PROPORTIONAL SUBDISTRIBUTION HAZARDS REGRESSION WITH INTERVAL-CENSORED COMPETING RISKS DATA

by

Yi Ren

M.S., University of North Carolina at Greensboro, 2007

B.S., Beijing Information Science & Technology University, China, 2003

Submitted to the Graduate Faculty of

the Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2015

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This dissertation was presented

by

Yi Ren

It was defended on

October 31, 2014

and approved by

Chung-Chou H. Chang, Ph.D.
Professor
Departments of Medicine and Biostatistics
School of Medicine and Graduate School of Public Health
University of Pittsburgh

Ruosha Li, Ph.D.
Assistant Professor
Department of Biostatistics
Graduate School of Public Health
University of Pittsburgh

Abdus S. Wahed, Ph.D.
Associate Professor
Department of Biostatistics
Graduate School of Public Health
University of Pittsburgh

Jonathan G. Yabes, Ph.D.
Assistant Professor
Department of Medicine

School of Medicine
University of Pittsburgh

Dissertation Advisors: Chung-Chou H. Chang, Ph.D.
Professor
Departments of Medicine and Biostatistics
School of Medicine and Graduate School of Public Health
University of Pittsburgh,

Ruosha Li, Ph.D.
Assistant Professor
Department of Biostatistics
Graduate School of Public Health
University of Pittsburgh

PROPORTIONAL SUBDISTRIBUTION HAZARDS REGRESSION WITH INTERVAL-CENSORED COMPETING RISKS DATA

Yi Ren, PhD

University of Pittsburgh, 2015

ABSTRACT

In survival analysis, the failure time of an event is interval-censored when the event is only known to occur between two consecutive observation times. Most existing methods for interval-censored data only account for a single cause of failure. However, in many situations a subject may fail due to more than one type of event. Such data scenarios are called competing risks data. Competing events may preclude the occurrence of the event of interest. In the analysis of competing risks, the conventional methods should be used with caution and may lead to nonsensical interpretation. With covariates, the proportional subdistribution hazards model is widely used to model the cumulative incidence function (also known as the subdistribution) of a particular event. This semiparametric regression model has a straightforward interpretation for estimators as it is akin to the Cox proportional hazards model. For interval-censored competing risks data, however, estimation procedures based on the proportional subdistribution hazards model has not been investigated. In this dissertation, we propose estimation and inference procedures that account for both interval censoring and competing risks by adopting the modeling framework of the proportional subdistribution hazards model. The objective is to examine the effects of covariates on the subdistribution of event of interest. The proposed estimating equations effectively utilize the ordering of event time pairs. The technique of inverse probability weighting is used to account for the missing mechanism. Simulation studies show that the proposed methods perform well under realistic scenarios. A lymphoma data set is used to illustrate the performance of the proposed method in comparison to the proportional subdistribution hazards model using the data imputed by midpoint of the observed time interval.

Public health significance: Interval-censored competing risks data are often encountered in biomedical research. The method we proposed serves a useful tool for exploring the covariate effects on the event of interest under this challenging censoring mechanism. The information on the effects of covariates has implications for proper clinical management of the different cohorts of patients. It quantifies the relationship between public health strategies and measurement of health status, and determines the efficacy information for possible improvement of interventions.

Keywords: Interval censoring; competing risks; proportional subdistribution hazards; inverse probability weighting.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

1.1 MOTIVATION AND CHALLENGES

In survival analysis, failure time is not always exactly observed or right-censored. Interval censoring naturally occurs in the research areas such as medical, financial and sociological studies. It is more common in clinical trials and longitudinal studies, especially in dentistry and HIV/AIDS studies. According to the scheduled times, a patient may have several clinical visits to observe possible changes in disease progression. Under the circumstance, the exact time of occurrence of the change may not be observed, instead it is only known to lie between the visit where the disease progression has already been detected and the previous visit where the disease progression has not. In these situations, interval censoring may arise in several cases. For example, a patient may miss one or more clinical visits and then return with a diagnosis of disease progression. Or it could be, patients make visits at times that are convenient to themselves rather than at predetermined observation times. In both situations, the data are subject to interval censoring. Interval-censored data is different from missing data in the nature that it provides incomplete data regarding failure times.

With complete or right-censored data, a variety of methods and statistical software packages are available for estimating the distribution function of failure time and covariate effects. However, the standard methods for right-censored data are not recommended to directly apply on the interval-censored data due to nature of censorships. With general right censoring, the observed event time is the time to first failure, which is the minimum of failure time and censoring time. For interval-censored data, however, we cannot observe the exact time when the failure occurs. Therefore, treatment to interval censoring requires additional cares.

In practice, subjects may be at risk of failure due to more than one cause of failure, which is called competing risks. In the presence of competing risks, there are two basic quantities that can be estimated without any assumptions on the dependence structure among latent failure times for competing risks. Cause-specific hazards function (CSHF) defines the instantaneous failure rate of a type of event in the presence of other competing events. Compared to the CSHF, it is often of interest to estimate the cumulative incidence function (CIF), which is also known as the subdistribution. CIF is straightforward to interpret and it represents the probability of experiencing a particular type of event by a given time in the presence of competing risks.

When subjects are followed up periodically and may fail from more than one cause, interval-censored competing risks data arises. For example, in HIV vaccine clinical trials, time of HIV infection is subject to interval censoring and only known to lie between the visit where the patient tested HIV positive and the previous visit tested negative. Like many viruses, HIV exhibits significant genotypic variation so it can be distinguished into several subtypes. Given that simultaneous infections with several subtypes are rare, the subtypes of the infecting virus are often analysed as competing risks. The challenges for such data are two-fold. First, making inference from incomplete information that caused by interval censoring and reducing the selection bias. Second, as a step forward, understanding a more complicated setting where both interval censoring and competing risks are present.

1.2 OVERVIEW OF PREVIOUS WORK

In practice, the methods developed for interval-censored data often imitate the statistical methods of right-censored data. Imputation approaches are commonly used due to simplicity. It imputes unobserved failure times using the observed data, then apply standard methods for right-censored data to make inference. For example, the idea of simple imputation is to replace the failure times by assuming the failures occurred at the endpoints (or midpoint) of time intervals. However, such method assumes that failures occur at artificially fixed time points. It may lead to invalid estimates and tend to underestimate the standard errors [Hsu et al., 2007].

Nonparametric methods have been well established in the literature on the estimation of distribution function or survival curve for interval-censored data. Peto [Peto, 1973] first proposed a method to compute the nonparametric maximum likelihood estimator (NPMLE) of survival function as an analog to the estimates from Kaplan-Meier life-table technique. The total likelihood only depends on the decreased sizes of survival curve at the endpoints of each time interval. A constraint Newton-Raphson method was then used to maximize the strictly convex log likelihood function. Turnbull [Turnbull, 1976] extended Petos method to account for truncated and interval-censored data and proposed a self-consistency algorithm to estimate NPMLE of failure time distribution. Maathuis [Maathuis, 2005]later generalized the NPMLE of the distribution function to bivariate interval-censored data on a finite dimension. The reduction algorithm was used to maximize the sum of the probability masses of observation rectangle that contains the pair of time variables.

Many literatures have been explored under the framework of Cox proportional hazards model [Cox, 1972]. Under the proportional hazards assumption, the conditional hazard function is assumed to be proportional to an unspecified nonnegative baseline hazard function, where the multiplicative constant involves a vector of covariates and the unknown regression coefficients. The model gives a natural estimate for the failure associated with a vector of covariate. The partial likelihood function only depends on covariate effects, so there is no need to deal with baseline hazard function. The method can yield efficient and asymptotic estimation [Cox, 1975]. To extend the Cox model to interval-censored data, the baseline hazard function can be specified parametrically in order to apply the standard likelihood estimation. However, model mis-specification may induce bias. In 1986, Finkelstein [Finkelstein, 1986] proposed to use discrete baseline survival with finite mass points to estimate the maximum likelihood estimates (MLEs) of survival function and regression coefficients. Finkelstein revised the Newton-Raphson algorithm by using the inverse Hessian matrix at each iteration which is computationally intensive. Based on the work of Finkelstein, Pan [Pan, 1999] maximized joint likelihood of Cox regression coefficient and baseline survival using iterative convex minorant (ICM) algorithm, which only uses diagonal elements rather than full Hessian matrix. Thus, the numerical method improved the computation. Pan

[Pan, 2000] also proposed a multiple imputation method to impute exact failure times for interval-censored data. After imputing by data augmentation, it used standard Cox partial likelihood to estimate the regression coefficients and the baseline survival. A profile likelihood approach [Huang, 1996] was later defined to estimate the MLE of regression coefficients and cumulative baseline hazard for current status data, which is the simplest form of interval-censored data where there is exactly one observation time for each subject. It maximized the partial likelihood with respect to the baseline hazard distribution at each of certain fixed values of the regression coefficient. Smoothing methods can also be used to smoothly estimate baseline hazard as a balance between parametric functions and completely unspecified functions. Regression splines [Kooperberg and Clarkson, 2000] and local likelihood smoothing method [Cai and Betensky, 2003] are suggested. Cai and Betensky have proposed a flexible locally parametric method to model the baseline hazard function and to obtain maximum likelihood estimates under Cox model for interval-censored data. Specifically, they parameterized the log hazard function with a piecewise-linear spline and estimated the hazard function from penalized likelihood. In addition to nonparametric and parametric methods above, semiparametric approaches are also explored by treating hazard function as a nuisance parameter. Zhang [Zhang et al., 2005] used estimating equation under linear transformation model to estimate covariate effects on interval-censored failure times. But only categorical covariates were considered. Recently, Heller [Heller, 2011] proposed a semiparametric method to estimate regression coefficients under Cox model. The method uses weighted estimating equations based on rank information of event time pairs, providing unbiased estimated regression coefficients under a surrogacy condition.

Competing risks make the analysis of interval censoring more complicated. Several scholars studied nonparametric maximum likelihood estimators (NPMLEs) of failure time distribution. Hudgens et al. [Hudgens et al., 2001] extended Turnbull's method for data subject to interval censoring and truncation to a competing risks setting. It used an EM algorithm to estimate the NPMLE of CIF. They also proposed a pseudo-likelihood estimator that conditions on the NPMLE of the survival function ignoring failure type. Jewell [Jewell et al., 2003] studied the NPMLE of CIF for current status data in the presence of competing risks. They also proposed a naive estimator that only uses a subset of the observed data in the ab-

sence of competing risks. Instead of cumulative incidence, Frydman and Liu [Frydman and Liu, 2013] studied the NPMLE of the cumulative intensities in a competing risks model by parameterizing the NPMLE of CIF from Hudgens et al. [Hudgens et al., 2001]. The method assumes mixed case interval censoring model with discrete and finite scheduled times. These nonparametric methods require dividing the sample into subsamples corresponding with each specific value of the covariates and computing the NPMLE for each subsample. This may result in practical issues when the data only contains few observations in the subsample and continuous covariates.

In contrast to nonparametric methods, Hudgens et al. [Hudgens et al., 2014] demonstrated a parametric method which directly estimated CIF under improper parametric models, specifically Gompertz cure-type models, for competing risks data subject to interval censoring. The full likelihood estimator takes account into all causes of failure, while the naive likelihood estimator ignores failures from competing risks and treats them as right censoring. Full likelihood estimator performed well when the model is correctly specified. The naive likelihood estimator may have lost efficiency due to information loss. Based on the likelihood formulation, the naive estimator is only valid in mixed case interval censoring setting.

1.3 OVERVIEW OF OUR WORK

As previously mentioned, Heller has presented an estimating equation method based on the rank of failure time pairs and used the inverse probability weighting (IPW) technique to account for the missingness results from interval censoring. Unlike likelihood estimation which requires estimation at each time point, it only uses the rank information of those intervals at which failures occur. Compared to other methods, it is simple and straightforward. The simulations in his work also showed that the methodology works well in practical situation. The statistical method was developed under the proportional hazards specification and only adapted to a general interval censoring condition in which patients may fail from a single cause of failure. However, in certain research patients may experience multiple potential causes just as the HIV example described.

When analyzing competing risks data, investigators may be interested in the effects of covariates on the cumulative incidence of a particular failure type. Given the regression covariates under proportional hazards assumption, the CIF can be estimated by combining the estimates of CSHF for all types of events derived from partial likelihood [Prentice et al., 1978]. The likelihood function can be easily expressed in terms of CSHF for right-censored data. However, interval censoring cannot simplify the likelihood estimation. Also, a discrepancy on the effect of a covariate may exist between the CSHF and the corresponding CIF [Gray, 1988, Pepe, 1991]. Compared to cause-specific hazards function which gives instantaneous rate of a failure type, the CIF would have more intuitive and useful interpretation in terms of cumulative probability. Therefore, we prefer to model CIF rather than CSHF. The proportional hazards subdistribution model [Fine and Gray, 1999] is a semiparametric model for modeling the cumulative incidence function and is akin to the Cox proportional hazards model. Under proportional hazards assumption, it provides a well-justified empirical representation on the CIF of the event of interest for complete and/or right-censored data. The estimators of covariate effects and baseline hazard have a straightforward interpretation. Nevertheless, this model is not directly applicable to estimate the covariate effects on CIF in the presence of interval censoring.

