12-17-2022

Regression Modeling of Complex Survival Data based on Pseudo-Observations

Rong Rong
rrong@smu.edu

REGRESSION MODELING OF COMPLEX SURVIVAL DATA BASED ON PSEUDO-OBSERVATIONS

Approved by:

Dr. Hong Zhu
Associate Professor in Peter O'Donnell Jr. School of Public Health, UTSW

Dr. Daniel F. Heitjan
Professor in Department of Statistical Science, SMU & Peter O'Donnell Jr. School of Public Health, UTSW

Dr. S. Lynne Stokes
Professor in Department of Statistical Science, SMU

Dr. Sandi L. Pruitt
Associate Professor in Peter O'Donnell Jr. School of Public Health, UTSW

Dr. Chul Moon
Assistant Professor in Department of Statistical Science, SMU

REGRESSION MODELING OF COMPLEX SURVIVAL DATA BASED ON PSEUDO-OBSERVATIONS

A Dissertation Presented to the Graduate Faculty of the

Dedman College

Southern Methodist University

in

Partial Fulfillment of the Requirements

for the degree of

Doctor of Philosophy

with a

Major in Biostatistics

by

Rong Rong

B.S., Statistics and Actuarial Science, University of Illinois at Urbana-Champaign
M.A., Statistics, Bowling Green State University

December 17, 2022

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor, Dr. Hong Zhu, for her guidance and help. Pursuing a doctoral degree is tough, but Dr. Zhu has always been there, torching my research path in true darkness. I am beyond fortunate to have Dr. Zhu as my advisor, without whom I could not have gotten this far.

I am also deeply grateful to Dr. Daniel Heitjan for bringing me to the program and teaching me all the valuable Biostatistics courses. I admire his knowledge and wisdom, and I am thankful for all he has done for the students in our department.

Dr. Lynne Stokes is my role model to look up to. I would like to offer my special thanks to her. The first-year methods course with Dr. Stokes has landed me a great foundation in statistics, and her sampling course inspired me to further work on research using weighting method to account for biased sampling scheme.

There is no word that can express my appreciation to Dr. Sandi Pruitt for her huge support and for providing me opportunities to work on various projects that sharpened my skills to apply statistics in different disciplines. During the journey of working with Dr. Pruitt, I acquired knowledge, gained experience, and have strengthened my desire to work as a biostatistician to make positive contributions within the public health communities, and, ultimately, enhance the lives of others.

Additionally, I would like to extend my sincere thanks to Dr. Chul Moon for his time and insightful comments to improve this dissertation.

I would also like to thank my great parents and darling husband for their lots of love. This was indeed a family endeavor, and together, we made it!

Rong, Rong
B.S., Statistics and Actuarial Science, University of Illinois at Urbana-Champaign
M.A., Statistics, Bowling Green State University

Regression Modeling of Complex Survival Data based on Pseudo-observations

Advisor: Dr. Hong Zhu
Doctor of Philosophy degree conferred December 17, 2022
Dissertation completed Oct 21, 2022

The restricted mean survival time (RMST) is a clinically meaningful summary measure in studies with survival outcomes. Statistical methods have been developed for regression analysis of RMST to investigate impacts of covariates on RMST, which is a useful alternative to the Cox regression analysis. However, existing methods for regression modeling of RMST are not applicable to left-truncated right-censored data that arise frequently in prevalent cohort studies, for which the sampling bias due to left truncation and informative censoring induced by the prevalent sampling scheme must be properly addressed. Meanwhile, statistical methods have been developed for regression modeling of the cumulative incidence function for left-truncated right-censored competing risks data. Nevertheless, existing methods typically involve complicated weighted estimating equations or nonparametric conditional likelihood function and often require a restrictive assumption that censoring and/or truncation times are independent of failure time. Andersen et al. introduced an approach of using pseudo observations (POs) in regression analysis of right-censored data [4, 5]. In this dissertation, we develop statistical methods for regression modeling of complex survival data based on POs.

In Chapter 1, we propose to directly model RMST as a function of baseline covariates based on POs for left-truncated right-censored data under general censoring mechanisms. We adjust for the potential covariate-dependent censoring or dependent censoring by the

inverse probability of censoring weighting method. We establish large sample properties of the proposed estimators and assess their finite sample performances by simulation studies under various scenarios. We apply the proposed methods to a prevalent cohort of women diagnosed with stage IV breast cancer identified from Surveillance, Epidemiology, and End Results-Medicare linked database.

In Chapter 2, we extend the PO approach to left-truncated right-censored competing risks data and propose to directly model the cumulative incidence as a function of baseline covariates based on POs, under general truncation and censoring mechanisms. We adjust for potential covariate-dependent truncation and/or covariate-dependent censoring by incorporating covariate-adjusted weights into the inverse probability weighted estimator of the cumulative incidence function. We derive large sample properties of the proposed estimators under reasonable model assumptions and regularity conditions and assess their finite sample performances by simulation studies under various scenarios. We apply the proposed methods to a cohort study on HIV disease progression and a cohort study on pregnancy exposed to coumarin derivatives for illustration.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

I dedicate this dissertation to my parents

CHAPTER 1

REGRESSION MODELING OF RESTRICTED MEAN SURVIVAL TIME

1.1. Introduction

The restricted mean survival time (RMST) is a clinically relevant summary measure in studies with survival outcomes. Unlike the uncensored data, the mean survival time may not be estimable due to censoring. As an alternative measure to the mean survival time or median survival time, the RMST defined as the expected survival time up to a fixed time point τ has been suggested [4, 9, 65]: $\mu(\tau) = \mathrm{E}(T \wedge \tau) = \int_0^\tau S(t)\,\mathrm{d}t$, which is the area under the survival curve over a specified time interval $[0, \tau]$. The RMST can be consistently estimated even when the largest observed time is censored as long as τ is no larger than the largest failure time. The difference or ratio of RMST characterizes the absolute magnitude of risk or benefit in survival and is a useful alternative measure to the hazard ratio from the Cox regression analysis. Although the Cox proportional hazards model is commonly used for exploring the relationship between survival and covariates, the validity of the proportional hazards assumption is often questionable and can be hard to be checked analytically for certain types of survival data, such as left-truncated right-censored data. In contrast, RMST is an easily interpretable measure of average survival time over a fixed followed-up time period and does not have any assumption requirement. Therefore, analysis based on RMST is more desirable in clinical settings, especially when the proportional hazards assumption is violated. Existing methods for estimating RMST include the indirect and direct estimations. The indirect methods estimate RMST through the Cox proportional hazards model [9, 28, 63, 67]. Such indirect RMST

estimation is inconvenient and requires the proportional hazards assumption to some extent. Hence, directly modeling RMST itself is more appealing. For right-censored data, Andersen et al. [4] proposed a regression analysis of RMST given baseline covariates using pseudo-observations (POs). Tian et al. [53] modeled the relationship between RMST and baseline covariates through a link function under covariate-independent censoring. Wang and Schaubel [58] developed generalized estimating equation methods to model RMST as a function of baseline covariates under general censoring mechanisms.

Left-truncated right-censored data are frequently encountered in prevalent cohort studies, in which diseased patients who have not yet experienced the disease-related failure event (e.g., death) are sampled and prospectively followed for the subsequent failure event [57]. One motivational example of such data is from a prevalent cohort of late-stage breast cancer patients identified from Surveillance, Epidemiology, and End Results (SEER)-Medicare linked database. The study cohort consists of patients diagnosed with Stage IV breast cancer before the sampling time and are still alive at the sampling time, and the goal is to investigate the impact of covariates on RMST among patients with Stage IV breast cancer. In addition to right censoring, survival data from a prevalent cohort are subject to left truncation because patients who died before the sampling time are not included in the study cohort. Statistical methods must account for the sampling bias due to left truncation and informative censoring induced by the prevalent sampling scheme. Although much research has been conducted into both regression analysis of left-truncated right-censored data [2, 24, 32, 66] and direct regression analysis of RMST for right-censored data [4, 53, 58], relatively little work is available on direct regression modeling of RMST for left-truncated right-censored data. To our knowledge, only one paper by Lee et al. [33] studied direct regression analysis of RMST for length-biased right-censored data, a special type of left-truncated right-censored data that assumes a constant disease incidence rate. That paper was mostly concerned with covariate-independent censoring and constructed unbiased estimating equations to obtain consistent estimators of covariate effects on RMST. In observational cohort studies, covariate-dependent censoring or dependent

censoring occurs frequently when the censoring time and failure time are correlated through common baseline covariates or possibly time-varying covariates, respectively. For example, in AIDS studies, patients who have low CD4 counts (an indicator of immune function in patients living with HIV) are more likely to drop out of the study, resulting in overestimation of the overall survival if covariate-independent censoring is assumed. Methods for regression modeling of RMST need to take into account covariate-dependent censoring and dependent censoring. The inverse probability of censoring weighting (IPCW) method, discussed by Robins and Rotnitzky [49], Robins [47], and Robins and Finkelstein [48] among others, can be used to correct for the bias due to covariate-dependent censoring and dependent censoring. For right-censored data, Xiang and Murray [61] developed a model for the log of RMST given baseline covariates by using POs that account for dependent censoring. In this paper, we will consider general censoring mechanisms in regression modeling of RMST for left truncated right-censored data.

POs are jackknife estimates that represent the contribution of each subject to the estimator of the parameter of interest [5]. POs are usually used to study the bias and precision of the parameter estimator. Andersen et al. introduced an approach of using POs in the regression analysis of right-censored data [4, 5]. The PO approach has also been used in regression modeling of competing risks data [30, 31] and the Cox regression analysis of left-truncated right-censored data [20]. In this paper, we propose to extend the PO approach in Andersen et al. [4] to left truncated right-censored data and directly model RMST as a function of baseline covariates based on POs under general censoring mechanisms. The PO approach has the advantage of handling complex issues related to left truncation and right censoring in the first step of generating POs and then using POs as responses in a generalized linear model for uncensored data. The remainder of this paper is organized as follows. In Section 1.2, we introduce the left-truncated right-censored data structure with notations and describe the regression model of RMST. In Section 1.3, we first present the proposed method for regression modeling of RMST given covariates using POs under covariate-independent censoring. Then, we relax the

covariate-independent censoring assumption to incorporate covariate-dependent censoring and dependent censoring. We investigate the finite sample performances of proposed estimators by simulation studies under various scenarios in Section 2.3. As an illustration, we apply the proposed methods to a prevalent cohort of women diagnosed with late-stage breast cancer identified from SEER-Medicare linked database in Section 2.5. We provide concluding remarks in Section 2.6. Technical details can be found in the Appendix.

1.2. Data, Notations, and Regression Model

In a prevalent cohort study, patients with a certain disease are sampled or enrolled and then followed prospectively till the occurrence of a failure event or censoring. We are interested in studying the underlying relationship between the RMST and baseline covariates through regression modeling based on the PO approach. Let $\tilde{T}$ be the time from the disease onset to the failure event (unbiased failure time). Let $\tilde{A}$ be the time from disease onset to study enrollment. Under the prevalent sampling, the failure time $\tilde{T}$ is not randomly sampled from the target population because patients who experienced the failure event prior to the enrollment are not included. Hence, patients in the prevalent cohort all have $\tilde{A} < \tilde{T}$. Let T be the sampled failure time from the disease onset (biased failure time) and A be the corresponding truncation time. For the sampled patients, let V be the time from enrollment to the failure event, and we have $T = A + V$, where V is subject to right censoring. Let C be the residual censoring time from enrollment. Let $Y = \min(T, A + C)$ be the follow-up time till failure event or censoring and $\delta = I(V < C)$ be the failure indicator. Let $\boldsymbol{X}$ be a $p \times 1$ vector of baseline covariates. The observed data are $(Y_i, A_i, \delta_i, \boldsymbol{X}_i), i = 1, 2, ..., n$. Let τ be a pre-specified time point of interest from the disease onset and $\tilde{T}_\tau = \min(\tilde{T}, \tau)$ be the restricted survival time for a fixed τ. The RMST is then defined as $\mu(\tau) = E[\tilde{T}_\tau]$. Throughout this paper, we use these notations and assume that $\tilde{T}$ and $\tilde{A}$ are conditionally independent given covariates $\boldsymbol{X}$, which is a standard assumption for left-truncated right-censored data. Note that the biased failure time T is correlated

with censoring time from the disease onset $A + C$ through a common variable A, which is referred to as informative censoring induced from prevalent sampling. Our goal is to directly model the relationship between RMST and covariates through a generalized linear model:

$$g[\mu(\tau \mid \boldsymbol{Z})] = \boldsymbol{Z}^\top \boldsymbol{\beta}_\tau, \tag{1.1}$$

where $g(\cdot)$ is a differentiable, strictly increasing link function, $\boldsymbol{Z} = (1, \boldsymbol{X}^\top)^\top$, and $\boldsymbol{\beta}_\tau$ is a $(p+1) \times 1$ coefficient vector specific to τ. Examples of common link functions include the linear link $g(m) = m$ and log link $g(m) = log(m)$. The linear link function leads to a simple linear regression of RMST, where the covariate effects can be interpreted as differences in the RMST. However, since the linear model may produce negative responses that are not meaningful for RMST, the log-linear model under the log link function would be a natural alternative, where the covariate effects can be interpreted as ratios in the RMST [22].

1.3. Regression Modeling of RMST based on Pseudo-Observations

POs for regression analysis of RMST can be defined by using a consistent estimator $\hat{\mu}(\tau)$ for the parameter of interest $\mu(\tau)$ [4]. For conventional right-censored data, a consistent estimator of $\mu(\tau)$ is $\hat{\mu}(\tau) = \hat{E}[\tilde{T}_\tau] = \int_0^\tau \hat{S}(t)\mathrm{d}t$, where $\hat{S}(t)$ is the Kaplan-Meier estimator for the survival function $P(\tilde{T} > t)$. Then, the ith PO is computed as

$$\hat{\mu}_i(\tau) = n\hat{\mu}(\tau) - (n-1)\hat{\mu}^{-i}(\tau), \tag{1.2}$$

where $\hat{\mu}^{-i}(\tau)$ is the jackknife leave-one-out estimator for $\mu(\tau)$ based on data leaving out subject i. The rationale behind the PO approach is that any estimator of $\mu(\tau) = E[\tilde{T}_\tau]$ is also implicitly an estimator of $E_{\boldsymbol{Z}}\left[E(\tilde{T}_\tau \mid \boldsymbol{Z})\right]$, where the inner expectation is the quantity of interest in the regression model (1.1) and the outermost expectation is taken with respect to the empirical distribution of $\boldsymbol{Z}$. Let $\tilde{\mu}(\tau) = \frac{1}{n}\sum_{i=1}^n E(\tilde{T}_\tau \mid \boldsymbol{Z}_i)$ be a consistent

estimator of $E_{\boldsymbol{Z}}\left[E(\tilde{T}_\tau \mid \boldsymbol{Z})\right]$. Then, the corresponding ith PO is $n\tilde{\mu}(\tau) - (n-1)\tilde{\mu}^{-1}(\tau) = n\left[\frac{1}{n}\sum_{i=1}^{n} E(\tilde{T}_\tau \mid \boldsymbol{Z}_i)\right] - (n-1)\left[\frac{1}{n-1}\sum_{j=1,j\neq i}^{n} E(\tilde{T}_\tau \mid \boldsymbol{Z}_j)\right] = E(\tilde{T}_\tau \mid \boldsymbol{Z}_i)$, which is the quantity of interest in regression modeling. As described in Anderson et al. [4], since both $\hat{\mu}(\tau)$ and $\tilde{\mu}(\tau)$ are consistent estimators for $\mu(\tau)$ and they will be approximately equal when n is large, $\hat{\mu}(\tau)$ that is estimable from censored survival data can be used to replace $\tilde{\mu}(\tau)$, and formula (1.2) can be used to generate POs that have the same conditional mean of interest for regression modeling as the original individual level data. In other words, models based on POs generated by (1.2) will have regression parameters similar to a model fit to the values of $\tilde{T}_\tau$ if all these values were uncensored. Thus, the POs, $\mathcal{PO} = \{\hat{\mu}_1(\tau), \hat{\mu}_2(\tau), ..., \hat{\mu}_n(\tau)\}$ obtained from (1.2) can be used as responses of the regression model (2.1) to estimate β_τ under a generalized estimating equation framework [4]. Graw et al. [21] and Overgaard et al. [40, 41] provide the formal theoretical justification of PO approach and asymptotic properties of parameter estimators. We extend the method in Anderson et al. [4] to left-truncated right-censored data where only the biased failure times are observable, and propose a modified PO approach to estimate and analyze RMST.

First, we consider covariate-independent censoring, that is, the residual censoring C is independent of (A, V). The Kaplan-Meier estimator would result in overestimation of the survival function for left-truncated right-censored data [60], and thus, overestimation of the RMST. The survival function $S(t)$ for such data can be consistently estimated by a product-limit estimator $\hat{S}_{PL}(t)$ with risk set $R(t) = \{i : A_i \leq t \leq Y_i\}$ [54], and

$$\hat{S}_{PL}(t) = \prod_{j:t_{(j)}\leq t}\left[1 - \frac{d_j}{r_j}\right],$$

where $\{t_{(1)}, ..., t_{(K)}\}$ denotes the set of K distinct ordered failure times from uncensored Y_i in the sample, $d_j = \sum_{i=1}^{n} I\left\{Y_i = t_{(j)}, \delta_i = 1\right\}$ is the number of failures at t_j, and $r_j = \sum_{i=1}^{n} I\left\{A_i < t_{(j)} \leq Y_i\right\}$ is the number of subjects "at risk" right before the jth failure time. The risk set $R(t)$ at any time t consists of subjects who have entered the study and have not failed or been censored by that time. Note that the difference between the Kaplan-Meier

estimator for right-censored data and the product-limit estimator $\hat{S}_{PL}(t)$ is the definition of risk set. For left-truncated right-censored data, $\hat{S}_{PL}(t)$ is similar to the Kaplan-Meier estimator, after replacing the risk set $R_{KM}(t) = \{i : t \leq Y_i\}$ with $R(t) = \{i : A_i \leq t \leq Y_i\}$. The product-limit estimator $\hat{S}_{PL}(t)$ is the nonparametric maximum likelihood estimator of $S(t)$ [56]. A consistent estimator $\hat{\mu}(\tau)$ can be obtained by integrating the product-limit estimator of the survival function over the time interval $[0, \tau]$, $\hat{\mu}(\tau) = \int_0^\tau \hat{S}_{PL}(t)\mathrm{d}t$, and is used to construct POs, $\mathcal{PO} = \{\hat{\mu}_1(\tau), \hat{\mu}_2(\tau), ..., \hat{\mu}_n(\tau)\}$, based on (1.2). The POs are then used as responses in the generalized linear model (2.1) with a suitable link function to estimate regression parameters β_τ and predict $\mu(\tau \mid \boldsymbol{Z}_i)$. The regression coefficients, β_τ, can be estimated by the generalized estimating equations

$$U\left(\beta_\tau\right) = \sum_{i=1}^{n} U_i\left(\beta_\tau\right) = \sum_{i=1}^{n} \left\{ \frac{\partial}{\partial \beta} g^{-1}\left(\boldsymbol{Z}_i^\top \beta_\tau\right) \right\} \mathcal{V}_i^{-1} \left\{ \hat{\mu}_i(\tau) - g^{-1}\left(\boldsymbol{Z}_i^\top \beta_\tau\right) \right\} = 0, \tag{1.3}$$

where $\mathcal{V}_i$ is a working variance of $\hat{\mu}_i(\tau)$ [34, 62] with a simple choice of $\mathcal{V}_i = 1$. Anderson et al. [6] showed that the estimates obtained from generalized estimating equations using POs are consistent for right-censored data. Since the nonparametric estimator $\hat{\mu}(\tau)$ based on the product-limit estimator is consistent for left-truncated right-censored data, we can also use (1.3) to obtain consistent estimates of the regression coefficients in model (2.1). Let $\hat{\beta}_\tau$ be the solution to (B.1) and $\beta_{\tau 0}$ be the true value of β_τ. The asymptotic properties of $\hat{\beta}_\tau$ are summarized in Theorem 1 with the proof and regularity conditions provided in the Appendix.

Theorem 1. *Under some regularity conditions, $\hat{\beta}_\tau$ is consistent to $\beta_{\tau 0}$, and $\sqrt{n}\left(\hat{\beta}_\tau - \beta_{\tau 0}\right)$ is asymptotically normal with mean zero and a covariance matrix that can be consistently estimated using a standard 'sandwich' estimator, which takes the form*

$$\hat{\Sigma} = \mathcal{I}\left(\hat{\beta}_\tau\right)^{-1} \hat{\text{var}}\left\{U\left(\beta_\tau\right)\right\} \mathcal{I}\left(\hat{\beta}_\tau\right)^{-1},$$

where

$$\mathcal{I}\left(\hat{\beta}_\tau\right) = \sum_i \left\{\frac{\partial g^{-1}\left(\boldsymbol{Z}_i^\top \hat{\beta}_\tau\right)}{\partial \hat{\beta}_\tau}\right\}^\top \mathcal{V}_i^{-1} \left\{\frac{\partial g^{-1}\left(\boldsymbol{Z}_i^\top \hat{\beta}_\tau\right)}{\partial \hat{\beta}_\tau}\right\},$$

$$\hat{\text{var}}\left\{U\left(\beta_\tau\right)\right\} = \sum_i U_i\left(\hat{\beta}_\tau\right) U_i\left(\hat{\beta}_\tau\right)^{\mathrm{T}}.$$

Second, covariate-independent censoring may be implausible in practice and covariate-dependent censoring often occurs in observational cohort studies, where the censoring time and failure time are only conditionally independent given baseline covariates. Furthermore, the censoring time may be correlated with the failure time through a mutual association with possibly time-varying covariates, which is referred to as dependent censoring. We relax the covariate-independent censoring assumption and model RMST under more general censoring mechanisms. The product-limit estimator is a consistent estimator of the survival function for left-truncated right-censored data under covariate-independent censoring. It is crucial to account for covariate-dependent censoring or dependent censoring to consistently estimate the survival function. The IPCW approach can be used to adjust for covariate-dependent censoring or dependent censoring by assigning extra weight to subjects who are not censored or who are observed [47–49, 59]. Each subject is assigned a weight inversely proportional to the estimated probability of remaining uncensored until time t given covariates. The Cox proportional hazards model for censoring is frequently used to model the relationship between censoring time and covariates and estimate such probability. For simplicity of discussion, we assume that the residual censoring C is conditionally independent of (A, V), given baseline covariates $\boldsymbol{X}$, although the

covariate-dependent censoring assumption can be easily relaxed to dependent censoring by incorporating time-varying covariates $M(t)$ into the Cox model [48]. The Cox model for the residual censoring time C given covariates $\boldsymbol{X}$ is:

$$\lambda_C\left(t|\boldsymbol{X}\right) = \lambda_{C_0}(t)\exp\left\{\boldsymbol{\alpha}^\top \boldsymbol{X}\right\},$$

where λ_{C_0} is the baseline hazard function for censoring and $\boldsymbol{\alpha}$ is the vector of model parameters. Let $\hat{\boldsymbol{\alpha}}$ be the partial likelihood estimate of $\boldsymbol{\alpha}$ and $\rho_{(j)}$'s denote the distinct ordered residual censoring times. A consistent estimator of the conditional probability that subject i remains uncensored through time t given $\boldsymbol{X}$ is provided by:

$$\hat{K}_i(t) = \prod_{\left\{j:\rho_{(j)}<t, \delta_j=0\right\}} \left[1 - \hat{\lambda}_{C_0}\left(\rho_{(j)}\right)\exp\left\{\hat{\boldsymbol{\alpha}}^\top \boldsymbol{X}_i\right\}\right],$$

where

$$\hat{\lambda}_{C_0}\left(\rho_{(j)}\right) = \frac{(1-\delta_j)}{\sum_{i=1}^n \exp\left\{\hat{\boldsymbol{\alpha}}^\top \boldsymbol{X}_i\right\} I\left(\rho_{(j)} \leq Y_i\right)}$$

is the Cox estimator of the baseline hazard function for censoring, λ_{C_0}, with $I\left(\rho_{(j)} \leq Y_i\right)$ being the at-risk indicator and δ being the failure indicator [48]. The subject-specific IPCW weight is $\hat{W}_i(t) = 1/\hat{K}_i(t)$. The contribution of subject i at risk at any time $t_{(j)}$ is weighted by the subject-specific weight $\hat{W}_i(t_{(j)})$. The IPCW version of the product-limit estimator for $S(t)$ for left-truncated right-censored data is then given by

$$\hat{S}_{IPCW}(t) = \begin{cases} 1 & \text{if } t < t_1 \\ \prod_{j:t_{(j)}\leq t}\left[1 - \frac{\sum_{i=1}^n I\{Y_i=t_{(j)}, \delta_i=1\}\hat{W}_i\left(t_{(j)}\right)}{\sum_{i=1}^n I\{(A_i,Y_i)\in R(t_{(j)})\}\hat{W}_i\left(t_{(j)}\right)}\right] & \text{if } t \geq t_1. \end{cases}$$

In the presence of covariate-dependent censoring, we can use $\hat{S}_{IPCW}(t)$ to consistently estimate the survival function $S(t)$ and further to obtain estimated RMST and corresponding POs. Then, the POs can be used in the generalized linear model (1.1) to estimate β_τ,

similar to the case under covariate-independent censoring. For right-censored data, Robins and Finkelstein [48] provided proof of the consistency of $\hat{S}_{IPCW}(t)$ for $S(t)$ under dependent censoring. The consistency of the IPCW estimator also holds for the left-truncated right-censored data when the risk set is properly adjusted. Therefore, the resulting estimator of β_τ is consistent and asymptotically normal, which can be proved similarly to Theorem 2 under covariate-independent censoring.

