il}\right)+\right.\right.$$

$$\left.\left.+(s_k-s_{k-1})\left(\sum_{l=1}^{k-1}d_{1il}\right)\left(1-\sum_{l=1}^{k}d_{2il}\right)+(Y_{i2}-s_{k-1})\left(\sum_{l=1}^{k-1}d_{1il}\right)d_{2ik}\right\}\right]$$

$$+\sum_{i=1}^{n}\delta_{i1}\delta_{i2}(a_{i12}+u_i)^{-1}, \qquad k=2,\ldots,K,$$

and

$$U_\theta = \theta^{-2}\left[n\psi(\theta^{-1})-n(1-\log(\theta))-\sum_{i=1}^{n}\{\log(u_i)-u_i\}\right].$$

All second partial derivatives for the parameters are shown in Appendix D. Since EM algorithm does not provide the information matrix for the observed data likelihood directly, Louis' s formula (Louis, 1982) is used to obtain it. The observed information matrix is given by

$$\widehat{I^{-1}} = -\hat{E}\left\{\frac{\partial^2 l}{\partial\eta\partial\eta'}|O,\hat{\eta}\right\}-\hat{E}\left\{\frac{\partial l}{\partial\eta}\frac{\partial l}{\partial\eta'}|O,\hat{\eta}\right\}+\hat{E}\left\{\frac{\partial l}{\partial\eta}|O,\hat{\eta}\right\}\hat{E}\left\{\frac{\partial l}{\partial\eta'}|O,\hat{\eta}\right\}.$$

All of these terms are evaluated at the last iteration of the EM algorithm. The first term is directly calculated as complete information, whereas the remaining terms as negative missing information are computed by the equation (7) from Friedl and Kauermann (2000).

Our numerical approach to solve $U(\eta)=0$ converges and can be summarized as

follows:

Step 0. Provide initial estimates,

$$\eta^{(0)} = \left\{\theta^{(0)}, \beta_1'^{(0)}, \beta_2'^{(0)}, \beta_3'^{(0)}, \lambda_{01}'^{(0)}, \lambda_{02}'^{(0)}, \lambda_{12}'^{(0)}\right\}',$$

Step 1. (E-step) Compute $E(u_i|O_i)$, $E(\log(u_i)|O_i)$ and $E((const.+u_i)^{-1}|O_i)$, $i=1,...,n$, based on the Gauss-Laguerre quadrature method, given the current values of the parameters,

Step 2. (M-step) Update the maximum likelihood estimates of parameters using Newton-Raphson method, and

Step 3. Iterate between Steps 1 and 2 until the estimates converge.

The estimate of the asymptotic variance of $(\widehat{\beta_1}', \widehat{\lambda_{01}}')'$ was estimated using the relevant submatrix from the inverse of the observed information matrix, $\hat{I}^{-1}$ for η. Similarly, the asymptotic covariance functions of $(\widehat{\beta_2}', \widehat{\lambda_{02}}')'$ and $(\widehat{\beta_3}', \widehat{\lambda_{12}}')'$ were obtained from the submatrices of $\hat{I}^{-1}$ associated with $(\beta_2', \lambda_{02}')'$ and $(\beta_3', \lambda_{12}')'$, respectively. Finally, the asymptotic covariance function of $\hat{\theta}$ was obtained from the submatrices of $\hat{I}^{-1}$ associated with θ.

IV. Simulation

4.1 Performance for proposed models

In this section we conducted simulations to evaluate the performance of the proposed model under five different scenarios. Observations were generated from the model (18)-(20). For simplicity, we considered three intervals $S_1 = (0, 0.15]$, $S_2 = (0.15, 0.77]$ and $S_3 = (0.77, \infty)$ for piecewise baseline hazards. Let $\lambda_{01} = \{\lambda_{011}, \lambda_{012}, \lambda_{013}\} = (0.2, 0.3, 0.4)$, $\lambda_{02} = \{\lambda_{021}, \lambda_{022}, \lambda_{023}\} = (0.2, 0.3, 0.4)$ and $\lambda_{12} = \{\lambda_{121}, \lambda_{122}, \lambda_{123}\} = (0.3, 0.4, 0.5)$. We set up a single binary covariate Z taking a value 0 or 1, each with probability 0.5 and $\beta_1 = 1$, $\beta_2 = 1$, $\beta_3 = 1$ and sample size n=300. The censoring time was simulated from a uniform distribution on (0, 5). Frailties were generated from a gamma distribution with mean 1 and variance θ=0.3 and θ=0.8. The observable data $\{Y_{i1}, Y_{i2}, \delta_{i1}, \delta_{i2}, z_i,\ i = 1, \ldots n\}$ were determined by the following steps.

Step 1: Generate a relapse time t_{i1} from the hazard $\lambda_1(t_{i1}|u_i, z_i) = \lambda_{01}(t_{i1}) + u_i + \beta_1 z_i$, a death time without the relapse t_{i2} from the hazard $\lambda_2(t_{i2}|u_i, z_i) = \lambda_{02}(t_{i2}) + u_i + \beta_2 z_i$, and a censoring time c_i.

Step 2: In the case of $c_i \leq \min(t_{i1}, t_{i2})$, observable data are $Y_{i1} = Y_{i2} = c_i$, $\delta_{i1} = 0$ and $\delta_{i2} = 0$.

Step 3: In the case of $t_{i2} \leq \min(t_{i1}, c_i)$, observable data are $Y_{i1} = Y_{i2} = t_{i2}$,

$\delta_{i1} = 0$ and $\delta_{i2} = 1$.

Step 4: In the case of $t_{i1} \leqq \min(t_{i2}, c_i)$, generate a death time with the relapse t_{i2}' from the hazard $\lambda_{12}(t_{i2}|t_{i1}, u_i, z_i) = \lambda_{12}(t_{i2}) + u_i + \beta_3 z_i$.

Step 5: In the case of $t_{i1} < c_i \leq t'_{i2}$, observable data are $Y_{i1} = t_{i1}$, $Y_{i2} = c_i$, $\delta_{i1} = 1$ and $\delta_{i2} = 0$.

Step 6: In the case of $t_{i1} < t'_{i2} \leq c_i$, observable data are $Y_{i1} = t_{i1}$, $Y_{i2} = t'_{i2}$, $\delta_{i1} = 1$ and $\delta_{i2} = 1$.

Results are evaluated based on the average estimate (Estimate), average bias (Bias), bias rate over the true value (Bias%), the empirical standard deviation (SD), the mean value of the estimated standard errors (MSE), and the coverage probability (CP) of the nominal 95% confidence intervals. Estimate is the mean of the parameter estimates (based on 1,000 replicates), Bias is the mean of the parameter estimates minus the true value, Bias (%) is the Bias rate (%) over the true value, |Bias|/True× 100, SD is the empirical standard deviation of the parameter estimate, MSE is the mean value of the estimated standard errors and CP is the coverage probability of the nominal 95% confidence intervals.

The results obtained from 1,000 replications are summarized in Table 1 ($\theta = 0.3$) and Table 2 ($\theta = 0.8$). In the setting of the censoring data, average 15% of total patients was censored. Average 40% of total patients experience relapse and death with relapse, average 45% only death without relapse and the rest 15% are censored in which 7% have only relapse and 8% have no event.

As shown in Table 1, the magnitudes of the empirical biases of the

estimates are close to zero and the CPs reach a nominal level except for the baseline hazard parameters in the last intervals having 10.3~18.5% of bias and 92~95% of the CPs when θ is 0.3. The variance of the frailty has a low CP when θ is 0.3. The biases and CPs of the baseline hazard parameters are most likely due to the small numbers of observations in the intervals. For the variance of the frailty, CP is relatively low and the MSE does not agree with the empirical SD. It is clear that the SE of θ is underestimated in the setting of $\theta = 0.3$ since the MSE are lower than the SD. In contrast, the SE of θ is overestimated in the setting of $\theta = 0.8$ since the MSE are higher than the SD. There seem to have decreases in biases, SDs and MSEs for all parameters in the setting of $\theta = 0.8$. The magnitudes of the empirical biases and MSEs of the estimates are larger in 15% censoring, comparing to those in no censoring. The patterns drawn with regards to the biases, MSEs and CPs of parameters do not change when the proportion of censoring is from 0% to 15%.

Figure 2 shows box-plots for $\widehat{\beta_1}$, $\widehat{\beta_2}$, $\widehat{\beta_3}$, $\widehat{\lambda_{01}}$,$\widehat{\lambda_{02}}$ and $\widehat{\lambda_{12}}$ in the setting of $\theta = 0.8$ with 15% censoring and $n = 300$. In the distributions of all parameters, they are approximately symmetric and normal although their distribution are a little skewed to the right. The histogram for $\hat{\theta}$ is shown in Figure 3. $\hat{\theta}$ is approximately normal distributed with a little skewed to the left.