In this dissertation, we propose a semiparametric method by extending Heller's estimating equation method under the framework of proportional subdistribution hazards regression. The proposed method well balances the model flexibility and efficiency for interval-censored competing risks data. Our goal is to estimate the regression coefficients of proportional subdistribution hazards model for interval-censored competing risks data. In Section 2 we present our inverse probability estimating equation method. In Section 3 we carry out two simulation studies to assess and compare the performance of the proposed estimators to Fine and Gray's method based on midpoint imputation. The first study involves two settings for exponential and Weibull subdistributions. The second study is under a Gompertz parametrization. In Section 4 we apply our method on a follicular cell lymphoma data and compare the performance to previous Fine and Gray's method. Finally in Section 5 we conclude with a discussion regarding the issues presented in this article.

2.0 MODEL AND ESTIMATION

2.1 INTRODUCTION

In longitudinal clinical studies, if at the end of the study period the event has not been observed, the time to the event is subject to right censoring. However, in certain areas of medical research such as dentistry and HIV/AIDS cohort studies, the occurrence of the event of interest can often be recorded only at the predetermined times, which give rise to interval-censored data. A typical example is breast cancer study (Finkelstein and Wolfe, 1985). In the study, early patients were randomly assigned to two treatments, radiation therapy alone and radiation therapy plus chemotherapy, to compare the treatment effects with respect to cosmetic deterioration. Patients were evaluated the cosmetic appearance every 4 to 6 months, so that the time to cosmetic deterioration is interval-censored.

The conventional approaches for handling interval-censored data under proportional hazards model generally include several types. First approach uses standard parametric maximum likelihood to derive the regression parameters and assumes the baseline hazard function to be any nondecreasing function. The misspecification can cause bias. Imputation methods [Pan, 2000] can be used, however, multiple imputation is computationally intensive. Another method uses smooth estimates for baseline hazard [Cai and Betensky, 2003] as a balance of the goodness of fit and smoothness. It allows for a straightforward maximum likelihood estimation of the regression parameters. Unfortunately, none of these approaches are available to directly estimate the covariate effects on the cumulative incidence of a particular event which is also know as subdistribution [Fine and Gray, 1999]. In contrast, semiparametric methods have a good trade-off between the model flexibility and efficiency.

There are two challenges for interval-censored competing risks data. First, interval censoring results in information loss that event time is only known to lie between two endpoints. The second challenge is that competing risks may preclude from observing main event. This makes the estimation of the marginal probability of ever having a main event become a difficult problem. The standard analysis for competing risks data involves modeling cause-specific hazards functions (CSHF) or cumulative incidence function (CIF). Compared to CSHF, CIF has a more natural interpretation in terms of cumulative probability of having a main event by certain time.

Previous work on the cumulative incidence function has focused on directly specifying parametric model and deriving the likelihood estimators. In this dissertation, an alternative method is proposed to estimate the regression coefficients on the cumulative incidence of the event of interest. It is organized as follows. In Section 2.2, we review the proportional subdistribution hazards model and present our notation and assumptions. In Section 2.3 we develop the methodology that allows estimation of regression coefficients under interval censoring.

2.2 MODEL AND ASSUMPTIONS

Let $h(\cdot)$ be an unknown strictly increasing function. The linear transformation model [Cheng et al., 1995] is

$$h(T) = \boldsymbol{X}^T \boldsymbol{\beta_1} + \epsilon \tag{2.1}$$

where T is the time to event, $\mathbf{X}$ is a vector of covariates, $\boldsymbol{\beta_1}$ is a vector of regression coefficients, and random error ϵ has a completely known continuous distribution function F. The linear transformation model generalizes a class of commonly used regression models as special cases. In particular, if $F(s) = 1 - \exp\{-\exp(s)\}$, an extreme value distribution, then (2.1) is the Cox proportional hazards model. To estimate this model under competing risks setting without simultaneously estimating cumulative incidence functions for all competing risks, the proportional hazards model for subdistribution is utilized.

Let T denote potential failure time, let $\kappa \in (1, ..., K)$ be cause of failure, where $\kappa = 1$ is of our interest. We assume that the causes are always observable. Let $\boldsymbol{X}$ be a vector of time-independent covariates. The proportional subdistribution hazards model for cause of interest [Gray, 1988, Fine and Gray, 1999] is in the form of

$$\begin{aligned}\lambda_1(t; \boldsymbol{X}) &= \lim_{\Delta t \to 0} \frac{\Pr\{t \leq T \leq t + \Delta t, \kappa = 1 | T \geq t \cup (T \leq t \cap \kappa \neq 1)\}}{\Delta t} \\ &= \frac{dF_1(t; \boldsymbol{X})/dt}{1 - F_1(t; \boldsymbol{X})} \\ &= \lambda_{10}(t) \exp(\boldsymbol{X}^T \boldsymbol{\beta_1}),\end{aligned} \tag{2.2}$$

where $F_1(t; \boldsymbol{X})$ is the cumulative incidence function for the cause of interest, and $\lambda_{10}(t)$ is the baseline subdistribution hazards for the cause of interest. It is an unspecified nonnegative function. The subdistribution hazards for event of interest implies such an improper failure time random variable $T^* = I(\kappa = 1) \times T + [1 - I(\kappa = 1)] \times \infty$. When $t < \infty$, the subdistribution for failure time T^* is $F_1(t; \boldsymbol{X})$. When $t = \infty$, it has a point mass equal to $1 - F_1(\infty; \boldsymbol{X})$. Under model (2.2), the subdistribution for event of interest conditional on the covariates is then

$$\begin{aligned}F_1(t; \boldsymbol{X}) &= \Pr(T \leq t, \kappa = 1 | \boldsymbol{X}) \\ &= 1 - \exp\left\{ - \int_0^t \lambda_{10}(s) \exp(\boldsymbol{X}^T \boldsymbol{\beta_1}) ds \right\}.\end{aligned} \tag{2.3}$$

Distinguished from the standard Cox model, the estimation of the regression coefficients and baseline hazard is based on a modified risk set, where subjects are retained at risk after they experienced a competing risk.

With random right censoring, Fine and Gray utilized the inverse probability censoring weighting (IPCW) techniques [Robins and Rotnitzky, 1992] on those subjects who have failed from competing risks so that they do not fully participate in the partial likelihood, and the weights are decreasing over time. The regression coefficients $\boldsymbol{\beta_1}$ are then estimated by maximizing the partial likelihood. Unfortunately, incorporation of interval censoring does not enable canceling the baseline hazard function in the partial likelihood function under

the proportional subdistribution hazards model. So their method is not directly applicable to interval-censored competing risks data.

Let C be the right censoring time. For the usual right-censored data, we can observe $\min(T, C)$. For interval-censored data, time to failure T may not be exactly observed but only known to lie in an interval between two observation time points $(L, R]$. If the event time is exactly observed, e.g. death, then $L = R$. If the event time is right-censored, then $R = \infty$ and thus the event time interval is (L, ∞). Here L refers to the last observation time point prior to the censoring time C.

Again, κ represents cause of failure. For simplicity, we only describe two competing risks in this paper, i.e. $\kappa \in (1, 2)$, even though the idea is easily extended to accommodate the situation where there are more than two possible causes. Let type of event $\varepsilon = \delta\kappa$ where $\delta = I(T \leq C)$, so that $\varepsilon \in (0, 1, 2)$ where 0 denotes right censoring, value of 1 denotes main event, and 2 denotes competing event. Assume that the scheduling process is independent of event time and cause of failure conditional on covariates. The observable data $\{(L_i, R_i], \varepsilon_i, \boldsymbol{X}_i\}$ are independent and identically distributed for subject $i = 1, ..., n$.

2.3 INVERSE PROBABILITY WEIGHTED ESTIMATING EQUATIONS

Given the fact that the potential failure time is interval-censored, in order to estimate the effects of covariates we focus on the ordering of event time pairs. Without accounting for censoring, the dichotomized ordering indicator of main event times for subjects i and j is defined as $S_{ij} = I(T_i > T_j, \kappa_i = \kappa_j = 1)$ for $i \neq j$. This is saying, suppose that the observation time is long enough to capture all possible events for all subjects, whether the main event for subject j occurs earlier than the main event for subject i or not. This quantity is invariant under model (2.1) and can provide efficient inference for regression coefficients under the framework of proportional subdistribution hazards model.

Nevertheless, the statistical quantity S_{ij} is not always observed under interval-censored competing risks setting. According to model (2.1), under proportional subdistribution hazards model $\lambda_1(t) = \lambda_{10}(t) \exp(\boldsymbol{X}^T \boldsymbol{\beta}_1)$, where baseline hazard for the subdistribution $\lambda_{10}(t)$

is an unspecified nondecreasing function of t, we have shown that

$$
\begin{aligned}
E(S_{ij}|\boldsymbol{X}_i,\boldsymbol{X}_j) &= E\{I(T_i > T_j, \kappa_i = \kappa_j = 1)|\boldsymbol{X}_i,\boldsymbol{X}_j\} \\
&= \Pr(T_i > T_j, \kappa_i = \kappa_j = 1|\boldsymbol{X}_i,\boldsymbol{X}_j) \\
&= \int_{-\infty}^{\infty} \{1 - F_1(t;X_i)\} f_1(t;X_j) dt \\
&= \int_{-\infty}^{\infty} \exp\{-\Lambda_{10}(t)\exp(\boldsymbol{X}_i\boldsymbol{\beta})\}\exp\{-\Lambda_{10}(t)\exp(\boldsymbol{X}_j\boldsymbol{\beta}_1)\}\lambda_{10}(t)\exp(\boldsymbol{X}_j\boldsymbol{\beta}_1)dt \\
&= \int_{-\infty}^{\infty} \exp\Big[-\Lambda_{10}(t)\{\exp(\boldsymbol{X}_i\boldsymbol{\beta}_1)+\exp(\boldsymbol{X}_j\boldsymbol{\beta}_1)\}\Big]\lambda_{10}(t)\exp(\boldsymbol{X}_j\boldsymbol{\beta}_1)dt \\
&= \frac{\exp(\boldsymbol{X}_j^T\boldsymbol{\beta}_1)}{\exp(\boldsymbol{X}_i^T\boldsymbol{\beta}_1)+\exp(\boldsymbol{X}_j^T\boldsymbol{\beta}_1)} \\
&= \frac{1}{1+\exp(\boldsymbol{X}_{ij}^T\boldsymbol{\beta}_1)},
\end{aligned}
\tag{2.4}
$$

where

$$
\begin{aligned}
f_1(t;X_j) &= \lambda_1(t;X_j)\{1 - F_1(t;X_j)\}, \\
\Lambda_{10}(t) &= \int_0^t \lambda_{10}(s)ds, \\
\boldsymbol{X}_{ij} &= \boldsymbol{X}_i - \boldsymbol{X}_j, i, j = 1, ..., n.
\end{aligned}
$$