1.4. Simulations

We conduct a series of simulations to assess the performance of the proposed methods for left-truncated right-censored data, under various scenarios. The failure time data are generated both under proportional hazards and under non-proportional hazards. For each scenario, the simulation was repeated 1000 times with a sample size of $n = 350$ or 500.

1.4.1. Simulations under Left Truncation and Covariate-Independent Censoring

First, we evaluate the performance of the proposed method under proportional hazards and covariate-independent censoring. We randomly assign each subject to two groups, A and B, with equal probability. Group A is treated as the reference. The assumed model for RMST is $\mu_\tau(x_1) = \mathrm{E}\left[\tilde{T}_\tau \mid X_1 = x_1\right] = \beta_{\tau 0} + \beta_{\tau 1} x_1$ with the linear link function. The failure time $\tilde{T}$ follows the Cox proportional hazards model and is generated from a distribution with hazard function $\lambda(t|X_1 = x_1) = exp(\gamma x_1)$, where $\gamma = 0.5$. The covariate X_1 is binary and equals to 1 for subjects in group B and equals to 0 for subjects in group A. The residual censoring time C is generated from an exponential distribution with parameter λ_C, allowing for various levels of censoring (i.e., censoring rates of 30% and 45%). The truncation time $\tilde{A}$ follows a Weibull distribution with scale parameter λ_l and shape parameter α_l, where λ_l is such that the truncation rate is 30% when $\alpha_l = 1$. Regression parameters in the RMST model are estimated at two values of $\tau = (0.69, 1.39)$, which are approximately

the 60th and 80th percentiles of the failure time $\tilde{T}$, respectively. The upper panel of Table 1.1 summarizes the simulation results. Next, we investigate how the proposed method performs under non-proportional hazards and covariate-independent censoring. Let $\boldsymbol{Z} = (1, X_1)^\top$. The failure time is generated from a distribution with hazard function $\lambda(t \mid \boldsymbol{Z} = \boldsymbol{z}) = \exp\{-(\boldsymbol{z}^\top \boldsymbol{\gamma}) + \boldsymbol{z}^\top \boldsymbol{\zeta} \log(8t)\}$, where $\boldsymbol{\gamma} = (0.5, 1)^\top$ and $\boldsymbol{\zeta} = (1, -0.3)^\top$. Apparently, the proportional hazards assumption is not valid because the hazard ratio of the two groups varies over time. The rest of the data generating and estimation procedures are similar to the case under the proportional hazards. The values of τ are set to be 0.69 and 1.39, which are approximately the 50th and 80th percentiles of the failure time, respectively. The lower panel of Table 1.1 summarizes the simulation results. Moreover, we carry out simulations with the log link function and the details of simulations are described in the Appendix with results summarized in Supplementary Table A.1.

From Table 1.1, the estimation procedure performs well with generally very small relative biases and the estimated model-based standard errors (SEs) computed as the GEE sandwich estimator being close to the empirical standard deviations (SDs) in all scenarios. Increasing the censoring rate from 30% to 45% does not affect the relative bias much but tends to increase SEs and SDs slightly. Additionally, the estimated SEs and SDs decrease as the sample size increases. The coverage probabilities are generally close to the nominal level of 95%, with some slight undercoverage in estimating the intercept, β_0. Nevertheless, the estimation of the regression coefficient, which is often the main focus, is reliable. It is noted that relative biases are larger in the non-proportional hazards scenario than those in the proportional hazards scenario, however, SDs and SEs under non-proportional hazards are considerably smaller than those under proportional hazards. Overall, there is no obvious directional trend when comparing the mean squared errors under non-proportional hazards and under proportional hazards, and coverage probabilities under these two scenarios are comparable. The results in Table A.1 also suggest a good performance of the proposed method under various scenarios with the log link function.

In both Table 1.1 and Table A.1, the relative bias of parameter estimate increases as τ increases. Such effect is more pronounced for the estimation of intercept and the bias can be relatively large in some cases. For example, the relative bias for β_0 is as large as 0.069 at τ = 1.39 in the lower panel of Table 1.1. This is possibly caused by the presence of the extremely small negative-valued POs that behave as outliers in the subsequent GEE analysis. Similar problems were observed in the simulations for regression analysis of RMST with right-censored data using POs in Anderson et al. [4] where the bias increases considerably as τ increases, especially under a high censoring rate and in the simulations for estimating regression parameters in the Cox model with left-truncated right-censored data using POs in Grand et al. [20] The bias observed at larger τ may be due to the low precision of the product-limit estimator at the tail part of the survival function and is enhanced for the RMST estimator by the integration. Another important explanation is that under left truncation, there can be very few subjects at risk at the beginning so that the information in the data is too sparse. Grand et al. [20] suggested select a set of time points where the information is less sparse to improve the estimation procedure. Here, we use a "conditional survival function" approach to address this issue. This approach is to only include subjects whose failure times are greater than k, where k is the failure time corresponding to the cutoff value of the POs that separates out the outliers. Therefore, this approach is essentially modeling the RMST based on the conditional survival function given surviving beyond k, i.e., $E\left(\tilde{T}_\tau \mid \tilde{T}_\tau > k\right) = \int_k^\tau P\left(\tilde{T} > t \mid \tilde{T} > k\right) \mathrm{d}t$. The extremely small negative POs are identified by using a cutoff value obtained by subtracting 2 times interquartile range from the first quartile. Supplementary Table A.3 in the Appendix includes the estimation results at $\tau = 1.39$ by using the "conditional survival function" approach with linear and log link functions. It shows that the estimation performance is improved remarkably comparing to those in the lower panel of Table 1.1 and the lower panel of Table A.1, respectively. The absolute value of relative bias after using the "conditional survival function" approach is no more than 0.012 across different simulation settings. In addition, the coverage probability is closer to the nominal level of 95% after the adjustment,

especially for β_0.

1.4.2. Simulations under Left Truncation and Covariate-Dependent Censoring

We conduct the following two simulation studies to evaluate the proposed methods under covariate-dependent censoring and compare the parameter estimates obtained by using the traditional PO approach and the IPCW-adjusted PO approach. We first generate the survival data under a proportional hazards model. We randomly assign each subject to three groups, A, B, and C, with unequal proportions of 40%, 30%, and 30%. Group A is treated as the reference. We set the two covariates X_1 and X_2 as dummy variables indicating groups B and C, respectively. The assumed model of RMST is $\mu_\tau(\boldsymbol{x}) = \mathrm{E}\left[\tilde{T}_\tau \mid \boldsymbol{X} = \boldsymbol{x}\right] = \beta_{\tau 0} + \beta_{\tau 1} x_1 + \beta_{\tau 2} x_2$. The failure time is generated from a distribution with hazard function $\lambda(t|\boldsymbol{X}) = \exp(\boldsymbol{\gamma}\boldsymbol{X})$, where $\boldsymbol{\gamma} = (0.5, 1)^\top$ and $\boldsymbol{X} = (X_1, X_2)^\top$. The residual censoring time is generated from an exponential distribution with parameter $\lambda_C = \lambda_{C_0}\exp(4X_1 + 5X_2)$, which depends on the two covariates. Varying λ_{C_0} allows for various levels of censoring (i.e., censoring rates of 30% and 45%). The truncation variable follows the same Weibull distribution as in Section 2.3.1, with a truncation rate of 30%. After the data are generated, we compute the IPCW-adjusted POs for the RMST at two values of $\tau = (0.69, 1.39)$, which are approximately the 65th and 85th percentiles of the failure time, respectively. Table 2.5 summarizes the simulation results. Next, we generate the survival data under non-proportional hazards. The assumed model for RMST is $\mu_\tau(\boldsymbol{x}) = \mathrm{E}\left[\tilde{T}_\tau \mid \boldsymbol{X} = \boldsymbol{x}\right] = \beta_{\tau 0} + \beta_{\tau 1} x_1 + \beta_{\tau 2} x_2$. Let $\boldsymbol{Z} = (1, X_1, X_2)^\top$. The failure time is generated from a distribution with hazard function $\lambda\left(t \mid \boldsymbol{Z} = \boldsymbol{z}\right) = \exp\left\{-\left(\boldsymbol{z}^\top\boldsymbol{\gamma}\right) + \boldsymbol{z}^\top\boldsymbol{\zeta}\log(8t)\right\}$, where $\boldsymbol{\gamma} = (0.5, 1, 1)^\top$ and $\boldsymbol{\zeta} = (1, -0.3, 0)^\top$. The residual censoring time is generated from an exponential distribution with parameter $\lambda_C = \lambda_{C_0}\exp\left(X_1 + 2X_2\right)$, where λ_{C_0} is such that the censoring rate is 30% or 45%. The truncation variable remains the same with a truncation rate of 30%. As previous, the values of τ are set to be 0.69 and 1.39, which are approximately the 45th and 80th percentiles of the failure time, respectively. Table

2.1 summarizes the simulation results. Likewise, we conduct simulations with the log link function and under covariate-dependent censoring. The details are provided in the Appendix with results summarized in Supplementary Table B.2.

Tables 1.2 and 1.3 show that the IPCW-adjusted PO approach can substantially reduce the bias after accounting for covariate-dependent censoring, especially when the censoring rate is high or when the RMST is computed at a larger τ. For example, in Table 1.2, under the scenario of $n = 500$, censoring rate of 45% and $\tau = 1.39$, the relative bias is reduced from 0.169 to 0.104 for the estimate of the intercept β_0, from 0.176 to 0.073 for the estimate of β_1, and from 0.206 to 0.132 for the estimate of β_2. Overall, the estimated standard errors (SEs) are close to the empirical standard deviations (SDs) in all scenarios. As the sample size increases from 350 to 500, SEs and SDs decrease. Moreover, the coverage probabilities under the IPCW-adjusted PO approach are closer to the nominal level of 95%, with some undercoverage in the estimation of the intercept, β_0. The estimation of the regression coefficients is generally reliable. Similarly, Table A.2 shows that the IPCW-adjusted PO approach greatly improves the estimation under various scenarios with the log link function. The substantial bias reduction by using the IPCW-adjusted PO approach suggests that when the censoring mechanism is more complicated than covariate-independent censoring (e.g., covariate-dependent censoring), which is often the case in many applications, the proposed method with IPCW adjustment outperforms the unadjusted PO approach. Although the IPCW-adjusted PO approach substantially reduces the bias, the bias is still relatively large at τ = 1.39 in Tables 1.3 and A.2, especially for the intercept β_0, similar to that observed in the lower panels of Tables 1.1 and A.1. This is probably due to the low precision of RMST estimator at the tail and the sparse data information at early event times under left truncation.

1.5. Application

The Surveillance, Epidemiology, and End Results (SEER)-Medicare linked database is a population-based cancer registry that provides data for prevalent cohorts, which include patients who have already been diagnosed with cancers. We identified a prevalent cohort from the SEER-Medicare linked database that consists of patients diagnosed with stage IV breast cancer from 2002 to 2006 and survived beyond 2006 with a last follow-up date of December 31, 2010 [66]. This prevalent cohort included 933 patients with complete information on receptors for either estrogen (ER) or progesterone (PR) in the tumor, receipt of chemotherapy, age at diagnosis, vital status, and death/last contact dates. The truncation time was the time from the breast cancer diagnosis to study enrollment, and the failure time was the overall survival after the breast cancer diagnosis. We apply the proposed method to directly model RMST and investigate the impact of chemotherapy, ER/PR status, and age at diagnosis on RMST among patients with stage IV breast cancer.

Among the 933 patients, 707 (75.8%) experienced failure events and 226 (24.2%) were censored by the end of the study, 465 (49.8%) received chemotherapy, and 791 (84.8%) patients were ER/PR positive. Figure 1.1 presents the survival function estimated by the product-limit estimator for left-truncated right-censored data as well as the corresponding nonparametric RMST estimator by integrating the survival curve at varying values of τ, among patients with and without chemotherapy. Figure 1.2 presents the estimated survival curve and RMST curve by ER/PR status. In summary, these figures show that the receipt of chemotherapy and positive ER/PR status tend to result in a longer RMST. Since there is no formal analytical tool for testing the proportional hazards assumption for left-truncated right-censored data in literature, the validity of the proportional hazards assumption cannot be rigorously checked and the analysis based on direct modeling of RMST is more applicable. Moreover, a preliminary Cox regression analysis of residual censoring time suggests that the residual censoring time is independent of covariates, and thus, the proposed method

that assumes covariate-independent censoring is used in the analysis.

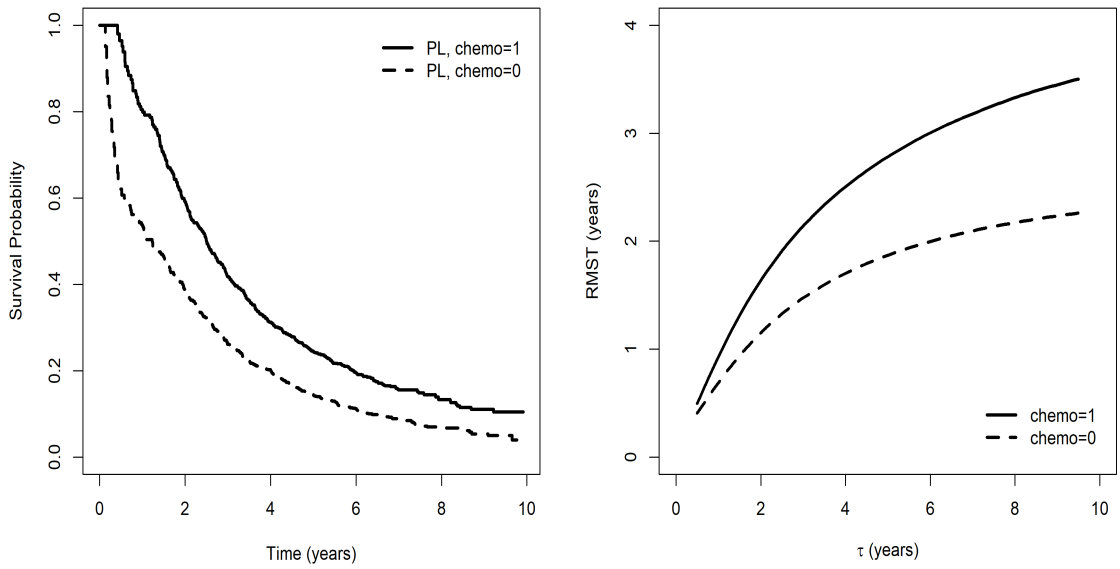

Figure 1.1: Product-limit estimator of survival function (left panel) and the nonparametric RMST estimator (right panel) by receipt of chemotherapy (chemo=1, receiving chemotherapy; chemo=0, not receiving chemotherapy).

The regression model of RMST at $\tau = 2$, 5, and 8 years post-diagnosis are considered, which are reasonable time points for this study of stage IV breast cancer. We use both the linear and log link functions for the regression model and include two binary variables for the receipt of chemotherapy and ER/PR status and one continuous variable for age at diagnosis as covariates. Table 1.4 summaries the regression analysis results. Overall, the covariate effects demonstrate similar trends between the two link functions. In the model with the linear link, the receipt of chemotherapy and positive ER/PR status are significantly associated with a longer average post-diagnosis survival time for all the values of τ. Older age at diagnosis tends to be associated with a shorter survival time and such association becomes significant at a later time point (τ = 8 years). In the model with the log link function, chemotherapy is marginally associated with an increase of the average post-diagnosis survival time, positive ER/PR status is significantly associated with an increase of the

survival time, and older age at diagnosis is significantly associated with a decrease of the survival time during the next 8 years post-diagnosis. Specifically, it is estimated that

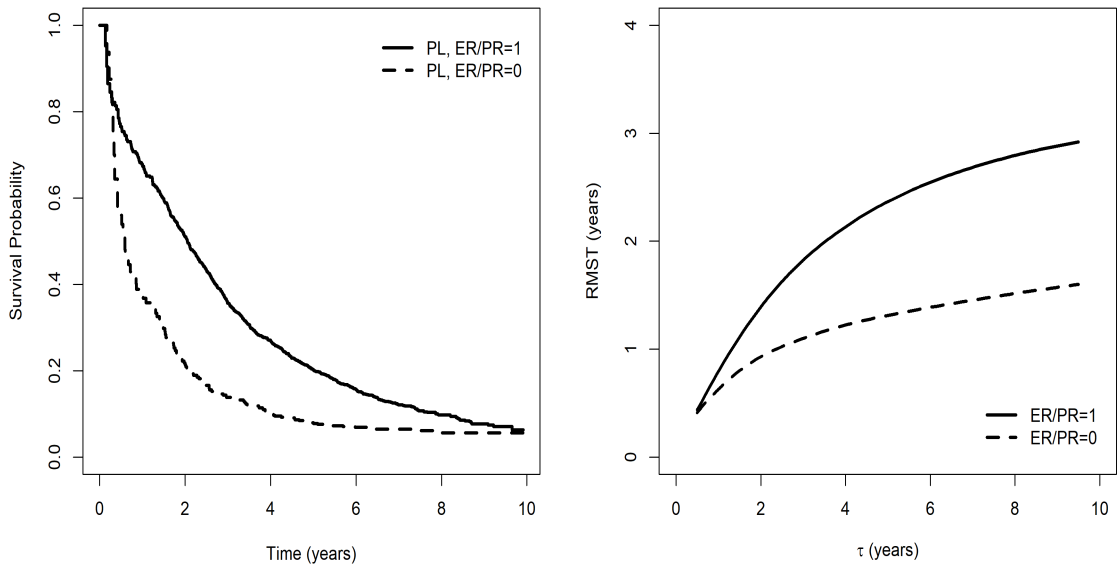

Figure 1.2: Product-limit estimator of survival function (left panel) and the nonparametric RMST estimator (right panel) for patients by ER/PR status (ER/PR=1, positive; ER/PR=0, negative).

the receipt of chemotherapy is associated with an increase of the survival time by 0.92 years (95% CI: 0.22-1.61) on average during the next 5 years post-diagnosis, using the linear link, and chemotherapy is associated with an increase of the survival time by a factor of 1.31 (95% CI: 0.96-1.79), using the log link. The positive ER/PR status is estimated to be associated with an increase of the average survival time by 1.78 years (95% CI: 0.84-2.72) during the next 5 years post-diagnosis using the linear link and is associated with an increase of the survival time by a factor of 2.04 (95% CI: 1.21-3.44), using the log link. During the next 5 years post-diagnosis, it is estimated that every one year increase in the age at diagnosis is associated with a decrease of the survival time by 0.05 years (95% CI: 0.00, 0.10) on average with the linear link and is associated with a decrease of the survival time by a factor of 0.98 (95% CI: 0.96, 1.00) with the log link. Additionally,

we can estimate the average post-diagnosis survival time based on these models. For patient without chemotherapy, having negative ER/PR status, and with an age at diagnosis of 50, the average post-diagnosis survival time out of the next 5 years is estimated to be approximately 1.42 years (17.0 months) and 1.62 years (19.5 months), using the linear link and the log link, respectively. For another patient receiving chemotherapy, having positive ER/PR status, and with the same age at diagnosis, the average post-diagnosis survival time out of the next 5 years is estimated as 4.11 years (49.3 months) and 4.33 years (52.0 months), using the linear link and the log link, respectively.

To compare the estimation results between the two link functions visually, Figure 1.3 presents the estimated RMST curve at $\tau = 5$ years against the age at diagnosis, by using the two link functions. For both link functions, RMST decreases as age at diagnosis increases. It is also observed that the discrepancy of the estimated RMST between the two link functions is overall small. Figure 1.4 presents the estimated RMST curves by the combination of receipt of chemotherapy, ER/PR status, and age at diagnosis, using the nonparametric method, multivariable regression model of RMST with the linear link, multivariable regression model of RMST with the log link, and by integrating the survival curve estimated from the multivariable Cox model. To facilitate the comparison with the nonparametric method, age at diagnosis is dichotomized into a binary covariate (< 70 and ≥ 70) in the multivariable models. Left truncation is adjusted in all the methods. The number of patients without chemotherapy and with negative ER/PR status is very small: only 4 patients with age at diagnosis < 70 and 24 patients with age at diagnosis ≥ 70. The nonparametric method would not be reliable in these cases, and thus, they are not included in Figure 1.4 for the comparison. Figure 1.4 shows that the direct modeling of RMST with the linear link and log link functions gives similar results in general. When the number of patients is relatively large, RMST estimated by the direct regression model is in better agreement with the nonparametric estimate, comparing with RMST estimated by the Cox model. Between the two link functions, although the linear link may be more appealing due to its straightforward interpretation, it does not always lead to estimated RMST values

within an admissible range (0, τ] [58], as shown in Figure 1.3 where negative values of estimated RMST appear as age at diagnosis increases, for patients without chemotherapy and with negative ER/PR status. This suggests that the regression model with the log link function may be a better fit for the observed data in this study.

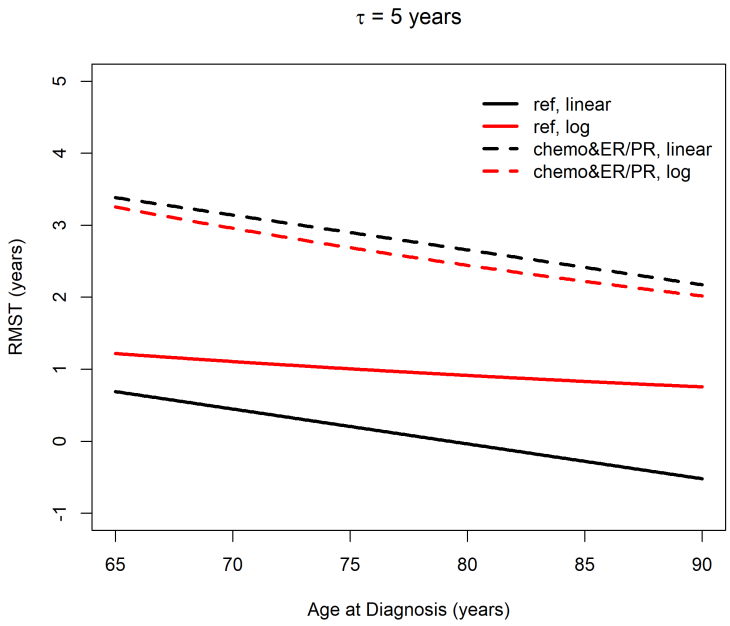

Figure 1.3: Estimated RMST during the next five years post diagnosis ($\tau = 5$ years) against the age at diagnosis using two link functions. 'Ref' represents the reference patients without chemotherapy and negative ER/PR status, and 'chemo&ER/PR' represents patients with chemotherapy and positive ER/PR status.

