이학박사 학위논문

Difference in restricted mean survival time in observational studies: A review of estimation methods and a development of sensitivity analysis

제한된 평균 생존시간 추정 방법 고찰 및 새로운 민감도 분석 방법 개발

2023년 2월

서울대학교 대학원

응용바이오공학과

이 승 재

Difference in restricted mean survival time in observational studies: A review of estimation methods and a development of sensitivity analysis

제한된 평균 생존시간 추정 방법 고찰 및 새로운 민감도 분석 방법 개발

지도 교수 박 지 훈

이 논문을 이학박사 학위논문으로 제출함

2022년 11월

서울대학교 대학원

응용바이오공학과

이 승 재

이승재의 이학박사 학위논문을 인준함

2023년 1월

위 원 장 이 경 호 (인)

부위원장 박 지 훈 (인)

위 원 이 우 주 (인)

위 원 전 종 준 (인)

위 원 장 원 (인)

Abstract

Seungjae Lee

The Department of Applied Bioengineering

The Graduate School

Seoul National University

The difference in restricted mean survival time (RMST) has been increasingly used as an alternative measure to hazard ratio in survival analysis. Unlike relative effect measure such as hazard ratio, RMST difference provides information about an intuitively interpretable absolute risk and is known to be robust regardless of the proportional hazards assumption.

In experimental studies such as a randomized controlled trial, the RMST is calculated by integrating the area under the Kaplan-Meier curve up to a specific time point, and the difference in RMST between the two groups is used as a causal effect of exposure. However, in observational studies, the standard Kaplan-Meier estimator cannot be directly used for calculating the RMST because of confounding bias due to non-random exposure assignment. The difference in RMST adjusted for potential confounders can be estimated using methods such as direct RMST regression, inverse probability weighting, G-computation, etc. Through multiple simulation studies in which all the models were correctly specified, we confirmed that all the methods being considered provided the unbiased estimates with the percentile bootstrap confidence intervals achieving near nominal coverage probability.

Although several methods have been developed for evaluating the difference in RMST adjusted for potential confounders in the observational study, there

is no study on the sensitivity analysis of unmeasured confounding. Therefore, we propose a novel sensitivity analysis method that considers unmeasured confounding for evaluating the estimate of the difference in adjusted RMST. Given a user-specified sensitivity parameter, one can obtain the sensitivity range and confidence interval of bias-adjusted difference in RMST. It is necessary to solve a complex optimization problem to obtain the sensitivity range and confidence interval, but there is no analytic solution except in special cases. While the optimization problem can be directly solved by using an optimization algorithm such as L-BFGS-B (hereafter referred to this method as the direct optimization method), it takes considerable computational time. Therefore, we propose an approximate optimization method comparable to the direct optimization method in terms of bias, achieving substantial reduction in the computational time. Through intensive Monte Carlo simulation studies, we showed that the proposed approximate optimization method can be a practical alternative. When applying our sensitivity analysis method in practice, we recommend using the approximate optimization method in case that the censoring rate is less than 0.7. Otherwise, one may use the direct optimization method using an optimization algorithm.

Keywords: restricted mean survival time, causal inference, survival analysis, observational study, sensitivity analysis, unmeasured confounding

Student Number: 2020-30064

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The difference in restricted mean survival time (RMST) has been increasingly used as an alternative measure to hazard ratio in survival analysis (Royston and Parmar, 2013; Uno et al., 2014; Trinquart et al., 2016; Kim, Uno, and Wei, 2017; Pak et al., 2017; Kloecker et al., 2020; Han and Jung, 2022). Although the hazard ratio is the most commonly used measure of exposure in survival analysis, its causal interpretation is often risky because the risk set is updated without randomization (Hernán, 2010; Stensrud et al., 2019; Stensrud and Hernán, 2020). Unlike relative effect measure such as hazard ratio, the difference in RMST provides information about an intuitively interpretable absolute risk and is known to be robust regardless of the proportional hazards assumption (Hernán, 2010; Royston and Parmar, 2011). In experimental studies such as a randomized controlled trial, the RMST is calculated by integrating the area under the Kaplan-Meier curve up to a specific time point, and the difference in RMST between the exposure groups is used as a causal effect of exposure.