Table 1. Simulation results I: Parameter estimates for θ=0.3 and n=300

		Proposed Model					
Parameter	True	Estimate	Bias	(Bias%)	SD	MSE	CP
			Censoring=0%				
β_1	1	1.023	0.023	(2.3)	0.297	0.288	0.952
β_2	1	1.012	0.012	(1.2)	0.287	0.289	0.955
β_3	1	1.030	0.030	(3.0)	0.336	0.328	0.947
θ	0.3	0.315	0.015	(4.9)	0.168	0.124	0.863
λ_{011}	0.2	0.195	-0.005	(2.6)	0.217	0.219	0.945
λ_{012}	0.3	0.313	0.013	(4.5)	0.186	0.169	0.922
λ_{013}	0.4	0.455	0.055	(13.8)	0.331	0.276	0.941
λ_{021}	0.2	0.204	0.004	(1.8)	0.218	0.222	0.948
λ_{022}	0.3	0.318	0.018	(6.0)	0.189	0.170	0.922
λ_{023}	0.4	0.450	0.050	(12.4)	0.307	0.275	0.937
λ_{121}	0.3	0.307	0.007	(2.4)	0.727	0.687	0.925
λ_{122}	0.4	0.405	0.005	(1.3)	0.264	0.247	0.935
λ_{123}	0.5	0.545	0.045	(8.9)	0.207	0.180	0.934
			Censoring=15%				
β_1	1	1.031	0.031	(3.1)	0.314	0.296	0.952
β_2	1	1.020	0.020	(2.0)	0.302	0.297	0.947
β_3	1	1.034	0.034	(3.4)	0.373	0.358	0.943
θ	0.3	0.316	0.016	(5.5)	0.161	0.142	0.892
λ_{011}	0.2	0.195	-0.005	(2.7)	0.217	0.220	0.940
λ_{012}	0.3	0.313	0.013	(4.4)	0.187	0.176	0.929
λ_{013}	0.4	0.464	0.064	(16.1)	0.408	0.327	0.951
λ_{021}	0.2	0.195	-0.005	(2.4)	0.221	0.222	0.932
λ_{022}	0.3	0.318	0.018	(5.8)	0.196	0.177	0.923
λ_{023}	0.4	0.474	0.074	(18.5)	0.390	0.330	0.950
λ_{121}	0.3	0.291	-0.009	(3.0)	0.781	0.702	0.928
λ_{122}	0.4	0.398	-0.002	(0.6)	0.277	0.262	0.936
λ_{123}	0.5	0.552	0.052	(10.3)	0.254	0.227	0.934

Estimate is the mean of the parameter estimates (based on 1,000 replicates); Bias is the mean of the parameter estimates minus the true value; Bias(%) is the Bias rate (%) over the true value, |Bias|/True×100; SD is the empirical standard deviation of the parameter estimate; MSE is the mean value of the estimated standard errors; CP is the coverage probability of the nominal 95% confidence intervals.

Table 2. Simulation results I: Parameter estimates for θ=0.8 and n=300

			Proposed Model				
Parameter	True	Estimate	Bias	(Bias%)	SD	MSE	CP
			Censoring=0%				
β_1	1	1.013	0.013	(1.3)	0.263	0.263	0.958
β_2	1	1.011	0.011	(1.1)	0.273	0.265	0.938
β_3	1	1.028	0.028	(2.8)	0.318	0.308	0.949
θ	0.8	0.780	-0.020	(2.5)	0.031	0.130	0.999
λ_{011}	0.2	0.209	0.009	(4.7)	0.190	0.199	0.929
λ_{012}	0.3	0.311	0.011	(3.6)	0.141	0.144	0.958
λ_{013}	0.4	0.428	0.028	(6.9)	0.180	0.173	0.945
λ_{021}	0.2	0.206	0.006	(3.0)	0.213	0.204	0.896
λ_{022}	0.3	0.308	0.008	(2.7)	0.151	0.147	0.943
λ_{023}	0.4	0.435	0.035	(8.7)	0.183	0.175	0.948
λ_{121}	0.3	0.411	0.111	(37.0)	0.735	0.743	0.929
λ_{122}	0.4	0.412	0.012	(3.0)	0.250	0.251	0.951
λ_{123}	0.5	0.531	0.031	(6.2)	0.140	0.140	0.952
			Censoring=15%				
β_1	1	1.013	0.013	(1.3)	0.269	0.272	0.954
β_2	1	1.016	0.016	(1.6)	0.291	0.273	0.939
β_3	1	1.029	0.029	(2.9)	0.355	0.345	0.939
θ	0.8	0.779	-0.021	(2.6)	0.024	0.134	1.000
λ_{011}	0.2	0.207	0.007	(3.5)	0.193	0.201	0.922
λ_{012}	0.3	0.311	0.011	(3.7)	0.151	0.151	0.956
λ_{013}	0.4	0.443	0.043	(10.7)	0.217	0.214	0.958
λ_{021}	0.2	0.203	0.003	(1.4)	0.216	0.204	0.881
λ_{022}	0.3	0.305	0.005	(1.7)	0.159	0.153	0.940
λ_{023}	0.4	0.444	0.044	(11.0)	0.258	0.216	0.947
λ_{121}	0.3	0.404	0.104	(34.8)	0.765	0.760	0.929
λ_{122}	0.4	0.417	0.017	(4.2)	0.268	0.268	0.948
λ_{123}	0.5	0.533	0.033	(6.6)	0.197	0.193	0.954

Estimate is the mean of the parameter estimates (based on 1,000 replicates); Bias is the mean of the parameter estimates minus the true value; Bias(%) is the Bias rate (%) over the true value, |Bias|/True×100; SD is the empirical standard deviation of the parameter estimate; MSE is the mean value of the estimated standard errors; CP is the coverage probability of the nominal 95% confidence intervals.

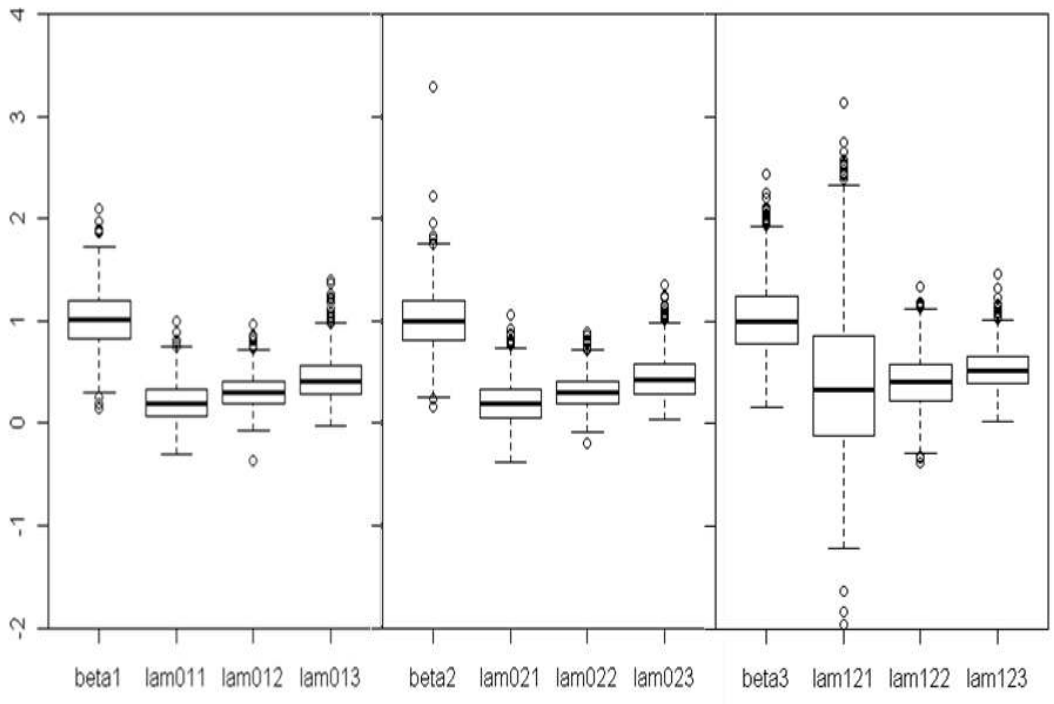

Figure 2. Box-plots for parameter estimates ($\theta = 0.8$, censoring=15% and n=300)