1.6. Discussion

The RMST is an appealing summary measure for survival data due to its simple and clinically meaningful interpretation, therefore, the analysis of RMST has attracted a growing research interest. However, little work is available on regression analysis of RMST for left-truncated right-censored data. As discussed in Lee et al. [33], the generalization of existing methods based on weighted estimating equations to left-truncated right-censored data is more challenging and complex. The estimation of weight functions would involve

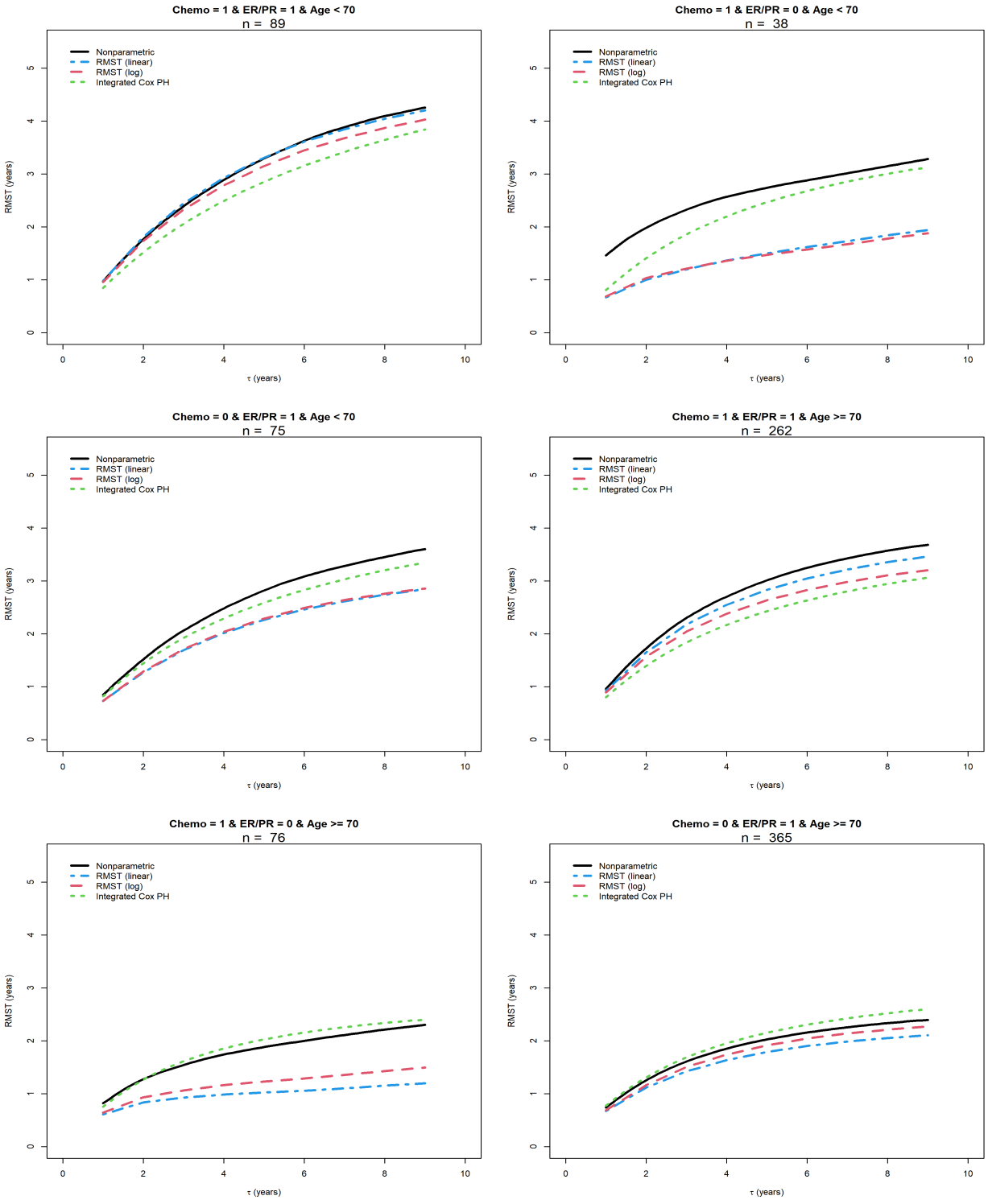

Figure 1.4: Estimated RMST by the combination of receipt of chemotherapy, ER/PR status, and age at diagnosis, using the nonparametric method, multivariable regression model of RMST with the linear link, multivariable regression model of RMST with the log link, and integrated multivariable Cox model survival curve. n is the number of patients in each combination.

estimating the survival function of failure time, distribution of truncation time, and survival function of residual censoring time. The PO approach has been used in regression analysis of RMST for right-censored data and competing risks data, but it has not been extended to the analysis of RMST under left truncation and right censoring, probably because left truncation and induced informative censoring further complicate such analysis. In this paper, we fill the methodological gap by proposing direct regression modeling of RMST under left truncation and general censoring mechanisms and using the PO approach to develop estimation equations for model parameters. The proposed methods have several attractive features. First, we directly model RMST as function of baseline covariates through a generalized linear model with a link function, rather than imposing any restrictive structural assumption such as the proportional hazards. This provides a flexible and robust way to investigate the association between RMST and covariates and to predict patient's expected survival time in the next τ years. Second, by using the PO approach, left truncation and right censoring are handled in the first step of generating the POs, and standard statistical programs/software can be used in the subsequent GEE analysis once the POs are obtained. Thus, the proposed methods can be readily implemented in practice. Third, we consider various censoring mechanisms and use the IPCW method to properly adjust for potential covariate-dependent censoring or dependent censoring. Lastly, we establish the asymptotic properties of proposed estimators, whereas important theoretical justification is often lacking in many existing work using the POs [4, 20, 61].

For the method under covariate-dependent censoring or dependent censoring, the Cox proportional hazards model is used as a working model for the residual censoring time while other semiparametric models, such as generalized transformation models or accelerated failure time model, may be used to compute the censoring distribution given covariates and subject-specific weight function. Extreme weights may present when using IPCW and such a problem can be handled by weight truncation [29], as pointed out by a reviewer. The bias of a parameter estimate will increase and its variance will decrease, as the weights are progressively truncated [10]. The range of weights in our simulations is (1,

45.0) under proportional hazards and (1,40.5) under non-proportional hazards. Although we didn't encounter highly extreme values of weights, we have conducted additional simulations with weights truncated at the 1st and 99th percentiles and evaluated the bias-variance tradeoff. The result shows a modest decrease in variance estimates and a relatively large increase in relative bias. For example, under proportional hazards and covariate-dependent censoring with a censoring rate of 45%, the relative bias changes from -0.020 to -0.122, SD changes from 0.177 to 0.150, and SE changes from 0.168 to 0.134, for the estimation of β_2 at $\tau = 1.39$ under the linear link. The small improvement in variance reduction due to weight truncation appears to be out-weighted by the relatively large bias induced. Nevertheless, with extreme weights, we recommend using weight truncation and exploring the bias-variance tradeoff. The values of τ are typically pre-specified based on the clinical relevance. In the simulations, considerable bias is observed in estimating the intercept in some cases, likely due to the presence of extremely small negative-valued POs that behave as outliers in the subsequent GEE analysis. Similar problems were observed in the simulations in Anderson et al. [4] and Grand et al. [20] The bias problem may be due to the low precision of product-limit estimator at the tail. Moreover, there can be very few subjects at risk at the beginning under left truncation. The sparse information at early event times could cause problems in estimating the survival function, which further affects the estimation of RMST by the integration. Specifically, having fewer subjects at the beginning would result in a big drop of the complete-sample estimates of the survival function $\hat{S}_{PL}(t)$ at early event times. Each corresponding leave-one-out estimate, $\hat{S}^{-1}_{PL}(t)$, on the other hand, is much larger than $\hat{S}_{PL}(t)$ by excluding the subject with small risk set right before his/her event time. These contrasts generate negative POs of RMST with potentially large absolute values, leading to bigger bias in regression parameter estimates. We use a "conditional survival function" approach to adjust for such bias and our simulation results show that this approach can substantially reduce the bias.

In the data application, regression models with linear link and log link functions are considered and compared with the nonparametric method for estimating RMST by graphical displays. Other link functions, such as the logistic link, may be considered. The choice of link function would depend on the scientific question of interest and actual data in a specific application. For example, the linear link or log link may be selected if the difference or ratio of RMST is of interest, respectively. Besides the comparison with nonparametric method, we can also use the Akaike's information criterion, the Bayesian information criterion, and cumulative sum of residuals [35] to assess the performances of models with different link functions. In particular, POs are defined for each subject and can therefore be used to construct cumulative sum of residuals analogous to that in a general linear model [35]. For right-censored data, Perme and Anderson [42] proposed methods for checking hazard regression models using POs, where pseudo-residuals are defined and used for checking the goodness-of-fit of a chosen model. Model diagnostics is crucial in survival analysis and applications. It is our intention in future research to develop rigorous residual-based goodness-of-fit tests for selecting an appropriate link function in regression modeling of RMST for left-truncated right-censored data. The proposed methods are also applicable to regression analysis of RMST for other types of complex survival data, where few alternative methods are available. For example, for regression modeling of the survival function or RMST with clustered survival data, we can compute leave-one-out POs and then use them as outcomes in generalized estimating equations to obtain consistent estimators of model parameters. The GEE sandwich variance estimator can be used to properly adjust for the within cluster correlation. Logan et al. proposed a method for modeling the marginal cumulative incidence function for clustered competing risks data using the PO approach [36]. In future research, we plan to extend the proposed methods to clustered survival data, such as clustered left-truncated right-censored data, clustered competing risks data under left truncation, and recurrent event data. Moreover, causal inference methods such as the propensity score method could be incorporated into the proposed methods to adjust for confounding. Left truncation and right censoring are handled in generating the POs

and then the POs can be used as a replacement for the possibly incompletely observed outcomes when applying standard causal inference methods, such as the propensity score method. Anderson et al. proposed to use POs for estimating the average causal effect with right-censored data [3]. Incorporating casual inference methods (e.g., inverse probability of treatment weighting with propensity scores) into regression modeling of RMST with left-truncated right-censored data based on POs is another interesting direction for our future research.

Table 1.1: Simulation results under covariate-independent censoring and with linear link function.

Proportional Hazards												
				30% Censoring Rate					45% Censoring Rate				
n	τ		True	RB[1]	SD[2]	SE[3]	CP[4]	MSE[5]	RB	SD	SE	CP	MSE
350	0.69	β_0	0.500	-0.004	0.041	0.033	0.900	0.001	-0.008	0.048	0.034	0.893	0.001
		β_1	-0.087	-0.023	0.067	0.057	0.956	0.003	-0.034	0.077	0.059	0.950	0.003
	1.39	β_0	0.750	-0.005	0.074	0.056	0.917	0.003	-0.017	0.119	0.058	0.920	0.004
		β_1	-0.205	0.005	0.116	0.092	0.954	0.008	-0.015	0.127	0.099	0.960	0.010
500	0.69	β_0	0.500	-0.008	0.040	0.028	0.927	0.001	-0.008	0.039	0.028	0.922	0.001
		β_1	-0.087	-0.011	0.072	0.049	0.960	0.002	-0.003	0.065	0.050	0.958	0.003
	1.39	β_0	0.750	-0.005	0.056	0.045	0.927	0.002	-0.011	0.071	0.047	0.912	0.002
		β_1	-0.205	0.029	0.094	0.079	0.954	0.006	0.020	0.106	0.082	0.959	0.007

Non-proportional Hazards												
				30% Censoring Rate					45% Censoring Rate				
n	τ		True	RB	SD	SE	CP	MSE	RB	SD	SE	CP	MSE
350	0.69	β_0	0.497	-0.022	0.028	0.030	0.974	0.001	-0.022	0.030	0.031	0.967	0.001
		β_1	0.126	0.024	0.032	0.034	0.979	0.001	0.032	0.037	0.036	0.971	0.001
	1.39	β_0	0.568	-0.065	0.043	0.048	0.945	0.004	-0.063	0.045	0.051	0.956	0.004
		β_1	0.437	0.021	0.060	0.060	0.950	0.004	0.016	0.064	0.064	0.953	0.004
500	0.69	β_0	0.497	-0.022	0.024	0.025	0.957	0.001	-0.022	0.024	0.026	0.970	0.001
		β_1	0.126	0.024	0.028	0.029	0.970	0.001	0.024	0.028	0.030	0.965	0.001
	1.39	β_0	0.568	-0.063	0.035	0.041	0.918	0.003	-0.069	0.038	0.043	0.903	0.003
		β_1	0.437	0.021	0.048	0.050	0.954	0.003	0.021	0.052	0.053	0.947	0.003

[1] RB is the relative bias, defined as bias/true.
[2] SD is the empirical standard deviation of 1000 parameter estimates.
[3] SE is the average of estimated standard errors across 1000 iterations.
[4] CP is the empirical coverage probability.
[5] MSE is the mean squared error, defined as bias^2 + SE^2.

Table 1.2: Simulation results under proportional hazards and covariate-dependent censoring, and with linear link function. Estimates obtained by using the traditional pseudo-observation (PO) approach and the inverse probability of censoring weighting (IPCW)-adjusted PO approach are compared.

					Unadjusted PO Method					IPCW-adjusted PO Method				
n	cen%	τ	True		RB	SD	SE	CP	MSE	RB	SD	SE	CP	MSE
350	30%	0.69	β_0	0.500	0.001	0.043	0.033	0.908	0.001	-0.008	0.037	0.030	0.910	0.001
			β_1	-0.087	0.023	0.083	0.064	0.950	0.004	-0.080	0.061	0.056	0.933	0.003
			β_2	-0.188	-0.059	0.086	0.077	0.931	0.006	0.048	0.086	0.074	0.959	0.006
		1.39	β_0	0.750	0.012	0.072	0.061	0.927	0.004	-0.019	0.064	0.052	0.926	0.003
			β_1	-0.205	-0.020	0.116	0.102	0.961	0.010	-0.083	0.108	0.093	0.945	0.009
			β_2	-0.391	-0.118	0.161	0.122	0.903	0.017	0.046	0.130	0.125	0.980	0.016
350	45%	0.69	β_0	0.500	0.016	0.042	0.037	0.905	0.001	-0.006	0.048	0.034	0.900	0.001
			β_1	-0.087	0.011	0.076	0.065	0.958	0.004	0.023	0.090	0.068	0.957	0.005
			β_2	-0.188	-0.213	0.087	0.078	0.839	0.008	-0.027	0.110	0.087	0.942	0.008
		1.39	β_0	0.750	0.028	0.106	0.072	0.941	0.006	-0.004	0.074	0.058	0.914	0.003
			β_1	-0.205	-0.078	0.152	0.116	0.951	0.014	-0.001	0.136	0.115	0.957	0.013
			β_2	-0.391	-0.312	0.142	0.129	0.785	0.032	-0.020	0.177	0.168	0.929	0.028
500	30%	0.69	β_0	0.500	-0.001	0.047	0.028	0.911	0.001	-0.012	0.039	0.027	0.926	0.001
			β_1	-0.087	-0.003	0.070	0.052	0.952	0.003	-0.069	0.068	0.052	0.935	0.003
			β_2	-0.188	-0.048	0.106	0.067	0.933	0.005	0.043	0.090	0.068	0.956	0.005
		1.39	β_0	0.750	0.015	0.060	0.049	0.914	0.003	-0.016	0.056	0.044	0.906	0.002
			β_1	-0.205	-0.024	0.091	0.084	0.956	0.007	-0.073	0.083	0.079	0.945	0.006
			β_2	-0.391	-0.113	0.123	0.102	0.891	0.012	0.054	0.103	0.105	0.965	0.011
500	45%	0.69	β_0	0.500	0.014	0.045	0.032	0.909	0.001	0.002	0.038	0.028	0.916	0.001
			β_1	-0.087	0.023	0.086	0.058	0.950	0.003	0.003	0.059	0.053	0.955	0.003
			β_2	-0.188	-0.197	0.079	0.070	0.850	0.006	-0.021	0.086	0.071	0.950	0.005
		1.39	β_0	0.750	0.025	0.091	0.062	0.929	0.004	-0.005	0.061	0.047	0.911	0.002
			β_1	-0.205	-0.127	0.111	0.095	0.946	0.010	0.010	0.109	0.099	0.946	0.010
			β_2	-0.391	-0.315	0.136	0.112	0.712	0.028	-0.023	0.161	0.149	0.928	0.022

Table 1.3: Simulation results under non-proportional hazards and covariate-dependent censoring, and with linear link function. Estimates obtained by using the traditional pseudo-observation (PO) approach and the inverse probability of censoring weighting (IPCW)-adjusted PO approach are compared.

					Unadjusted PO Method					IPCW-adjusted PO Method				
n	cen%	τ	True		RB	SD	SE	CP	MSE	RB	SD	SE	CP	MSE
350	30%	0.69	β_0	0.497	-0.038	0.035	0.036	0.962	0.002	-0.028	0.031	0.034	0.976	0.001
			β_1	0.126	0.103	0.041	0.042	0.955	0.002	0.056	0.038	0.040	0.967	0.002
			β_2	0.109	0.110	0.040	0.042	0.963	0.002	0.073	0.039	0.041	0.966	0.002
		1.39	β_0	0.568	-0.125	0.051	0.055	0.808	0.008	-0.086	0.047	0.054	0.922	0.005
			β_1	0.437	0.117	0.076	0.072	0.895	0.008	0.055	0.077	0.071	0.927	0.006
			β_2	0.310	0.165	0.075	0.071	0.887	0.008	0.135	0.074	0.074	0.921	0.007
350	45%	0.69	β_0	0.497	-0.048	0.036	0.037	0.960	0.002	-0.038	0.033	0.036	0.971	0.002
			β_1	0.126	0.151	0.044	0.043	0.952	0.002	0.095	0.042	0.042	0.958	0.002
			β_2	0.109	0.119	0.043	0.043	0.961	0.002	0.101	0.041	0.043	0.960	0.002
		1.39	β_0	0.568	-0.171	0.054	0.059	0.665	0.013	-0.104	0.053	0.057	0.894	0.007
			β_1	0.437	0.185	0.076	0.077	0.829	0.012	0.073	0.078	0.076	0.937	0.007
			β_2	0.310	0.206	0.077	0.073	0.858	0.009	0.152	0.081	0.079	0.905	0.008
500	30%	0.69	β_0	0.497	-0.034	0.030	0.029	0.954	0.001	-0.032	0.028	0.029	0.961	0.001
			β_1	0.126	0.087	0.036	0.035	0.963	0.001	0.063	0.036	0.035	0.961	0.001
			β_2	0.109	0.101	0.035	0.035	0.955	0.001	0.110	0.034	0.035	0.957	0.001
		1.39	β_0	0.568	-0.127	0.043	0.046	0.701	0.007	-0.090	0.042	0.045	0.856	0.005
			β_1	0.437	0.112	0.062	0.061	0.872	0.006	0.050	0.062	0.060	0.936	0.004
			β_2	0.310	0.181	0.062	0.059	0.837	0.007	0.139	0.061	0.061	0.906	0.005
500	45%	0.69	β_0	0.497	-0.046	0.029	0.031	0.952	0.001	-0.018	-0.036	0.030	0.969	0.001
			β_1	0.126	0.135	0.036	0.036	0.957	0.002	0.087	0.033	0.035	0.964	0.001
			β_2	0.109	0.101	0.035	0.036	0.959	0.001	0.101	0.033	0.036	0.962	0.001
		1.39	β_0	0.568	-0.169	0.045	0.049	0.495	0.012	-0.104	0.042	0.049	0.858	0.006
			β_1	0.437	0.176	0.066	0.064	0.767	0.010	0.073	0.062	0.063	0.931	0.005
			β_2	0.310	0.206	0.064	0.060	0.800	0.008	0.132	0.065	0.067	0.912	0.006

Table 1.4: Estimated covariate effects with 95% confidence intervals (CIs) and p-values at various values of τ (years) for the prevalent cohort from the SEER-Medicare data. The linear link function (estimates are additive effects on RMST) and the log link function (estimates are multiplicative effects on RMST) are used.

	Linear Link								
	Chemo			ER/PR			Age		
τ	$\hat{\beta}$*	CI†	p‡	$\hat{\beta}$	CI	p	$\hat{\beta}$	CI	p
2	0.49	(0.11, 0.87)	0.01	0.80	(0.29, 1.32)	<0.01	-0.02	(-0.04, 0.01)	0.20
5	0.92	(0.22, 1.61)	0.01	1.78	(0.84, 2.72)	<0.01	-0.05	(-0.10, 0.00)	0.05
8	1.13	(0.28, 1.97)	0.01	2.16	(1.02, 3.31)	<0.01	-0.07	(-0.13, -0.01)	0.02
	Log Link								
	Chemo			ER/PR			Age		
τ	$e^{\hat{\beta}}$	CI	p	$e^{\hat{\beta}}$	CI	p	$e^{\hat{\beta}}$	CI	p
2	1.31	(0.98, 1.75)	0.07	1.64	(1.09, 2.48)	0.02	0.99	(0.97, 1.01)	0.25
5	1.31	(0.96, 1.79)	0.09	2.04	(1.21, 3.44)	0.01	0.98	(0.96, 1.00)	0.07
8	1.32	(0.96, 1.82)	0.09	2.04	(1.20, 3.48)	0.01	0.98	(0.96, 1.00)	0.04

* regression parameter estimate.

† 95% confidence interval.

‡ p-value.

CHAPTER 2

REGRESSION MODELING OF CUMULATIVE INCIDENCE FUNCTION FOR COMPETING RISKS DATA

2.1. Introduction

In medical studies, competing risks data commonly arise when patients are subject to failures from more than one causes or participants are at risk for multiple types of events; therefore, we often can only observe the event that occurs first. For example, a diabetic patient's death due to diabetes will be unobservable if the patient dies of cardiovascular disease, given the fact that diabetic patients are at a higher risk for cardiovascular disease comparing to the normal population. Here, death due to cardiovascular disease is a competing event that prevents the observation of the event of interest, death due to diabetes. Ignoring the competing risk would lead to biased estimations of the incidence/risk of event of interest and covariate effects on the incidence/risk. To examine covariate effects on a specific cause of failure in competing risks data, two types of regression models are often employed. One is to fit the Cox proportional hazards model for the cause-specific hazard rate $\lambda_j(s)$, where $\lambda_j(s)$ is the cause-j hazard. The standard estimation procedure for the Cox regression model can be used, with subjects failed from causes other than the cause of interest being treated as censored. Another approach is to perform regression modeling of the cumulative incidence function. The cumulative incidence function is a proper statistic for competing risks data that shows the cumulative probabilities of occurrence of a particular event over time while taking competing risks into account. As a function of cause-specific hazard rates of all causes, the cumulative incidence function of cause j is defined as $F_j(t) = \int_0^t \lambda_j(s)S(s)\mathrm{d}s$, where $S(s)$ is the overall survival function, the probability of not

having failed from any cause at time s. Apparently, regression modeling targeting the cause-specific hazard would produce highly nonlinear covariate effects on the corresponding cumulative incidence function. Fine and Gray [13] introduced an approach of modeling the subdistribution hazard, $\tilde{\lambda}_j(t)$ that is directly associated to the corresponding cumulative incidence function through $\tilde{\lambda}_j(t) = -\frac{\mathrm{d}}{\mathrm{d}t}\log\{1 - F_j(t)\}$, and developed the estimation procedure based on the inverse probability of censoring weighting, which weights subjects who experienced a competing event according to their event times and an estimate of the censoring distribution. Covariate effects based on the subdistribution hazards model are interpretable with respect to the cumulative incidence function. One difference in the estimation procedure between the cause-specific hazards model and the subdistribution hazards model is the risk set. For the cause-specific hazards model, the size of the risk set decreases every time when there is a failure of another cause, whereas subjects fail from other causes remain in the risk set for the subdistribution hazards model.