However, in observational studies, the standard Kaplan-Meier estimator cannot be directly used for calculating the RMST because of confounding bias

due to non-random exposure assignment, and thus the difference in RMST based on the Kaplan-Meier estimator may be interpreted as an association. In other words, some strong assumptions should be required to interpret the association from an observational study as the causation. One of the key assumptions is a conditional exchangeability assumption (also referred to as no unmeasured confounding or ignorability assumption) (De Finetti, 1974; Rubin, 1978), which means that given a set of covariates, the potential outcome is independent of the exposure. Under the conditional exchangeability assumption, various statistical methods such as propensity score (PS) matching (Rosenbaum and Rubin, 1983), inverse probability (IP) weighting (Horvitz and Thompson, 1952), G-estimation (Robins, 1989), and G-computation (Robins, 1986) have been developed to estimate the causal estimand of interest. By applying these methods to the RMST, the difference in RMST adjusted for confounders can be estimated using methods such as direct RMST regression (Andersen, Hansen, and Klein, 2004; Tian, Zhao, and Wei, 2014), IP weighting (Hernán, Brumback, and Robins, 2000; Cole and Hernán, 2004; Xie and Liu, 2005; Hernán, 2010), G-computation (Chatton et al., 2022), etc. In order to have a causal interpretation, these methods require a full adjustment for all possible confounders. In this thesis, several simulation studies are performed to assess whether the methods being considered provide the unbiased estimates of the difference in adjusted RMST along with confidence intervals achieving near nominal coverage probability when all the models are correctly specified.

Even if a researcher makes effort to adjust for all possible potential confounders, it is generally unverifiable by observed data that these confounders satisfy the conditional exchangeability. When a researcher is concerned about

possible unmeasured confounding (i.e., a violation of conditional exchangeability), it is necessary to investigate how the analysis results would be vary by the degree of violation of the conditional exchangeability assumption. This is so-called sensitivity analysis for unmeasured confounding (Cornfield et al., 1959; Lin, Psaty, and Kronmal, 1998; Robins, 1999; Scharfstein, Rotnitzky, and Robins, 1999; Robins, Rotnitzky, and Scharfstein, 2000; Rosenbaum, 2002; Brumback et al., 2004; Tan, 2006; Zhao, Small, and Bhattacharya, 2019; Dorn and Guo, 2022). There is a long history of sensitivity analysis in observational studies. Cornfield et al. (1959) is considered a monumental paper in sensitivity analysis for unmeasured confounding. Lin, Psaty, and Kronmal (1998) assessed the sensitivity of regression analysis results including binary response and censored time data as special cases to the residual confounding by an unmeasured confounder. Robins (1999) and Brumback et al. (2004) performed an interesting sensitivity analysis using the difference of the expectation of a potential outcome between with and without the exposure, given the measured confounders. Similarly, Robins, Rotnitzky, and Scharfstein (2000) performed a sensitivity analysis using the difference of the conditional distribution of exposure between with and without the potential outcome, given the measured confounders. For matched observational study, Rosenbaum (2002) developed a sensitivity analysis considering that two individuals with the same measured confounders can have different odds of exposure due to unmeasured confounder. Recently, based on Rosenbaum (2002) and Tan (2006), Zhao, Small, and Bhattacharya (2019) proposed a parametric sensitivity model, and this approach allowed each individual to have the true odds of exposure that can differ from the odds of exposure estimated by measured confounders.

Most of the sensitivity analysis methods for unmeasured confounding in observational survival data have been developed for Cox proportional hazards-type models (Klungsøyr et al., 2009; Lin, Logan, and Henley, 2013; Huang, Xu, and Dulai, 2020) and studied how much the hazard ratio changes with respect to the amount of unmeasured confounding. Similarly, RMST can also suffer from unmeasured confounding because of complexity of data generating mechanism. To best our knowledge, there is few studies on the sensitivity analysis for RMST regarding the degree of unmeasured confounding. Therefore, by adapting Zhao, Small, and Bhattacharya (2019)'s parametric sensitivity model for non-survival data and using the estimate of adjusted RMST up to specific time point obtained by integrating Xie and Liu (2005)'s adjusted Kaplan-Meier curve, we propose a novel sensitivity analysis method for the estimate of the difference in RMST adjusted only for measured confounders when unmeasured confounding is suspected. By using our sensitivity analysis method, one can obtain the sensitivity range of the point estimates for the difference in bias-adjusted RMST along with confidence interval for theirs partially identified region. To facilitate its use in practice, we made an R package, entitled `RMSTSens`, to perform sensitivity analysis of unmeasured confounding for the estimate of the difference in adjusted RMST and plot the results of the sensitivity analysis (https://github.com/seungjae2525/RMSTSens).