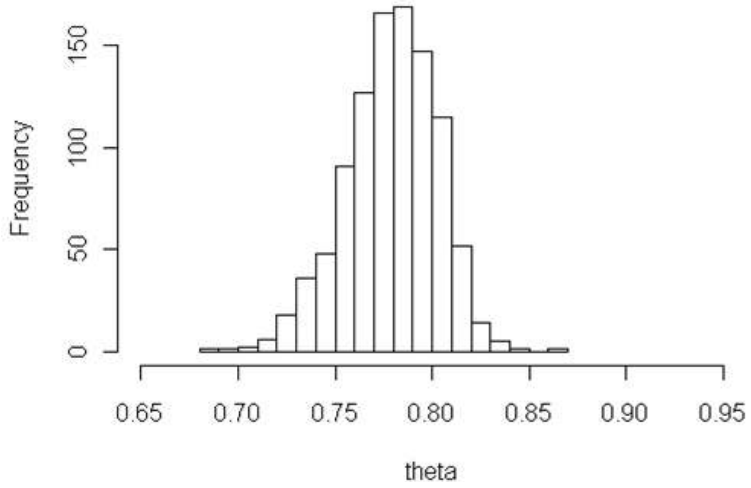

Figure 3. The histogram for $\theta = 0.8$ (censoring=15% and n=300)

4.2 Adaptation of frailty

To evaluate how frailty affects the parameter estimation in the proposed model, We compared the proposed model with the proposed model without considering frailty in the setting of frailty existence (setting II). The results are shown in Table 3 ($\theta = 0.3$) and 4 ($\theta = 0.8$). As expected, it is clear that ignoring the dependence among the hazards can result in significant biases, especially for baseline hazards. In the model without frailty, the biases in regression estimates are not largely increased but in the piecewise baseline parameters. However, MSEs in the model without frailty are a little increased, compared to the model with frailty in the both cases of $\theta = 0.3$ and 0.8.

Table 3. Simulation results II: Parameter estimates for θ=0.3 and n=300

		Censoring=0%														Censoring=15%											
		Proposed Model(A)						Proposed Model without Frailty(B)						Comparison		Proposed Model(A)						Proposed Model without Frailty(B)						Comparison	
	True	Est.	Bias	(Bias %)	SD	MSE	CP	Est.	Bias	(Bias %)	SD	MSE	CP	Bias% (B/A)	MSE (B/A)	Est.	Bias	(Bias %)	SD	MSE	CP	Est.	Bias	(Bias %)	SD	MSE	CP	Bias% (B/A)	MSE (B/A)
β_1	1	1.02	0.02	(2.3)	0.30	0.29	0.95	1.03	0.03	(3.4)	0.29	0.29	0.95	1.45	1.00	1.03	0.03	(3.1)	0.31	0.30	0.95	1.04	0.04	(3.8)	0.30	0.30	0.95	1.20	1.00
β_2	1	1.01	0.01	(1.2)	0.29	0.29	0.96	1.03	0.03	(2.5)	0.29	0.29	0.96	2.08	1.00	1.02	0.02	(1.9)	0.30	0.30	0.95	1.03	0.03	(2.7)	0.30	0.30	0.96	1.40	1.00
β_3	1	1.03	0.03	(3.0)	0.34	0.33	0.95	0.99	-0.01	(1.0)	0.32	0.33	0.95	0.34	0.99	1.03	0.03	(3.4)	0.37	0.36	0.94	0.97	-0.03	(3.1)	0.36	0.36	0.95	0.91	1.00
θ	0.3	0.31	0.01	(4.9)	0.17	0.12	0.86									0.32	0.02	(5.5)	0.16	0.14	0.89								
λ_{011}	0.2	0.19	-0.01	(2.6)	0.22	0.22	0.95	0.15	-0.05	(23.9)	0.22	0.22	0.94	9.25	1.01	0.19	-0.01	(2.8)	0.22	0.22	0.94	0.15	-0.05	(23.9)	0.22	0.22	0.95	8.47	1.02
λ_{012}	0.3	0.31	0.01	(4.5)	0.19	0.17	0.92	0.13	-0.17	(56.7)	0.18	0.17	0.78	12.67	1.01	0.31	0.01	(4.5)	0.19	0.18	0.93	0.13	-0.17	(56.3)	0.18	0.18	0.79	12.50	1.01
λ_{013}	0.4	0.46	0.06	(13.8)	0.33	0.28	0.94	0.04	-0.36	(89.3)	0.30	0.28	0.62	6.49	1.01	0.46	0.06	(16.0)	0.41	0.33	0.95	0.07	-0.33	(82.8)	0.35	0.33	0.69	5.19	1.00
λ_{021}	0.2	0.20	0.00	(1.8)	0.22	0.22	0.95	0.16	-0.04	(20.9)	0.22	0.22	0.93	11.86	1.00	0.20	0.00	(2.2)	0.22	0.22	0.93	0.15	-0.05	(24.2)	0.22	0.22	0.94	11.06	1.01
λ_{022}	0.3	0.32	0.02	(6.0)	0.19	0.17	0.92	0.13	-0.17	(55.2)	0.18	0.17	0.78	9.21	1.00	0.32	0.02	(5.7)	0.20	0.18	0.92	0.14	-0.16	(54.7)	0.19	0.18	0.78	9.52	1.00
λ_{023}	0.4	0.45	0.05	(12.4)	0.31	0.27	0.94	0.04	-0.36	(90.9)	0.29	0.28	0.63	7.32	1.01	0.47	0.07	(18.6)	0.39	0.33	0.95	0.08	-0.32	(80.3)	0.35	0.33	0.72	4.32	1.00
λ_{121}	0.3	0.31	0.01	(2.4)	0.73	0.69	0.93	0.50	0.20	(65.3)	0.66	0.70	0.97	27.16	1.02	0.29	-0.01	(2.5)	0.78	0.70	0.93	0.51	0.21	(69.0)	0.68	0.71	0.97	27.84	1.01
λ_{122}	0.4	0.41	0.01	(1.3)	0.26	0.25	0.94	0.41	0.01	(3.6)	0.26	0.25	0.94	2.82	0.99	0.40	0.00	(0.6)	0.28	0.26	0.94	0.42	0.02	(4.7)	0.27	0.26	0.95	7.25	1.00
λ_{123}	0.5	0.54	0.04	(8.9)	0.21	0.18	0.93	0.27	-0.23	(45.5)	0.21	0.18	0.66	5.09	1.01	0.55	0.05	(10.4)	0.25	0.23	0.93	0.32	-0.18	(36.1)	0.25	0.23	0.79	3.48	1.01