In addition to the typical right-censoring, competing risks data are often observed subject to left truncation in medical studies, where patients who have not yet experienced any type of event are sampled and prospectively followed for the subsequent first-occurring event. One example of left-truncated competing risks data is from a cohort study on pregnancy exposed to coumarin derivatives [38]. The study cohort consists of 1186 pregnant women who contacted an information service several weeks after conception, and the goal is to investigate the impact of coumarin exposure on the risk of spontaneous abortion while pregnancy may end in induced abortion or live birth (competing risks). Time to spontaneous abortion was left truncated by the first time of contacting the service, thus pregnant women who had early spontaneous abortions were not included in the study. Statistical analysis must appropriately account for both competing risks and left truncation. Under the assumption that failure time is independent of truncation and censoring times, several methods have been introduced for regression modeling of the cumulative incidence function for left-truncated right-censored competing risks data. Geskus [16], Shen [51], and Zhang et al. [64] extended Fine and Gray's method [13] to left-truncated right-censored

competing risks data and proposed estimation procedures by using alternative weighting techniques. Recently, Bellach et al. [7] considered a general direct regression model for the subdistribution hazard for situations when the proportional subdistribution hazards assumption is violated, based on a conditional nonparametric maximum likelihood procedure. In many observational studies, covariate-dependent truncation and/or covariant-dependent censoring occurs frequently when truncation time and/or censoring time are correlated with failure time through common baseline covariate. For instance, in the cohort study on pregnancy exposed to coumarin derivatives, the dependence between left-truncation time and event time was noted, because women who considered an induced abortion were more likely to contact and seek advice from the information service center [52]. Furthermore, such dependence between truncation and event times was likely due to their common correlations with the covariate of exposure to coumarin derivatives [39]. In AIDS studies, patients who have low CD4 counts (an indicator of immune function in patients living with HIV) are more likely to drop out of the study, resulting in covariate-dependent censoring. For right-censored competing risks data, Binder et al. [8] proposed a modified pseudo-observation (PO) approach to account for covariate-dependent censoring. For left-truncated competing risks data, Stegherr et al. [52] investigated dependent left truncation by using inverse probability of left-truncation weights obtained from the cause-specific Cox proportional hazards model with truncation time as a covariate. This modeling approach, however, imposed a restrictive assumption that truncation and failure times are correlated through a proportional hazards model, which may be hard to satisfy in real applications. For left-truncated right-censored competing risks data, Zhang et al. [64] handled covariate-dependent truncation and covariate-dependent censoring in the Fine-Gray subdistribution hazards model by utilizing stratified nonparametric weight or covariate-adjusted weight. In this paper, we will consider general truncation and censoring mechanisms where censoring and/or truncation time may depend on covariates in flexible regression modeling of cumulative incidence function for left-truncated right-censored competing risks data.

Andersen et al. [5], Klein and Andersen [31], and Klein [30] introduced an approach of using POs in the regression modeling of competing risks data, where POs are jackknife estimates that represent the contribution of each subject to the estimator of the parameter of interest [5]. Graw et al. [21] provided the formal theoretical justification of PO approach in regression modeling of right-censored competing risks data under covariate-independent censoring. Graw et al. [20] studied the performance of POs in the Cox regression analysis of left-truncated right-censored data. The advantage of PO approach is that it handles complex competing risks data subject to left truncation and right censoring in the first step of generating POs, and then the POs can be used as responses in a generalized linear model for uncensored data and analyzed by standard statistical software. In this paper, we propose to directly model the cumulative incidences as a function of baseline covariates based on POs, for left-truncated right-censored competing risks data under general truncation and censoring mechanisms. The remainder of this paper is organized as follows. In Section 2.2, we introduce the left-truncated right-censored competing risks data structure with notations and describe a general regression model of the cumulative incidence function. In Section 2.3, we first present the proposed method for regression modeling of cumulative incidence function given covariates using POs, under covariate-independent truncation and covariate-independent censoring. We then relax the model assumption to incorporate covariate-dependent censoring and/or covariate-dependent truncation. We investigate the finite sample performances of proposed estimators by simulation studies under various scenarios in Section 2.4. We illustrate the proposed methods by applications to a cohort study on HIV disease progression and a cohort study on pregnancy exposed to coumarin derivatives in Section 2.5. We provide concluding remarks in Section 2.6. Technical details can be found in the Appendix.

2.2. Data, Notations, and Regression Model

In a competing risks setting with left truncated right censored data, let $\varepsilon = 1, 2, \ldots, K$ indicates the cause of failure for a total of K competing risks. Only the first cause of failure is observable or is of interest. Let T be the failure time from the disease onset to the first-occurring failure event. Let L be the left truncation time and C be the right censoring time. Let $X = \min(T, C)$ be the follow-up time till failure event or censoring and $\delta = I(T \leqslant C)$ be the failure indicator. Due to left truncation, X is observable only if $L < T$. Let $\boldsymbol{Z}$ be a $p \times 1$ vector of baseline covariates. The observed data are $(X_i, L_i, \delta_i, \delta_i\varepsilon_i, \mathbf{Z_i}), i = 1, 2, \ldots, n$. Let F_j denote the cumulative incidence function for cause j, defined as $F_j(t \mid \mathbf{Z}) = \Pr(T \leqslant t, \varepsilon = j \mid \mathbf{Z})$, which is the conditional probability of failing from cause j at or before time t given covariates. It can be expressed in terms of the cause-specific hazard and the overall survival function as

$$F_j(t \mid \mathbf{Z}) = \int_0^t \lambda_j(u \mid \mathbf{Z}) \Pr(T \geqslant u \mid \mathbf{Z}) \, \mathrm{d}u = \int_0^t \lambda_j(u \mid \mathbf{Z}) S(u \mid \mathbf{Z}) \, \mathrm{d}u,$$

where $\lambda_j(u \mid \mathbf{Z}) = \lim_{\Delta t \to 0} \Pr(u \leqslant T < u + \Delta t, \varepsilon = j \mid T \geqslant u, \mathbf{Z}) / \Delta t$ is the cause-specific hazard for cause j at time u conditional on covariates and $S(u \mid \mathbf{Z}) = \exp\left\{-\int_0^u \sum_{j=1}^K \lambda_j(s \mid \mathbf{Z}) \mathrm{d}s\right\}$ is the overall survival function. Our goal is to directly model the relationship between the cumulative incidence function of cause j and covariates through a generalized linear model:

$$g\{F_j(t \mid \mathbf{Z})\} = \alpha(t) + \mathbf{Z}^{\mathrm{T}}\boldsymbol{\gamma}, \tag{2.1}$$

where $g(\cdot)$ is a known differentiable link function, $\alpha(t)$ determines the baseline failure probability when $\mathbf{Z} = \mathbf{0}$, and $\boldsymbol{\gamma}$ is a $p \times 1$ regression coefficient vector. Model (2.1) is identical to the regression model of cumulative incidence function considered in Fine [12], where an extension of a least-squares technique of Fine et al. [11] is used for

estimation. Examples of common link functions include the logit link, $g(b) = \log\{b/(1-b)\}$, complementary log-log link on b, $g(b) = -\log\{-\log(b)\}$, and complementary log-log link on $1-b$, $g(b) = -\log\{-\log(1-b)\}$. Specifically, the complementary log-log link on $1-F_j(t)$ gives the Fine-Gray proportional subdistribution hazards model [13].

2.3. Regression Modeling of Cumulative Incidence Function based on Pseudo-Observations

For right-censored competing risks data, under the assumption that censoring time C is independent of failure time and cause (T, ε) (covariate-independent censoring), the cumulative incidence function $F_j(t)$ of cause j at t can be consistently non-parametrically estimated by the Aalen-Johansen estimator [1]:

$$\hat{F}_j(t) = \int_0^t \hat{S}(u-)\,\mathrm{d}\hat{\Lambda}_j(u),$$

where $\hat{\Lambda}_j(t) = \int_0^t \sum_{i=1}^n \mathrm{d}N_{ij}(u)/Y(u)$ is the Nelson-Aalen estimator for the integrated cause-j specific hazard, $N_{ij}(u) = I(X_i \leqslant u, \varepsilon_i = j, \delta_i = 1)$ indicates whether subject i has experienced a cause-j event prior to time u, and $Y(u) = \sum_{i=1}^n I(X_i \geqslant u)$ is the observed number of subjects at risk at time $u-$, and $\hat{S}(t)$ is the Kaplan-Meier estimator of survival from any failure. Klein and Andersen [31] proposed a PO approach for direct modeling of covariate effects on the cumulative incidence function under covariate-independent censoring. The PO for the cumulative incidence function of cause j for the ith subject at time t is computed as:

$$\hat{F}_{ij}(t) = n\hat{F}_j(t) - (n-1)\hat{F}_j^{(-i)}(t), \tag{2.2}$$

where $\hat{F}_j^{(-i)}(t)$ is the jackknife leave-one-out estimator for $F_j(t)$ based on data leaving out subject i. The POs, $\mathcal{PO} = \left\{\hat{F}_{1j}(t), \hat{F}_{2j}(t), \ldots, \hat{F}_{nj}(t)\right\}$ obtained from (2.2) can then be used as responses of the regression model (2.1) to estimate γ under a generalized estimating equation (GEE) framework [31]. Graw et al. [21] provided the formal theoretical justification

of PO approach and asymptotic properties of regression parameter estimators obtained from GEE, for competing risks data under covariate-independent censoring. Binder et al. [8] investigated the bias of using POs generated based on the Aalen-Johansen estimator when censoring time depends on covariates and introduced modified PO values based on alternative estimators to reduce the bias in the presence of covariate-dependent censoring. We extend the methods in Klein and Andersen [31] and Binder et al. [8] to left-truncated right-censored competing risks data and propose a modified PO approach for regression modeling of cumulative incidence function under general truncation and censoring mechanisms where censoring and/or truncation time may depend on covariates.

2.3.1. Covariate-Independent Truncation and Covariate-Independent Censoring

We consider covariate-independent truncation and covariate-independent censoring, that is, (L, C) and (T, ε) are assumed to be independent. The cumulative incidence function $F_j(t)$ of cause j for left-truncated right-censored competing risks data can be consistently estimated by the Aalen-Johansen type estimator after properly adjusting for the risk set. Define the counting process $N_{ij}^L(u) = I\{L_i < X_i \leqslant u, \varepsilon_i = j, \delta_i = 1\}$ and $Y^L(u) = \sum_{i=1}^n I\{L_i < u \leqslant X_i\}$. A consistent estimator of $F_j(t)$ is

$$\hat{F}_j^{AJ}(t) = \int_0^t \hat{S}_{PL}\left(u^-\right) \mathrm{d}\hat{\Lambda}_j^L(u),$$

where $\hat{\Lambda}_j^L(t) = \int_0^t \sum_{i=1}^n dN_{ij}^L(u)/Y^L(u)$ is the left-truncated version of Nelson-Aalen estimator and $\hat{S}_{PL}(t)$ is the product-limit estimator of survival function for left-truncated right-censored data, with risk set $r(t) = \{i : L_i < t \leqslant X_i\}$, adjusting for left truncation [25, 54]. Geskus [16] introduced an alternative representation for the cumulative incidence function, an inverse probability weighted (IPW) estimator, denoted as $\hat{F}_j^{IPW}$ for cause j. Let $t_{(1)} < \cdots < t_{(i)} < \cdots < t_{(N)}$ denote the ordered distinct observed event times, $c_{(1)} < \cdots < c_{(w)} < \cdots < c_{(W)}$ denote the ordered distinct observed censoring times, and $l_{(1)} < \cdots < l_{(m)} < \cdots < l_{(M)}$ are the ordered distinct observed truncation times. Let $\mathrm{d}\left\{t_{(i)}\right\}$

be the number of observed events of any type at $t_{(i)}$, $\mathrm{d}\left\{c_{(w)}\right\}$ be the number of censorings at $c_{(w)}$, $\mathrm{d}\left\{l_{(m)}\right\}$ be the number of truncation times at $l_{(m)}$, and $R\left(t\right)$ be the observed number at risk at time t. As a weighted empirical estimator of the cumulative incidence function, $\hat{F}_j^{IPW}$ takes the form:

$$\hat{F}_j{}^{IPW}\left(t\right)=\frac{1}{\hat{n}}\sum_{i=1}^{n}\frac{N_{ij}^{L}\left(t\right)}{\hat{S}_C\left(X_i-\right)\times\hat{F}_L\left(X_i-\right)},$$

where $\hat{S}_C\left(X_i-\right)$ is the estimator of the survival function for censoring time C, $\hat{F}_L\left(X_i-\right)$ is the estimator of the cumulative distribution function (CDF) for truncation time L, and $\hat{n}=\sum_{i=1}^{N}\left[\mathrm{d}\left\{t_{(i)}\right\}\Big/\hat{F}_L\left\{t_{(i)}-\right\}\right]+\sum_{w=1}^{W}\left[\mathrm{d}\left\{c_{(w)}\right\}\Big/\hat{F}_L\left\{c_{(w)}-\right\}\right]$. Reversing the role of T and C yields the estimator of the survival function for C: $\hat{S}_C(t)=\prod_{c_{(w)}\leqslant t}\left[1-\frac{\mathrm{d}\left\{c_{(w)}\right\}}{R\left\{c_{(w)}\right\}}\right]$. Since L is right truncated by $X=\min(T,C)$, the product-limit statistic for truncation time L can be obtained through reversal of time such that $-L$ is left truncated by $-X$: $\hat{F}_L\left(t\right)=\prod_{-l_{(m)}<-t}\left[1-\frac{\mathrm{d}\left\{l_{(m)}\right\}}{R\left\{l_{(m)}\right\}}\right]=\prod_{l_{(m)}>t}\left[1-\frac{\mathrm{d}\left\{l_{(m)}\right\}}{R\left\{l_{(m)}\right\}}\right]$. Geskus [16] proved the equivalence of $\hat{F}_j^{AJ}(t)$ and $\hat{F}_j^{IPW}(t)$ under covariate-independent truncation and covariate-independent censoring. Thus, either $\hat{F}_j^{AJ}(t)$ or $\hat{F}_j^{IPW}(t)$ can be used to consistently estimate $F_j(t)$ and then construct POs. As noted in Geskus [16], only when L and C are independent, $\hat{S}_C$ is the estimator of the survival function for C and $\hat{F}_L$ is the estimator of the CDF for L. Nevertheless, the result that $\hat{F}_j{}^{IPW}$ and $\hat{F}_j{}^{AJ}$ are equivalent holds irrespective of the relationship between L and C. For ease of discussion, we assume that L and C are independent or conditionally independent given covariates Z in this paper, but the proposed methods would generally work when L and C are dependent.

To construct POs, we start with selecting a grid of points τ_1, τ_2, ..., τ_H. Suppose that cause-1 event is the event of interest. Define $\theta_{ih}=F_1\left(\tau_h\mid\mathbf{Z}_i\right)$ as the conditional cumulative incidence function that we are intended to model. Based on (2.2), the PO for the ith subject at time τ_h is computed as: $\hat{\theta}_{ih}=n\hat{F}_1\left(\tau_h\right)-\left(n-1\right)\hat{F}_1^{(-i)}\left(\tau_h\right)$, where $\hat{F}_1$ is either the Aalen-Johansen estimator adjusting for left truncation or the IPW estimator by Geskus [16].

Then, these POs are used as responses in the generalized linear model with a suitable link function

$$g\left(\theta_{ih}\right)=\alpha(\tau_h)+\mathbf{Z}_i^{\mathrm{T}}\boldsymbol{\gamma}=\mathbf{Z}_{ih}^{\mathrm{T}}\boldsymbol{\beta},\quad i=1,\ldots,n,\quad h=1,\ldots,H.$$

Let $\hat{\boldsymbol{\theta}}_i=\left(\hat{\theta}_{i1},\hat{\theta}_{i2},\ldots,\hat{\theta}_{iH}\right)$ and $\boldsymbol{\theta}_i=\left\{g^{-1}\left(\mathbf{Z}_{i1}^{\mathrm{T}}\boldsymbol{\beta}\right),g^{-1}\left(\mathbf{Z}_{i2}^{\mathrm{T}}\boldsymbol{\beta}\right),\ldots,g^{-1}\left(\mathbf{Z}_{iH}^{\mathrm{T}}\boldsymbol{\beta}\right)\right\}$. The regression coefficients, $\boldsymbol{\beta}$, can be estimated by the GEE:

$$U\left(\boldsymbol{\beta}\right)=\sum_{i=1}^{n}U_i\left(\boldsymbol{\beta}\right)=\sum_{i=1}^{n}\left\{\frac{\partial\boldsymbol{\theta_i}}{\partial\boldsymbol{\beta}}\right\}^{\mathrm{T}}\boldsymbol{\mathcal{V}}_i^{-1}\left(\boldsymbol{\beta}\right)\left\{\hat{\boldsymbol{\theta}}_i-\boldsymbol{\theta_i}\right\}=0, \tag{2.3}$$

where $\boldsymbol{\mathcal{V}}_i\left(\boldsymbol{\beta}\right)$ is a working covariance matrix of $\hat{\boldsymbol{\theta}}_i$ [34, 62]. Klein and Andersen [31] suggested three possible choices for the working covariance matrix. The simplest choice is the independent working covariance matrix. Another option is the 'exact' working covariance matrix with elements $v_{ihq}=\mathrm{cov}\left(\hat{\theta}_{ih},\hat{\theta}_{iq}\right)=F_1\left(\tau_h\mid\mathbf{Z}_i\right)\times\left\{1-F_1\left(\tau_q\mid\mathbf{Z}_i\right)\right\}$, for $\tau_h\leqslant\tau_q,\ \ i=1,\ldots,n,$ and $h,q=1,\ldots,M.$ However, the exact working covariance matrix results in complicated estimating equations as the matrix $\boldsymbol{\mathcal{V}}_i$ changes with each iteration in the root finding process. An alternative is to use a common working covariance matrix $\boldsymbol{\mathcal{V}}$ to estimate $\boldsymbol{\mathcal{V}}_i$, i.e. an empirical working covariance matrix defined by $\boldsymbol{\mathcal{V}}=(\hat{v}_{hq})=\frac{1}{n}\sum_i\left(\hat{\theta}_{ih}-\bar{\theta}_h\right)\left(\hat{\theta}_{iq}-\bar{\theta}_q\right)$, where $\bar{\theta}_h=\frac{1}{n}\sum_i\hat{\theta}_{ih}$. We will use the independent working covariance matrix and the empirical working covariance matrix in simulations. For right-censored competing risks data, Klein and Andersen [31] suggested that the regression parameter estimates obtained from GEE using POs based on Aalen-Johansen estimator are consistent, under covariate-independent censoring. Graw et al. [21] provided theoretical justification of the asymptotic properties. Since the nonparametric estimator $\hat{F}_1(t)$ based on either the Aalen-Johansen estimator adjusting for left truncation or the IPW estimator by Geskus [16] is consistent for left-truncated right-censored competing risks data, similarly we can use (2.3) to obtain consistent estimates of regression parameters in model (2.1). Let $\hat{\boldsymbol{\beta}}$ be the solution to (2.3) and $\boldsymbol{\beta}_0$ be the true value of $\boldsymbol{\beta}$. The asymptotic properties of $\hat{\boldsymbol{\beta}}$ are summarized in Theorem 2 with a proof and regularity conditions

provided in the Appendix.

Theorem 2. *Under regularity conditions* $C1 - C3$ *in the Appendix B,* $\hat{\boldsymbol{\beta}}$ *is consistent to* $\boldsymbol{\beta}_0$*, and* $\sqrt{n}\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\right)$ *is asymptotically normal with mean zero and a covariance matrix that can be consistently estimated using a standard 'sandwich' estimator, which takes the form*

$$\hat{\Sigma} = \mathcal{I}\left(\hat{\boldsymbol{\beta}}\right)^{-1} \hat{\text{var}}\left\{U\left(\hat{\boldsymbol{\beta}}\right)\right\} \mathcal{I}\left(\hat{\boldsymbol{\beta}}\right)^{-1},$$

where

$$\mathcal{I}\left(\hat{\boldsymbol{\beta}}\right) = \sum_i \left\{\frac{\partial \theta_i}{\partial \boldsymbol{\beta}}\right\}^\top \mathcal{V}_i\left(\boldsymbol{\beta}\right)^{-1} \left\{\frac{\partial \theta_i}{\partial \boldsymbol{\beta}}\right\} \Bigg|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}},$$

$$\hat{\text{var}}\left\{U\left(\hat{\boldsymbol{\beta}}\right)\right\} = \sum_i U_i\left(\hat{\boldsymbol{\beta}}\right) U_i\left(\hat{\boldsymbol{\beta}}\right)^{\text{T}}.$$

Moreover, Geskus [16] proposed to estimate regression parameters in the Fine-Gray subdistribution hazards model for left-truncated right-censored competing risks data through a "re-weighting approach", where each subject ξ is assigned a weight $\omega_\xi\left(t_{(i)}\right)$ that equals to 1 if the subject is at risk at $t_{(i)}$, equals to $\frac{\hat{S}_C\left(t_{(i)}-\right)}{\hat{S}_C\left(t_{(o)}-\right)} \frac{\hat{F}_L\left(t_{(i)}-\right)}{\hat{F}_L\left(t_{(o)}-\right)}$ if the subject had a competing event observed at $t_{(o)} < t_{(i)}$, and equals to 0 otherwise. These weights are equal to the ones used in the Fine-Gray model without left truncation by setting $\hat{F}_L \equiv 1$. In section 2.4, we conduct simulation studies to compare the performance of the proposed PO approach and IPW regression estimator by Geskus [16], where the complementary log-log link function on $1 - F_j(t)$ is used in model (2.1).