To account for finite follow-up period, we define the bounded ordering of event time pairs as $S_{ij}^{\tau} = I[T_i > T_j, T_j < \tau, \varepsilon_i = \varepsilon_j = 1]$, where τ is the maximum follow-up time of study. The expectation of S_{ij}^{τ} given $\boldsymbol{X}_i$ and $\boldsymbol{X}_j$ under models (2.2, 2.3) is

$$
\begin{aligned}
E(S_{ij}^{\tau}|\boldsymbol{X}_i,\boldsymbol{X}_j) =& E\{I(T_i > T_j, T_j < \tau, \varepsilon_i = \varepsilon_j = 1)|\boldsymbol{X}_i,\boldsymbol{X}_j\} \\
=& \Pr(T_i > T_j, T_j < \tau, \varepsilon_i = \varepsilon_j = 1|\boldsymbol{X}_i,\boldsymbol{X}_j) \\
=& \int_0^{\tau} \Big\{1 - F_1(t;X_i)\Big\} f_1(t;X_j)dt \\
=& \int_0^{\tau} \exp\Big\{-\Lambda_{10}(t)\exp(\boldsymbol{X}_i\boldsymbol{\beta}_1)\Big\}\exp\Big\{-\Lambda_{10}(t)\exp(\boldsymbol{X}_j\boldsymbol{\beta}_1)\Big\}\lambda_{10}(t)\exp(\boldsymbol{X}_j\boldsymbol{\beta}_1)dt \\
=& \frac{1}{1+\exp(\boldsymbol{X}_{ij}\boldsymbol{\beta}_1)}\int_0^{\tau} \exp\Big[-\Lambda_{10}(t)\Big\{\exp(\boldsymbol{X}_i\boldsymbol{\beta}_1)+\exp(\boldsymbol{X}_j\boldsymbol{\beta}_1)\Big\}\Big] \\
& \times \lambda_{10}(t)\{\exp(\boldsymbol{X}_i\boldsymbol{\beta}_1)+\exp(\boldsymbol{X}_j\boldsymbol{\beta}_1)\}dt \\
=& \frac{1}{1+\exp(\boldsymbol{X}_{ij}\boldsymbol{\beta}_1)}\left(1 - \exp\Big[-\int_0^{\tau}\lambda_{10}(t)\Big\{\exp(\boldsymbol{X}_i\boldsymbol{\beta}_1)+\exp(\boldsymbol{X}_j\boldsymbol{\beta}_1)\Big\}dt\Big]\right) \\
=& \frac{1}{1+\exp(\boldsymbol{X}_{ij}\boldsymbol{\beta}_1)}\left(1 - \exp\Big[-\beta_2\Big\{\exp(\boldsymbol{X}_i\boldsymbol{\beta}_1)+\exp(\boldsymbol{X}_j\boldsymbol{\beta}_1)\Big\}\Big]\right).
\end{aligned}
$$

Let $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^T, \beta_2)^T$ where $\boldsymbol{\beta}_1$ is the vector of covariate coefficients on the subdistribution (2.3) for event of interest under model (2.2), and $\beta_2 = \int_0^\tau \lambda_{10}(t)dt$ indicates the cumulative baseline subdistribution hazards up to time τ. To estimate $\boldsymbol{\beta}$ under the proportional hazards specification with usual right censoring, we can apply Fine and Gray's IPCW method and solve the estimating equation from score function of partial likelihood. With interval censoring, we can construct an estimating equation for statistical quantity S_{ij}^τ to solve for $\boldsymbol{\beta}$. However, it may not be observed under some cases where event time intervals overlap. We adapt inverse probability weighting techniques [Horvitz and Thompson, 1952] to construct an unbiased estimating equation. The inverse probability weight is used to account for missingness introduced by unknown ordering. Let Δ_{ij} be the complete-case indicator and it is defined as

$$\begin{aligned}\Delta_{ij} =& I[L_i \geq R_j, \varepsilon_i \neq 2, \varepsilon_j = 1] + I[\varepsilon_i = 2, \varepsilon_j = 1] \\ &+ I[L_j \geq R_i, \varepsilon_i = 1, \varepsilon_j \neq 2] + I[\varepsilon_i = 1, \varepsilon_j = 2] + I[\varepsilon_i = \varepsilon_j = 2].\end{aligned}$$

As a sufficient condition of observing bounded ordering, the complete-case indicator Δ_{ij} is equal to 1 if S_{ij}^τ is observed, and 0 otherwise. Given the observed S_{ij}^τ and covariates, we define the complete-case probability as

$$\pi^*(S_{ij}^\tau, \boldsymbol{X}_i, \boldsymbol{X}_j) = \Pr(\Delta_{ij} = 1 | S_{ij}^\tau, \boldsymbol{X}_i, \boldsymbol{X}_j).$$

By incorporating the inverse probability weight, data pairs which are less likely to capture the ordering of main event times will have inflated weight. Based on the selected event time pairs, IPW technique can reduce the selection bias that results from interval censoring and competing risks. The unbiased inverse probability weighted estimating equation is

$$\sum_i \sum_{j \neq i} \frac{\Delta_{ij}}{\pi^*(S_{ij}^\tau, \boldsymbol{X}_i, \boldsymbol{X}_j)} W_{ij}(\boldsymbol{\beta}) \left\{ S_{ij}^\tau - E(S_{ij}^\tau | \boldsymbol{X}_i, \boldsymbol{X}_j) \right\} = 0, \tag{2.5}$$

where

$$
\begin{aligned}
W_{ij}(\boldsymbol{\beta}) =& \frac{\partial E(S_{ij}^{\tau}|\boldsymbol{X}_i, \boldsymbol{X}_j)}{\partial \boldsymbol{\beta}} \\
=& \left\{ \frac{\partial E(S_{ij}^{\tau}|\boldsymbol{X}_i, \boldsymbol{X}_j)}{\partial \boldsymbol{\beta}_1}, \frac{\partial E(S_{ij}^{\tau}|\boldsymbol{X}_i, \boldsymbol{X}_j)}{\partial \beta_2} \right\},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial E(S_{ij}^{\tau}|\boldsymbol{X}_i, \boldsymbol{X}_j)}{\partial \boldsymbol{\beta}_1} =& \frac{1}{\{1+\exp(\boldsymbol{X}_{ij}\boldsymbol{\beta}_1)\}^2} \Big\{ \{1+\exp(\boldsymbol{X}_{ij}\boldsymbol{\beta}_1)\} \Big[\beta_2 \{ \boldsymbol{X}_i \exp(\boldsymbol{X}_i\boldsymbol{\beta}_1) + \boldsymbol{X}_j \exp(\boldsymbol{X}_j\boldsymbol{\beta}_1) \} \Big] \\
& \times \exp\Big[-\beta_2 \{ \exp(\boldsymbol{X}_i\boldsymbol{\beta}_1) + \exp(\boldsymbol{X}_j\boldsymbol{\beta}_1) \} \Big] \\
& - \boldsymbol{X}_{ij} \exp(\boldsymbol{X}_{ij}\boldsymbol{\beta}_1) \Big(1 - \exp\Big[-\beta_2 \{ \exp(\boldsymbol{X}_i\boldsymbol{\beta}_1) + \exp(\boldsymbol{X}_j\boldsymbol{\beta}_1) \} \Big] \Big) \Big\},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial E(S_{ij}^{\tau}|\boldsymbol{X}_i, \boldsymbol{X}_j)}{\partial \beta_2} =& \frac{1}{1+\exp(\boldsymbol{X}_{ij}\boldsymbol{\beta}_1)} \exp\Big[-\beta_2 \{ \exp(\boldsymbol{X}_i\boldsymbol{\beta}_1) + \exp(\boldsymbol{X}_j\boldsymbol{\beta}_1) \} \Big] \{ \exp(\boldsymbol{X}_i\boldsymbol{\beta}_1) + \exp(\boldsymbol{X}_j\boldsymbol{\beta}_1) \} \\
=& \exp\Big[\boldsymbol{X}_j\boldsymbol{\beta}_1 - \beta_2 \{ \exp(\boldsymbol{X}_i\boldsymbol{\beta}_1) + \exp(\boldsymbol{X}_j\boldsymbol{\beta}_1) \} \Big].
\end{aligned}
$$

If π^* is known, the inference for $\boldsymbol{\beta}$ can be derived by estimating equation (2.5). If π^* is unknown and S_{ij}^{τ} is missing at random (MAR) [Little and Rubin, 2002], then $\pi^*(S_{ij}^{\tau}, \boldsymbol{X}_i, \boldsymbol{X}_j) = \pi^*(\boldsymbol{X}_i, \boldsymbol{X}_j)$. Using a working model for π^* would produce an unbiased estimating equation. However, π^* depends on both of observed and unobserved S_{ij}^{τ}.

$$
\begin{aligned}
\pi^*(S_{ij}^{\tau}, \boldsymbol{X}_i, \boldsymbol{X}_j) =& \Pr(\Delta_{ij} = 1 | S_{ij}^{\tau} = s_{ij}^{\tau}, \boldsymbol{X}_i = \boldsymbol{x}_i, \boldsymbol{X}_j = \boldsymbol{x}_j) \\
=& I(s_{ij}^{\tau} = 1) \Pr(\Delta_{ij} = 1 | S_{ij}^{\tau} = 1, \boldsymbol{X}_i = \boldsymbol{x}_i, \boldsymbol{X}_j = \boldsymbol{x}_j) \\
& + I(s_{ji}^{\tau} = 1) \Pr(\Delta_{ij} = 1 | S_{ji}^{\tau} = 1, \boldsymbol{X}_i = \boldsymbol{x}_i, \boldsymbol{X}_j = \boldsymbol{x}_j) \\
& + I(s_{ij}^{\tau} = s_{ji}^{\tau} = 0) \Pr(\Delta_{ij} = 1 | S_{ij}^{\tau} = S_{ji}^{\tau} = 0, \boldsymbol{X}_i = \boldsymbol{x}_i, \boldsymbol{X}_j = \boldsymbol{x}_j) \\
=& I(s_{ij}^{\tau} = 1) \Pr(\Delta_{ij} = 1 | S_{ij}^{\tau} = 1, \boldsymbol{X}_i = \boldsymbol{x}_i, \boldsymbol{X}_j = \boldsymbol{x}_j) \\
& + I(s_{ji}^{\tau} = 1) \Pr(\Delta_{ij} = 1 | S_{ji}^{\tau} = 1, \boldsymbol{X}_i = \boldsymbol{x}_i, \boldsymbol{X}_j = \boldsymbol{x}_j)
\end{aligned}
$$

where S_{ji}^{τ} is opposed to S_{ij}^{τ}, i.e. $S_{ji}^{\tau} = I[T_j > T_i, T_i < \tau, \varepsilon_i = \varepsilon_j = 1]$. If $S_{ij}^{\tau} + S_{ji}^{\tau} = 0$, then $\Pr(\Delta_{ij} = 0) = 1$. Thus, S_{ij}^{τ} is not missing at random (NMAR). To obtain the valid inference under NMAR, a model for data and missing mechanism is required. To avoid the estimation of missingness, Ibriham [Ibrahim et al., 2001] introduced the use of auxiliary variable to reduce bias. The auxiliary variable should always be observed and associated with the unobserved variable. In our context, the unobserved variables are S_{ij}^{τ} and S_{ji}^{τ}. We use

an observable auxiliary variable A_{ij} to represent the ordering of main event time intervals. It is denoted as

$$\begin{aligned} A_{ij} =& I[L_i \geq L_j, R_i > R_j, \varepsilon_i = 1, \varepsilon_j = 1] + I[L_i \geq L_j, \varepsilon_i = 0, \varepsilon_j = 1] \\ &+ I[\varepsilon_i = 2, \varepsilon_j \neq 0]. \end{aligned}$$

When $A_{ij} + A_{ji} = 1$, the potential ordering of intervals is observed. Under two situations, the potential ordering cannot be observed, i.e. $A_{ij} + A_{ji} = 0$. One is due to right censoring. Since it is impossible to keep track of events after right censoring, the ordering of time pairs is obscured unless one's main event occurs prior to another subject's censoring. To determine the potential ordering of main event time, we compare the rank of left endpoints of two intervals instead. The other situation to observe an zero valued auxiliary variable is that, main event time interval for a subject is completely contained in the other's interval. Given that the auxiliary variable can provide potential ordering information and correlates to the unobserved variable, it may in place of S_{ij}^{τ} to model π^*.

If $A_{ij} + A_{ji} = 0$, this replacement is trivial since $\Pr(\Delta_{ij} = 1 | S_{ij}^{\tau}, A_{ij}, A_{ji}, \boldsymbol{X}_i, \boldsymbol{X}_j) = 0$. If $A_{ij} + A_{ji} = 1$, we assume that S_{ij}^{τ} is ignorable for predicting Δ_{ij} conditional on $(A_{ij}, A_{ji}, \boldsymbol{X}_i, \boldsymbol{X}_j)$. That is, $\Pr(\Delta | S_{ij}^{\tau}, A_{ij}, A_{ji}, \boldsymbol{X}_i, \boldsymbol{X}_j) \cong \Pr(\Delta | A_{ij}, A_{ji}, \boldsymbol{X}_i, \boldsymbol{X}_j)$.