Table 4. Simulation results II: Parameter estimates for θ=0.8 and n=300

		Censoring=0%														Censoring=15%											
		Proposed Model(A)						Proposed Model without Frailty(B)						Comparison		Proposed Model(A)						Proposed Model without Frailty(B)						Comparison	
	True	Est.	Bias	(Bias %)	SD	MSE	CP	Est.	Bias	(Bias %)	SD	MSE	CP	Bias% (B/A)	MSE (B/A)	Est.	Bias	(Bias %)	SD	MSE	CP	Est.	Bias	(Bias %)	SD	MSE	CP	Bias% (B/A)	MSE (B/A)
β_1	1	1.01	0.01	(1.3)	0.26	0.26	0.96	1.04	0.04	(3.9)	0.27	0.27	0.96	2.92	1.01	1.01	0.01	(1.3)	0.27	0.27	0.95	1.04	0.04	(3.6)	0.27	0.28	0.96	2.87	1.02
β_2	1	1.01	0.01	(1.1)	0.27	0.27	0.94	1.04	0.04	(3.7)	0.28	0.27	0.94	3.45	1.01	1.02	0.02	(1.6)	0.29	0.27	0.94	1.04	0.04	(3.7)	0.29	0.28	0.94	2.39	1.01
β_3	1	1.03	0.03	(2.8)	0.32	0.31	0.95	0.97	-0.03	(2.9)	0.30	0.30	0.95	1.03	0.99	1.03	0.03	(2.9)	0.35	0.34	0.94	0.93	-0.07	(7.4)	0.34	0.34	0.94	2.57	0.99
θ	0.8	0.78	-0.02	(2.5)	0.03	0.13	1.00									0.78	-0.02	(2.6)	0.02	0.13	1.00								
λ_{011}	0.2	0.21	0.01	(4.7)	0.19	0.20	0.93	0.10	-0.10	(48.8)	0.20	0.22	0.93	10.31	1.08	0.21	0.01	(3.5)	0.19	0.20	0.92	0.10	-0.10	(49.9)	0.20	0.22	0.92	14.05	1.08
λ_{012}	0.3	0.31	0.01	(3.6)	0.14	0.14	0.96	-0.06	-0.36	(120.4)	0.15	0.15	0.32	33.73	1.03	0.31	0.01	(3.7)	0.15	0.15	0.96	-0.06	-0.36	(118.9)	0.16	0.16	0.38	32.37	1.03
λ_{013}	0.4	0.43	0.03	(6.9)	0.18	0.17	0.95	-0.26	-0.66	(163.8)	0.19	0.18	0.13	23.75	1.02	0.44	0.04	(10.7)	0.22	0.21	0.96	-0.22	-0.62	(154.6)	0.23	0.22	0.23	14.51	1.01
λ_{021}	0.2	0.21	0.01	(3.0)	0.21	0.20	0.90	0.10	-0.10	(51.5)	0.22	0.21	0.90	17.17	1.05	0.20	0.00	(1.4)	0.22	0.20	0.88	0.10	-0.10	(52.4)	0.22	0.22	0.89	36.97	1.06
λ_{022}	0.3	0.31	0.01	(2.7)	0.15	0.15	0.94	-0.06	-0.36	(121.2)	0.16	0.15	0.33	44.60	1.01	0.31	0.01	(1.7)	0.16	0.15	0.94	-0.06	-0.36	(120.7)	0.16	0.15	0.36	71.23	1.01
λ_{023}	0.4	0.43	0.03	(8.7)	0.18	0.17	0.95	-0.25	-0.65	(161.4)	0.19	0.18	0.14	18.46	1.01	0.44	0.04	(11.0)	0.26	0.22	0.95	-0.21	-0.61	(153.1)	0.23	0.22	0.26	13.94	1.00
λ_{121}	0.3	0.41	0.11	(37.0)	0.73	0.74	0.93	0.78	0.48	(159.1)	0.71	0.75	0.96	4.30	1.02	0.40	0.10	(34.8)	0.77	0.76	0.93	0.80	0.50	(167.3)	0.73	0.77	0.96	4.81	1.01
λ_{122}	0.4	0.41	0.01	(3.0)	0.25	0.25	0.95	0.41	0.01	(1.4)	0.24	0.25	0.95	0.45	0.99	0.42	0.02	(4.2)	0.27	0.27	0.95	0.43	0.03	(8.2)	0.26	0.27	0.95	1.95	1.00
λ_{123}	0.5	0.53	0.03	(6.2)	0.14	0.14	0.95	0.02	-0.48	(96.3)	0.17	0.14	0.17	15.63	1.03	0.53	0.03	(6.6)	0.20	0.19	0.95	0.10	-0.40	(80.3)	0.21	0.20	0.44	12.08	1.02

4.3 Mis-specification of the baseline hazard function

In the real data, we need to select the appropriate intervals in time domains for the piecewise baseline hazards in the proposed model. We may suggest to plot the cumulative coefficient for the intercept on Lin and Ying (1994)'s additive hazard models for the times to events. However, it is a just guess. From the plots, we may just approximate one around the true point.

Thus, we simulated to explore the robustness of the regression estimates in the proposed model on the mis-specificated intervals for the baseline hazard (setting III). In the simulation setting, five piecewise hazards were assumed to have parameters $\lambda_{0R}^T = \{\lambda_{011}, \lambda_{012}, \lambda_{013}, \lambda_{014}, \lambda_{015}\} = (0.2, 0.23, 0.3, 0.36, 0.4)$, $\lambda_{0D}^T = \lambda_{0R}^T$ and $\lambda_{0RD}^T = \{\lambda_{121}, \lambda_{122}, \lambda_{123}, \lambda_{124}, \lambda_{125}\} = (0.3, 0.33, 0.4, 0.46, 0.5)$ in the five intervals, $S_1 = (0, 0.09]$, $S_2 = (0.09, 0.27]$, $S_3 = (0.27, 0.65]$, $S_4 = (0.65, 1.46]$ and $S_5 = (1.46, \infty)$. The parameters were from the weighted means of the simulation setting shown in Table 1. The results of the simulation are displayed in Table 5. From our simulation setting, bias rates of regression parameters are between 0.6~4.8% among both cases with no censoring and with 15% censoring. MSEs and CPs are similar to the results from the simulation setting I.

Table 5. Simulation results III: Mis-specification of the intervals for piecewise baseline hazards (n=300)

				Proposed Model		
Parameter	True	Estimate	Bias	(Bias%)	SD	MSE	CP
Censoring=0%			θ=0.3			
β_1	1	1.022	0.022	(2.2)	0.283	0.285	0.949
β_2	1	1.006	0.006	(0.6)	0.286	0.286	0.954
β_3	1	1.029	0.029	(2.9)	0.327	0.320	0.944
θ	0.3	0.286	-0.014	(4.8)	0.148	0.110	0.867
			θ=0.8				
β_1	1	1.008	0.008	(0.8)	0.261	0.260	0.957
β_2	1	1.007	0.007	(0.7)	0.268	0.263	0.943
β_3	1	1.023	0.023	(2.3)	0.304	0.297	0.948
θ	0.8	0.774	-0.026	(3.2)	0.044	0.128	0.997
Censoring=15%			θ=0.3				
β_1	1	1.026	0.026	(2.6)	0.308	0.294	0.950
β_2	1	1.014	0.014	(1.4)	0.311	0.294	0.950
β_3	1	1.025	0.025	(2.5)	0.363	0.352	0.940
θ	0.3	0.287	-0.013	(4.4)	0.133	0.131	0.912
			θ=0.8				
β_1	1	1.011	0.011	(1.1)	0.267	0.269	0.957
β_2	1	1.008	0.008	(0.8)	0.281	0.272	0.947
β_3	1	1.037	0.037	(3.7)	0.352	0.336	0.938
θ	0.8	0.773	-0.027	(3.3)	0.028	0.133	1.000

Estimate is the mean of the parameter estimates (based on 1,000 replicates); Bias is the mean of the parameter estimates minus the true value; Bias(%) is the Bias rate (%) over the true value, |Bias|/True×100; SD is the empirical standard deviation of the parameter estimate; MSE is the mean value of the estimated standard errors; CP is the coverage probability of the nominal 95% confidence intervals.

4.4 Mis-specification of frailty distribution

We adapted the gamma distribution which was the popular one used for the frailty model. In this simulation (setting IV), we examined the robustness of the regression and hazard function parameters against the mis-specification of the gamma frailty distribution. Hsu, Gorfine and Malone (2007) studied the effects of the frailty distribution mis-specification on the marginal regression estimates and hazards functions under the gamma distribution. We set up their simulation under the true Gamma distribution and possible alternative frailty distributions such as inverse Gaussian or positive stable distribution, and then compared with the simulation results of the proposed model. To allow for comparability across various distributions, we made Kendall's τ, which was the same for all distributions. Kendall's τ is a commonly used measure of dependency in characterizing the dependence function for bivariate survival times when comparing with various distributions. Table 6 displays the parameter values under the distributions that correspond to $\tau = 0.13$ and 0.28, corresponding to mild and moderate dependency between paired failure times. Their density functions and the distributions are presented in Appendix E.

Table 6. Parameter values used in the simulation

Kendall's τ	0.13	0.28
Gamma θ	0.30	0.80
Inverse Gaussian η	0.79	3.37
Positive Stable α	0.87	0.72

Table 7. Simulation results IV: Mis-specification of frailty distribution (n=300)

			Bias(%)			MSE			CP		
τ		True	G	IG	PS	G	IG	PS	G	IG	PS
Censoring=0%											
0.13	β_1	1	0.0	0.6	0.3	0.29	0.31	0.28	0.945	0.959	0.952
	β_2	1	1.2	2.5	1.3	0.29	0.31	0.29	0.958	0.953	0.947
	β_3	1	4.6	1.9	3.6	0.33	0.35	0.33	0.948	0.958	0.950
	θ	0.3	7.4			0.12			0.861		
0.28	β_1	1	0.1	2.4	4.9	0.26	0.32	0.26	0.954	0.947	0.935
	β_2	1	0.5	2.7	4.4	0.26	0.32	0.26	0.950	0.937	0.933
	β_3	1	1.3	3.6	1.1	0.30	0.36	0.29	0.939	0.947	0.952
	θ	0.8	2.4			0.13			0.998		
Censoring=15%											
0.13	β_1	1	0.2	0.7	4.1	0.29	0.32	0.30	0.947	0.960	0.946
	β_2	1	1.9	2.6	3.5	0.30	0.32	0.30	0.958	0.953	0.947
	β_3	1	4.1	0.7	3.6	0.36	0.38	0.36	0.939	0.952	0.954
	θ	0.3	9.5			0.14			0.874		
0.28	β_1	1	1.2	2.5	6.1	0.27	0.33	0.27	0.949	0.949	0.938
	β_2	1	1.0	3.1	3.2	0.27	0.33	0.26	0.949	0.945	0.939
	β_3	1	0.9	4.8	1.4	0.34	0.39	0.33	0.952	0.951	0.958
	θ	0.8	2.6			0.13			0.999		

Bias(%) is the Bias rate (%) over the true value, |Bias|/True×100; MSE is the mean value of the estimated standard errors; CP is the coverage probability of the nominal 95% confidence intervals.