2.3.2. Covariate-Independent Truncation and Covariate-Dependent Censoring

We relax the model assumption to consider the case that the truncation time is independent of covariates but the censoring time depends on covariates, that is covariate-independent truncation and covariate-dependent censoring. Covariate-dependent censoring often occurs in observational studies where the censoring time and failure time are only conditionally independent given covariates. Graw et al. [21] showed that the consistency

of the solution to (2.3) relies on the covariate-independent censoring assumption. Thus, it is crucial to account for covariate-dependent censoring to consistently estimate the cumulative incidence function and obtain corresponding POs. Assuming L is independent of (T, ε) and C is conditionally independent of (T, ε) given $\mathbf{Z}$, it is natural to estimate the cumulative incidence function by incorporating a covariate-adjusted weight for censoring C, $\hat{S}_C\left(X_i- \mid \mathbf{Z}\right)$ into the IPW estimator in Section 2.3.1:

$$\hat{F}_j^{\,IPW}(t) = \frac{1}{\hat{n}} \sum_{i=1}^{n} \frac{N_{ij}^L(t)}{\hat{S}_C\left(X_i- \mid \mathbf{Z}\right) \times \hat{F}_L\left(X_i-\right)}, \tag{2.4}$$

where $\hat{F}_L\left(X_i-\right)$ is a weight for truncation L, the nonparametric product-limit statistic as previously described. We further assume that L and C are conditionally independent given covariates Z, so that $S_C\left(t \mid \mathbf{Z}\right)$ is the estimator of conditional survival function of C given Z. Parametric regression model, or semiparametric regression models such as the Cox proportional hazards model, generalized transformation models, and accelerated failure time model, or local Kaplan-Meier estimator [56] can be adopted to estimate the conditional survival function of censoring time given covariates. Without loss of generality, we use the Cox model to characterize the relationship between censoring time and covariates, and estimate the conditional survival probability for censoring, $S_C\left(t \mid \mathbf{Z}\right) = \exp\left\{-B_{C_0}\left(t\right)\exp\left(\boldsymbol{\eta}^{\mathrm{T}}\mathbf{Z}\right)\right\}$, where $B_{C_0}(t)$ is the cumulative baseline hazard for censoring and $\boldsymbol{\eta}$ is the vector of model parameters, through $\hat{S}_C\left(t \mid \mathbf{Z}_i\right) = \exp\left\{-\hat{B}_{C_0}\left(t\right)\exp\left(\hat{\boldsymbol{\eta}}^{\mathrm{T}}\mathbf{Z}_i\right)\right\}$, where $\hat{\boldsymbol{\eta}}$ is the partial likelihood estimate of $\boldsymbol{\eta}$, and $\hat{B}_{C_0}(t) = \sum_{i=1}^{n} \frac{I(X_i \leqslant t)(1-\delta_i)}{\sum_{i=1}^{n}\exp(\hat{\boldsymbol{\eta}}^{\mathrm{T}}\mathbf{Z}_i)I(L_i < t \leqslant X_i)}$ is the Breslow estimator of the cumulative baseline hazard function for censoring $B_{C_0}\left(t\right)$, with $I\left(L_i < t \leqslant X_i\right)$ being the at-risk indicator and δ being the failure indicator. Plugging this Cox-model-based consistent estimator of conditional survival function for censoring in (2.4), we can obtain an IPW estimator of the cumulative incidence function $F_j(t)$ with the covariate-adjusted censoring weight. The consistency of the IPW estimator $\hat{F}_j^{IPW}(t)$ of $F_j(t)$ under covariate-dependent censoring is summarized in Theorem 3 with a proof provided in the Appendix.

Theorem 3. *The estimator $\hat{F}_j^{IPW}(t)$ is a consistent estimator of the true cumulative incidence function $F_j(t)$.*

Then, POs can be constructed based on the consistent IPW estimator $\hat{F}_j^{IPW}(t)$. Specifically, the jackknife leave-one-out estimator is computed by re-fitting the censoring model n times, eliminating each subject in turn to get $\hat{\boldsymbol{\eta}}^{(-i)}$ and $\hat{B}_{C_0}^{(-i)}(\cdot)$. Lastly, the POs can be used in the generalized linear model (2.1) to estimate $\boldsymbol{\gamma}$, similar to the case under covariate-independent truncation and covariate-independent censoring. The resulting estimators are consistent and asymptotically normal, under the assumption that the model for the censoring distribution is correctly specified and under the regularity conditions specified in the proof of Theorem 2. The proof of the large sample properties follows the lines in the proof of Theorem 2 and is omitted for brevity.

2.3.3. Covariate-Dependent Truncation and Covariate-Dependent Censoring

We further consider the case that both truncation and censoring times are allowed to depend on covariates, that is covariate-dependent truncation and covariate-dependent censoring. Covariate-dependent truncation occurs when the truncation time and failure time are only conditionally independent given covariates. Assuming that (L, C) and (T, ε) are conditionally independent given $\mathbf{Z}$, covariate-adjusted truncation and censoring weights can be utilized to consistently estimate the cumulative incidence function. Similarly, we assume that L and C are conditionally independent given Z. The covariate-adjusted weight for censoring $\hat{S}_C\left(X_i- \mid \mathbf{Z}\right)$ can be consistently estimated by $\hat{S}_C\left(t \mid \mathbf{Z}_i\right) = \exp\left\{-\hat{B}_{C_0}\left(t\right)\exp\left(\hat{\boldsymbol{\eta}}^{\mathrm{T}}\mathbf{Z}_i\right)\right\}$, as described in Section 2.3.2. The covariate-adjusted weight for truncation, $\hat{F}_L\left(X_i- \mid \mathbf{Z}_i\right)$, can be obtained based on appropriate regression models of L on Z. Without loss of generality, we use the Cox proportional hazards model to describe the relationship between truncation time L and covariates $\mathbf{Z}$: $S_L\left(t \mid \mathbf{Z}\right) = \exp\left\{-B_{L_0}\left(t\right)\exp\left(\boldsymbol{\phi}^{\mathrm{T}}\mathbf{Z}\right)\right\}$, where $B_{L_0}(t)$ is the cumulative baseline hazard for truncation and $\boldsymbol{\phi}$ is the vector of regression parameters. Note that the left truncation time L is subject to right truncation by $X = \min(T, C)$. Existing methods

for fitting the Cox model to right-truncated data include maximizing the likelihood with respect to both baseline hazard and regression parameters [2, 14] and solving weighted estimating equations where the weight is inversely proportional to the selection probability that a subject is observed in the sample [37, 45]. Most of the methods require the independence assumption between truncation and observed survival times, which does not hold here as L is correlated with X through $\mathbf{Z}$. Under covariate-dependent truncation, Rennert and Xie [46] proposed an expectation-maximization (EM) algorithm for estimating the Cox model for right-truncated data. By maximizing the log-likelihood of the observed survival times conditional on the observed truncation times and covariates, the EM algorithm provides a convenient estimation approach to obtain estimates of the baseline hazard and regression coefficients. Rennert and Xie [46] also showed the consistency and asymptotic normality of the proposed EM estimators. Therefore, under covariate-dependent truncation and covariate-dependent censoring, the cumulative incidence function can be estimated by:

$$\hat{F}_j^{IPW}(t) = \frac{1}{\hat{n}} \sum_{i=1}^{n} \frac{N_{ij}^L(t)}{\hat{S}_C\left(X_i- \mid \mathbf{Z}_i\right) \times \left\{1 - \hat{S}_L\left(X_i- \mid \mathbf{Z}_i\right)\right\}},$$

where $\hat{S}_L\left(X_i- \mid \mathbf{Z}_i\right)$ is the consistent estimator of the conditional survival function for truncation time L, based on the Cox model estimation using the EM algorithm. Since $\hat{S}_C\left(X_i- \mid \mathbf{Z}_i\right)$ and $\hat{S}_L\left(X_i- \mid \mathbf{Z}_i\right)$ consistently estimate $S_C\left(X_i- \mid \mathbf{Z}_i\right)$ and $S_L\left(X_i- \mid \mathbf{Z}_i\right)$, respectively, $\hat{F}_j^{IPW}(t)$ is a consistent estimator of the true cumulative incidence function $F_j(t)$, following similar arguments in the proof of Theorem 3. To construct POs, the jackknife leave-one-out estimator can be computed in the usual way, by re-fitting the truncation model n times, eliminating each subject in turn to get $\hat{\boldsymbol{\phi}}^{(-i)}$ and $\hat{B}_{L_0}^{(-i)}(\cdot)$. However, this approach is computationally intensive because it requires to fit the Cox model for truncation as many times as there are subjects in the dataset, which is especially time consuming when the EM algorithm is involved. One less computationally expensive approach suggested by Binder et al. [8] is to keep the EM estimates of regression coefficients $\hat{\boldsymbol{\phi}}$ from the full dataset and only to estimate the cumulative baseline hazard n times by the Breslow

estimator, leaving out subject i each time. The corresponding leave-i-out estimator for the baseline hazard given $\hat{\boldsymbol{\phi}}$ is: $\hat{B}_{L_0}^{(-i)}(t) = \sum_{\xi \neq i} \frac{I\left(l_\xi \leqslant t\right)}{\sum_{\xi \neq i} \exp(\hat{\boldsymbol{\eta}}^{\mathrm{T}} \mathbf{Z}_i) I\left(l_\xi \geqslant t\right)}$, with $I\left(l_\xi \geqslant t\right)$ being the at-risk indicator and l_ξ being the truncation time from the leave-i-out dataset. Then, the leave-i-out estimator of the cumulative incidence function is:

$$\frac{1}{\hat{n}^{(-i)}} \sum_{\xi \neq i} \frac{N_{rj}^{L}(t)}{\exp\left[-\hat{B}_{C_0}^{(-i)}(X_\xi -) \exp\left\{\hat{\boldsymbol{\eta}}^{(-i)\mathrm{T}} \mathbf{Z}_\xi\right\}\right] \times \left[1 - \exp\left\{-\hat{B}_{L_0}^{(-i)}(X_\xi -) \exp\left(\hat{\boldsymbol{\phi}}^{\mathrm{T}} \mathbf{Z}_\xi\right)\right\}\right]}.$$

Lastly, the POs, defined as $\hat{F}_j^{IPW}(t) - \hat{F}_j^{IPW(-i)}(t)$, can be used in the generalized linear model (2.1) to estimate $\boldsymbol{\gamma}$, similar to the cases in Section 2.3.1 and 2.3.2. Under the assumption that the models for the censoring and truncation distributions are correctly specified and under the usual regularity conditions, the resulting estimators are consistent and asymptotically normal, which can be proved similarly to Theorem 2.

2.4. Simulations

We conduct a series of simulations to assess the performance of the proposed methods for left-truncated right-censored competing risks data, under various scenarios. The competing risks failure time data are generated according to the proportional subdistribution hazards model [13] with two covariates Z_1 and Z_2: Z_1 is binary with equal probability of being 0 and 1, and Z_2 is continuous that follows a standard normal distribution. Our main interest is to model the cumulative incidence function for cause 1 event given the two covariates. The underlying cumulative incidence functions are given by $F_1(t; z_1, z_2) = 1 - \{1 - p\,(1 - e^{-0.5t})\}^{\exp(\beta_{11} z_1 + \beta_{12} z_2)}$ and $F_2(t; z_1, z_2) = (1-p)^{\exp(\beta_{11} z_1 + \beta_{12} z_2)} \times \{1 - e^{-0.5t \exp(\beta_{21} z_1 + \beta_{22} z_2)}\}$, where p, the probability of experiencing a cause 1 event when both covariate values are zero, is set to be 0.7. The true covariates effects are set to be $(\beta_{11}, \beta_{12}) = (0.5, 0.5)$ and $(\beta_{21}, \beta_{22}) = (-0.5, 0.5)$. Moreover, we carry out simulations with covariates effects set to be $(\beta_{11}, \beta_{12}) = (-1, 1)$ and $(\beta_{21}, \beta_{22}) = (-1, 1)$, and results

are summarized in tables in the Appendix. For each scenario, we compare parameter estimates obtained by Geskus's method [16], the PO approach based on Aalen-Johansen estimator of cumulative incidence function (PO-AJ), and the PO approach based on IPW estimator of cumulative incidence function (PO-IPW). Following the recommendation in Klein and Andersen [31], we calculate POs at 6 grid points that are equally spaced on the event scale (i.e., grid points are selected after 14, 28, 42, 56, 70, and 84% of the observed event status information), and a complementary log-log link function on $1 - F_j(t)$ is used. The simulation is repeated 1000 times with a sample size of $n = 350$ or 500.

2.4.1. Simulations under Covariate-Independent Truncation and Covariate-Independent Censoring

First, we evaluate the performance of the proposed method under covariate-independent truncation and covariate-independent censoring. The truncation time follows a uniform distribution on [0, U_L], allowing for various levels of truncation (i.e., truncation rates of 30% and 50%). The censoring time follows a uniform distribution on [0.5, U_C], allowing for various levels of censoring (i.e., censoring rates of 30% and 50%). As discussed in Section 2.3, left truncation is adjusted for in the Aalen-Johansen (AJ) and IPW estimators in the PO approach. From Table 2.3, the PO approach performs well with comparable results by using the AJ and IPW estimators. For the proposed PO-based estimators, the biases are generally very small, and the estimated model-based standard errors (SEs) computed as the GEE sandwich estimator are close to the empirical standard deviations (SDs) in all scenarios. Increasing the censoring rate or the truncation rate from 30% to 50% does not affect the bias much but tends to increase the estimated SEs and SDs. As the sample size increases, the SEs and SDs decrease. Additionally, the coverage probabilities are overall close to the nominal level of 95%, with some slight overcoverage. In contrast, the estimated SEs of Geskus's estimator are constantly smaller than the corresponding empirical SDs, indicating that the estimated Greenwood SEs by Geskus's method may underestimate the variations of regression parameter estimates, and this may further lead to relatively large

undercoverage. Results in the supplementary Table B.1 also suggest a good performance of the proposed PO approach.

2.4.2. Simulations under Covariate-Independent Truncation and Covariate-Dependent Censoring

Next, we evaluate the proposed method under covariate-independent truncation and covariate-dependent censoring. The truncation time follows the same uniform distribution as in Section 2.4.1, with a truncation rate of 30% or 50%. The censoring time now is generated from an exponential distribution with parameter $\lambda_C = \lambda_{C_0} \exp(1.5Z_1 + Z_2)$ that depends on the two covariates. λ_{C_0} is such that the censoring rate is 30% or 50%. Covariate-dependent censoring is adjusted for in the PO-IPW. For the PO-IPW estimator, after POs are obtained, we compute SEs under both the independence and the empirical working covariance matrices in the subsequent GEE analysis [31]. Table 2.4 shows that the PO-IPW can substantially reduce the bias after accounting for covariate-dependent censoring, especially for the estimation of β_{11} that is heavily biased by Geskus's method or the PO-AJ only adjusting for covariate-independent censoring and truncation. For example, in Table 2.4, under the scenario of $n = 350$, censoring rate of 30% and truncation rate of 30%, the absolute value of bias is reduced from 0.089 (by GeskusâĂŹs method) and 0.063 (by PO-AJ) to 0.001 (by PO-IPW with the independent working covariance matrix) in estimating β_{11} and from 0.040 and 0.037 to 0.025 in estimating β_{12}. In addition to large biases, the Greenwood variance estimator by Geskus's method yields substantial underestimation of the variations of regression parameter estimates, resulting in large undercoverage. The GEE sandwich variance estimator by the PO-IPW, on the other hand, provides satisfactory measure of the variations of estimates. The coverage probabilities of PO-IPW estimator are much closer to the nominal level of 95%. The independence and the empirical working covariance matrices give similar results, although the estimates are slightly more variable under the empirical working covariance matrix. Results in the supplementary table B.2 show a similar good performance of the PO-IPW.

2.4.3. Simulations under Covariate-Dependent Truncation and Covariate-Dependent Censoring

Lastly, we evaluate the performance of the proposed method under covariate-dependent truncation and covariate-dependent censoring. The truncation time is generated from an exponential distribution with hazard function $\lambda_L = \lambda_{L_0} \exp(0.5Z_1 + 0.3Z_2)$, where λ_{L_0} is set to different values to vary the truncation rate (i.e., truncation rates of 30% and 50%). The censoring time follows an exponential distribution with parameter $\lambda_C = \lambda_{C_0} \exp(1.5Z_1 + Z_2)$, where λ_{C_0} is such that the censoring rate is 30% or 50%.

Table 2.5 shows that the PO-IPW reduces the bias remarkably after accounting for covariate-dependent truncation and censoring. The Geskus's method and PO-AJ result in considerably larger bias, likely due to assuming that left truncation and right censoring distributions are independent of covariates. For instance, in Table 2.5, under the scenario of $n = 500$, censoring rate of 50% and truncation rate of 50%, the absolute value of the bias is reduced from 0.164 (by Geskus's method) and 0.168 (by PO-AJ) to 0.003 (by PO-IPW with the independent working covariance matrix) in estimating of β_{11} and from 0.111 and 0.113 to 0.001 in estimating β_{12}. The coverage probabilities by the PO-IPW are generally much closer to the nominal level of 95%, comparing to Geskus's method or PO-AJ. The independence and the empirical working covariance matrices give similar results. It is noticed that the estimated SEs and SDs by the PO-IPW can be relatively large, especially under heavy censoring and truncation. This is because the IPW estimator of the cumulative incidence function is affected by the variations in estimating truncation and censoring distributions given covariates. This variability is carried on in constructing POs, and thus the estimates of regression parameters become more variable. Such variation, however, can be decreased with the increased sample size. Results in the supplementary table B.3 also show a good performance of the PO-IPW.

2.5. Application

We illustrate the proposed methods by two data applications: a cohort study on HIV disease progression and a cohort study on pregnancy exposed to coumarin derivatives.

2.5.1. CCR5 Genotypes on HIV progression

We analyze the data from 329 homosexual men from the Amsterdam Cohort Studies on HIV infection and AIDS [15]. This dataset was used as an example of left-truncated right-censored competing risks data in Putter et al. [44] and in Geskus [17] and is publicly available in the `R` package `mstate`. During the process of HIV infection, the syncytium-inducing (SI) HIV phenotype often appears. The presence of the SI phenotype has been associated with rapid disease progression in HIV infected individuals because the SI variants are generally more cytopathic to T-cells and cause the infected T-cells to fuse to healthy ones [18]. AIDS and SI are the two event types in this data, where they compete to be the first to occur [17]. Clinical research has shown that a 32-bp deletion in the C-C chemokine receptor 5 gene (CCR5-$\triangle$32) is associated with delayed HIV disease progression [50]. As in Putter et al. [44], the primary goal is to investigate whether SI appears more rapidly in patients with CCR5-$\triangle$32 deletion. We assess the effect of CCR5 on the risk of "SI appearance before AIDS", where "AIDS before SI appearance" is a competing event. In the dataset, the CCR5 genotype is classified as "WW" (W stands for wild-type) for patients without the deletion and as "WM" for those who have the deletion on one of the chromosomes (M stands for mutation). Patients with the deletion on both chromosomes were not collected in the data [44]. The truncation time (in years) is from HIV infection to study enrollment, and the failure times are from HIV infection to SI appearance before AIDS and from HIV infection to AIDS before SI appearance. Right censoring is due to the end of study. We directly model the effect of CCR5 on cumulative incidence of SI appearance before AIDS and that on AIDS before SI appearance by the proposed method

and compare it with Geskus's method.

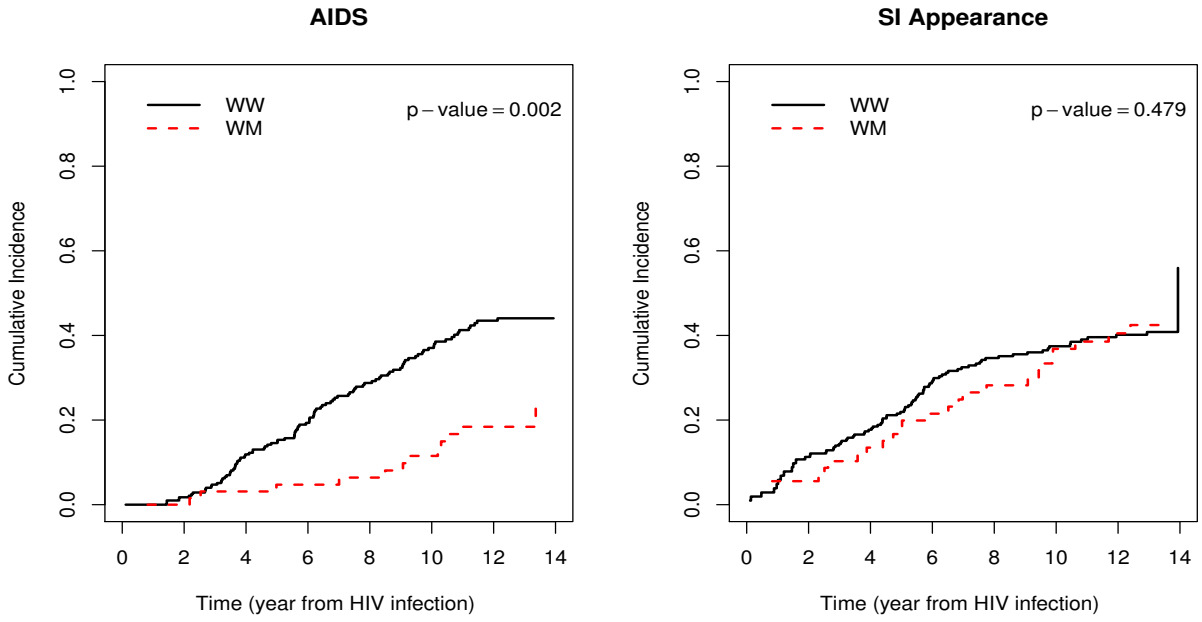

Figure 2.1: Estimated cumulative incidences of AIDS (left) and SI appearance (right) for wild-type (WW) and mutant (WM) CCR5 genotypes.

Among 324 patients with CCR5 genotype information, 259 (80%) had the wild-type variant (WW) and 65 (20%) had the mutant variant (WM). A total of 113 (35%) of patients experienced AIDS before SI appearance, 107(33%) experienced SI appearance before AIDS diagnosis, and 104 (32%) were censored by the end of the study. Preliminary Cox regression analyses of censoring and truncation given CCR5 genotype suggest that both the censoring time and truncation time are independent of CCR5 genotype. Thus, the proposed method that assumes covariate-independent truncation and covariate-independent censoring is used in the analysis. Figure 2.1 presents the nonparametric estimations of cumulative incidence functions by the CCR5 genotype for the time to AIDS before SI appearance and the time to SI appearance before AIDS, respectively, by the Aalen-Johansen estimator adjusting for left truncation. It is observed that the mutant genotype has a significant protective effect on AIDS before SI appearance ($P = 0.002$). The

effect of CCR5 on SI appearance before AIDS is not significant ($P = 0.479$). The cumulative incidence of SI appearance before AIDS is lower for the mutant genotype initially. Then, the two cumulative incidence curves overlapped after approximately 10 years and crossed at approximately 12 years.

We use the complementary log-log link function on $1 - F_j(t)$ for the regression models of cumulative incidence functions, corresponding to the Fine-Gray subdistribution hazards models, and include a binary variable of CCR5 as a covariate. Table 2.1 summarizes the regression analysis results by the PO-AJ and PO-IPW. Overall, the PO-AJ and PO-IPW give almost identical estimated covariate effects, probably because the adjusted Aalen-Johansen estimator and IPW estimator are equivalent under covariate-independent truncation and covariate-independent censoring [16]. These results suggest that the mutant CCR5 genotype is significantly associated with a lower incidence of AIDS before SI appearance, whereas the effect of CCR5 genotype on SI appearance before AIDS is not significant. Specifically, by the PO-IPW, the estimated subdistribution hazard ratio (SubHR) for AIDS before SI appearance is 0.3115 (P = 0.0023) comparing mutant CCR5 genotype to wile-type genotype, which indicates that the mutant CCR5 genotype is associated with an approximately 69% decrease in the expected subdistribution hazard of AIDS before SI appearance.

Table 2.1: Estimated subdistribution hazard ratio for the mutant (WM) effect on AIDS and syncytium-inducing (SI) appearance.

	PO-AJ			PO-IPW		
Cause	$e^{\hat{\beta}}$	CI	p	$e^{\hat{\beta}}$	CI	p
AIDS	0.3117	(0.1474, 0.6590)	0.0022	0.3115	(0.1473, 0.6588)	0.0023
SI appearance	0.8246	(0.4834, 1.4065)	0.4790	0.8247	(0.4835, 1.4065)	0.4792

$e^{\hat{\beta}}$: estimated subdistribution hazard ratio.

CI: 95% confidence interval.

p: p-value.