To completely transform the problem into a MAR framework, cases where right censoring obscure the ordering need to be removed. Let indicator ω_{ij} denoted as

$$\begin{aligned} \omega_{ij} =& I[\varepsilon_i \neq 0, \varepsilon_j \neq 0] + I[L_i \geq L_j, \varepsilon_i = 0, \varepsilon_j = 1] \\ &+ I[L_i \leq L_j, \varepsilon_i = 1, \varepsilon_j = 0] \end{aligned}$$

This defines $\omega_{ij} = 0$ when right censoring masks the potential ordering of the event time intervals under competing risk setting. This situation may only happen when $A_{ij} + A_{ji} = 0$. If $A_{ij} + A_{ji} = 1$, then $\Pr(\omega_{ij} = 1) = 1$. Finally, those observations with $\omega_{ij} = 0$ will be eliminated from the data. The estimating equation for the selected data pairs where there is definitive ordering of main event times is

$$\sum_i \sum_{j \neq i} \frac{\Delta_{ij}}{\pi(A_{ij}, A_{ji}, \boldsymbol{X}_i, \boldsymbol{X}_j)} W_{ij}(\boldsymbol{\beta}) \left\{ S_{ij}^{\tau} - E(S_{ij}^{\tau} | \boldsymbol{X}_i, \boldsymbol{X}_j) \right\} = 0, \tag{2.6}$$

where

$$\pi(A_{ij}, A_{ji}, \boldsymbol{X}_i, \boldsymbol{X}_j) = \Pr(\Delta_{ij} = 1 | A_{ij}, A_{ji}, \boldsymbol{X}_i, \boldsymbol{X}_j, \omega_{ij} = 1).$$

The selection probability π can be estimated by a working model. In our work, we assume a logistic regression model

$$\pi(A_{ij}, A_{ji}, \boldsymbol{X}_i, \boldsymbol{X}_j; \boldsymbol{\gamma}) = A_{ij} \frac{\exp(\gamma_0 + \boldsymbol{X}_{ji}^T \boldsymbol{\gamma}_1)}{1 + \exp(\gamma_0 + \boldsymbol{X}_{ji}^T \boldsymbol{\gamma}_1)} + A_{ji} \frac{\exp(\gamma_0 + \boldsymbol{X}_{ij}^T \boldsymbol{\gamma}_1)}{1 + \exp(\gamma_0 + \boldsymbol{X}_{ij}^T \boldsymbol{\gamma}_1)}.$$

Since potential ordering indicators A_{ij} and A_{ji} are pairwise outcomes, we cannot use the standard logistic regression package in R to solve for $\boldsymbol{\gamma} = (\gamma_0, \gamma_1)$. Instead, we propose the following estimating equation to estimate $\boldsymbol{\gamma}$.

$$n^{-3/2} \sum_i \sum_{j \neq i} \boldsymbol{D}_{ij}(\boldsymbol{\gamma}) \left\{ \frac{\Delta_{ij}}{\pi(A_{ij}, A_{ji}, \boldsymbol{X}_i, \boldsymbol{X}_j; \boldsymbol{\gamma})} - 1 \right\} = 0, \tag{2.7}$$

where

$$\boldsymbol{D}_{ij}(\boldsymbol{\gamma}) = \left\{ \frac{\partial \pi(a_{ij}, a_{ji}, x_i, x_j; \boldsymbol{\gamma})}{\partial \boldsymbol{\gamma}_0}, \frac{\partial \pi(a_{ij}, a_{ji}, x_i, x_j; \boldsymbol{\gamma})}{\partial \boldsymbol{\gamma}_1} \right\},$$

$$\frac{\partial \pi(a_{ij}, a_{ji}, x_i, x_j; \boldsymbol{\gamma})}{\partial \boldsymbol{\gamma}_0} = a_{ij} \frac{\exp(\gamma_0 + \boldsymbol{X}_{ji}^T \boldsymbol{\gamma}_1)}{\{1 + \exp(\gamma_0 + \boldsymbol{X}_{ji}^T \boldsymbol{\gamma}_1)\}^2} + a_{ji} \frac{\exp(\gamma_0 + \boldsymbol{X}_{ij}^T \boldsymbol{\gamma}_1)}{\{1 + \exp(\gamma_0 + \boldsymbol{X}_{ij}^T \boldsymbol{\gamma}_1)\}^2}$$

$$\frac{\partial \pi(a_{ij}, a_{ji}, x_i, x_j; \boldsymbol{\gamma})}{\partial \boldsymbol{\gamma}_1} = a_{ij} \frac{x_{ji} \exp(\gamma_0 + \boldsymbol{X}_{ji}^T \boldsymbol{\gamma}_1)}{\{1 + \exp(\gamma_0 + \boldsymbol{X}_{ji}^T \boldsymbol{\gamma}_1)\}^2} + a_{ji} \frac{x_{ij} \exp[\gamma_0 + \boldsymbol{X}_{ij}^T \boldsymbol{\gamma}_1]}{\{1 + \exp(\gamma_0 + \boldsymbol{X}_{ij}^T \boldsymbol{\gamma}_1)\}^2}.$$

Here π may have value of zero. We define $0/0$ as zero for any term in the estimating equations appeared in the article.

Based on estimated selection probability $\hat{\pi}$, we can then estimate the regression parameter $\boldsymbol{\beta}$ using the proposed weighted estimating equation

$$n^{-3/2} \sum_i \sum_{j \neq i} \frac{\Delta_{ij}}{\hat{\pi}(A_{ij}, A_{ji}, \boldsymbol{X}_i, \boldsymbol{X}_j; \hat{\gamma})} W_{ij}(\boldsymbol{\beta}) \left\{ S_{ij}^{\tau} - E(S_{ij}^{\tau} | \boldsymbol{X}_i, \boldsymbol{X}_j) \right\} = 0. \tag{2.8}$$

The regression coefficients $(\boldsymbol{\beta}, \boldsymbol{\gamma})$ do not have closed-form solutions. By assuming that estimating equations are continuous at $(\boldsymbol{\beta}, \boldsymbol{\gamma})$, the estimates of regression coefficients are estimated through Newton-Raphson iterative method using R package rootSolve.

We use bootstrap method [Efron and Tibshirani, 1986] to obtain the estimated standard error due to the complicated equation form involving unknown components. The bootstrap method was suggested for deriving the asymptotic variance for interval-censored data [Zhang, 2009]. The algorithm is based on the following steps:

1. Generate a sample data from the parameters given above and assume that they are independent and identically distributed.
2. Draw a bootstrap sample of size n from the actual sample with replacement.
3. Independently draw the bootstrap samples B times and estimated the regression coefficients for each bootstrap sample.
4. The sample standard deviation of the B estimates can then be calculated as the bootstrap estimate of standard error.
5. Repeat these steps m times and take the mean it will give the averaged estimated standard error.

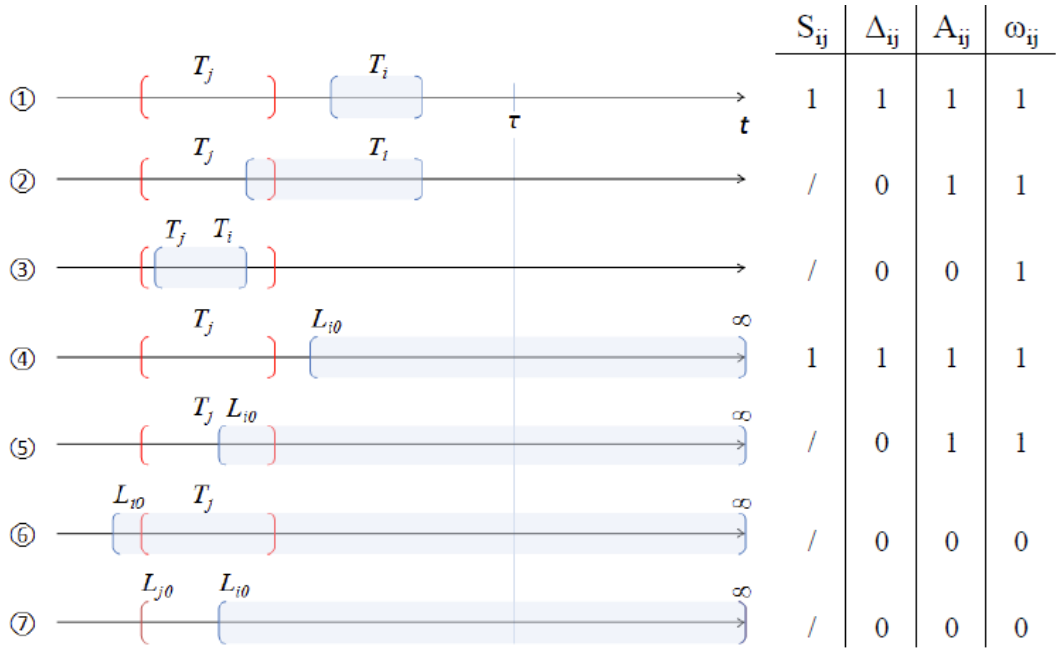
For a fairly large number of bootstrap samples, the bootstrap standard error will approach the standard error of the unknown sampling distribution. Under most cases, B in the range of 50 to 200 is considered adequate.

To illustrate the proposed estimation procedures, we compare the cases under regular setting (Figure 1) where subjects may only fail from a single cause with the cases under competing risks setting (Figure 2).

Auxiliary variable A_{ij} helps capture the ordering information of event time pairs when S_{ij}^{τ} is unobserved to predict selection probability. For subjects i and j under regular setting, there are only three conditions having $A_{ij} + A_{ji} = 0$:

1. If $(L_i, R_i], (L_j, R_j]$ are observed, and T_i's interval fully contains T_j's. Note that "fully contain" means $L_i < L_j$ and $R_i > R_j$, or vise versa.
2. If only T_i is right censored, and T_i's interval fully contains T_j's.
3. If both T_i and T_j are right censored.

Figure 1 covers all the possible cases under interval censoring without competing risks [Heller, 2011]. In cases 1 and 4, the ordering of event time pairs can be observed, so regression parameters can be directly estimated. For cases 2 and 5 the ordering cannot be directly



S_{ij}	Δ_{ij}	A_{ij}	ω_{ij}
1	1	1	1
/	0	1	1
/	0	0	1
1	1	1	1
/	0	1	1
/	0	0	0
/	0	0	0

Figure 1: Under regular setting (no competing risks), all the possible cases with the corresponding values of statistical quantities $S_{ij}, \Delta_{ij}, A_{ij}, \omega_{ij}$.

observed, but with auxiliary variables the selection probability can be estimated under MAR mechanism. In cases 3, 6, and 7, however, we cannot determine the ordering, even with the weighting technique. The missingness is considered not at random. Eventually, cases 6 and 7 will be eliminated as right censoring masks the potential ordering.

Situations under competing risks setting are more complex. With competing risks, either main event or competing event may be observed by the end of study, otherwise it is considered as censored. We summarize the definitions of all the ordering statistics from this section as follows.

1. S_{ij}^{τ} : If the main event of the jth subject occurred before that of the ith subject and the maximum follow-up time τ,

$$S_{ij}^{\tau} = I[T_i > T_j, T_j < \tau, \varepsilon_i = \varepsilon_j = 1]$$

2. Δ_{ij} : If the order of the main event times are accessible, in the sense that the event time intervals are non-overlapping. If the orders are available, the probability π is known, then we can use estimating equation (2.5) to estimate β.

$$\begin{aligned}\Delta_{ij} =& I[L_i \geq R_j, \varepsilon_i \neq 2, \varepsilon_j = 1] + I[\varepsilon_i = 2, \varepsilon_j = 1] \\ &+ I[L_j \geq R_i, \varepsilon_i = 1, \varepsilon_j \neq 2] + I[\varepsilon_i = 1, \varepsilon_j = 2] + I[\varepsilon_i = \varepsilon_j = 2].\end{aligned}$$

3. A_{ij} : When the ordering of main event times is unobserved ($\Delta_{ij} = 0$), if two event time intervals can provide any potential ordering information in estimating the selection probability.

$$\begin{aligned}A_{ij} =& I[L_i \geq L_j, R_i > R_j, \varepsilon_i = 1, \varepsilon_j = 1] + I[L_i \geq L_j, \varepsilon_i = 0, \varepsilon_j = 1] \\ &+ I[\varepsilon_i = 2, \varepsilon_j \neq 0].\end{aligned}$$

4. ω_{ij} : If potential ordering of event time intervals is unobservable ($A_{ij} + A_{ji} = 0$) due to right censoring.

$$\begin{aligned}\omega_{ij} =& I[\varepsilon_i \neq 0, \varepsilon_j \neq 0] + I[L_i \geq L_j, \varepsilon_i = 0, \varepsilon_j = 1] \\ &+ I[L_i \leq L_j, \varepsilon_i = 1, \varepsilon_j = 0]\end{aligned}$$

Data with $\omega_{ij} = 0$ will be removed from estimating equation (2.6). It turns out that ω_{ij} is trivial in the estimation.

Figure 2 illustrates the cases that main event for subject j is observed first. The observed intervals for subject j are indicated in red. Cases 1, 4, and 5 can observe ordering of time interval pairs with $\Delta_{ij} = 1$. Cases 2 and 6 are MAR with $A_{ij} = 1$. Cases 3 and 7 are NMAR, however, case 7 will be eliminated from estimating equation (2.6) while case 3 will be kept in.