Gamma distribution, inverse Gaussian distribution and positive stable distribution are abbreviated as G, IG and PS.

Table 7 shows a summary of biases (%), empirical standard errors and coverage probabilities of the proposed model. Biases (%) are generally under 10% and lower, even when the true frailty distribution deviates from the assumed gamma distribution. When the dependency between paired failure times is moderate ($\tau = 0.28$), the biases and MSEs increase, comparing to the results from the case of $\tau = 0.13$. For positive stable distribution, the biases are relatively high with 15% censoring. Similar patterns with the study findings by Hsu, Gorfine and Malone (2007) are determined in our study. As they have notated, this is probably because of the infinity dependence at the time of origin, which makes the assumed gamma model a poor fit for the very early ages at onset, and a problem which is more pronounced with higher censoring rates. In the simulation results, the regression estimates under the assumed gamma frailty model appear to be minimally affected when the true frailty distribution is inverse Gaussian or positive stable.

V. Colon cancer data

This data are from one of the first successful trials of adjuvant chemotherapy for colon cancer. This trial is a national intergroup trial in the 1980's to study the effectiveness of two adjuvant therapy regimens in improving surgical cure rates in stage III (Dukes stage C) colon cancer. Levamisole is a low-toxicity compound to treat worm infestations in animals; 5-FU is a moderately toxic chemotherapy agent. In the study, 315, 310 and 304 patients with stage C disease received observation care only, levamisole alone and levamisole plus fluorouracil, respectively. The time to cancer recurrence and the survival time were both considered important outcome measures. The results showed that levamisole plus fluorouracil was highly effective in prolonging survival and reducing the risk of cancer recurrence. The main report is given by Moertel et al.(1995). This data have been analysed by Lin (1994) to illustrate the marginal approach to multivariate time-to-event data, by Shen and Thall (1998) to illustrate the parametric likelihoods approach for multiple non-fatal competing risks and death and by He and Lawless (2003) to illustrate the flexible likelihood methods for bivariate proportional hazards models.

For this study, we focused only on the comparison between the observation and levamisole plus fluorouracil (Lev+5-FU) groups on the semi-competing risks problem with cancer recurrence and death. We treated cancer recurrence as the non-terminal event and death as the terminal event. Thus the total number of the patients were 619. Of the 315 patients receiving observation care alone, there were 177 recurrences and 168 deaths (of which 155 were after recurrence); of the 304 patients receiving levamisole plus fluorouracil, there were 119

recurrences and 123 deaths (of which 108 were after recurrence). The number of patients with no event was 295 (47.7% of all patients). Maximum follow-up was about 9 years. The summary of colon cancer data is shown in Table 8.

Table 8. A summary of colon cancer data (n=619)

	Observation only	Lev+5-FU
Total patients	315 (100)	304 (100)
Censoring with no event	125 (40)	170 (56)
Death without recurrence	13 (4)	15 (5)
Recurrence	177 (56)	119 (39)
Total patients with recurrence	177 (100)	119 (100)
Censoring with recurrence	22 (12)	11 (9)
Death with recurrence	155 (88)	108 (91)

(%)

For the proposed model, we firstly needed to select intervals for piecewise baseline hazard. To pick up the cut-points, we plotted the cumulative coefficient for intercept in Lin and Ying (1994)'s model that could give us where the additive hazard would be changed. We chose three intervals with two cut-points, 2 and 4 years, for colon cancer data. In Figure 4, the cumulative coefficient for intercept in Lin and Ying (1994)'s additive hazard model for the time of recurrence, time of death without recurrence and time of death with recurrence were evaluated. For the time of death with recurrence, patients with

recurrence (n=296) were only used in the Lin and Ying (1994)'s model. Table 9 displays the frequency of cancer patients in three intervals with two cut-points chosen in the time domain.

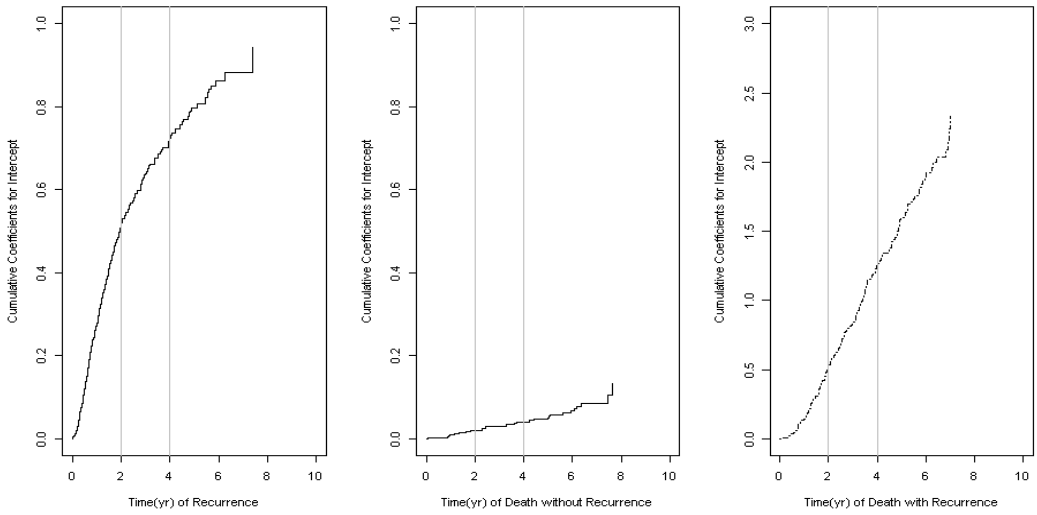

Figure 4. Cumulative coefficients for intercept in Lin and Ying (1994)'s model (colon cancer data)

Table 9. Frequency of colon cancer patients in three intervals

	Recurrence	Death without recurrence	Death with recurrence	No event
(0 , 2 yrs]	223	9	126	1
(2 yrs, 4 yrs]	50	7	92	2
(4 yrs, ∞)	23	12	45	292
Total	296	28	263	295

Table 10. Result of the proposed model for colon cancer data

	Excess risk	(SE)	95% CI	P-value
Lev+5FU				
Recurrence	-39.66	11.99	(-63.17, -16.15)	0.001
Death without recurrence	-1.76	4.41	(-10.40, 6.88)	0.690
Death with recurrence	152.68	81.58	(-7.21, 312.58)	0.061
Variance of frailty	0.78	0.11	(0.57, 0.99)	0.000
Baseline hazard				
Recurrence				
(0 , 2 yrs]	244.15	16.27	(212.26, 276.05)	0.000
(2 yrs, 4 yrs]	90.58	12.45	(66.18, 114.98)	0.000
(4 yrs, ∞)	56.88	11.93	(33.50, 80.26)	0.000
Death without recurrence				
(0 , 2 yrs]	9.24	4.17	(1.07, 17.42)	0.027
(2 yrs, 4 yrs]	9.52	4.07	(1.53, 17.50)	0.020
(4 yrs, ∞)	15.74	5.47	(5.02, 26.46)	0.004
Death with recurrence				
(0 , 2 yrs]	798.11	81.53	(638.31, 957.92)	0.000
(2 yrs, 4 yrs]	543.55	65.41	(415.34, 671.76)	0.000
(4 yrs, ∞)	330.05	58.91	(214.60, 445.51)	0.000

Unit: per a 1,000 person year

From the result of the proposed model, patients treated with fluorouracil plus levamisole had 39.66 patients risk reduction per a 1,000 person year in the hazard for the recurrence and 1.76 patients risk reduction per a 1,000 person year in the hazard for the death without recurrence when compared with patients assigned to have observation care only. However, with recurrence, patients treated with fluorouracil plus levamisole had 152.68 patients risk excess per a 1,000 person year in the hazard for the death despite no statistical significant result. Of the patients with recurrence, patients treated with fluorouracil plus levamisole had a somewhat shorter survival time than patients with observation care only (median times, 2.21 years and 2.54 years, respectively). Moertel (1995) indicated that patients failing to respond to fluorouracil plus levamisole had more virulent disease or, perhaps, that they had a resistance to fluorouracil-containing regimens, which were most often used for palliation after recurrence.

The hazard for recurrence and for death with recurrence in the interval from 0 to 2 years were relatively higher than that after 2 years. The hazard for death without recurrence increased over time. The hazard for death with recurrence was 20~86 times higher than the hazard for the death without recurrence. It is anticipated that recurrence is a relatively strong indicator of death. A correlation between recurrence and death was observed with the variance of the unobserved frailty 0.78 (95% CI 0.57-0.99) in a 1,000 person-year unit of the hazard.