Geskus's method gives a comparable result with slightly different magnitude of estimated CCR5 effect on AIDS (SubHR = 0.3663, P = 0.0004), while the effect of mutant CCR5 genotype on SI is reversed (SubHR = 1.0598, P = 0.7890) comparing to that by the PO-AJ and PO-IPW, although the effect is far from significant by any of the methods. Nevertheless, the protective trend of mutant CCR5 genotype on SI suggested by the PO-AJ and PO-IPW is more consistent with the nonparametric cumulative incidence curves in Figure 2.1.

2.5.2. Pregnancy Exposed to Coumarin Derivatives

We use the data from a prospective cohort study on pregnancy exposed to coumarin derivatives, collected by a Germany Teratology Information Service (TIS), and the data are publicly available in the `R` package `etm`. Coumarin derivatives are used to inhibit formation of blood clots, and TIS provides risk assessment and treatment recommendations to pregnant women and their health-care providers [38]. There are three competing risks in the data: induced abortion, spontaneous abortion, and live birth. Typically, women contact TIS when the pregnancy is recognized and a drug risk assessment is needed, which is usually several weeks after conception, and thus, the data are left truncated by the first time of contacting TIS and those who had early abortion were excluded from the study. Moreover, Stegherr et al. [52] noted a potential dependence between left-truncation time and event time because women who considered an induced abortion were more likely to contact and seek advice from TIS. Ning et al. [39] developed a formal conditional independence test of failure and truncation times and applied the test to this data. They found that the dependence between truncation and event times was likely due to their common correlations with the exposure to coumarin derivatives, and the truncation and event times were independent conditional on the medication intake. Therefore, the proposed method that accounts for covariate-dependent truncation would be more appropriate. We analyzed the data to investigate effects of exposure to coumarin derivatives on induced abortion and spontaneous abortion,

respectively, by directly modeling coumarin on cumulative incidences using the proposed method.

Since the pregnancy outcome data are available for all the woman, there is no right censoring. As in Stegherr et al. [52], we added a small uniform random noise to break ties because the original time scale in the data is gestational age (in weeks). Among 1186 pregnant women who contacted TIS during their pregnancies, 173 (15%) were exposed to coumarin derivatives and 1013 (85%) were not. A total of 58 (5%) of women had an induced abortion, 112 (9%) of women had a spontaneous abortion, and 1016 (86%) experienced a live birth. Figure 2.2 presents the nonparametric estimations of cumulative incidence functions by the coumarin exposure for the time to an induced abortion and the time to a spontaneous abortion, respectively, based on the IPW estimator adjusting for dependent truncation.

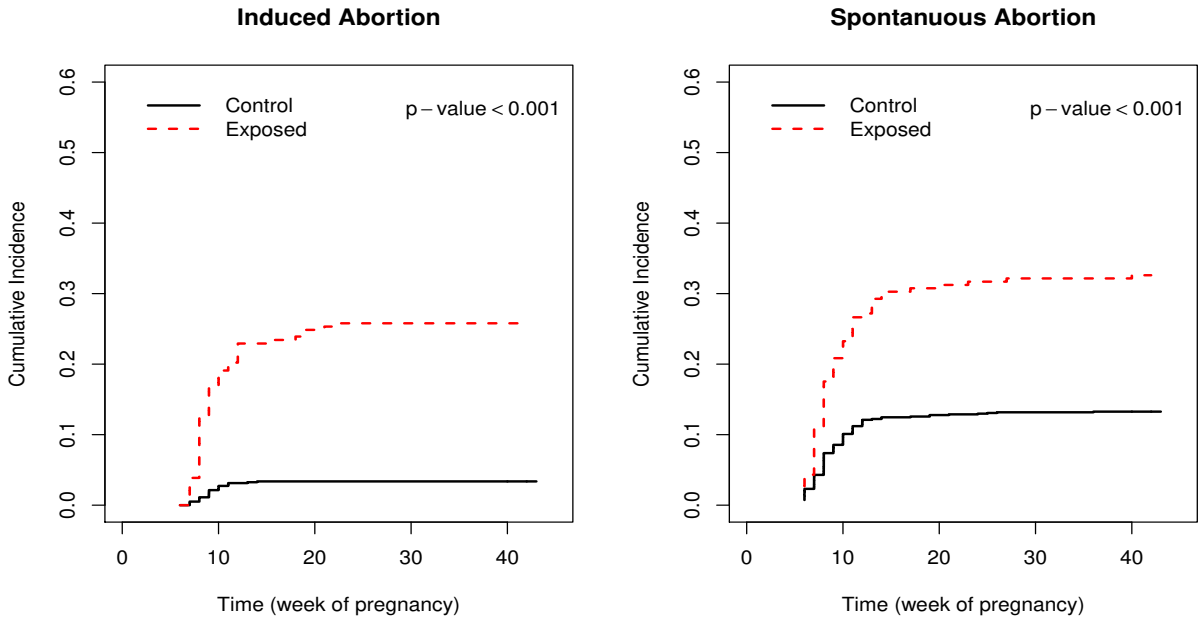

Figure 2.2: Estimated cumulative incidences of induced abortion (left) and spontaneous abortion (right) for controlled and exposed patients.

It shows that the cumulative incidences of induced abortion (P $<$ 0.001) and spontaneous abortion (P $<$ 0.001) are significantly higher for coumarin exposed women. The high incidence of induced abortion for exposed women mostly occurred during the first 12 weeks (first trimester), and the coumarin exposure appears to have a stronger impact induced abortion comparing to that on the spontaneous abortion. We use the complementary log-log link function on $1 - F_j(t)$ in the regression models of cumulative incidence functions for induced abortion and spontaneous abortion and include a binary variable of coumarin exposure as a covariate. Table 2.2 summarizes the regression analysis results by the PO-AJ and PO-IPW, in which dependent truncation is adjusted for by the PO-IPW but not PO-AJ. The results based on the PO-IPW suggest that the exposure to coumarin derivatives is significantly associated with higher incidences of induced abortion and spontaneous abortion. Specifically, by the PO-IPW, the estimated subdistribution hazard ratio for induced abortion is 9.3375 (P $<$ 0.0001) comparing the coumarin exposed women to unexposed women, which indicates that coumarin derivatives intake is associated with an approximately 8 times increase in the expected subdistribution hazard of induced abortion. The estimated subdistribution hazard ratio for spontaneous abortion is 2.7649 (P $<$ 0.0001) comparing the exposed women to unexposed ones, indicating that exposing to coumarin derivatives is associated with an approximately 1.8 times increase in the expected subdistribution hazard of spontaneous abortion. The PO-AJ that assuming covariate-independent truncation appears to overestimate coumarin effects on both induced abortion and spontaneous abortion.

Table 2.2: Estimated hazard ratio for the coumarin derivatives effect on pregnancies with different competing risks.

	PO-AJ			PO-IPW		
Cause	$e^{\hat{\beta}}$	CI	p	$e^{\hat{\beta}}$	CI	p
Induced abortion	9.7761	(5.1473, 18.5673)	$<$0.0001	9.3375	(5.0786, 17.1681)	$<$0.0001
Spontaneous abortion	3.0930	(1.6201, 5.9049)	$<$0.0001	2.7649	(1.4211, 5.3795)	$<$0.0001

2.6. Discussion

Much research has been conducted into regression modeling of cumulative incidence function for left-truncated right-censored competing risks data. However, existing work either extend the inverse probability of censoring-based method in Fine and Gray [13] to left-truncated data by constructing complicated weighted estimating equations, or involve sophisticated nonparametric conditional likelihood function. Moreover, many of these methods require the assumption that censoring and/or truncation times are independent of failure time. Alternatively, the PO approach has been employed in regression modeling of cumulative incidence function for right-censored competing risks data under covariate-independent censoring [31] or covariate-dependent censoring [8]. Yet this approach has not been extended to regression analysis of cumulative incidence function for left-truncated right-censored competing risks data, probably due to the additional analytical complexity posed by left truncation in the presence of competing risks In this paper, we address the methodological challenges by proposing direct regression modeling of cumulative incidence function based on POs, under general censoring and truncation mechanisms. The appealing features of the proposed methods include: first, we provide a flexible and robust way to investigate the association between cumulative incidence function and covariates for left-truncated right-censored competing risks data through a generalized linear model with a link function. Fine-Gray subdistribution hazards model is a special case with the complementary log-log link on $1 - F_j(t)$, but the proposed methods do not require the proportional subdistribution hazards assumption; second, the proposed methods handle the complex issues of potentially covariate-dependent left truncation and right censoring in competing risks setting in the first step of computing the POs, and allow one to apply GEE techniques with standard statistical software to obtain regression parameter estimates and corresponding standard errors in the subsequent analysis; third, we take into account various truncation and censoring mechanisms by modeling the conditional truncation and censoring distributions given covariates and incorporating appropriate

weights into the IPW estimator of cumulative incidence function; fourth, besides covariate effect estimates, the proposed methods also provide estimates of baseline cumulative incidence functions at the grid points; lastly, existing methods for regression analysis of competing risks data based on POs often lack rigorous theoretical justification [8, 31]. We fill the gap in literature and establish the asymptotic properties of proposed PO-based estimators for regression modeling of left-truncated right-censored competing risks data, under reasonable assumptions and regularity conditions.

Under covariate-independent truncation and covariate-independent censoring, Geskus [16] proved the equivalence of the adjusted Aalen-Johansen estimator and IPW estimator. Such equivalence is also shown by the results of our simulation studies and data application. Under covariate-dependent censoring and/or covariate-dependent truncation, the Cox proportional hazards model is used as a working model for censoring and/or truncation distributions given covariates to compute covariate-adjusted weights for replacing the product-limit statistics in the IPW estimator, although other parametric or semiparametric models (e.g., generalized transformation models, accelerated failure time model) may be used. The Cox model of the censoring time is straightforward by reversing the role of failure time and censoring time; however, modeling the truncation time is much more challenging due to right truncation. When data are left truncated, the standard Cox's regression method can be readily modified by only including subjects who have entered the study because $T \geqslant t$ is a subset of $T \geqslant L$ when $t \geqslant L$. However, this relationship does not hold for right-truncated data where sampling is restricted to subjects whose failure times are less than or equal to the right truncation times. Thus, specialized methods are required for analysis of right-truncated data [37]. We adopted the EM algorithm in Rennert and Xie [46] for estimation in the Cox model of the truncation time, under covariate-dependent right truncation. Under covariate-dependent censoring and/or covariate-dependent truncation, the proposed methods rely on correct specifications of models for censoring and truncation distributions given covariates. When the information concerning true models for censoring and truncation is absent, a sensitivity analysis would provide important insight in the validity

of the corresponding model assumptions and robustness of the proposed methods.

The choice of grid points depends on the pattern of the observed data. Klein and Andersen [31] recommended that between 5 and 10 grid points are sufficient to provide reliable estimation of the regression parameters. We select six grid points that are equally spaced on the event scale to compute POs in simulation studies and data applications. The simulations suggest a satisfactory performance with small biases, when POs are computed at these grid points. While the data reduction induced by using a limited number of grid points may cause a loss of efficiency of the proposed estimators [8], as shown by the relatively larger empirical standard deviations and estimated model-based standard errors produced by the PO-AJ and PO-IPW. The proposed methods accommodate various link functions, regression models, and working covariance matrices. As an illustration, in simulations and data applications, the regression models with a complementary log-log link function on $1 - F_j(t)$ are considered and compared with the GeskusâĂŹs method for the Fine-Gray model with independent left-truncated right-censored competing risks data. Other link functions, such as logit link function and complementary log-log link function on $F_j(t)$ can be used. The choice of a link function depends on the actual data and scientific relevance in a specific application. Methods based on pseudo-residuals have been proposed for graphical goodness-of-fit assessment of regression models for right-censored data [43]. Future research is intended to develop rigorous residual-based goodness-of-fit tests for selecting an appropriate link function in regression modeling of cumulative incidence function for left-truncated right-censored competing risks data. The independence and empirical working covariance matrices give similar results in the simulation studies; thus, we would generally recommend the simple independence working covariance matrix in the current setting. The proposed methods can be extended to regression analyses of other types of complex survival data, such as clustered left-truncated right-censored data, clustered left-truncated right-censored competing risks data, and recurrent event data. Logan et al. [36] proposed to model the marginal cumulative incidence function for clustered right-censored competing risks data based on POs. For

clustered left-truncated right-censored competing risks data, we could first generate POs while accounting for left truncation and right censoring in the presence of competing risks, and then use them as responses in generalized estimation equations, where the within cluster correlation is adjusted by the sandwich variance estimator.

Table 2.3: Simulation results under covariate-independent truncation and covariate-independent censoring ($\beta_{11} = 0.5, \beta_{12} = 0.5$).

				IPW				PO-AJ				PO-IPW			
n	cen%	trun%	Para	Bias	SD	SE	CP	Bias	SD	SE	CP	Bias	SD	SE	CP
350	30%	30%	β_{11}	0.001	0.171	0.132	0.864	-0.002	0.301	0.293	0.953	0.016	0.303	0.296	0.967
			β_{12}	0.005	0.091	0.073	0.888	0.014	0.174	0.178	0.953	0.014	0.166	0.175	0.948
		50%	β_{11}	0.009	0.224	0.133	0.795	-0.006	0.337	0.315	0.954	0.008	0.353	0.359	0.976
			β_{12}	0.012	0.104	0.074	0.842	0.009	0.180	0.188	0.961	0.010	0.196	0.216	0.963
350	50%	30%	β_{11}	-0.002	0.187	0.165	0.921	0.021	0.332	0.323	0.960	0.017	0.312	0.316	0.963
			β_{12}	0.003	0.100	0.088	0.923	0.028	0.193	0.194	0.958	0.027	0.189	0.193	0.957
		50%	β_{11}	0.002	0.206	0.164	0.901	0.025	0.331	0.326	0.968	0.023	0.350	0.335	0.962
			β_{12}	0.011	0.109	0.089	0.885	0.029	0.193	0.200	0.967	0.025	0.218	0.209	0.968
500	30%	30%	β_{11}	-0.001	0.157	0.110	0.837	0.020	0.252	0.248	0.957	0.003	0.261	0.252	0.960
			β_{12}	0.001	0.078	0.061	0.886	0.009	0.146	0.147	0.964	0.008	0.149	0.153	0.961
		50%	β_{11}	0.006	0.169	0.110	0.816	-0.007	0.272	0.267	0.958	0.003	0.290	0.270	0.970
			β_{12}	0.011	0.088	0.062	0.830	-0.004	0.150	0.161	0.952	-0.002	0.169	0.164	0.966
500	50%	30%	β_{11}	0.002	0.159	0.138	0.923	0.031	0.285	0.277	0.956	0.029	0.270	0.269	0.965
			β_{12}	0.008	0.086	0.074	0.913	0.027	0.180	0.172	0.965	0.035	0.157	0.163	0.958
		50%	β_{11}	0.009	0.178	0.137	0.881	0.039	0.269	0.277	0.970	0.016	0.277	0.280	0.966
			β_{12}	0.005	0.089	0.074	0.897	0.021	0.158	0.168	0.961	0.006	0.162	0.169	0.961

Table 2.4: Simulation results under covariate-independent truncation and covariate-dependent censoring ($\beta_{11} = 0.5, \beta_{12} = 0.5$).

				IPW				PO-AJ				PO-IPW				PO-IPW(empirical)			
n	cen%	trun%	Para	Bias	SD	SE	CP	Bias	SD	SE	CP	Bias	SD	SE	CP	Bias	SD	SE	CP
350	30%	30%	β_{11}	-0.089	0.183	0.144	0.846	-0.063	0.304	0.297	0.897	-0.001	0.320	0.364	0.970	0.003	0.380	0.361	0.956
			β_{12}	-0.040	0.103	0.081	0.846	-0.037	0.191	0.182	0.897	-0.025	0.240	0.259	0.952	-0.012	0.270	0.247	0.935
		50%	β_{11}	-0.112	0.222	0.153	0.792	-0.048	0.329	0.336	0.952	-0.008	0.405	0.461	0.970	-0.040	0.414	0.442	0.974
			β_{12}	-0.063	0.128	0.084	0.762	-0.035	0.228	0.212	0.924	-0.057	0.280	0.315	0.955	-0.039	0.283	0.303	0.946
350	50%	30%	β_{11}	-0.088	0.217	0.191	0.892	-0.103	0.427	0.404	0.913	0.004	0.546	0.589	0.964	-0.036	0.592	0.595	0.964
			β_{12}	-0.069	0.121	0.101	0.838	-0.041	0.275	0.255	0.893	-0.019	0.363	0.403	0.920	-0.010	0.426	0.410	0.924
		50%	β_{11}	-0.132	0.246	0.205	0.850	-0.083	0.463	0.446	0.926	-0.002	0.564	0.711	0.969	-0.054	0.689	0.770	0.976
			β_{12}	-0.094	0.129	0.106	0.801	-0.046	0.267	0.269	0.889	-0.059	0.401	0.488	0.932	-0.070	0.492	0.511	0.914
500	30%	30%	β_{11}	-0.072	0.154	0.119	0.827	-0.089	0.267	0.254	0.913	0.003	0.326	0.322	0.971	-0.001	0.288	0.292	0.972
			β_{12}	-0.047	0.086	0.066	0.803	-0.051	0.155	0.155	0.904	-0.011	0.269	0.230	0.958	-0.009	0.219	0.203	0.954
		50%	β_{11}	-0.120	0.187	0.127	0.745	-0.052	0.326	0.293	0.955	0.001	0.302	0.377	0.972	<0.001	0.478	0.441	0.956
			β_{12}	-0.073	0.111	0.069	0.719	-0.030	0.177	0.183	0.924	-0.028	0.191	0.265	0.960	-0.011	0.371	0.305	0.944
500	50%	30%	β_{11}	-0.091	0.188	0.159	0.861	-0.081	0.373	0.349	0.915	0.004	0.451	0.515	0.965	-0.013	0.549	0.524	0.959
			β_{12}	-0.059	0.106	0.085	0.833	-0.068	0.247	0.218	0.865	-0.023	0.333	0.374	0.909	0.006	0.403	0.367	0.937
		50%	β_{11}	-0.129	0.207	0.170	0.841	-0.071	0.397	0.400	0.933	-0.055	0.525	0.656	0.963	-0.067	0.594	0.678	0.959
			β_{12}	-0.095	0.109	0.088	0.727	-0.021	0.272	0.252	0.913	-0.058	0.356	0.453	0.935	-0.080	0.385	0.449	0.924

Table 2.5: Simulation results under covariate-dependent truncation and covariate-dependent censoring ($\beta_{11} = 0.5, \beta_{12} = 0.5$).

				IPW				PO-AJ				PO-IPW				PO-IPW(empirical)			
n	cen%	trun%	Para	Bias	SD	SE	CP	Bias	SD	SE	CP	Bias	SD	SE	CP	Bias	SD	SE	CP
350	30%	30%	β_{11}	-0.113	0.173	0.142	0.819	-0.107	0.250	0.257	0.916	-0.024	0.292	0.309	0.955	-0.040	0.309	0.309	0.959
			β_{12}	-0.059	0.095	0.080	0.822	-0.089	0.138	0.152	0.857	-0.035	0.184	0.220	0.943	-0.049	0.249	0.215	0.920
		50%	β_{11}	-0.141	0.199	0.143	0.747	-0.112	0.298	0.292	0.925	0.023	0.305	0.356	0.976	-0.006	0.483	0.399	0.962
			β_{12}	-0.081	0.113	0.082	0.757	-0.083	0.176	0.176	0.880	-0.009	0.225	0.261	0.966	-0.012	0.338	0.275	0.960
350	50%	30%	β_{11}	-0.132	0.222	0.188	0.858	-0.138	0.378	0.366	0.915	0.004	0.521	0.517	0.963	-0.035	0.556	0.519	0.964
			β_{12}	-0.086	0.120	0.102	0.814	-0.094	0.253	0.221	0.819	-0.029	0.362	0.367	0.919	-0.042	0.406	0.364	0.924
		50%	β_{11}	-0.159	0.236	0.195	0.816	-0.171	0.338	0.338	0.907	-0.003	0.532	0.589	0.981	-0.013	0.620	0.623	0.976
			β_{12}	-0.108	0.133	0.105	0.741	-0.132	0.189	0.201	0.785	-0.051	0.362	0.417	0.934	-0.025	0.459	0.415	0.922
500	30%	30%	β_{11}	-0.107	0.146	0.118	0.799	-0.117	0.218	0.216	0.901	-0.013	0.245	0.268	0.975	-0.024	0.263	0.262	0.960
			β_{12}	-0.064	0.086	0.067	0.755	-0.089	0.123	0.128	0.807	-0.028	0.189	0.197	0.962	-0.038	0.201	0.183	0.920
		50%	β_{11}	-0.138	0.166	0.119	0.722	-0.108	0.247	0.241	0.922	0.028	0.309	0.319	0.974	0.039	0.337	0.325	0.969
			β_{12}	-0.083	0.097	0.068	0.692	-0.081	0.147	0.150	0.864	$<$0.001	0.224	0.236	0.979	0.009	0.269	0.232	0.957
500	50%	30%	β_{11}	-0.129	0.173	0.157	0.847	-0.147	0.360	0.310	0.911	0.005	0.403	0.436	0.967	-0.011	0.487	0.460	0.972
			β_{12}	-0.086	0.102	0.084	0.766	-0.112	0.221	0.181	0.787	-0.038	0.297	0.321	0.913	-0.023	0.349	0.326	0.915
		50%	β_{11}	-0.164	0.183	0.162	0.807	-0.168	0.289	0.284	0.881	0.003	0.453	0.517	0.967	-0.033	0.332	0.369	0.948
			β_{12}	-0.111	0.109	0.087	0.722	-0.113	0.152	0.173	0.811	0.001	0.546	0.564	0.970	-0.005	0.427	0.395	0.939

APPENDIX A

APPENDIX of CHAPTER 1

In this Appendix, Appendix A.1 provides a proof and regularity conditions of Theorem 1 and Appendix A.2 presents additional simulations under various settings in Chapter 1.

A.1. Proof of Theorem 1

The following regularity conditions are introduced to establish the asymptotic properties of $\hat{\beta}_\tau$:

C1: For a prespecified time τ, $P\left(A + C \geq \tau\right) > 0$.

C2: The baseline covariates $\boldsymbol{X}$ are bounded almost surely.

C3: The matrix $\mathcal{I}\left(\beta_{\tau 0}\right)$ is positive definite.

In order to use the general theorem for GEE [34] to develop the asymptotic properties of $\hat{\beta}_\tau$, the equation (B.1) needs to be unbiased at the true parameter value $\beta_{\tau 0}$, $E\left[U\left(\beta_{\tau 0}\right)\right] = 0$, in addition to the above regularity conditions. This requires the following "asymptotic unbiasedness" of the POs [21]:

$$E\left[\hat{\mu}_i\left(\tau\right) \mid \boldsymbol{Z}_i\right] = g^{-1}\left(\boldsymbol{Z}_i^\top \boldsymbol{\beta}_\tau\right) + \mathrm{o}_p(1).$$

It holds without the remainder term if all failure times are uncensored and untruncated because the POs, $\hat{\mu}_i\left(\tau\right)$ $(i = 1, ..., n)$, are exactly equal to $\tilde{T}_i \wedge \tau$ when the data are complete. In the presence of left truncation and right censoring, we show the "asymptotic unbiasedness" of the POs and asymptotic properties of $\hat{\beta}_\tau$, using techniques similar to those in Graw et al [21].