Similarly, although we can see that two time intervals of competing risks are overlapped in case 8, according to the modeling framework the main events will both occur at infinity. So the selection probability can be directly estimated. The ordering of data pairs in cases 9 and 10 are obscured by right censoring, thus they will be eventually eliminated from estimating equation (2.6).

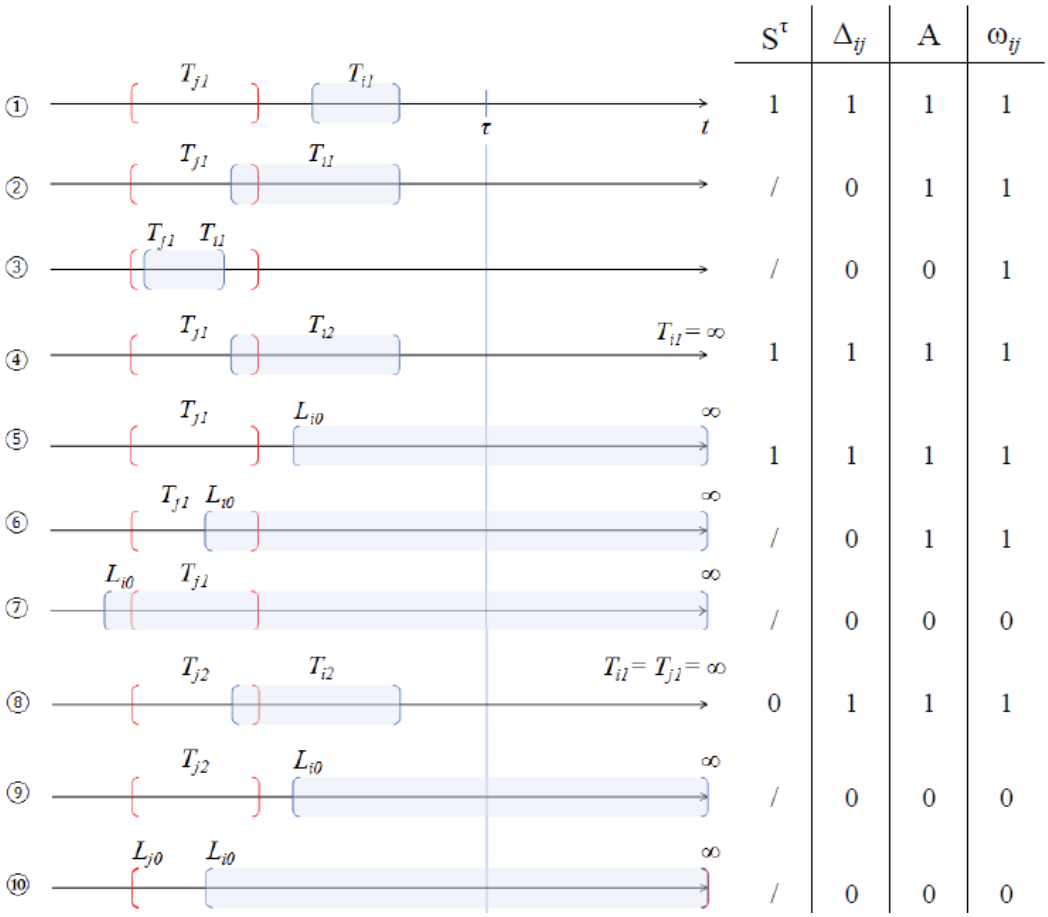


Figure 2: When event of interest for subject j is observed under competing risks setting, all cases with the corresponding values of statistical quantities $S_{ij}^{\tau}, \Delta_{ij}, A_{ij}, \omega_{ij}$.

3.0 SIMULATION STUDIES

In this section we carried out a number of simulation studies to evaluate the performance of the proposed estimation procedures. The first study consists of two settings of simulations. For each setting, extensive simulations were conducted to compare the estimators of the proposed estimating equations with Fine and Gray's method using midpoint of the observed time interval as the hypothetical event time. The details of implementation on Fine and Gray's method will be explained later.

For both settings in the first study, data were generated as follows. There were two covariates $X_i = (X_{i1}, X_{i2})$. The continuous covariate X_{i1} follows a standard normal distribution. The categorical variable X_{i2} follows Bernoulli(0.5). Assume that there were two types of events, a main event and a competing event. The subdistribution for main events was given by

$$\Pr(T_i \leq t, \varepsilon_i = 1 | \boldsymbol{X_i}) = 1 - \Big[1 - p\big\{1 - \exp(-\lambda_1 t^{\alpha_1})\big\}\Big]^{\exp(X_{i1}\beta_{11} + X_{i2}\beta_{12})}$$

where p is the proportion of main events without censoring. You may think of this subdistribution as a Weibull mixture. It has a Weibull distribution function with shape parameter α_1 and scale parameter λ_1 when $t < \infty$, and a point mass p at infinity when covariates are zero. The subdistribution for competing risk was then obtained by taking $\Pr(\varepsilon_i = 2|X_i) = 1 - \Pr(\varepsilon_i = 1|X_i)$. It follows an exponential distribution with rate of $\lambda_2 \exp(X_{i1}\beta_{21} + X_{i2}\beta_{22})$, that is,

$$\Pr(T_i \leq t | \varepsilon_i = 2, \boldsymbol{X_i}) = 1 - \exp\Big\{-\lambda_2 t \exp(X_{i1}\beta_{21} + X_{i2}\beta_{22})\Big\}.$$

Our interest is to estimate $\boldsymbol{\beta}_1$, the effects of covariates on the subdistribution for main event. We also investigate $F_{10} = 1 - \exp(-\beta_2)$, the baseline subdistribution for main event

by the end of study. Note that the nuisance parameter $\beta_2 = \Lambda_{10}(\tau) = \int_0^\tau \lambda_{10}(t)dt$ should be distinguished from the regression coefficients (β_{21}, β_{22}) for subdistribution of competing risk. The cumulative baseline subdistribution hazard function for main event, $\Lambda_{10}(t)$, depends on the percentage of main events p and the maximum follow-up time t. It is calculated as $-\log\Big[1 - p\big\{1 - \exp(-\lambda_1 t^{\alpha_1})\big\}\Big]$.

For all the simulation studies, we set β_{11} and β_{12} to be 0.5 and 1, respectively. In both settings of the first study, failure times to main events are assumed to follow an exponential distribution and a two-parameter Weibull distribution, respectively. The corresponding parameters were specified as follows:

1. $(\lambda_1, \lambda_2, \alpha_1, \beta_{21}, \beta_{22}) = (0.5, 0.5, 1, -0.8, -0.2)$;
2. $(\lambda_1, \lambda_2, \alpha_1, \beta_{21}, \beta_{22}) = (0.02, 0.35, 2, -1, -0.2)$.

We also considered a cure-type Gompertz distribution [Hudgens et al., 2014, Jeong and Fine, 2006] to model the subdistributions in the second study which will be discussed later.

The study visit times were generated as $\eta_{ik} = qk + U_{ik}$ where q indicates the time span of each scheduled visit; $k = 1, ..., K$, indicates the kth visit up to the maximum number of visit K. Random variable U_{ik} was uniformly distributed on $(-0.1q, 0.1q)$, indicating each visit may occur a certain time earlier or later from the scheduled times. Left and right endpoints for the failure time interval are computed as

$$\begin{aligned}(L_i, R_i) =& (0, \eta_{i1})I[T_i < \eta_{i1}] + \sum_{k=2}^{K}\{\eta_{i(k-1)}, \eta_{ik}\}I[\eta_{i(k-1)} < T_i < \eta_{ik}] \\ &+ (\eta_{iK}, \infty)I[T_i > \eta_{iK}].\end{aligned}$$

The underlying failure times are therefore interval-censored. Note that in many occasions, some events may be observed exactly while others are interval-censored. Our estimation procedures can handle both cases, but for simplicity of data generation, we considered only the latter case. The observed time information are the left and right endpoints where the first failure was observed.

Censoring time C_i was randomly generated from a uniform $[a, \tau]$ distribution, where τ is the maximum follow up time (qK); adjusting a produced averaged 15% and 30% censoring rates. If censoring time is less than the visit time where failure was observed, the follow-up

time will be censored at the time of the previous visit, that is, if $C_i < R_i$ where $T_i \in (L_i, R_i)$, then it will be censored at the visit time prior to C_i.

To assess the performance of our proposed estimation procedures, we created six scenarios for each setting. Each scenario was simulated with sample size $n = \{200, 400\}$ and 1000 replications. The numbers of intervals varied from 5, 20, to 100. For all setting, we chose values for a and τ to achieve censoring rates of 15% and 30% on average. In the presence of independent censoring, the percentages of main events were also given in Table 1. Under 15% censoring, there are 64% failures from main event and 21% failures from competing risk. For 30% censoring, there are 53% failures from main event and 17% failures from competing risk. These proportions remain the same across different numbers of intervals.

Table 1: Simulation setting I where subdistribution for main events follows an exponential distribution.

Number of intervals	Interval length	Censoring (a, τ)	Censoring %	Main event %
5	0.8	(2.2, 4)	15	64
5	0.8	(0.3, 4)	30	53
20	0.2	(2.2, 4)	15	64
20	0.2	(0.3, 4)	30	53
100	0.04	(2.2, 4)	15	64
100	0.04	(0.3, 4)	30	53

For each simulated data set, the regression coefficients (β_{11}, β_{12}) and the baseline subdistribution F_{10} were computed based on the proposed weighted estimating equations (2.7, 2.8). Based on 1000 data sets per scenario per sample size, we calculated the mean, empirical standard error (ESE), average estimated standard error (ASE) for those estimators. ESE's were the standard deviations of the simulation estimates. Due to the complicated equation form which involves unknown components, the ASE's were computed using bootstrap method [Efron and Tibshirani, 1986] to assess the efficiency. First we generated a sample data and assumed they were independent and identically distributed, then a bootstrap sample of size

n was drawn from the actual sample with replacement. Independently drew the bootstrap samples $B = 100$ times and estimated the regression coefficients for each bootstrap sample. The sample standard deviation of the 100 estimates was then calculated as the bootstrap estimate of standard error. Repeat these steps 1000 times and take the mean will give the ASE. This method was suggested for the asymptotic variance [Zhang, 2009] in the presence of interval censoring.

Table 2 presents the average estimates of $(\boldsymbol{\beta}_1, F_{10})$ from 1000 samples, empirical standard error of those estimates from 1000 samples, average estimated standard error based on bootstrap estimate of standard error, and coverage probabilities of 95% confidence intervals for $(\beta_{11}, \beta_{12}, F_{10})$. True parameter values for $\boldsymbol{\beta}_1 = (\beta_{11}, \beta_{12})$ were predefined to be the same in all simulations. In the absence of covariates the subdistribution F_{10} is equivalent to $1 - \exp(-\beta_2)$ which depends only on p and maximum follow-up time τ. For settings I and II, the true parameter values were $(\beta_{11}, \beta_{12}, F_{10}) = (.5, 1, .52)$.

In terms of the simulation results, intuitively there would be substantial bias when very few number of visits (K) were made, for example, 5 visits. Table 2 shows that the bias for both estimated regression coefficients β_{11} and β_{12} are within a reasonable range from 0.05 (under 15% censoring) to 0.07 (under 30% censoring). As sample size increases, the bias decreases. As number of intervals increases from 5 up to 100, the bias of estimated coefficients reduces to 0.01 when sample size reaches 400. We can see the same pattern for empirical standard errors. The averaged estimated standard errors provide a good approximation to ESE. Comparing to $\boldsymbol{\beta}_1$, estimated baseline subdistribution F_{10} gives relatively large bias and small standard error. The trend also suggests further investigation.

We performed additional simulations to compare our results to those estimates derived from Fine and Gray's weighted score function [Fine and Gray, 1999]. For right-censored competing risks data, Fine and Gray's model estimates the effects of covariates on the subdistribution of a particular type of failure from the partial likelihood. The subjects who experienced a competing risk that are remained in the risk set and the time-dependent IPCW is applied to the likelihood function based on the conditional probability of being followed up given main event have not occurred. Under interval censoring setting, we imputed the failure time using the midpoint of the time interval. More specifically, for interval-censored cases we

Table 2: Simulation results for our proposed method reporting the mean, empirical standard error (ESE), averaged estimated standard error (ASE), and 95% coverage probability (CP) based on sample sizes 200 and 400 with 1000 replications under Setting I (Exponential). True parameter values $(\beta_{11}, \beta_{12}, F_{10}) = (.5, 1, .52)$.