VI. Conclusion and discussion

In biomedical studies, the non-terminal event such as relapse or recurrence is an important event in additional information about the death process. In the example of the colon cancer study, it was of interest to show the efficacy of treatment on both mortality and morbidity, because the relapse is a relatively strong indicator of death. Therefore, it will be useful to know whether the treatment reducing morbidity will also reduce mortality with morbidity. This situation can be treated on the semi-competing risks or the illness-death model in which the terminal event, death can preclude relapse, but not vice versa.

In this study, we focused on the interest to quantify the efficacy of the treatment on the three hazards, such as relapse, death with relapse and death without relapse, and proposed an illness-death model with additive treatment and additive frailty effects. Frailty is considered as the unmeasured factors in a subject. In recent literatures, Xu, Kalbfleisch and Tai (2010) proposed a Cox-type illness-death model with multiplicative frailty for semi-competing risks problem. Chong et al. (2011) presented an illness-death model with additive-type hazard function with multiplicative frailty. In contrast to the literature, in the proposed model, frailty is inserted additively. The proposed model maintains an additive structure for natural relationships. This idea is from the research by Silva and Turkman (2004) that adapted an additive frailty model for cluster failure data with a Bayesian point of view.

For the inference of parameters in the proposed model, we considered a full maximum likelihood on the complete data. We incorporated EM algorithm to deal with missing frailty and Gauss-Laguerre quadrature method for calculating the

functions of frailty at the expectation step based on the piecewise baseline hazards. We approximately expected the observed information matrix for standard errors of the parameter estimates.

As the result of the simulation, the estimates of the regressions and the piecewise hazard parameters generally have a good performance in whatever the association between relapse and death is mild or moderate. However, in the mild association between relapse and death, the parameter of frailty seems to be underestimated since the MSEs are less than the SDs. In existing the dependency between two events, the proposed model without considering frailty may violate the estimates of regressions and hazard parameters unreasonable. The magnitude of the violation in the moderate association was larger than in the mild association. We also simulated the mis-specification of intervals and frailty distribution. The violation of the proposed model was evaluated in the simulation setting with alternative intervals or frailty distribution. We found that the estimates of the regressions are robust on the similar possible distributions with the association of Kendall's $\tau \leq 0.28$.

We illustrated the proposed models on the colon cancer study. From the result, the treatment of Lev+5FU decreases the risk of cancer recurrence and death without recurrence but decreases survival time with recurrence. However, the treatment effect on survival time with recurrence is not statistically significant. These conclusions are in agreement with those in the research of Lin (1994) and He and Lawless (2003).

There are limitations in the proposed models. The piecewise baseline hazards are assumed to have different constants in the same intervals. There is a problem to determine the numbers and the ranges of intervals in real data. To approximate them, we could suggest to plot the cumulative coefficient for the intercept on Lin and Ying (1994)'s additive hazard model for the time of event.

Another limitation is that additive proposed models do not guarantee that the estimated hazard rates are greater than zero. It was typically found in very low risk of effects on hazards. This is one of the reason that additive models have not been widely used. However, the problem is not unique to the additive model but also occurs when using a product-limit estimated for fitted probabilities in the Cox model (Klein, 2006).

In the future works, it is possible to be extended to a semi-parametric model with unspecified baseline hazards and to Aalen (1989)'s flexible additive model over time. The goodness of fit assessment of the models could be a valuable research topic for real data analyses.

Appendix

Appendix A

Marginal hazard function in equation (21) can be derived from the marginal survival function and the marginal density distribution (Duchateau and Janssen, 2007) as follows:

$$S(t_1) = \int_0^\infty \exp^{-\{\Lambda_{01}(t_1)+t_1u+\Lambda_{02}(t_1)+t_1u)\}} g(u)du = \int_0^\infty \exp^{-\{\Lambda_{01}(t_1)+\Lambda_{02}(t_1)+2t_1u)\}} g(u)du$$
$$= \exp^{-\{\Lambda_{01}(t_1)+\Lambda_{02}(t_1)\}} (1+2\theta t_1)^{-1/\theta},$$

$$f(t_1) = \int_0^\infty (\lambda_{01}(t_1)+u)\exp^{-\{\Lambda_{01}(t_1)+t_1u+\Lambda_{02}(t_1)+t_1u\}} g(u)du$$
$$= \exp^{-\{\Lambda_{01}(t_1)+\Lambda_{02}(t_1)\}} \left[\lambda_{01}(t_1)\int \exp^{-2t_1u} g(u)du + \int u\exp^{-2t_1u} g(u)du\right]$$
$$= \exp^{-\{\Lambda_{01}(t_1)+\Lambda_{02}(t_1)\}} (1+2\theta t_1)^{-1/\theta}[\lambda_{01}(t_1)+(1+2\theta t_1)^{-1}],$$

$$\lambda_1(t_1) = f(t_1)/S(t_1) = \lambda_{01}(t_1)+(1+2\theta t_1)^{-1}.$$

Similarly, $\lambda_2(t_2 \mid z) = \lambda_{02}(t_2)+\beta_2'z+(1+2\theta t_2)^{-1}$. Also, the marginal likelihood function in the equation (23) comes from the joint survival function $S(t_1,t_2)$ and the density function $f(t_1,t_2)$ of the times T_1 and T_2 as follows:

$$S(t_1,t_2) = \int P(T_1 > t_1, T_2 > t_2|u)g(u)du$$
$$= \int \exp^{-\{\Lambda_{01}(t_1)+t_1u+\Lambda_{02}(t_1)+t_1u\}} \exp^{-\{\Lambda_{12}(t_2-t_1)+(t_2-t_1)u\}} g(u)du$$
$$= \exp^{-\{\Lambda_{01}(t_1)+\Lambda_{02}(t_1)+\Lambda_{12}(t_2-t_1)\}} \{1+\theta(t_1+t_2)\}^{-1/\theta},$$

$$f(t_1,t_2) = \frac{\partial^{\,2} S(t_1,t_2)}{\partial t_1 \partial t_2}$$
$$= S(t_1,t_2)(\lambda_{01}(t_1)+(1+\theta(t_1+t_2)^{-1})$$
$$\times \left[\lambda_{12}(t_2)+(1+\theta(t_1+t_2))^{-1}\{\lambda_{01}(t_1)+(1+\theta(t_1+t_2))^{-1}\}^{-1}\right.$$
$$\left.\times\{\lambda_{01}(t_1)+(1+\theta)(1+\theta(t_1+t_2)^{-1}\}\right],$$

$$\lambda_{12}(t_2 \mid t_1, z) = \lambda_{12}(t_2)+\beta_3'z+\{1+\theta(t_1+t_2)\}^{-1}\left[\lambda_{01}(t_1)+\beta_1'z+\{1+\theta(t_1+t_2)\}^{-1}\right]^{-1}$$
$$\times\left[\lambda_{01}(t_1)+\beta_1'z+(1+\theta)\{1+\theta(t_1+t_2)\}^{-1}\right].$$

Appendix B

The marginal likelihood function in Section 3.2 is the product of all contributions from the possible situations (combination of δ_{i1} and δ_{i2}) for the i^{th} subject, such as

For $\delta_{i1}=0$ and $\delta_{i2}=0$,

$L_i = S(Y_{i1}, Y_{i2}) = \exp^{-\{\Lambda_{01}(Y_{i1})+\Lambda_{02}(Y_{i1})\}}\{1+\theta(2\,Y_{i1})\}^{-1/\theta}$, $Y_{i1} = Y_{i2}$,

For $\delta_{i1}=0$ and $\delta_{i2}=1$,

$L_i = S(Y_{i1}, Y_{i2})\lambda_2(Y_{i2}) = S(Y_{i2}, Y_{i2})\{\lambda_{02}(Y_{i2})+\{1+\theta(2\,Y_{i2})\}^{-1}\}$, $Y_{i1} = Y_{i2}$,

For $\delta_{i1}=1$ and $\delta_{i2}=0$,

$$L_i = S(Y_{i1}, Y_{i2})\lambda_1(Y_{i1}) = S(Y_{i1}, Y_{i2})\left\{\lambda_{01}(Y_{i1}) + (1+\theta(Y_{i1}+Y_{i2})^{-1}\right\},$$

and

For $\delta_{i1}=1$ and $\delta_{i2}=1$,

$$\begin{aligned} L_i = f(Y_{i1}, Y_{i2}) &= S(Y_{i1}, Y_{i2})\left\{\lambda_{01}(Y_{i1}) + (1+\theta(Y_{i1}+Y_{i2})^{-1}\right\} \\ &\times \left[\lambda_{12}(Y_{i2}) + (1+\theta(Y_{i1}+Y_{i2}))^{-1}\left\{\lambda_{01}(Y_{i1}) + (1+\theta(Y_{i1}+Y_{i2}))^{-1}\right\}^{-1}\right. \\ &\times \left.\left\{\lambda_{01}(Y_{i1}) + (1+\theta)(1+\theta(Y_{i1}+Y_{i2})^{-1}\right\}\right]. \end{aligned}$$