Proof. Let P denote the probability law of the vector of observed data Y_i and $P_n(\cdot) = n^{-1}\sum_{i=1}\mathcal{I}\left(Y_i \in \cdot\right), i = 1, ..., n$ denote the empirical law corresponding to the sample of left-truncated right-censored observations $Y_1, ...Y_n$. Further denote by $P_n^{(i)}$ the empirical distribution of the reduced sample $Y_1, ..., Y_{i-1}, Y_{i+1}, ..., Y_n$. The RMST functional ϕ operates on a set $\mathcal{P}$ of probability measures for Y_i that includes P and the empirical measures [19, 27]. It is defined such that $\phi(P) = \mu(\tau)$ is the parameter of interest and $\phi(P_n) = \hat{\mu}(\tau)$ the RMST estimate corresponding to the observed data $Y_1, ..., Y_n$. Thus, the POs can be expressed as $\hat{\mu}_i(\tau) = n\phi(P_n) - (n-1)\phi(P_n^{(i)})$. We use the von Mises expansion on a smooth statistical functional ϕ [19]:

$$\phi\left(P_n\right) = \phi(P) + n^{-1}\sum_{k=1}^{n}\dot{\phi}\left(Y_k\right) + \frac{1}{2}n^{-2}\sum_{k=1}^{n}\sum_{j=1}^{n}\ddot{\phi}\left(Y_k, Y_j\right) + \mathrm{O}_P\left(n^{-\frac{3}{2}}\right), \tag{A.1}$$

where $\dot{\phi}$ and $\ddot{\phi}$ are the first and second order influence functions [23] of the functional ϕ. The first order influence function is centered, $E\{\dot{\phi}(Y_i)\} = 0$ [26]. The second order influence function is symmetric, $\ddot{\phi}(Y_i, Y_j) = \ddot{\phi}(Y_j, Y_i)$, and should satisfy for every t [55]

$$E\left\{\ddot{\phi}\left(Y_i, t\right)\right\} = \int \ddot{\phi}(y, t)\mathrm{d}P(y) = 0. \tag{A.2}$$

From the equation (A.1), we have:

$$n\phi\left(P_n\right) - (n-1)\phi\left(P_n^{(i)}\right) = \phi(P) + \dot{\phi}\left(Y_i\right) + \frac{1}{n-1}\sum_{k=1}^{n}\ddot{\phi}\left(Y_k, Y_i\right) + \mathrm{o}_P(1),$$

as shown in Graw et al. [21] By the law of large numbers, $\frac{1}{n-1}\sum_{k=1}^{n}\ddot{\phi}\left(Y_k, Y_i\right)$ converges to $E\left\{\ddot{\phi}\left(Y_i, t\right)\right\}$, which equals to 0 by the equation (A.2). Thus, for the smooth statistical functional ϕ with a second order von Mises expansion as in (A.1) such that the equation (A.2) holds, the POs can be represented by:

$$n\phi\left(P_n\right) - (n-1)\phi\left(P_n^{(i)}\right) = \phi(P) + \dot{\phi}\left(Y_i\right) + \mathrm{o}_P(1).$$

For left-truncated right-censored data, the RMST estimate obtained by integrating out the product-limit estimator of the survival function is consistent, under the assumption that the residual censoring time C is independent of $(\tilde{T}, A, \boldsymbol{X})$. James [27] discussed the property of the second order influence function of the Kaplan-Meier functional and by similar arguments, we can show that the RMST functional ϕ also has the second order von Mises expansion as in (A.1). Essentially, we have shown that, under the regularity conditions and covariate-independent censoring assumption, the POs of RMST for left-truncated right-censored data can be represented as

$$\hat{\mu}_i(\tau) = \dot{\phi}\left(Y_i\right) + \mu(\tau) + \mathrm{o}_P(1) \tag{A.3}$$

This leads to important properties of POs as follows:

1. $\hat{\mu}_i(\tau)$ $(i = 1, ..., n)$ can be approximated by independent and identically distributed variables.
2. $E[\hat{\mu}_i(\tau)] = \mu(\tau) + \mathrm{o}_P(1)$, for all $i = 1, ..., n$.

Since any estimator of $\mu(\tau) = E[\tilde{T}_\tau]$ is also implicitly an estimator of $E_{\boldsymbol{Z}}\left[E(\tilde{T}_\tau \mid \boldsymbol{Z})\right]$, similarly, we have

$$E[\hat{\mu}_i(\tau) \mid \boldsymbol{Z}_i] = \mu(\tau \mid \boldsymbol{Z}_i) + \mathrm{o}_P(1), \text{ for all } i = 1, ..., n.$$

Therefore, $U\left(\beta_{\tau 0}\right)$ is an asymptotically unbiased estimation equation. Based on the equation (A.3), $U\left(\beta_\tau\right)$ can be approximated by a sum of independent and identical distributed random variables. Following the arguments in Graw et al. [21] and by Liang and Zeger [34], $\hat{\beta}_\tau$ is consistent to $\beta_{\tau 0}$, and $\sqrt{n}\left(\hat{\beta}_\tau - \beta_{\tau 0}\right)$ is asymptotically normal with mean zero and a covariance matrix that can be estimated using a standard 'sandwich' estimator, which takes the form

$$\hat{\Sigma} = \mathcal{I}\left(\hat{\beta}_\tau\right)^{-1} \widehat{\mathrm{var}}\left\{U\left(\beta_\tau\right)\right\} \mathcal{I}\left(\hat{\beta}_\tau\right)^{-1},$$

with

$$\mathcal{I}\left(\hat{\beta}_\tau\right) = \sum_i \left\{ \frac{\partial g^{-1}\left(\boldsymbol{Z}_i^\top \hat{\beta}_\tau\right)}{\partial \hat{\beta}_\tau} \right\}^\top \mathcal{V}_i^{-1} \left\{ \frac{\partial g^{-1}\left(\boldsymbol{Z}_i^\top \hat{\beta}_\tau\right)}{\partial \hat{\beta}_\tau} \right\},$$

$$\widehat{\text{var}}\left\{U\left(\beta_\tau\right)\right\} = \sum_i U_i\left(\hat{\beta}_\tau\right) U_i\left(\hat{\beta}_\tau\right)^{\text{T}}.$$

□

A.2. Additional Simulations and Supplementary Tables

We conduct additional simulations to assess the performance of the proposed methods for RMST model with the log link function. We randomly assign each subject to two groups, A and B, with equal probability. Group A is treated as the reference. The covariate X_1 is binary and equals to 1 for subjects in group B and equals to 0 for subjects in group A. The assumed model for RMST is $\log\{\mu_\tau(x_1)\} = \log\left\{\mathrm{E}\left[\tilde{T}_\tau \mid X_1 = x_1\right]\right\} = \beta_{\tau 0} + \beta_{\tau 1} x_1$ with the log link function. The failure time data are generated both under proportional hazards and under non-proportional hazards. For each scenario, the simulation was repeated 1000 times. First, we evaluate the performance of the proposed method under covariate-independent censoring and with the log link function. The data generating process is similar to the case with the linear link in Section 2.3.1. Table B.1 summarizes the simulation results. Second, we evaluate the proposed methods under covariate-dependent censoring and with the log link function. Under proportional hazards, the failure time $\tilde{T}$ is generated from a distribution with hazard function $\lambda\left(t \mid X_1 = x_1\right) = exp\left(\gamma x_1\right)$, where $\gamma = 0.5$. The residual censoring time is generated from an exponential distribution with parameter $\lambda_C = \lambda_{C_0}\exp(4X_1)$. Varying λ_{C_0} allows for various levels of censoring (i.e., censoring rates of 30% and 45%). Under non-proportional hazards, the failure time is generated from a distribution with hazard function $\lambda\left(t \mid \boldsymbol{Z} = \boldsymbol{z}\right) = \exp\left\{-\left(\boldsymbol{z}^\top \boldsymbol{\gamma}\right) + \boldsymbol{z}^\top \boldsymbol{\zeta} \log(8t)\right\}$, where $\boldsymbol{Z} = (1, X_1)^\top$, $\boldsymbol{\gamma} = (0.5, 1)^\top$ and $\boldsymbol{\zeta} = (1, -0.3)^\top$. The residual censoring time is generated from an exponential distribution with parameter $\lambda_C = \lambda_{C_0}\exp\left(X_1\right)$, where λ_{C_0} is such that the censoring rate is 30% or 45%. The truncation variable follows the same Weibull distribution as in Section 2.3.1, with a truncation rate of 30%. Table B.2 summarizes the simulation results.

Table B.3 presents the simulation results where the "conditional survival function" approach is used to adjust for the bias observed at a larger $\tau = 1.39$.

Table A.1: Simulation results under covariate-independent censoring and with log link function.

Proportional Hazards												
				30% Censoring Rate					45% Censoring Rate				
n	τ		True	RB	SD	SE	CP	MSE	RB	SD	SE	CP	MSE
350	0.69	β_0	-0.693	0.015	0.089	0.068	0.900	0.005	0.009	0.085	0.068	0.904	0.005
		β_1	-0.191	-0.015	0.147	0.144	0.967	0.021	0.019	0.154	0.144	0.957	0.021
	1.39	β_0	-0.288	0.059	0.102	0.079	0.928	0.007	0.042	0.099	0.078	0.917	0.006
		β_1	-0.320	0.021	0.213	0.184	0.967	0.034	0.031	0.207	0.201	0.963	0.040
500	0.69	β_0	-0.693	0.012	0.074	0.058	0.921	0.003	0.010	0.074	0.058	0.916	0.003
		β_1	-0.191	0.042	0.162	0.136	0.962	0.019	-0.001	0.120	0.120	0.964	0.014
	1.39	β_0	-0.288	0.043	0.082	0.066	0.924	0.005	0.045	0.082	0.065	0.928	0.004
		β_1	-0.320	0.034	0.168	0.146	0.964	0.021	0.013	0.161	0.146	0.955	0.021

Non-proportional Hazards												
				30% Censoring Rate					45% Censoring Rate				
n	τ		True	RB	SD	SE	CP	MSE	RB	SD	SE	CP	MSE
350	0.69	β_0	-0.700	0.032	0.059	0.062	0.965	0.004	0.033	0.062	0.064	0.969	0.005
		β_1	0.225	0.040	0.064	0.068	0.972	0.005	0.040	0.070	0.071	0.956	0.005
	1.39	β_0	-0.566	0.122	0.081	0.095	0.982	0.014	0.120	0.083	0.098	0.979	0.014
		β_1	0.571	0.075	0.092	0.103	0.984	0.012	0.071	0.093	0.106	0.980	0.013
500	0.69	β_0	-0.700	0.037	0.052	0.053	0.974	0.003	0.033	0.054	0.055	0.975	0.004
		β_1	0.225	0.059	0.057	0.059	0.975	0.004	0.048	0.060	0.060	0.974	0.004
	1.39	β_0	-0.566	0.125	0.067	0.079	0.950	0.011	0.131	0.095	0.088	0.950	0.013
		β_1	0.571	0.074	0.074	0.085	0.969	0.009	0.078	0.104	0.094	0.969	0.011

Table A.2: Simulation results under covariate-dependent censoring and with log link function ($n = 500$). Estimates obtained by using the traditional pseudo-observation (PO) approach and the inverse probability of censoring weighting (IPCW)-adjusted PO approach are compared.

Proportional Hazards													
				Unadjusted PO Method					IPCW-adjusted PO Method				
cen%	τ	True		RB	SD	SE	CP	MSE	RB	SD	SE	CP	MSE
30%	0.69	β_0	-0.693	-0.004	0.069	0.059	0.907	0.003	0.012	0.073	0.059	0.900	0.004
		β_1	-0.191	-0.022	0.128	0.120	0.957	0.014	-0.003	0.118	0.123	0.963	0.015
	1.39	β_0	-0.288	-0.022	0.081	0.070	0.916	0.005	0.014	0.077	0.064	0.903	0.004
		β_1	-0.320	-0.113	0.161	0.144	0.928	0.022	0.023	0.139	0.146	0.957	0.021
45%	0.69	β_0	-0.693	-0.014	0.076	0.065	0.909	0.004	-0.012	0.063	0.059	0.891	0.004
		β_1	-0.191	-0.139	0.129	0.121	0.932	0.015	0.001	0.118	0.122	0.959	0.015
	1.39	β_0	-0.288	-0.056	0.083	0.081	0.932	0.007	-0.022	0.092	0.071	0.889	0.005
		β_1	-0.320	-0.331	0.144	0.146	0.811	0.033	0.003	0.192	0.201	0.946	0.040

Non-proportional Hazards													
				Unadjusted PO Method					IPCW-adjusted PO Method				
cen%	τ	True		RB	SD	SE	CP	MSE	RB	SD	SE	CP	MSE
30%	0.69	β_0	-0.700	0.042	0.052	0.053	0.967	0.004	0.033	0.051	0.052	0.958	0.003
		β_1	0.225	0.053	0.056	0.058	0.968	0.004	0.034	0.058	0.057	0.955	0.003
	1.39	β_0	-0.566	0.185	0.069	0.079	0.839	0.017	0.148	0.067	0.078	0.942	0.013
		β_1	0.571	0.108	0.078	0.085	0.944	0.011	0.084	0.074	0.085	0.976	0.010
45%	0.69	β_0	-0.700	0.046	0.051	0.054	0.969	0.004	0.042	0.050	0.054	0.982	0.004
		β_1	0.225	0.045	0.057	0.059	0.976	0.004	0.057	0.056	0.059	0.974	0.004
	1.39	β_0	-0.566	0.233	0.076	0.084	0.751	0.024	0.158	0.068	0.083	0.925	0.015
		β_1	0.571	0.128	0.087	0.090	0.922	0.013	0.081	0.078	0.090	0.974	0.010

Table A.3: Simulation results under non-proportional hazards and covariate-independent censoring, with adjustment by the "conditional survival function" approach.

Linear Link												
				30% Censoring Rate					45% Censoring Rate				
n	τ		True	RB	SD	SE	CP	MSE	RB	SD	SE	CP	MSE
350	1.39	β_0	0.568	0.003	0.036	0.039	0.960	0.002	0.004	0.039	0.042	0.957	0.002
		β_1	0.437	0.007	0.048	0.050	0.967	0.003	0.008	0.051	0.053	0.961	0.003
500	1.39	β_0	0.568	0.007	0.033	0.032	0.945	0.001	0.006	0.034	0.035	0.961	0.001
		β_1	0.437	$<$0.001	0.042	0.041	0.942	0.002	0.001	0.043	0.045	0.956	0.002

Log Link												
				30% Censoring Rate					45% Censoring Rate				
n	τ		True	RB	SD	SE	CP	MSE	RB	SD	SE	CP	MSE
350	1.39	β_0	-0.566	-0.003	0.067	0.068	0.958	0.005	-0.012	0.070	0.074	0.956	0.006
		β_1	0.571	0.003	0.072	0.075	0.960	0.006	-0.004	0.077	0.081	0.966	0.007
500	1.39	β_0	-0.566	-0.007	0.054	0.057	0.957	0.003	-0.002	0.059	0.062	0.959	0.004
		β_1	0.571	-0.001	0.058	0.063	0.969	0.004	0.004	0.062	0.068	0.965	0.005

APPENDIX B

APPENDIX of CHAPTER 2

In this Appendix, Appendix B.1 provides a proof and regularity conditions of Theorem 2, Appendix B.2 provides a proof and regularity conditions of Theorem 3, and Appendix B.3 presents additional simulation results under various settings in Chapter 2.

B.1. Proof of Theorem 2

The following regularity conditions are introduced to establish the asymptotic properties of $\hat{\boldsymbol{\beta}}$:

C1: $\Pr(C > t) > 0$ (Positivity).

C2: The baseline covariates $\mathbf{Z}$ are bounded almost surely.

C3: The matrix $\mathcal{I}(\boldsymbol{\beta}_0)$ is positive definite.

In addition to the above regularity conditions, the equation (2.3) needs to be unbiased at the true parameter value $\boldsymbol{\beta}_0$, $E[U(\boldsymbol{\beta}_0)] = 0$, to be able to use the general theorem for GEE [34] to develop the asymptotic properties of $\hat{\boldsymbol{\beta}}$, which requires the following "asymptotic unbiasedness" of the POs [21]:

$$E\left[\hat{F}_{ij}(t) \mid \boldsymbol{Z}_i\right] = g^{-1}\left(\boldsymbol{Z}_i^{\top}\boldsymbol{\beta}\right) + \mathrm{o}_P(1).$$

It holds without the remainder term if all failure times are uncensored and untruncated because the POs, $\hat{F}_{ij}(t)$ $(i = 1, \ldots n)$, are exactly equal to $I(T_i \leqslant t, \varepsilon_i = j)$ when the data are complete. In the presence of left truncation and right censoring, we show the

"asymptotic unbiasedness" of the POs and asymptotic properties of $\hat{\beta}$, using techniques similar to those in [21].

Proof. Let P denote the probability law of the vector of observed data X_i and $P_n(\cdot) = n^{-1}\sum_{i=1} I\left(X_i \in \cdot\right), i = 1, ..., n$ denote the empirical law corresponding to the sample of left-truncated right-censored observations $X_1, ...X_n$. Further, let $P_n^{(-i)}$ denote the empirical distribution of the reduced sample $X_1, ..., X_{i-1}, X_{i+1}, ..., X_n$. The IPW functional Ω_j: $\mathcal{P} \to \mathcal{F}$ operates on a set $\mathcal{P}$ of probability measures for X_i that includes P and the empirical measures [19, 27] and maps into the set $\mathcal{F}$ of all sub-distribution functions. It is defined such that $\Omega_j(P) = F_j$ is the parameter of interest and $\Omega_j(P_n) = \hat{F}_j^{IPW}$ is the IPW estimate corresponding to the observed data $X_1, ..., X_n$. Thus, the POs can be expressed as $\hat{F}_{ij} = n\Omega_j(P_n) - (n-1)\Omega_j(P_n^{(-i)})$. Von Mises expansion is used on a smooth statistical functional Ω [19]:

$$\Omega\left(P_n\right) = \Omega(P) + n^{-1}\sum_{k=1}^{n}\dot{\Omega}\left(X_k\right) + \frac{1}{2}n^{-2}\sum_{k=1}^{n}\sum_{\zeta=1}^{n}\ddot{\Omega}\left(X_k, X_\zeta\right) + \mathrm{O}_P\left(n^{-\frac{3}{2}}\right), \tag{B.1}$$

where $\dot{\Omega}$ and $\ddot{\Omega}$ are the first and second order influence functions [23] of the functional Ω. The first order influence function is centered, $E\{\dot{\Omega}(X_i)\} = 0$ [26]. The second order influence function is symmetric, $\ddot{\Omega}(X_i, X_k) = \ddot{\Omega}(X_k, X_i)$, and should satisfy for every y [55]

$$E\left\{\ddot{\Omega}\left(X_k, y\right)\right\} = \int \ddot{\Omega}(x, y)\mathrm{d}P(x) = 0. \tag{B.2}$$

As shown in [21], from the equation (B.1), we have:

$$n\Omega\left(P_n\right) - (n-1)\Omega\left(P_n^{(-i)}\right) = \Omega(P) + \dot{\Omega}\left(X_i\right) + \frac{1}{n-1}\sum_{k=1}^{n}\ddot{\Omega}\left(X_k, X_i\right) + \mathrm{o}_P(1).$$

By the law of large numbers, $\frac{1}{n-1}\sum_{k=1}^{n}\ddot{\Omega}\left(X_k, X_i\right)$ converges to $E\left\{\ddot{\Omega}\left(X_k, y\right)\right\}$, which equals to 0 by the equation (B.2). Thus, for the smooth statistical functional Ω with a second order von Mises expansion as in (B.1) such that the equation (B.2) holds, the POs

can be represented by:

$$n\Omega\left(P_n\right)-(n-1)\Omega\left(P_n^{(-i)}\right)=\Omega(P)+\dot{\Omega}\left(X_i\right)+\mathrm{o}_P(1).$$

For left-truncated right-censored data, the IPW estimate is equivalent to the Aalen-Johansen estimate and is $n^{\frac{1}{2}}$-consistent, under the assumptions that (L, C) and (T, ε) are independent and that L and C are independent [16]. [27] discussed the property of the second order influence function of the product-limit Kaplan-Meier estimator functional and by similar arguments, we can show the IPW functional Ω_j also has the second order von Mises expansion as in (B.1). Essentially, we have shown that, under the regularity conditions and the assumptions of independence of (L, C) and (T, ε) and independence of L and C, the POs of IPW estimate for left-truncated right-censored data can be represented as

$$\hat{F}_{ij}(t)=\dot{\Omega}_j\left(X_i\right)+F_j(t)+\mathrm{o}_P(1) \tag{B.3}$$

This leads to important properties of POs as follows:

1. $\hat{F}_{ij}(t)\ (i=1,...,n)$ can be approximated by independent and identically distributed variables.
2. $E[\hat{F}_{ij}(t)]=F_j(t)+\mathrm{o}_P(1)$, for all $i=1,...,n$.

Since any estimator of $F_j\left(t\right)=E\left\{I\left(T\leqslant t,\varepsilon=j\right)\right\}$ is also implicitly an estimator of $E_{\boldsymbol{Z}}\left[E\left\{I\left(T\leqslant t,\varepsilon=j\right)|\boldsymbol{Z}\right\}\right]$, similarly, we have

$$E[\hat{F}_{ij}(t)\mid\boldsymbol{Z}_i]=F_j(t\mid\boldsymbol{Z}_i)+\mathrm{o}_P(1),\ \text{for all } i=1,...,n.$$

Therefore, $U\left(\boldsymbol{\beta}_0\right)$ is an asymptotically unbiased estimation equation. Based on equation (B.3), $U\left(\boldsymbol{\beta}\right)$ can be approximated by a sum of independent and identical distributed random variables. Following the arguments in [21] and by [34], $\hat{\boldsymbol{\beta}}$ is consistent to $\boldsymbol{\beta_0}$, and $\sqrt{n}\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0\right)$ is asymptotically normal with mean zero and a covariance matrix that can be

estimated using a standard 'sandwich' estimator, which takes the form

$$\hat{\Sigma} = \mathcal{I}\left(\hat{\boldsymbol{\beta}}\right)^{-1} \hat{\text{var}}\left\{U\left(\hat{\boldsymbol{\beta}}\right)\right\} \mathcal{I}\left(\hat{\boldsymbol{\beta}}\right)^{-1},$$

where

$$\mathcal{I}\left(\hat{\boldsymbol{\beta}}\right) = \sum_i \left\{\frac{\partial \theta_i}{\partial \boldsymbol{\beta}}\right\}^{\top} \mathcal{V}_i\left(\boldsymbol{\beta}\right)^{-1} \left\{\frac{\partial \theta_i}{\partial \boldsymbol{\beta}}\right\}\Bigg|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}},$$

$$\hat{\text{var}}\left\{U\left(\hat{\boldsymbol{\beta}}\right)\right\} = \sum_i U_i\left(\hat{\boldsymbol{\beta}}\right) U_i\left(\hat{\boldsymbol{\beta}}\right)^{\text{T}}.$$

□

B.2. Proof of Theorem 3

Regularity conditions C1 - C2 are assumed to show $\hat{F}_j^{IPW}(t)$ is a consistent estimator of the true CIF $F_j(t)$:

Proof. The counting process is defined as: $N_{ij}^L(t) = I\left\{L_i < X_i \leqslant t, \varepsilon_i = j, \delta_i = 1\right\}$, where $X = \min(T, C)$ and $\delta = I(T \leqslant C)$. Assuming that L is independent of (T, ε), C is conditionally independent of (T, ε) given $\mathbf{Z}$, and L is independent of C, we have

$$\begin{aligned} E\left\{N_{ij}^L(t) \mid T = t, \mathbf{Z} = \mathbf{z}\right\} &= E\left\{I\left\{L_i < X_i \leqslant t, \varepsilon_i = j, \delta_i = 1\right\} \mid T = t, \mathbf{Z} = \mathbf{z}\right\} \\ &= \Pr\left(\delta = 1 \mid T = t, \mathbf{Z} = \mathbf{z}\right) \times \Pr\left(T \leqslant t, \varepsilon = j \mid \mathbf{Z} = \mathbf{z}\right) \times \\ &\quad \Pr\left(L < t \mid \mathbf{Z} = \mathbf{z}\right). \end{aligned}$$

In the above expression, $\Pr\left(\delta = 1 \mid T = t, \mathbf{Z} = \mathbf{z}\right)$ = $\Pr\left(C \geqslant T \mid T = t, \mathbf{Z} = \mathbf{z}\right) = \Pr\left(C \geqslant t \mid \mathbf{Z} = \mathbf{z}\right) = S_C(t- \mid \mathbf{z})$, $\Pr\left(T \leqslant t, \varepsilon = j \mid \mathbf{Z} = \mathbf{z}\right) = F_j\left(t \mid \mathbf{z}\right)$, and $\Pr\left(L < t\right) = F_L(t-)$. Thus,

$$E\left\{N_{ij}^L(t) \mid T = t, \mathbf{Z} = \mathbf{z}\right\} = S_C(t- \mid \mathbf{z}) \times F_j\left(t \mid \mathbf{z}\right) \times F_L(t-) \tag{B.4}$$

since $\hat{S}_C\left(t \mid \mathbf{Z}\right)$ is a consistent estimator of $S_C\left(t \mid \mathbf{Z}\right)$ and $\hat{F}_L\left(t\right)$ is a consistent estimator of

F_L, together with equation (B.4), for any t, the limit of estimator $\hat{F}_j^{IPW}(t)$ is

$$\begin{aligned}
\lim_{n\to\infty} \hat{F}_j^{IPW}(t) &= \lim_{n\to\infty} \frac{1}{\hat{n}} \sum_{i=1}^{n} \frac{N_{ij}^L(t)}{\hat{S}_C\left(X_i- \mid \mathbf{Z}\right) \times \hat{F}_L\left(X_i-\right)} \\
&= \mathrm{E}_{\mathbf{Z}} \left\{ \int_0^t \frac{S_C(u- \mid \mathbf{Z}) F_L(u-)}{S_C(u- \mid \mathbf{Z}) F_L(u-)} F_j(\, \mathrm{d}u \mid \mathbf{Z}) \right\} \\
&= \mathrm{E}_{\mathbf{Z}} \left\{ F_j(t \mid \mathbf{Z}) \right\} \\
&= F_j(t)
\end{aligned}$$

Hence, we have proved that as $n \to \infty$, $F_j\left(t \mid \mathbf{Z}\right)$ converges in probability to the true cumulative incidence function $F_j(t)$ uniformly for any t. In other words, $\hat{F}_j^{IPW}(t)$ is a consistent estimator of $F_j(t)$.