			Mean			ESE[§]			ASE[♮]			CP[♯]		
n	K	PRC[†]	β_{11}	β_{12}	$F_{10}^{\ddagger}$	β_{11}	β_{12}	F_{10}	β_{11}	β_{12}	F_{10}	β_{11}	β_{12}	F_{10}
200	5	15	0.551	1.069	0.573	0.124	0.217	0.059	0.132	0.228	0.058	0.963	0.952	0.826
		30	0.570	1.085	0.565	0.144	0.251	0.066	0.178	0.290	0.073	0.966	0.950	0.896
	20	15	0.534	1.055	0.611	0.113	0.221	0.058	0.117	0.216	0.057	0.953	0.943	0.629
		30	0.546	1.069	0.617	0.127	0.246	0.065	0.137	0.257	0.067	0.958	0.950	0.666
	100	15	0.512	1.015	0.628	0.108	0.200	0.054	0.110	0.210	0.056	0.948	0.957	0.518
		30	0.525	1.029	0.639	0.121	0.229	0.062	0.125	0.241	0.064	0.953	0.962	0.522
400	5	15	0.545	1.055	0.571	0.086	0.151	0.039	0.088	0.154	0.040	0.938	0.945	0.755
		30	0.562	1.076	0.562	0.100	0.177	0.046	0.103	0.177	0.045	0.935	0.933	0.846
	20	15	0.518	1.024	0.612	0.076	0.138	0.037	0.079	0.147	0.039	0.947	0.957	0.325
		30	0.530	1.035	0.618	0.084	0.157	0.041	0.089	0.166	0.044	0.954	0.961	0.384
	100	15	0.510	1.020	0.623	0.076	0.145	0.038	0.076	0.144	0.039	0.953	0.945	0.246
		30	0.523	1.033	0.634	0.085	0.162	0.042	0.085	0.164	0.043	0.948	0.942	0.255

[§]ESE: empirical standard error [♮]ASE: average estimated standard error [♯]CP : 95% coverage probability

[†]PRC: percent of right censoring [‡]F_{10}: baseline subdistribution

used the midpoint ($\frac{L_i+R_i}{2}$) of the observed time interval as the hypothetical event time. For right-censored cases, subjects censored at the left endpoint (L_i) of the time interval where the censoring occurs. In Fine and Gray, the estimate of F_{10} is computed from $1-\exp\{-\sum_i^n h_i(\cdot)\}$ where $h(\cdot)$ is the jump size in the Breslow-type estimate of underlying subdistribution hazard. We calculated the average estimated regression coefficients of (β_{11}, β_{12}) and average estimated baseline subdistribution F_{10} by the end of study from 1000 samples using R package *cmprsk* [Gray, 2014].

Overall, the results from Fine and Gray's midpoint imputation method (Table 3) demonstrate a good estimation with slight bias toward small sample. The variance estimators are consistent in all scenarios. In comparison to their results, our estimates exhibit a higher small-sample bias but when sample size reaches 400 both methods produce unbiased estimates.

As previously described, in the second setting we assumed a Weibull distribution for subdistribution of main event. Similarly, six scenarios (Table 7) were created for sample sizes $n = \{200, 400\}$. Each scenario was simulated and estimated 1000 times. There are three numbers of intervals, varying from 5 to 100. The parameters gave similar percentages of failures from main event and competing risk. The maximum follow-up time now extended to 10, so did the time span for each interval.

The results for Weibull setting in Table 5 have a pattern resemble that of the exponential setting. As number of intervals increases from 5 up to 100, the bias decreases. As censoring rate increases from 15% to 30%, the bias of $\boldsymbol{\beta}_1$ increases. Similarly, the estimates tend to be unbiased for large sample size. Compare to setting I, there were slightly larger bias and estimated standard errors, especially for smaller number of intervals K and heavier censored sample. Estimated F_{10} exhibits an increasing trend. The ASE agreed well with ESE. Estimates of regression coefficient β_{12} have the highest variability amongst those three.

Interestingly, the bias from Fine and Gray's method (Table 6) increases as numbers of intervals increase for sample size 200. For large sample size ($n = 400$), the performance is getting stable and the estimates are unbiased. Given the results from the first two settings, we consider that our method can provide unbiased estimates for regression coefficient $\boldsymbol{\beta}_1$ under large sample while it does not give a very reliable prediction for baseline hazard.

Table 3: Simulation results for Fine and Gray's midpoint imputation method reporting the mean, empirical standard error (ESE), averaged estimated standard error (ASE), and 95% coverage probability (CP) based on sample sizes 200 and 400 with 1000 replications under Setting I (Exponential). True parameter values $(\beta_{11}, \beta_{12}, F_{10}) = (.5, 1, .52)$.

			Mean			ESE[§]			ASE[♮]		CP[‡]	
n	K	PRC[†]	β_{11}	β_{12}	F_{10}	β_{11}	β_{12}	F_{10}	β_{11}	β_{12}	β_{11}	β_{12}
200	5	15	0.493	0.988	0.573	0.097	0.192	0.071	0.092	0.185	0.935	0.935
		30	0.494	0.987	0.564	0.106	0.205	0.074	0.099	0.205	0.925	0.955
	20	15	0.506	1.016	0.523	0.102	0.195	0.062	0.093	0.186	0.933	0.942
		30	0.510	1.005	0.523	0.104	0.213	0.073	0.102	0.206	0.943	0.957
	100	15	0.513	1.015	0.513	0.097	0.187	0.062	0.095	0.187	0.940	0.947
		30	0.503	1.012	0.509	0.108	0.205	0.069	0.102	0.206	0.928	0.948
400	5	15	0.486	0.986	0.581	0.097	0.192	0.071	0.065	0.131	0.947	0.962
		30	0.497	0.989	0.574	0.106	0.205	0.074	0.070	0.145	0.950	0.933
	20	15	0.504	1.002	0.533	0.064	0.134	0.056	0.066	0.132	0.957	0.947
		30	0.506	1.013	0.528	0.076	0.134	0.055	0.072	0.145	0.927	0.973
	100	15	0.498	1.002	0.520	0.064	0.133	0.045	0.066	0.132	0.943	0.950
		30	0.506	0.996	0.521	0.073	0.149	0.056	0.072	0.145	0.945	0.952

[§]ESE: empirical standard error [♮]ASE: average estimated standard error

[†]PRC: percent of right censoring [‡]CP : 95% coverage probability

Table 4: Simulation setting II where subdistribution for main events follows a Weibull distribution.

Number of intervals	Interval length	Censoring (a, τ)	Censoring %	Main event %
5	2	(6.5, 10)	15	61
5	2	(2.5, 10)	30	47
20	0.5	(6.5, 10)	15	61
20	0.5	(2.5, 10)	30	47
100	0.1	(6.5, 10)	15	61
100	0.1	(2.5, 10)	30	47

To examine whether the proposed method is robust to the choice of distribution function, we considered Gompertz parametrization with an exponentially increasing hazard rate for the second simulation study. The subdistribution for main events were generated as

$$\Pr(T_i \leq t, \varepsilon_i = 1|\boldsymbol{X_i}) = 1 - \Big\{1 - p\Big(1 - \exp\big[\eta_1\{1 - \exp(\alpha_1 t)\}/\alpha_1\big]\Big)\Big\}^{\exp(X_{i1}\beta_{11} + X_{i2}\beta_{12})}$$

The subdistribution for competing risks was given by

$$\Pr(T_i \leq t, \varepsilon_i = 2|\boldsymbol{X_i}) = 1 - \exp\big[\eta_2\{1 - \exp(\alpha_2 t)\}\exp(X_{i1}\beta_{21} + X_{i2}\beta_{22})/\alpha_2\big],$$

where $(\beta_{21}, \beta_{22}) = (-0.2, -0.2)$. To ensure an improper distribution function, we set Gompertz parameters $(\eta_1, \eta_2, \alpha_1, \alpha_2) = (0.1, 0.25, -0.06, -0.035)$. If one unlimitedly prolongs the study time t to infinity, the probability of failing from the main event would be $p\{1 - \exp(\eta_1/\alpha_1)\}$. The baseline subdistribution by any time t is

$$F_{10}(t) = 1 - p\Big(1 - \exp\Big[\eta_1\big\{1 - \exp(\alpha_1 t)\big\}/\alpha_1\Big]\Big).$$

The true parameter values were $(\beta_{11}, \beta_{12}, F_{10}) = (0.5, 1, 0.38)$.

Still, there are six scenarios (Table 7) for sample sizes $n = \{200, 400\}$. Each scenario was simulated 600 times. For each scenario, regression coefficients (β_{11}, β_{12}) and baseline

Table 5: Simulation results for our proposed method reporting the mean, empirical standard error (ESE), averaged estimated standard error (ASE), and 95% coverage probability (CP) based on sample sizes 200 and 400 with 1000 replications under Setting II (Weibull). True parameter values $(\beta_{11}, \beta_{12}, F_{10}) = (.5, 1, .52)$.

			Mean			ESE[§]			ASE[♮]			CP[♯]		
n	K	PRC[†]	β_{11}	β_{12}	$F_{10}^{\ddagger}$	β_{11}	β_{12}	F_{10}	β_{11}	β_{12}	F_{10}	β_{11}	β_{12}	F_{10}
200	5	15	0.578	1.120	0.536	0.130	0.229	0.060	0.135	0.240	0.060	0.946	0.930	0.928
		30	0.618	1.159	0.490	0.158	0.281	0.068	0.168	0.299	0.072	0.931	0.944	0.938
	20	15	0.536	1.059	0.574	0.116	0.226	0.060	0.121	0.226	0.062	0.952	0.937	0.827
		30	0.563	1.094	0.532	0.136	0.267	0.067	0.145	0.275	0.068	0.942	0.947	0.940
	100	15	0.513	1.017	0.590	0.110	0.205	0.056	0.113	0.214	0.058	0.949	0.957	0.765
		30	0.539	1.050	0.550	0.129	0.247	0.063	0.135	0.259	0.066	0.945	0.956	0.929
400	5	15	0.567	1.113	0.534	0.095	0.162	0.041	0.092	0.164	0.041	0.906	0.901	0.943
		30	0.605	1.158	0.486	0.111	0.196	0.046	0.112	0.199	0.047	0.881	0.881	0.909
	20	15	0.519	1.029	0.576	0.076	0.136	0.037	0.080	0.151	0.040	0.949	0.975	0.749
		30	0.544	1.059	0.535	0.091	0.166	0.041	0.095	0.181	0.045	0.952	0.965	0.95
	100	15	0.511	1.019	0.584	0.079	0.149	0.039	0.077	0.148	0.039	0.953	0.945	0.64
		30	0.537	1.051	0.544	0.094	0.171	0.043	0.092	0.180	0.045	0.946	0.951	0.928

[§]ESE: empirical standard error [♮]ASE: average estimated standard error [♯]CP : 95% coverage probability

[†]PRC: percent of right censoring [‡]F_{10}: baseline subdistribution

Table 6: Simulation results for Fine and Gray's midpoint imputation method reporting the mean, empirical standard error (ESE), averaged estimated standard error (ASE), and 95% coverage probability (CP) based on sample sizes 200 and 400 with 1000 replications under Setting II (Weibull). True parameter values $(\beta_{11}, \beta_{12}, F_{10}) = (.5, 1, .52)$.

			Mean			ESE[§]			ASE[♮]		CP[‡]	
n	K	PRC[†]	β_{11}	β_{12}	F_{10}	β_{11}	β_{12}	F_{10}	β_{11}	β_{12}	β_{11}	β_{12}
200	5	15	0.505	0.999	0.580	0.098	0.197	0.059	0.094	0.188	0.935	0.940
		30	0.507	1.007	0.567	0.115	0.219	0.067	0.106	0.215	0.920	0.943
	20	15	0.507	1.019	0.531	0.105	0.199	0.067	0.095	0.190	0.932	0.940
		30	0.513	1.014	0.520	0.112	0.227	0.076	0.107	0.218	0.943	0.955
	100	15	0.514	1.017	0.511	0.099	0.191	0.065	0.096	0.191	0.942	0.950
		30	0.505	1.006	0.505	0.110	0.223	0.075	0.107	0.218	0.937	0.952
400	5	15	0.498	0.998	0.585	0.066	0.130	0.041	0.067	0.133	0.953	0.962
		30	0.509	1.012	0.573	0.075	0.161	0.050	0.075	0.152	0.942	0.930
	20	15	0.505	1.005	0.544	0.065	0.136	0.057	0.067	0.134	0.958	0.945
		30	0.506	1.021	0.531	0.081	0.141	0.056	0.076	0.153	0.920	0.975
	100	15	0.499	1.002	0.520	0.065	0.134	0.047	0.067	0.134	0.947	0.940

[§]ESE: empirical standard error [♮]ASE: average estimated standard error

[†]PRC: percent of right censoring [‡]CP : 95% coverage probability

CIF F_{10} were estimated by proposed estimating equation method and compared to Fine and Gray's midpoint approach. Numbers of intervals are the same as those in previous study. The parameters generated less type-1 failures than the previous study.