The remainder of this thesis is organized as follows. In Chapter 2, we present notation and assumptions and describe the definition of the difference in RMST. Chapter 3 reviews the methods for estimating the difference of RMST adjusted for confounders, compares the performance between the methods being considered through simulation studies, and applies each method to real data. We propose a novel sensitivity analysis method in Chapter 4. We also show

the Monte Carlo simulation results to demonstrate that our proposed method performs well under a variety of settings and illustrate the proposed method with two real data. Finally, a discussion and summary are presented in Chapter 5. The overall flow diagram of thesis structure is shown in Figure 1.1.

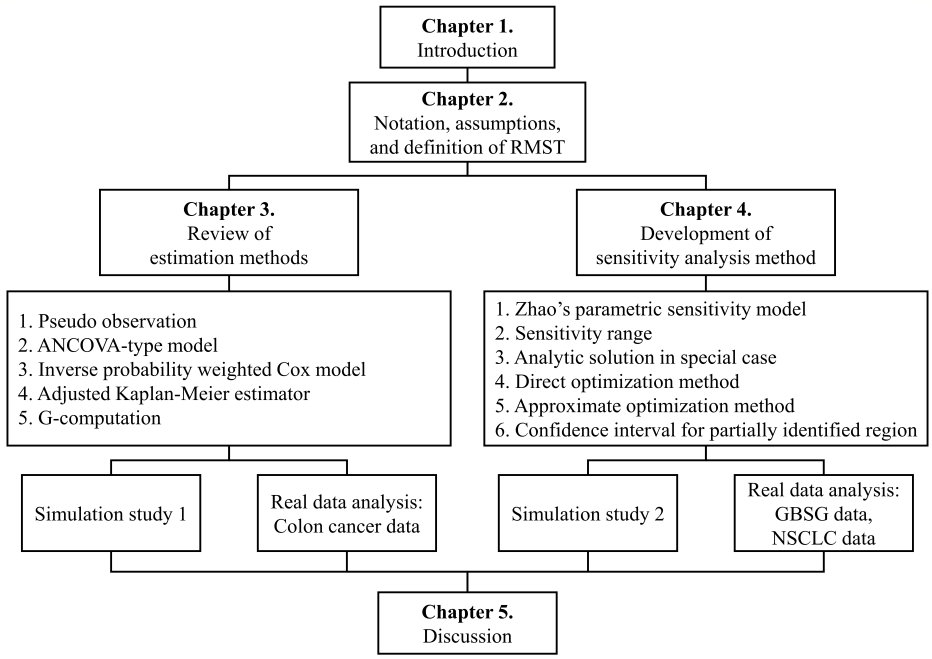


Figure 1.1 Overall flow diagram of thesis structure

Chapter 2

Restricted Mean Survival Time

2.1 Notation and assumptions

Let $(\widetilde{T}_i, \delta_i, A_i, \boldsymbol{L}_i)$ denote an independent sample of right-censored survival data for subject i $(i = 1, \ldots, N)$. N is a sample size. $\widetilde{T}_i = \min\{T_i, C_i\}$ is an observed survival time where T_i is the event time and C_i is the censoring time. $\delta_i = I(T_i \leq C_i)$ is an event indicator where $I(\cdot)$ denotes the indicator function taking the value 1 when the condition is true and 0 otherwise. A_i is a binary exposure indicator which is 1 for the exposed group and 0 for the unexposed group. $\boldsymbol{L}_i$ is a vector of relevant prognostic factors. Define

$$P(A_i = 1 \mid \boldsymbol{L}_i) \tag{2.1}$$

as the propensity score (PS) which is the conditional probability of receiving exposure given $\boldsymbol{L}_i$ (Rosenbaum and Rubin, 1983). In practice, $P(A_i = 1 \mid \boldsymbol{L}_i)$ can be estimated by using maximum likelihood estimation. Let $\widehat{P}(A_i = 1 \mid \boldsymbol{L}_i)$ be a consistent estimate of $P(A_i = 1 \mid \boldsymbol{L}_i)$.