□

B.3. Supplementary Table

Table B.1: Simulation results under covariate-independent truncation and covariate-independent censoring ($\beta_{11} = -1, \beta_{12} = 1$).

				IPW				PO-AJ				PO-IPW			
n	cen%	trun%	Para	Bias	SD	SE	CP	Bias	SD	SE	CP	Bias	SD	SE	CP
350	30%	30%	β_{11}	-0.020	0.242	0.157	0.808	0.002	0.437	0.428	0.972	-0.031	0.441	0.446	0.966
			β_{12}	0.024	0.128	0.098	0.856	0.031	0.338	0.327	0.971	0.032	0.324	0.328	0.973
		50%	β_{11}	-0.076	0.363	0.205	0.751	0.022	0.612	0.566	0.980	0.042	0.523	0.527	0.971
			β_{12}	0.052	0.198	0.131	0.796	-0.017	0.485	0.429	0.954	-0.032	0.419	0.385	0.952
350	50%	30%	β_{11}	-0.076	0.363	0.205	0.751	0.022	0.612	0.566	0.980	0.010	0.557	0.559	0.978
			β_{12}	0.052	0.198	0.131	0.796	-0.017	0.485	0.429	0.954	-0.015	0.463	0.431	0.966
		50%	β_{11}	-0.008	0.292	0.191	0.839	0.013	0.522	0.523	0.975	0.005	0.487	0.509	0.976
			β_{12}	0.022	0.162	0.118	0.863	0.007	0.450	0.400	0.963	0.010	0.379	0.386	0.955
500	30%	30%	β_{11}	-0.016	0.200	0.130	0.808	-0.046	0.353	0.372	0.980	-0.038	0.393	0.378	0.976
			β_{12}	0.018	0.144	0.081	0.832	0.052	0.271	0.284	0.988	0.043	0.270	0.283	0.981
		50%	β_{11}	-0.041	0.305	0.168	0.798	0.038	0.455	0.471	0.980	0.049	0.443	0.433	0.970
			β_{12}	0.033	0.168	0.108	0.850	-0.030	0.404	0.353	0.961	-0.037	0.342	0.325	0.952
500	50%	30%	β_{11}	-0.041	0.305	0.168	0.798	0.038	0.455	0.471	0.980	0.035	0.432	0.455	0.978
			β_{12}	0.033	0.168	0.108	0.850	-0.030	0.404	0.353	0.961	-0.025	0.385	0.352	0.950
		50%	β_{11}	-0.020	0.245	0.159	0.833	-0.037	0.444	0.455	0.973	-0.009	0.445	0.456	0.982
			β_{12}	0.022	0.136	0.099	0.855	0.035	0.351	0.341	0.973	0.003	0.368	0.345	0.971

Table B.2: Simulation results under covariate-independent truncation and covariate-dependent censoring ($\beta_{11} = -1, \beta_{12} = 1$).

				IPW				PO-AJ				PO-IPW				PO-IPW(empirical)			
n	cen%	trun%	Para	Bias	SD	SE	CP	Bias	SD	SE	CP	Bias	SD	SE	CP	Bias	SD	SE	CP
350	30%	30%	β_{11}	-0.149	0.247	0.188	0.810	-0.141	0.467	0.467	0.963	-0.013	0.445	0.477	0.959	0.038	0.444	0.463	0.969
			β_{12}	-0.046	0.142	0.105	0.837	-0.005	0.317	0.356	0.961	-0.007	0.359	0.393	0.954	-0.002	0.390	0.370	0.957
		50%	β_{11}	-0.206	0.299	0.204	0.758	-0.111	0.566	0.539	0.964	0.051	0.488	0.550	0.970	0.041	0.548	0.591	0.970
			β_{12}	-0.050	0.174	0.114	0.790	-0.037	0.430	0.413	0.947	-0.085	0.428	0.476	0.940	-0.060	0.491	0.497	0.955
350	50%	30%	β_{11}	-0.207	0.302	0.271	0.885	-0.216	0.495	0.465	0.945	-0.095	0.642	0.791	0.983	-0.126	0.767	0.803	0.973
			β_{12}	-0.065	0.137	0.120	0.866	-0.034	0.301	0.336	0.939	-0.053	0.487	0.469	0.922	0.009	0.544	0.487	0.920
		50%	β_{11}	-0.277	0.372	0.299	0.838	-0.250	0.666	0.577	0.947	$<$0.001	0.659	0.944	0.981	-0.013	0.745	0.975	0.985
			β_{12}	-0.079	0.161	0.127	0.835	-0.028	0.395	0.425	0.962	-0.100	0.498	0.586	0.916	-0.080	0.618	0.604	0.908
500	30%	30%	β_{11}	-0.161	0.206	0.158	0.773	-0.164	0.407	0.391	0.961	0.018	0.409	0.417	0.977	0.015	0.388	0.395	0.959
			β_{12}	-0.050	0.119	0.087	0.823	0.015	0.281	0.297	0.973	0.004	0.358	0.359	0.980	0.008	0.344	0.329	0.958
		50%	β_{11}	-0.207	0.270	0.170	0.696	-0.090	0.473	0.451	0.970	0.044	0.414	0.461	0.967	0.046	0.543	0.566	0.958
			β_{12}	-0.055	0.155	0.094	0.785	-0.034	0.367	0.360	0.975	-0.075	0.357	0.409	0.947	-0.084	0.476	0.482	0.937
500	50%	30%	β_{11}	-0.201	0.262	0.223	0.846	-0.200	0.433	0.400	0.930	-0.077	0.666	0.708	0.978	-0.075	0.612	0.680	0.979
			β_{12}	-0.060	0.118	0.100	0.849	-0.011	0.287	0.299	0.945	-0.005	0.425	0.438	0.943	-0.009	0.375	0.407	0.949
		50%	β_{11}	-0.274	0.305	0.248	0.793	-0.280	0.563	0.506	0.937	-0.025	0.610	0.870	0.984	-0.062	0.721	0.907	0.978
			β_{12}	-0.078	0.140	0.106	0.796	0.028	0.297	0.362	0.970	-0.091	0.451	0.532	0.931	-0.046	0.565	0.562	0.929

Table B.3: Simulation results under covariate-dependent truncation and covariate-dependent censoring ($\beta_{11} = -1, \beta_{12} = 1$).

				IPW				PO-AJ				PO-IPW				PO-IPW(empirical)			
n	cen%	trun%	Para	Bias	SD	SE	CP	Bias	SD	SE	CP	Bias	SD	SE	CP	Bias	SD	SE	CP
350	30%	30%	β_{11}	-0.210	0.227	0.176	0.721	-0.190	0.380	0.387	0.960	0.055	0.376	0.384	0.949	0.022	0.460	0.420	0.952
			β_{12}	-0.078	0.131	0.103	0.799	-0.056	0.256	0.280	0.945	-0.033	0.312	0.322	0.960	-0.042	0.386	0.330	0.945
		50%	β_{11}	-0.240	0.255	0.182	0.686	-0.160	0.527	0.522	0.973	0.095	0.391	0.429	0.953	0.080	0.405	0.445	0.952
			β_{12}	-0.083	0.159	0.112	0.780	-0.078	0.404	0.394	0.941	-0.019	0.358	0.385	0.968	-0.008	0.420	0.386	0.960
350	50%	30%	β_{11}	-0.249	0.141	0.247	0.813	-0.308	0.415	0.387	0.898	-0.053	0.604	0.683	0.980	-0.102	0.716	0.702	0.979
			β_{12}	-0.088	0.141	0.120	0.819	-0.105	0.274	0.273	0.921	-0.051	0.447	0.451	0.928	-0.059	0.446	0.417	0.917
		50%	β_{11}	-0.295	0.324	0.254	0.766	-0.370	0.563	0.533	0.920	-0.008	0.711	0.756	0.972	-0.043	0.723	0.805	0.971
			β_{12}	-0.099	0.162	0.125	0.799	-0.046	0.362	0.374	0.949	-0.035	0.513	0.514	0.935	-0.014	0.525	0.504	0.943
500	30%	30%	β_{11}	-0.202	0.194	0.146	0.699	-0.174	0.340	0.326	0.941	0.064	0.304	0.319	0.953	0.052	0.317	0.324	0.943
			β_{12}	-0.076	0.109	0.086	0.767	-0.064	0.227	0.240	0.940	-0.032	0.240	0.271	0.954	-0.020	0.252	0.262	0.938
		50%	β_{11}	-0.231	0.222	0.151	0.733	-0.154	0.409	0.426	0.977	0.110	0.325	0.351	0.949	0.067	0.400	0.393	0.944
			β_{12}	-0.088	0.130	0.092	0.733	-0.071	0.337	0.327	0.954	-0.019	0.325	0.322	0.973	0.012	0.418	0.348	0.958
500	50%	30%	β_{11}	-0.238	0.240	0.206	0.786	-0.316	0.361	0.327	0.844	-0.055	0.566	0.583	0.974	-0.115	0.620	0.627	0.977
			β_{12}	-0.084	0.124	0.101	0.798	-0.091	0.219	0.226	0.903	-0.037	0.400	0.397	0.917	0.005	0.414	0.393	0.926
		50%	β_{11}	-0.295	0.254	0.210	0.702	-0.341	0.493	0.453	0.902	0.019	0.590	0.674	0.974	-0.012	0.484	0.501	0.955
			β_{12}	-0.111	0.132	0.104	0.750	-0.055	0.303	0.320	0.947	-0.027	0.720	0.741	0.968	0.012	0.536	0.481	0.938

BIBLIOGRAPHY

[1] Odd O. Aalen and S/oren Johansen. An empirical transition matrix for nonhomogeneous markov chains based on censored observations. *Scandinavian Journal of Statistics*, 5(3):141–150, 1978.

[2] Ahmadou Alioum and Daniel Commenges. A proportional hazards model for arbitrarily censored and truncated data. *Biometrics*, 52(2):512–524, 1996.

[3] Per K. Andersen, Elisavet Syriopoulou, and Erik T. Parner. Causal inference in survival analysis using pseudo-observations. *Statistics in Medicine*, 36(17):2669–2681, 2017.

[4] Per Kragh Andersen, Mette Gerster Hansen, and John P. Klein. Regression analysis of restricted mean survival time based on pseudo-observations. *Lifetime Data Analysis*, 10:335–350, 2004.

[5] Per Kragh Andersen, John P. Klein, and Susanne Rosthøj. Generalised linear models for correlated pseudo-observations, with applications to multi-state models. *Biometrika*, 90:15–27, 2003.

[6] Per Kragh Andersen and Maja Pohar Perme. Pseudo-observations in survival analysis. *Statistical Methods in Medical Research*, 19:71–99, 2010.

[7] A. Bellach, M. R. Kosorok, P. B. Gilbert, and J. P. Fine. General regression model for the subdistribution of a competing risk under left-truncation and right-censoring. *Biometrica*, 107(4):949–964, 2020.

[8] Nadine Binder, Thomas A. Gerds, and Per Kragh Andersen. Pseudo-observations for competing risks with covariate dependent censoring. *Lifetime Data Anal*, 20(2):303–315, 2014.

[9] Pei-Yun Chen and Anastasios A. Tsiatis. Causal inference on the difference of the restricted mean lifetime between two groups. *Biometrics*, 57:1030–1038, 2001.

[10] Stephen R Cole and Miguel A Hernán. Constructing inverse probability weights for marginal structural models. *American Journal of Epidemiology*, 168(6):656–664, 2008.

[11] J. P. Fine, Z. Ying, and L. J. Wei. On the linear transformation model for censored data. *Biometrika*, 85(4):980–986, 1998.

[12] Jason P. Fine. Regression modeling of competing crude failure probabilities. *Biostatistics*, 2(1):85–97, 2001.

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

[14] Dianne M. Finkelstein, Dirk F. Moore, and David A. Schoenfeld. A proportional hazards model for truncated aids data. *Biometrics*, 49(3):731–740, 1993.

[15] Ronald B. Geskus. On the inclusion of prevalent cases in hiv/aids natural history studies through a marker-based estimate of time since seroconversion. *Statistics in Medicine*, 19:1753–1769, 2000.

[16] Ronald B. Geskus. Cause-specific cumulative incidence estimation and the fine and gray model under both left truncation and right censoring. *Biometrics*, 67(1):39–49, 2011.

[17] Ronald B. Geskus. *Data analysis with competing risks and intermediate states*. Chapman and Hall/CRC, Philadelphia, PA, 2016.

[18] Ronald B. Geskus, Frank A. Miedema, Jaap Goudsmit, Peter Reiss, Hanneke Schuitemaker, and Roel A. Coutinho. Prediction of residual time to aids and death based on markers and cofactors. *Journal of Acquired Immune Deficiency Syndromes*, 32:514–521, 2003.

[19] Richard D. Gill. Non- and semi-parametric maximum likelihood estimators and the von mises method (part 1). *Scand J Stat*, 16:97–128, 1989.

[20] Mia K. Grand, Hein Putter, Arthur Allignol, and Per K. Andersen. A note on pseudo-observations and left-truncation. *Biometrical Journal*, 61:290–298, 2019.

[21] Frederik Graw, Thomas A. Gerds, and Martin Schumacher. On pseudo-values for regression analysis in competing risks models. *Lifetime Data Anal*, 15:241–255, 2009.

[22] Changbin Guo and Yu Liang. Analyzing restricted mean survival time using sas/stat. *SAS Institute Inc., Cary, NC*, 2019.

[23] Frank R. Hampel. The influence curve and its role in robust estimation. *Journal of the American Statistical Association*, 69:383–393, 1974.

[24] Chiung-Yu Huang, Jing Ning, and Jing Qin. Semiparametric likelihood inference for left-truncated and right-censored data. *Biometrics*, 14:785–798, 2015.

[25] Ying Huang and Mei-Cheng Wang. Estimating the occurrence rate for prevalent survival data in competing risks models. *Journal of the American Statistical Association*, 90(432):1406–1415, 1995.

[26] Peter J. Huber. *Robust statistical procedures*. Society for Industrial and Applied Mathematics, Philadelphia, 1977.

[27] Lancelot F. James. A study of a class of weighted bootstraps for censored data. *The Annals of Statistics*, 25(4):1595–1621, 1997.

[28] Theodore Karrison. Restricted mean life with adjustment for covariates. *Journal of the American Statistical Association*, 82:1169–1176, 1987.

[29] Leslie Kish. Weighting for unequal p_i. *Journal of Official Statistics*, 8(2):183–200, 1992.

[30] John P. Klein. Modelling competing risks in cancer studies. *Statistics in Medicine*, 25:1015–1034, 2006.

[31] John P. Klein and Per Kragh Andersen. Regression modeling of competing risks data based on pseudovalues of the cumulative incidence function. *Biometrics*, 61:223–229, 2005.

[32] Tze Leung Lai and Zhiliang Ying. Rank regression methods for left-truncated and right-censored data. *The Annals of Statistics*, 19:531–556, 1991.

[33] Chi Hyun Lee, Jing Ning, and Yu Shen. Analysis of restricted mean survival time for length-biased data. *Biometrics*, 74:575–583, 2018.

[34] Kung-Yee Liang and Scott L. Zeger. Longitudinal data analysis using generalized linear models. *Biometrika*, 78:13–22, 1986.

[35] D. Y Lin, L. J. Wei, and Z. Ying. Model-checking techniques based on cumulative residuals. *Biometrics*, 58(1):1–12, 2002.

[36] Brent R. Logan, Mei-Jie Zhang, and John P. Klein. Marginal models for clustered time to event data with competing risks using pseudovalues. *Biometrics*, 67(1):1–7, 2011.

[37] Micha Mandel, Jacobo de UÃśa-ÃĄlvarez, David K. Simon, and Rebecca A. Betensky. Inverse probability weighted cox regression for doubly truncated data. *Biometrics*, 74(2):481–487, 2018.

[38] Reinhard Meister and Christof Schaeferb. Statistical methods for estimating the probability of spontaneous abortion in observational studiesâĂŤanalyzing pregnancies exposed to coumarin derivatives. *Reproductive Toxicology*, 26:31–35, 2008.

[39] Jing Ning, Daewoo Pak, Hong Zhu, and Jing Qin. Conditional independence test of failure and truncation times: Essential tool for method selection. *Computational Statistics and Data Analysis*, 168(107402):1–11, 2022.

[40] Morten Overgaard, Erik Thorlund Parner, and Jan Pedersen. Asymptotic theory of generalized estimating equations based on jack-knife pseudo-observations. *The Annals of Statistics*, 45:1988âĂŞ2015, 2017.

[41] Morten Overgaard, Erik Thorlund Parner, and Jan Pedersen. Pseudo-observations under covariate-dependent censoring. *Journal of Statistical Planning and Inference*, 202:112–122, 2019.

[42] Maja Pohar Perme and Per Kragh Andersen. Checking hazard regression models using pseudo-observations. *Statistics in Medicine*, 27(25):5309–5328, 2008.

[43] Maja Pohar Perme and Per Kragh Andersen. Checking hazard regression models using pseudo-observations. *Statistics in Medicine*, 27(25):5309–5328, 2008.

[44] H. Putter, M. Fiocco, and R. B. Geskus. Tutorial in biostatistics: competing risks and multi-state models. *Statistics in Medicine*, 26:2389–2430, 2007.

[45] Lior Rennert and Sharon X. Xie. Cox regression model with doubly truncated data. *Biometrics*, 74(2):725–733, 2018.

[46] Lior Rennert and Sharon X. Xie. Cox regression model under dependent truncation. *Biometrics*, Early View:1–14, 2021.

[47] James M. Robins. Information recovery and bias adjustment in proportional hazards regression analysis of randomized trials using surrogate markers. *Proceedings of the Biopharmaceutical Section, American Statistical Association*, pages 24–33, 1993.

[48] James M. Robins and Dianne M. Finkelstein. Correcting for noncompliance and dependent censoring in an aids clinical trial with inverse probability of censoring weighted (ipcw) log-rank tests. *Biometrics*, 56:779–788, 2000.

[49] James M. Robins and Andrea Rotnitzky. Recovery of information and adjustment for dependent censoring using surrogate markers. *AIDS Epidemiology*, pages 297–331, 1992.

[50] Col Partha Roy and Sekhar Chakrabarti. Mutation in aids restriction gene affecting hiv infection and disease progression in a high risk group from northeastern india. *Medical Journal Armed Forces India*, 72:111–115, 2016.

[51] Pao-Sheng Shen. Proportional subdistribution hazards regression for left-truncated competing risks data. *Journal of Nonparametric Statistics*, 23(4):885–895, 2011.

[52] Regina Stegherr, Arthur Allignol, Reinhard Meister, Christof Schaefer, and Jan Beyersmann. Estimating cumulative incidence functions in competing risks data with dependent left-truncation. *Statistics in Medicine*, 39:481–493, 2020.

[53] Lu Tian, Lihui Zhao, and L.J.Wei. Predicting the restricted mean event time with the subjectâĂŹs baseline covariates in survival analysis. *Biostatistics*, 15:222–233, 2014.

[54] Wei-Yann Tsai, Nicholas P. Jewell, and Mei-Cheng Wang. A note on the product-limit estimator under right censoring and left truncation. *Biometrika*, 74:883–886, 1987.

[55] Vander Vaart. *Asymptotic Statistics*. Cambridge University Press, Cambridge, 1998.

[56] Huixia Judy Wang and Lan Wang. Quantile regression analysis of length-biased survival data. *Statistics*, 3:31–47, 2014.

[57] Mei-Cheng Wang, Ron Brookmeyer, and Nicholas P. Jewell. Statistical models for prevalent cohort data. *Biometrics*, 49:1–11, 1993.

[58] Xin Wang and Douglas E. Schaubel. Modeling restricted mean survival time under general censoring mechanisms. *Lifetime Data Anal*, 24:176–199, 2018.

[59] SJW Willems, A Schat, MS van Noorden, and M Fiocco. Correcting for dependent censoring in routine outcome monitoring data by applying the inverse probability censoring weighted estimator. *Statistical Methods in Medical Research*, 27:323âĂŞ335, 2018.

[60] Christina Wolfson, David B. Wolfson, Masoud Asgharian, Cyr Emile M'Lan, Truls Østbye, Kenneth Rockwood, and D.B. Hogan. A reevaluation of the duration of survival after the onset of dementia. *The New England Journal of Medicine*, 344:1111–1116, 2001.

[61] Fang Xiang and Susan Murray. Restricted mean models for transplant benefit and urgency. *Statistics in Medicine*, 31:561–576, 2012.

[62] Scott L. Zeger and Kung-Yee Liang. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*, 42:121–130, 1986.

[63] Min Zhang and Douglas E. Schaubel. Estimating differences in restricted mean lifetime using observational data subject to dependent censoring. *Biometrics*, 64:740–749, 2011.

[64] Xu Zhang, Mei-Jie Zhang, and Jason Fine. A proportional hazards regression model for the subdistribution with right-censored and left-truncated competing risks data. *Statistics in Medicine*, 30(16):1933–1951, 2011.

[65] Hongwei Zhao and Anastasios A. Tsiatis. A consistent estimator for the distribution of quality adjusted survival time. *Biometrika*, 84:339–348, 1997.

[66] Hong Zhu, Jing Ning, Yu Shen, and Jing Qin. Semiparametric density ratio modeling of survival data from a prevalent cohort. *Biostatistics*, 18:62–75, 2017.

[67] David M. Zucker. Restricted mean life with covariates: modification and extension of a useful survival analysis method. *Journal of the American Statistical Association*, 93:702–709, 1998.