Table 7: Simulation setting III where subdistribution for main events follows a Gompertz distribution.

Number of intervals	Interval length	Censoring (a, τ)	Censoring %	Main event %
5	3	(11, 15)	15	54
5	3	(2, 15)	30	44
20	0.75	(11, 15)	15	54
20	0.75	(2, 15)	30	44
100	0.15	(11, 15)	15	54
100	0.15	(2, 15)	30	44

Table 8 presents the average estimates from 600 samples, empirical standard error (ESE) of those estimates from 600 samples, and average estimated standard error (ASE) of 600 bootstrap estimates of standard error based on 100 times bootstrapping. The coverage probabilities of 95% confidence intervals were also reported for estimates of β_{11} and β_{12}.

Results using our method exhibit reasonable estimates for regression coefficients β_{11} and β_{12}. Under light censoring (15%), it is approximately unbiased. The average estimated standard errors of estimators are consistent with those simulation standard errors, although the estimated standard errors under sample size 200 are slightly larger compared to the those from Settings I and II. The coverage probabilities of 95% confidence intervals are between 92% to 97.6%. Fine and Gray's results shown in Table 9 indicate that for small time intervals ($K = 5, 10$) the estimators have smallest relative bias and estimated standard errors. The 95% confidence intervals of the estimators achieve an overall nice coverage. The coverage probabilities range from 92.7% to 96%, which are close to but slightly tighter than ours.

In summary, the simulation studies have shown that our method provides a reasonable estimate of regression coefficient $\boldsymbol{\beta}_1$ with a plausible trend. As sample size and number

Table 8: Simulation results for our proposed method reporting the mean, empirical standard error (ESE), averaged estimated standard error (ASE), and 95% coverage probability (CP) based on sample sizes 200 and 400 with 600 replications under Setting III (Gompertz). True parameter values $(\beta_{11}, \beta_{12}, F_{10}) = (.5, 1, .38)$.

			Mean			ESE[§]			ASE[♮]			CP[♯]		
n	K	PRC[†]	β_{11}	β_{12}	$F_{10}^{\ddagger}$	β_{11}	β_{12}	F_{10}	β_{11}	β_{12}	F_{10}	β_{11}	β_{12}	F_{10}
200	5	15	0.566	1.102	0.433	0.135	0.242	0.059	0.163	0.296	0.074	0.963	0.960	0.860
		30	0.593	1.143	0.418	0.158	0.277	0.066	0.178	0.311	0.084	0.930	0.953	0.930
	20	15	0.525	1.061	0.465	0.124	0.241	0.060	0.136	0.248	0.065	0.947	0.957	0.718
		30	0.542	1.078	0.465	0.139	0.274	0.068	0.146	0.283	0.072	0.960	0.952	0.751
	100	15	0.513	1.030	0.479	0.113	0.229	0.058	0.125	0.239	0.064	0.963	0.968	0.628
		30	0.527	1.038	0.479	0.121	0.248	0.064	0.134	0.265	0.069	0.971	0.976	0.695
400	5	15	0.552	1.085	0.435	0.089	0.168	0.043	0.095	0.173	0.042	0.945	0.923	0.720
		30	0.580	1.127	0.417	0.105	0.193	0.047	0.112	0.203	0.048	0.923	0.920	0.878
	20	15	0.518	1.024	0.467	0.077	0.150	0.040	0.084	0.163	0.041	0.970	0.963	0.442
		30	0.535	1.054	0.465	0.085	0.171	0.043	0.095	0.186	0.047	0.960	0.956	0.553
	100	15	0.509	1.010	0.477	0.077	0.164	0.041	0.082	0.158	0.041	0.960	0.942	0.363
		30	0.521	1.028	0.480	0.088	0.181	0.045	0.093	0.180	0.046	0.963	0.940	0.421

[§]ESE: empirical standard error [♮]ASE: average estimated standard error [♯]CP : 95% coverage probability

[†]PRC: percent of right censoring [‡]F_{10}: baseline subdistribution

Table 9: Simulation results for Fine and Gray's midpoint imputation method reporting the mean, empirical standard error (ESE), averaged estimated standard error (ASE), and 95% coverage probability (CP) based on sample sizes 200 and 400 with 600 replications under Setting III (Gompertz). True parameter values $(\beta_{11}, \beta_{12}, F_{10}) = (.5, 1, .38)$.

			Mean			ESE[§]			ASE[♮]		CP[‡]	
n	K	PRC[†]	β_{11}	β_{12}	F_{10}	β_{11}	β_{12}	F_{10}	β_{11}	β_{12}	β_{11}	β_{12}
200	5	15	0.492	0.994	0.453	0.104	0.214	0.101	0.100	0.205	0.938	0.945
		30	0.499	1.016	0.451	0.112	0.243	0.131	0.110	0.229	0.943	0.943
	20	15	0.503	0.990	0.393	0.108	0.205	0.078	0.102	0.206	0.927	0.950
		30	0.506	0.998	0.387	0.117	0.232	0.088	0.112	0.229	0.938	0.960
	100	15	0.507	1.011	0.381	0.104	0.218	0.062	0.102	0.206	0.945	0.933
		30	0.509	1.014	0.378	0.117	0.231	0.073	0.112	0.229	0.933	0.953
400	5	15	0.497	0.999	0.440	0.073	0.144	0.059	0.071	0.145	0.937	0.952
		30	0.501	1.008	0.464	0.078	0.164	0.118	0.077	0.161	0.955	0.952
	20	15	0.503	1.000	0.403	0.074	0.149	0.079	0.071	0.146	0.935	0.947
		30	0.503	1.003	0.395	0.079	0.166	0.078	0.078	0.162	0.952	0.950
	100	15	0.509	1.002	0.378	0.076	0.156	0.041	0.072	0.146	0.935	0.940
		30	0.512	1.009	0.375	0.083	0.169	0.054	0.079	0.162	0.933	0.948

[§]ESE: empirical standard error [♮]ASE: average estimated standard error

[†]PRC: percent of right censoring [‡]CP : 95% coverage probability

of intervals increase, the bias of estimate decreases. The estimate tends to be unbiased (bias < 0.01) when right censoring is light (15%), while the bias tends to be larger with higher proportion (30%) of right censoring. We have tried even higher censoring rates, such as 40%, which turned out to be fine as well. Regarding the performance of estimated baseline subdistribution F_{10} by the maximum follow-up time, or β_2 in other forms, neither increasing sample size nor having more frequent visits guarantee an accurate estimate. In other words, our method does not provide a very reliable estimate for baseline subdistribution (or hazard) as Fine and Gray's estimator given the large bias and standard error. This can be understandable as F_{10} is a function of nuisance parameter β_2 and it is estimated simultaneously with other coefficients in the estimating equation, while Fine and Gray use a modified Breslow estimator for baseline hazard of subdistribution.

4.0 EXAMPLE

In this chapter we apply the proposed method to a real follicular cell lymphoma data [Pintilie, 2007]. The lymphoma patient data was collected at the Princess Margaret Hospital in Toronto, Canada since 1967. The subset of 541 patients includes all patients identified as having follicular type lymphoma, registered for treatment at the hospital between 1967 and 1996, with early stage disease (I or II) and treated with radiation alone or with radiation and chemotherapy. The goal of this study was to report the long-term outcome in this group of patients.

The recorded outcome includes response to treatment, first relapse (local, distant or both) and death. The response to treatment includes complete response and no response. Those with a complete response may have relapsed later locally, distantly or both locally and distantly. Those with no response were never disease-free and are considered local failures. The time to first failure is calculated in years from the date of diagnosis. For the patients with no response, the time to first failure is taken to be 1 day. For those with complete response but without relapse, the time to first failure is calculated up to the last follow-up date. The median follow-up time was 5.5 years.

In our application, main event $(272, 50.3\%)$ is disease relapse or no response. Death without relapse is the competing risk $(76, 14.0\%)$. Those subjects who have complete response without relapse and are alive by the end of study are censored $(193, 35.7\%)$. The patients ages in years (mean = 57 and sd = 14) and haemoglobin levels (mean = 138 and sd = 15) were two continuous variables. There is no significant correlation (p=0.16) between those two variables. The subdistribution curves for two types of failure (Figure 3) are calculated from the Kaplan-Meier estimator without covariates. Within the first 15 years, the inci-dence of lymphoma recurrence grows exponentially and then smooths out to reach about the probability of 0.6.

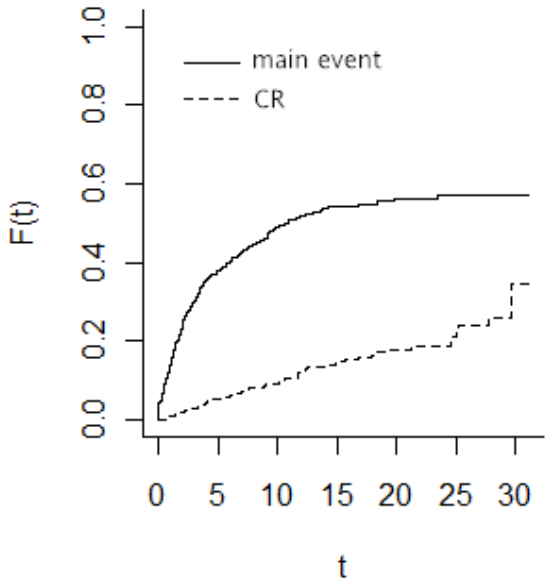


Figure 3: Cumulative incidence curves for main event and competing risk in the follicular cell lymphoma study.

We focus the analysis on the effect of age and haemoglobin levels on the cumulative incidence of lymphoma recurrence. Given the fact that the data are right-censored with competing risks, we create a situation where the underlying failure time is assumed to be interval-censored with a variety of time interval lengths. First, we generate a sequence of consecutive intervals (120, 60, 30, 15, and 6) with equal length (every 3 months, 6 months, 1 year, 2 years, and 5 years) during the period of 30 years. The hypothetical left and right endpoints of the interval were determined by the time from diagnosis to first failure (death, relapse or no response) or censoring, whichever occurs first. The time to death is usually exactly observed. However, for simplicity we treat it as interval-censored as well. Our method can easily adapt to a situation where both interval censoring and direct observation are present. Under proportional subdistribution hazards framework, if one is censored the right endpoint will be set to infinity. Both covariates were centered to the mean. The distribution

of centered age is approximately normal, while the distribution of centered haemoglobin levels is slightly left skewed. We assume independent and identically distributed covariates and constant covariate effects over time (i.e. proportional hazards). We also assume non-informative censoring. There is no missing values.

We began the analysis by considering the effect of a single covariate age in the model. To create a hypothetical interval-censored competing risks setting, the data are artificially interval censored at five time interval lengths, $\{0.25, 0.5, 1, 2, 5\}$ years. Table 10 gives the estimated regression coefficient, estimated standard error (SE) based on 100 times bootstrap, 95% confidence interval (CI), and p-value from Wald-type test assuming that the estimate for regression coefficient is asymptotically normal with mean zero, that is, $(\hat{\beta} - \beta_0)/SE(\hat{\beta}) \sim N(0, 1)$. For comparison, these performance characteristics were computed based on our proposed estimating equations and Fine and Gray's method using midpoint imputation. The true parameter values were $(\beta_{age}, F_{10}) = (0.0134, 0.5719)$, computed by Fine and Gray's method based on the observed right-censored data.

Table 10 summarizes the results of our proposed method and Fine and Gray's midpoint imputation method given five intervals. The proposed method exhibits an increasing trend for covariate effects and an opposing trend for F_{10} as time interval increases from 0.25 (3 months) up to 5 (6 years). In contrast, Fine and Gray's method shows a decreasing pattern for both estimators. Interestingly, although the trends for estimated age effect are different between these two methods, bias consistently decreases when patients visit more frequently. This is what we expected to see. The interpretation may attenuate the difference between two methods in some cases. Take our results as an example, if lymphoma patients go to hospital every 3 months, the patient with 1-year older is 1.02 (≈ 1) times more likely to experience lymphoma recurrence or no response to treatment. In comparison, the hazard ratio from Fine and Gray's midpoint imputation method is 1.01, which means the hazard for recurrence is nearly 1 and thus can be considered as no difference among age groups. The 95% confidence intervals cover the true parameters throughout all cases while our estimates are more conservative. Age has a significant effect ($p < 0.05$) on the cumulative incidence of lymphoma recurrence for both methods.

Table 10: Application data analysis results comparing the estimated effect of age and baseline subdistribution (Coefficient), estimated standard error (SE), 95% confidence interval (CI), and p-value for the follic cell lymphoma study. Based on Fine and Gray's method on the right-censored data, the estimated $(\beta_{age}, F_{10}) = (0.0134, 0.5719)$.