Thus, the marginal likelihood is

$$\begin{aligned} &\exp^{-\{\Lambda_{01}(Y_{i1})+\Lambda_{02}(Y_{i1})+\Lambda_{12}(Y_{i2}-Y_{i1})\}}\left\{1+\theta(Y_{i1}+Y_{i2})\right\}^{-1/\theta} \\ &\times\left[\lambda_{01}(Y_{i1}) + \left\{1+\theta(Y_{i1}+Y_{i2})\right\}^{-1}\right]^{\delta_{i1}}\left[\lambda_{02}(Y_{i2}) + \left\{1+\theta(Y_{i1}+Y_{i2})\right\}^{-1}\right]^{\delta_{i2}(1-\delta_{i1})} \\ &\times[\lambda_{12}(Y_{i2}) + \left\{1+\theta(Y_{i1}+Y_{i2})\right\}^{-1}\left\{\lambda_{01}(Y_{i1}) + \left\{1+\theta(Y_{i1}+Y_{i2})\right\}^{-1}\right\}^{-1} \\ &\times\left\{\lambda_{01}(Y_{i1}) + (1+\theta)\left\{1+\theta(Y_{i1}+Y_{i2})\right\}^{-1}\right\}]^{\delta_{i1}\delta_{i2}}. \end{aligned}$$

Appendix C

Gauss-Laguerre quadrature is a numerical approximation method for the value of the following kind of the integral (Abramowitz and Stegun, 1972),

$$\int_0^{\infty} e^{-x} f(x)dx \approx \sum_{i=1}^{n} w_i f(x_i),$$

where x_i is the i^{th} root of Laguerre polynomial $L_n(x)$ (abscissas) and the weight $w_i = \frac{x_i}{(n+1)^2[L_{n+1}(x_i)]^2}$.

Appendix D

All second partial derivatives are following

$$\frac{\partial^2 l}{\partial^2 \beta_1} = \Sigma_{i=1}^{n} -\delta_{i1}(a_{i1}+u_i)^{-2} z_i^2,$$

$$\frac{\partial^2 l}{\partial^2 \beta_2} = \Sigma_{i=1}^{n} -\delta_{i2}(1-\delta_{i1})(a_{i2}+u_i)^{-2} z_i^2,$$

$$\frac{\partial^2 l}{\partial^2 \beta_3} = \Sigma_{i=1}^{n} -\delta_{i2}\delta_{i1}(a_{i12}+u_i)^{-2} z_i^2,$$

$$\frac{\partial^2 l}{\partial^2 \lambda_{01k}} = \sum_{i=1}^{n} -\delta_{i1}(a_{i1}+u_i)^{-2} d_{1ik}, \quad k=1,2,...,K,$$

$$\frac{\partial^2 l}{\partial^2 \lambda_{02k}} = \sum_{i=1}^{n} -(1-\delta_{i1})\delta_{i2}(a_{i2}+u_i)^{-2} d_{1ik}, \quad k=1,2,..,K,$$

$$\frac{\partial^2 l}{\partial^2 \lambda_{12k}} = \sum_{i=1}^{n} -\delta_{i1}\delta_{i2}(a_{i12}+u_i)^{-2} d_{2ik}, \quad k=1,2,...,K,$$

$$\frac{\partial^2 l}{\partial \beta_1 \partial \lambda_{01k}} = \Sigma_{i=1}^{n} -\delta_{i1}(a_{i1}+u_i)^{-2} d_{1ik} z_i, \; k=1,2,...,K,$$

$$\frac{\partial^2 l}{\partial\beta_2\partial\lambda_{02k}} = \sum_{i=1}^{n} -\delta_{i2}(1-\delta_{i1})(a_{i2}+u_i)^{-2} d_{1ik} z_i,\ k=1,2,...,K,$$

$$\frac{\partial^2 l}{\partial\beta_3\partial\lambda_{12k}} = \sum_{i=1}^{n} -\delta_{i2}\delta_{i1}(a_{i12}+u_i)^{-2} d_{2ik} z_i,\ k=1,2,...,K,$$

and

$$\frac{\partial^2 l}{\partial^2\theta} = -2n\theta^{-3}\psi(\theta^{-1}) - n\theta^{-4}\psi'(\theta^{-1}) + n\theta^{-3}(3-2\log(\theta)) + 2\theta^{-3}\sum_{i=1}^{n}\{\log(u_i)-u_i\}.$$

All other second partial derivatives are zero.

Appendix E

The density function for the inverse Gaussian is

$$f(x) = (\pi\eta)^{-1/2}\exp(2/\eta)x^{-3/2}\exp^{\{-x/\eta-1/(x\eta)\}}.$$

The inverse Gaussian random variable is generated by Rinvgauss function in R by setting parameters $\mu=1$ and $\lambda=2/n$.

The second frailty distribution is positive stable, for which the Laplace transform is given by $\exp(-x^{\alpha})$, $0<\alpha<1$. The density function is

$$f(x) = -\frac{1}{\pi x}\sum_{k=1}^{\infty}\frac{\Gamma(k\alpha+1)}{k!}(-x^{-\alpha})^k\sin(\alpha k\pi),$$

for $0<\alpha<1$. The positive stable random variable is generated by using the method of Hsu, Gorfine and Malone (2007).

The histogram and summary of simulated data (n=10,000) generated from three frailty distributions are presented in Figure 5 and Table 11. Three

distributions are very similar for τ=0.13 and a little different for τ=0.28. For τ=0.28, the positive stable variable is the most wildly distributed and inverse Gaussian is distributed close to 0.

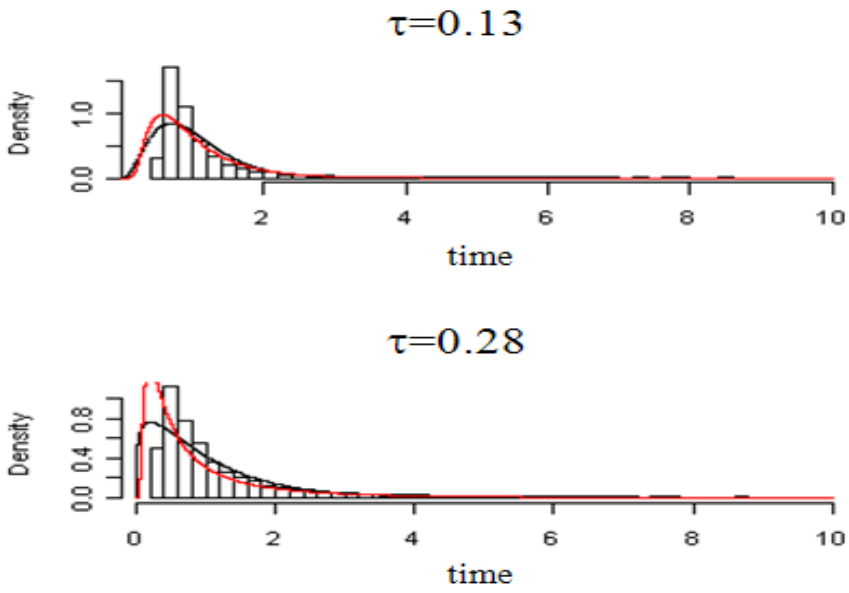

Figure 5. Histogram or density of data (n=10,000) generated from three frailty distributions (black line for Gamma, red or gray line for Inverse Gaussian and histogram for Positive stable)

Table 11. A summary of Sample data (n=10,000) generated from three frailty distributions

	Min	1st Q	Median	Mean	3rd Q	Max	Variance	IQR
				τ=0.13				
Gamma	0.014	0.599	0.903	0.999	1.297	5.817	0.299	0.697
Inverse Gaussian	0.093	0.564	0.841	1.001	1.258	7.997	0.392	0.695
Positive Stable	0.445	0.709	0.876	4.303	1.270	12490	17120	0.561
				τ=0.28				
Gamma	0.000	0.355	0.752	1.002	1.388	10.690	0.797	1.033
Inverse Gaussian	0.026	0.282	0.557	0.999	1.171	27.440	1.706	0.889
Positive Stable	0.208	0.557	0.918	109.400	1.866	398000	24553632	1.310

Bibliography

Aalen, O. O. (1989). A linear regression model for the analysis of life times. *Statistics in Medicine* 8, 907-925.