In the counterfactual framework by Rubin (1974) and Holland (1986), denote T^a and C^a as the potential outcomes for event time and censoring time

had subject received exposure a, respectively. Also, define $\delta^a = I(T^a \leq C^a)$. To estimate causal effects from observational survival studies, the following identifiability conditions are assumed in this thesis:

For $a \in \{0, 1\}$,

(A1) Consistency: $T^a = T$ and $C^a = C$, if $A = a$.

(A2) Independent censoring: $T^a \perp\!\!\!\perp C^a \mid (A, \boldsymbol{L})$.

(A3) Conditional exchangeability: $T^a \perp\!\!\!\perp A \mid \boldsymbol{L}$ and $C^a \perp\!\!\!\perp A \mid \boldsymbol{L}$.

(A4) Positivity: $0 < P(A = a \mid \boldsymbol{L}) < 1$.

Note that the assumption (**A1**) connects the potential outcomes to the observable outcomes. The assumption (**A2**) means that in each exposure arm, the prognostic factors $\boldsymbol{L}$ suffice to explain the dependence between the event and censoring times. And, the assumption (**A3**) means that $\boldsymbol{L}$ can block all backdoor paths between the exposure and the survival time (Pearl, 1995), so that there is no unmeasured confounding. Under the assumptions (**A1**)–(**A3**), the assumption (**A1**) implies that $\delta^a = \delta$ if $A = a$ and the assumption (**A3**) implies that $\delta^a \perp\!\!\!\perp A \mid \boldsymbol{L}$. The assumption (**A4**) guarantees that there should be existed both the exposed and unexposed subjects at all possible combinations of the values of the prognostic factors (Westreich and Cole, 2010).

We additionally assume that one subject's potential outcome under exposure a does not rely on the others' exposure values (Cox, 1958) and there is no multiple version of exposure value. These assumptions are collectively referred to as the stable unit treatment value assumption (STUVA) (Rubin, 1980).

2.2 Difference in RMST

Let $S(t) = P(T > t)$ denote the (factual) survival function and $S^a(t) = P(T^a > t)$ denote the survival function of the potential (counterfactual) outcome for

event time under exposure a. Suppose that we are interested in comparing RMST up to τ between the exposure groups where τ (> 0) is the fixed truncation time point which is pre-specified at the study design stage based on clinical domain knowledge. When supposing $P(T^a \geq \tau) > 0$, the RMST up to τ under exposure a (i.e., μ_τ^a) (Irwin, 1949; Chen and Tsiatis, 2001) is defined as

$$\mu_\tau^a = E(T_\tau^a) = \int_0^\tau S^a(t)\,\mathrm{d}t \tag{2.2}$$

where $T_\tau^a = \min\{T^a, \tau\}$. Therefore, the difference in RMST (i.e., average causal effect on RMST) can be defined as

$$\mu_\tau^1 - \mu_\tau^0 = E(T_\tau^1) - E(T_\tau^0) = \int_0^\tau S^1(t)\,\mathrm{d}t - \int_0^\tau S^0(t)\,\mathrm{d}t.$$

Chapter 3

Methods for Estimation of Difference in RMST

3.1 Difference in RMST in randomized controlled trial

Suppose that the events occur at D distinct times in the whole sample. Without loss of generality, assume that the first n subjects are exposed ($A = 1$) and the rest $N - n$ are unexposed ($A = 0$) in total N subjects and that the first m of n subjects experience the events of interest in the exposed group and the first $D - m$ of $N - n$ subjects experience the events of interest in the unexposed group. Also, assume that in each exposure group, the event times are ordered increasingly (i.e., $t_1 < \ldots < t_m$ are event times in the exposed group and $t_{n+1} < \ldots < t_{n-m+D}$ are event times in the unexposed group). At time t_j, $j = 1, \ldots, m, n+1, \ldots, n-m+D$, there are d_j^a events out of Y_j^a subjects at risk under exposure $a \in \{0, 1\}$. Then, we can write $d_j^a = \sum_{i:T_i=t_j} \delta_i I(A_i = a)$ and $Y_j^a = \sum_{i:T_i \geq t_j} I(A_i = a)$.