Time Interval	Method	Coefficient		SE		CI		P-value
(year)		age	F_{10}	age	F0	age		age
0.25	FG	0.0133	0.5711	0.0044		0.0046	0.0220	0.003
	Proposed	0.0144	0.8872	0.0052	0.0533	0.0043	0.0246	0.005
0.5	FG	0.0131	0.5689	0.0044		0.0046	0.0217	0.003
	Proposed	0.0145	0.8863	0.0052	0.0526	0.0042	0.0248	0.006
1	FG	0.0128	0.5664	0.0043		0.0045	0.0212	0.003
	Proposed	0.0149	0.8873	0.0054	0.0598	0.0043	0.0254	0.005
2	FG	0.0121	0.5589	0.0040		0.0042	0.0200	0.003
	Proposed	0.0150	0.8865	0.0056	0.0743	0.0040	0.0260	0.008
5	FG	0.0097	0.5467	0.0037		0.0026	0.0169	0.008
	Proposed	0.0155	0.8724	0.0057	0.0740	0.0044	0.0266	0.006

Next, we added haemoglobin levels along with age into the model. Comparative results are given in Table 11. The performance of estimated coefficient for age between Table 10 and Table 11 are almost identical for both methods. In contrast, the estimated regression coefficients for age based on our method only differentiate from those of Fine and Gray's method in 0.001, which is very small, although Fine and Gray's estimates are closer to the estimated values based on observed data. The estimates for haemoglobin levels are approximately zero. According to Wald test from our method, age is still significant while haemoglobin is not ($p > 0.6$). In terms of baseline subdistribution, our estimates are larger than those from Fine and Gray's model. One explanation is that the estimating equations find the roots simultaneously using Newton-Raphson, and only use the information from observations which have definitive main event time orderings. Instead, Fine and Gray's method estimates baseline subdistribution hazard separately using Breslow-type estimator. In addition, the inverse probability censoring weighting technique adjusts for missingness for those who have a competing risks event even if ambiguous ordering exists.

Table 11: Application data analysis results comparing the estimated effect of age and haemoglobin levels and baseline subdistribution (Coefficient), estimated standard error (SE), 95% confidence interval (CI), and p-value for the follic cell lymphoma study. Based on Fine and Gray's method on the right-censored data, the estimated $(\beta_{age}, \beta_{hgb}, F_{10}) = (0.0134, 0.0005, 0.5720)$.

Time Interval	Method	Coefficient			SE			CI				P-value	
(year)		age	hgb	F_{10}	age	hgb	F0	age		hgb		age	hgb
0.25	FG	0.0134	0.0006	0.5687	0.0044	0.0040		0.0047	0.0220	-0.0072	0.0083	0.003	0.890
	Proposed	0.0144	-0.0021	0.8887	0.0046	0.0043	0.0587	0.0054	0.0234	-0.0106	0.0064	0.002	0.627
0.5	FG	0.0132	0.0007	0.5691	0.0044	0.0039		0.0046	0.0217	-0.0071	0.0084	0.003	0.870
	Proposed	0.0145	-0.0021	0.8877	0.0047	0.0047	0.0651	0.0052	0.0237	-0.0113	0.0071	0.002	0.656
1	FG	0.0129	0.0009	0.5666	0.0043	0.0039		0.0046	0.0213	-0.0066	0.0085	0.002	0.810
	Proposed	0.0149	-0.0018	0.8888	0.0045	0.0041	0.0545	0.0060	0.0238	-0.0098	0.0062	0.001	0.657
2	FG	0.0122	0.0014	0.5592	0.0040	0.0037		0.0043	0.0201	-0.0059	0.0086	0.003	0.710
	Proposed	0.0150	-0.0011	0.8875	0.0053	0.0046	0.0629	0.0046	0.0254	-0.0102	0.0080	0.005	0.812
5	FG	0.0099	0.0024	0.5474	0.0037	0.0033		0.0028	0.0171	-0.0040	0.0088	0.007	0.460
	Proposed	0.0157	0.0001	0.8724	0.0054	0.0051	0.0605	0.0051	0.0263	-0.0099	0.0102	0.004	0.979

5.0 DISCUSSION AND CONCLUSIONS

Interval-censored competing risks data have gained increasing attention, especially in longitudinal clinical studies. Nonparametric methods have been explored extensively. The most commonly used methods include imputation methods (midpoint imputation, multiple imputation) and partial likelihood estimation under proportional hazards assumption. However, estimation procedures based on the proportional subdistribution hazards model are seldom investigated. In this dissertation, we have demonstrated the estimation and inference on the regression parameters under the proportional subdistribution hazards regression model. We proposed an inverse probability weighted estimating equations method to account for missingness due to interval censoring and competing risks.

In Chapter 2, we briefly reviewed the proportional subdistribution hazards model and modified the model to adapt to the situation where failure time is not exactly observed but censored within interval. The interval-censored competing risks data have faced two issues with respect to incomplete data. One comes from unobserved failure times resulted from interval censoring. The second issue arises when competing risks preclude from observing the event of interest. By introducing an auxiliary variable and assuming missing at random mechanism, our proposed semiparametric method can efficiently reduce the potential bias that is caused by missingness.

Simulation studies were conducted with various settings and scenarios in Chapter 3 in order to examine and evaluate the accuracy and efficiency of proposed estimators. The performance was also compared to a set of IPCW partial likelihood estimators which proved to be consistent and efficient under right censoring setting. The results have shown that our method can provide reasonable estimates of regression coefficients when censoring rate is light to moderate. It is easy to implement and straightforward to interpret. Two simulation

studies with three settings show that proposed estimators are robust to certain choices of cumulative incidence functions. Extensive simulations also show that for convergence purpose main event should be at least 25% of total sample size or 100 whichever is greater. For baseline subdistribution, our estimator is not optimal. The performance of baseline hazard estimator for usual interval-censored data is not reported in Heller's work. Although it is worthwhile to consider an alternative method to improve the estimation performance, regression coefficient $\boldsymbol{\beta}_1$ is of our primary interest in the presence of competing risks.

There are several issues that can be addressed in the future work. In this article, we only considered two competing risks. This framework can be generalized to more than two competing risks. We can also consider the case where patients may miss several visits during the study and then return. Effects of time-varying covariates could be our next interest. Furthermore, doubly robust estimating equations can be used to provide unbiased estimators even under model misspecification. The properties of the estimators could also be established.

BIBLIOGRAPHY

T. Cai and R.A. Betensky. Hazard regression for interval-censored data with penalized spline. *Biometrics*, 59(3):570–579, 2003.

S.C. Cheng, L.J. Wei, and Z. Ying. Analysis of transformation models with censored data. *Biometrika*, 82(4):835–845, 1995.

D.R. Cox. Regression models and life-tables. *Journal of the Royal Statistical Society. Series B*, 34(2):187–220, 1972.

D.R. Cox. Partial likelihood. *Biometrika*, 62(2):269–276, 1975.

I.D. Diamond, J.W. McDonald, and I.H. Shah. Proportional hazards models for current status data: application to the study of differentials in age at weaning in pakistan. *Demography*, 23:607–620, 1986.

B. Efron and R. Tibshirani. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, 1(1):54–75, 1986.

J.P. Fine and R.J. Gray. A proportional hazards model for the subdistribution of a competing risk. *Journal of American Statistical Association*, 94(446):496–509, 1999.

D.M. Finkelstein. A proportional hazards model for interval-censored failure time data. *Biometrics*, 42(4):845–854, 1986.

H. Frydman and J. Liu. Nonparametric estimation of the cumulative intensities in an interval censored competing risks model. *Lifetime Data Analysis*, 19:79–99, 2013.

B. Gray. Package 'cmprsk' - cran. *cran.r-project.org/web/packages/cmprsk*, 2014.

R.J. Gray. A class of k-sample tests for comparing the cumulative incidence of a competing risk. *The Annals of Stat.*, 16:1141–1154, 1988.

P. Groeneboom, M.H. Maathuis, and J.A. Wellner. Current status data with competing risks: Consistency and rates of convergence of the mle. *The Annals of Statistics*, 36(3): 1031–1063, 2008.

G. Heller. Proportional hazards regression with interval censored data using an inverse probability weight. *Lifetime Data Anal.*, 17:373–385, 2011.

D.G. Horvitz and D.J. Thompson. A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47(260):663–685, 1952.

C. Hsu, J.M. Taylor, S. Murray, and D. Commenges. Multiple imputation for interval censored data with auxiliary variables. *Statistics in Medicine*, 26:769–781, 2007.

J. Huang. Efficient estimation for the proportional hazards model with interval censoring. *The Annals of Statistics*, 24(2):540–568, 1996.

J. Huang and J.A. Wellner. Asymptotic normality of the npmle of linear functionals for interval censored data, case 1. *Statistica Neerlandica*, 49:153–163, 1995.

M. Hudgens, G. Satten, and I.M. Longini. Nonparametric maximum likelihood estimation for competing risks survival data subject to interval censoring and truncation. *Biometrics*, 57:74–80, 2001.

M.G. Hudgens, C. Li, and J. Fine. Parametric estimation of cumulative incidence functions for interval censored competing risks data. *Biometrics*, 70:1–9, 2014.

J.G. Ibrahim, S.R. Lipsitz, and N. Horton. Using auxiliary data for parameter estimation with non-ignorably missing outcomes. *Journal of the Royal Statistical Society. Series C*, 50(3):361–373, 2001.

J. Jeong and J.P. Fine. Direct parametric inference for the cumulative incidence function. *Journal of the Royal Statistical Society, Series C*, 55:187–200, 2006.

N.P. Jewell, M. van der Laan, and T. Henneman. Nonparametric estimation from current status data with competing risks. *Biometrika*, 90(1):183–197, 2003.

C. Kooperberg and D.B. Clarkson. Hazard regression with interval-censored data. *Biometrics*, 53:1485–1494, 2000.

J. Lawless. Statistical models and methods for lifetime data. 2003.

R.J. Little and D.B. Rubin. Statistical analysis with missing data. second edition. *Wiley Series in Probability and Statistics*, 2002.

Marloes H. Maathuis. Nonparametric estimation for current status data with competing risks. *Dissertation*, 2006.

M.H. Maathuis. Reduction algorithm for the npmle for the distribution function of bivariate interval censored data. *Journal of Computational and Graphical Statistics*, 14(2):352–362, 2005.

M.H. Maathuis and M.G. Hudgens. Nonparametric inference for competing risks current status data with continuous, discrete or grouped observation times. *Biometrika*, 98(2): 325–340, 2011.

W. Pan. Extending the iterative convex minorant algorithm to the cox model for interval-censoreddata. *Journal of Computational and Graphical Statistics*, 8(1):109–120, 1999.

W. Pan. A multiple imputation approach to cox regression with interval-censored data. *Biometrics*, 56:199–203, 2000.

M.S. Pepe. Inference for events with dependent risks in multiple endpoint studies. *Journal of the American Statistical Association*, 86:770–778, 1991.

R. Peto. Experimental survival curves for interval-censored data. *Journal of the Royal Statistical Society. Series C*, 22(1):86–91, 1973.

M. Pintilie. Competing risks: A practical perspective. *John Wiley & Sons, New York*, 2007.

R. L. Prentice, J. D. Kalbfleisch, A. V. Peterson, N. Flournoy, V. T. Farewell, and N. E. Breslow. The analysis of failure times in the presence of competing risks. *Biometrics*, 34: 541–554, 1978.

J.M. Robins and A. Rotnitzky. Recovery of information and adjustment for dependent censoring using surrogate markers. *AIDS Epidemiology*, pages 297–331, 1992.

T.H. Scheike and M. Zhang. Analyzing competing risk data using the r timereg package. *Journal of Statistical Software*, 38(2):1–15, 2011.

T.H. Scheike, M.J. Zhang, and T.A. Gerds. Predicting cumulative incidence probability by direct binomial regression. *Biometrika*, 95:205–220, 2008.

B.W. Turnbull. The empirical distribution function with arbitrarily grouped, censored and truncated data. *Journal of the Royal Statistical Society. Series B*, 38(3):290–295, 1976.

M.J. Van Der Laan and A. Hubbard. Estimation with interval censored data and covariates. *Lifetime Data Analysis*, 3:77–91, 1997.

Z. Zhang. Linear transformation models for interval-censored data: prediction of survival probability and model checking. *Statistical Model*, 9:321–343, 2009.

Z. Zhang, L. Sun, X. Zhao, and J. Sun. Regression analysis of in terval-censored failure time data with linear transformation models. *The Canadian Journal of Statistics*, 33(1):61–70, 2005.