Aalen, O. O., Borgan, O. and Fekjar, H. (2001). Covariate adjustment of event histories estimated from Markov chains: the additive approach. *Biometrics* 57, 993-1001.

Abramowitz, M. and Stegun, I. A. (1972). Handbook of mathematical functions, 10th printing with corrections, Dover, Equation 25.4.45.

Chong, E. Y. C., Tai, B. C., Kalbfleisch, J. D. and Xu, J. (2011). A frailty model with conditional additive hazards for semicompeting risks. *presented in International Society for Computational Biology (ISCB).*

Dejardin, D., Lesaffre, E., and Verbeke, G. (2010). Joint modeling of progression-free survival and death in advanced cancer clinical trials, *Statistics in Medicine* 29, 1724-1734.

Duchateau, L., Janssen, P., Lindsey, P., Legrand, C., Nguti, R. and Sylvester R. (2002). The shared frailty model and the power for heterogeneity tests in multicenter trials. *Computational Statistics and Data Analysis* 40, 603-630.

Duchateau, L. and Janssen, P. (2007). The frailty model. New York, Springer.

Fine, J. P. (2001). Regression modelling of competing crude failure probabilities. *Biostatistics* 2, 85-97.

Fine, J. P., Jiang, H. and Chappell, R. (2001). On semi-competing risks data. *Biometrika* 88, 907-919.

Friedl, H. and Kauermann, G. (2000). Standard errors for EM estimates in generalized linear models with random effects. *Biometrics* 56, 761-767.

Frydman, H. and Szarek, M. (2009). Nonparametric estimation in a Markov illness-death process from interval censored observations with missing intermediate transition status. *Biometrics* 65, 143-151.

Ghosh, D. (2009). On assessing surrogacy in a single trial setting using a semicompeting risks paradigm, *Biometrics* 65, 521-529.

He, W. and Lawless, J. F. (2003). Flexible maximum likelihood methods for bivariate proportional hazards models. *Biometrics* 59, 837-848.

Hsu, L., Gorfine, M. and Malone, K. (2007). On robustness of marginal regression coefficient estimates and hazard functions in multivariate survival analysis of family data when the frailty distribution is mis-specified, *Statistics in Medicine* 26, 4657-4678.

Jiang, H., Fine, J. and Chappell R. J. (2004). Semiparmetric methods for semi-competing risks problem with censoring and truncation. *Harvard University*

Biostatistics Working Paper Series, Working Paper 15.

Klein, J. P. (2006). Modelling competing risks in cancer studies, *Statistics in Medicine* 25, 1015-1034.

Lin, D. Y. and Ying, Z. (1994). Semiparametric analysis of the additive risk model, *Biometrika* 81, 61-71.

Lin, D. Y. (1994). Cox regression analysis of multivariate failure time data: the marginal approach, *Statistics in Medicine* 13, 2233-2247.

Lin, D. Y. and Ying, Z. (1997). Additive hazards regression models for survival data. In *Proceedings of the First Seattle Symposium in Biostatistics: Survival Analysis*, Ed. D. Y. Lin and T. R. Fleming, pp. 185-98. New York: Springer.

Louis, T. A. (1982). Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society: Series B* 44, 226-233.

Martinussen, T., Scheike, T. H. and Zucker, D. M. (2011). The aalen additive gamma frailty hazards model, *Biometrika* 98, 831-843.

Mckeague, I. W. and Sasieni, P. D. (1994). A partially parametric additive risk model. *Biometrika* 81, 501-514.

Moertel, C. G., Fleming, T. R., Macdonald, J. S., Haller, D. G., Laurie, J. A., Tangen, C. M.,Ungerleider, J. S., Emerson, W. A., Tormey, D. C., Glick, J. H., Veeder, M. H. and Mailliard, J. A. (1995). Fluorouracil plus levamisole as

effective adjuvant therapy after resection of stage III colon carcinoma: a final report. *Annals of Internal Medicine* 122, 321-326.

Murphy, S. A. (1995). Asymptotic theory for the frailty model. *Annals of Statistics* 23, 182-198.

Peng, L., and Fine, J. P. (2007). Regression modeling of semicompeting risks data *Biometrics*, 63, 96-108.

Pipper, C. B. and Martinussen, T. (2004). An estimating equation for parametricshared frailty models with marginal additive hazards. *Journal of the Royal Statistical Society: Series B* 66, 207-220.

Rosenkranz G. K. (2010). Another view on the analysis of cardiovascular morbidity/mortality trials *Pharmaceutical Statistics*, 10, 196-202.

Shen, Y. and Thall, P. F. (1998). Parametric likelihoods for multiple non-fatal competing risks and death. *Statistics in Medicine* 17, 999-1015.

Shu, Y. and Klein, J. P. (2005). Additive hazards Markov regression models illustrated with bone marrow transplant data. *Biometrika* 92, 283-301.

Silvar, G. L., Amaral-Turkman M. A. (2004). Bayesian analysis of an additive survival model with frailty, Communications in Statististics, 33, 2517-2533.

Vaupel J. W., Manton K. G. and Stallard E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality, *Demography* 16, 439-454.

Xu, J., Kalbfleisch, J. D. and Tai B. (2010). Statistical analysis of illness-death processes and semicompeting risks data. *Biometrics* 66, 716-725.

Yin G. and Cai J. (2004). Additive hazards model with multivariate failure time data. *Biometrika* 91, 801-18.

Zhang, X., Akcin, H. and Lim, H. J. (2011). Regression analysis of competing risks data via semi-parametric additive hazard model, *Statistical Methods and Applications* 20, 357-381.

국 문 요 약

준-경쟁구조 자료에서 가법적인 프레일티와 가법적인 위험모형에 대한 연구

본 논문은 의학연구에서 질병의 재발과 사망의 준-경쟁구조에 있는 생존 자료에 대한 회귀분석으로써 가법적인 프레일티를 고려한 가법적 위험함수 모형을 제안하였다.

예제로 적용한 대장암환자 임상연구에서는 대장암의 재발과 사망은 서로 경쟁관계에 있으면서 사망을 하면 대장암 재발시간을 관찰할 수 없지만 반대로 대장암이 재발하여도 이후 사망시간은 관찰할 수 있는 관계에 있다. 이와 같은 자료를 준-경쟁구조 자료라고 말한다. 제안모형에서 도입한 가법적인 위험함수는 콕스의 비례위험함수와 비교하여 실제 현상을 반영하는 표현이라고 Klein(2006)은 지적한 바 있다. 또한 준-경쟁구조는 다중상태모형의 한 모형에 해당하는 질병-사망 모형이라고도 볼 수 있는데 이 질병-사망 모형에서 가법적인 위험함수를 가정하는 것이 모형의 내재적 일치성(internal consistency)을 유지할 수 있다. 제안모형에서와 같이 가법적인 프레일티에 대한 연구로서 Lin and Ying (1997)은 짝을 이룬 자료에 대한 생존 자료 분석에서 가법적인 프레일티를 도입하였고 샌드위치 공식에 의해 표준오차만을 조정하였고 Silva and Turkman(2004)은 베이지안 측면에서 다기관 생존 자료에 대한 생존 자료 분석에 적용한 바 있다.

제안한 모형은 준-경쟁관계에 있는 각 사건으로의 위험함수에 대해 공통의 공변량 변수를 다루었다. 모형의 기저위험함수는 추정을 용이하게 하기 위해 구간 고정상수 기저함수(piecewise constant baseline hazard)로 가정하였다. 기저위험함수는 구간별 고정상수 값은 다르지만 구간의 범위와 수는 모든 위험함수에 대해서 동일하게 적용하였다. 이것은 제안한 모형의 제한점이다. 하지만, 기저위험함수의 구간에 대한

모의실험 결과에 따르면 실제 구간에 근사한 범위를 결정한 상황에서는 모형의 회귀계수의 추정에는 로버스트 하였다. 향후에는 비모수적인 기저위험함수를 고려한 회귀모형으로의 확장을 고려할 수 있겠다.

모수추정은 frailty의 조건우도함수와 frailty의 사전분포를 이용하여 완전우도함수를 구성하고 frailty에 대한 결측값을 해결하기 위해 EM-알고리즘을 이용하였다. E-과정에서는 Gauss-Laguerre Quadrature 방법으로 프레일티의 함수의 기댓값을 근사적으로 계산하였다. 이 논문에서는 시뮬레이션을 통해 제안모형에 대한 유효성을 살펴보았고 대장암 임상연구의 한 사례에 적용하였다.

핵심 되는 말 : 준-경쟁구조 자료, 가법적인 프레일티, 가법적인 위험함수, 질병-사망 모형, 내재적 일치성, 감마분포, EM 알고리즘, Gauss-Laguerre Quadrature, 구간 고정상수 기저함수