In randomized controlled trials, subjects are randomly assigned to either exposed or unexposed group, so that both assumptions (**A3**) and (**A4**) are satisfied without conditioning on the prognostic factors $\boldsymbol{L}$. Therefore, Kaplan-Meier estimate of the survival function at time t under exposure a (Kaplan and Meier, 1958) is

$$\widehat{S}^a(t) = \begin{cases} 1 & \text{if } t < t_1^a \\ \prod_{j:t_j \leq t} \left(1 - \dfrac{d_j^a}{Y_j^a}\right) & \text{if } t \geq t_1^a \end{cases} \tag{3.1}$$

if $Y_j^a > 0$ and where t_1^a is the first event time for the exposure group a. Therefore, the estimate of the difference in RMST based on Kaplan-Meier estimate is

$$\widehat{\mu}_\tau^1 - \widehat{\mu}_\tau^0 = \widehat{E}(T_\tau^1) - \widehat{E}(T_\tau^0) = \int_0^\tau \widehat{S}^1(t)\,\mathrm{d}t - \int_0^\tau \widehat{S}^0(t)\,\mathrm{d}t.$$

Using Greenwood formula (Greenwood, 1926), the asymptotic variance of $\widehat{\mu}_\tau^a$ (Klein and Moeschberger, 2003; Cox and Oakes, 2018) is estimated by

$$\widehat{V}\big(\widehat{\mu}_\tau^a\big) = \sum_{j:t_j^a \leq \tau} \left(\int_{t_j^a}^\tau \widehat{S}^a(t)\,\mathrm{d}t\right)^2 \frac{d_j^a}{Y_j^a(Y_j^a - d_j^a)}$$

where t_j^a denotes the event time for the exposure group a. Then, the variance for the estimate of the difference in RMST is

$$\widehat{V}\big(\widehat{\mu}_\tau^1 - \widehat{\mu}_\tau^0\big) = \widehat{V}\big(\widehat{\mu}_\tau^1\big) + \widehat{V}\big(\widehat{\mu}_\tau^0\big).$$

The estimate of the difference in RMST can be obtained using `survfit` function in `survival` package (Therneau, 2022) and `rmst` function in `RISCA` package (Foucher et al., 2022) in the software environment R (R Core Team, 2021). Otherwise, one can obtain the estimate with its asymptotic variance by using `rmst2` function without "covariates" argument in `survRM2` package (Uno

et al., 2022) in R. Whichever method is used, the values of the estimate are the same. We provide example code in Section A.1.1.

3.2 Difference in RMST in observational study

In observational studies, we need to adjust for the potential confounders $\boldsymbol{L}$ to obtain a consistent causal effect of exposure. Given $\boldsymbol{L}$ that satisfy assumptions (**A2**)–(**A4**), there have been proposed several methods available to estimate the difference in RMST adjusted for $\boldsymbol{L}$ (hereinafter referred to as the difference in adjusted RMST) in observational survival studies. In this Section, we focus on five methods described in below Subsections because they are the most commonly used in applied practice and are available in standard statistical software packages. Because we only compare these methods, this thesis is not a comprehensive evaluation of methods for estimating the difference in adjusted RMST. Also, the results for simulation study and real data analysis cannot be exploited out of context as a pretext for preferring one estimation method over the others.

3.2.1 Direct regression

3.2.1.1 Pseudo-observation

Andersen, Hansen, and Klein (2004) proposed the estimation method using regression model for the difference in adjusted RMST based on the pseudo-observations. Let $\widehat{\mu}_{\tau,\text{pseudo}}$ be the estimate of RMST up to pre-specified time point τ from the Kaplan-Meier estimator and $\widehat{\mu}^{(-i)}_{\tau,\text{pseudo}}$ be the leave-one-out estimate of RMST up to τ from the Kaplan-Meier estimator obtained by

eliminating the i-th subject. For i-th subject, the pseudo-observation is defined as $\mu_{\tau,\text{pseudo},i} = N \times \widehat{\mu}_{\tau,\text{pseudo}} - (N-1) \times \widehat{\mu}^{(-i)}_{\tau,\text{pseudo}}$. See Andersen and Perme (2010) for the details of pseudo-observations in various survival analyses.

Andersen, Hansen, and Klein (2004) considered the regression model for pseudo-observations which corresponds to a specification of the relation between $\mu_{\tau,\text{pseudo},i}$ and $\boldsymbol{L}_i^* = (1, A_i, \boldsymbol{L}_i^\intercal)^\intercal$. To access the effect of exposure on the RMST adjusted for confounders $\boldsymbol{L}$, we can use a generalized linear model with identity link function, as follows:

$$E(\mu_{\tau,\text{pseudo},i} \mid A_i, \boldsymbol{L}_i) = \beta_0 + \beta_1 A_i + \boldsymbol{\beta_l}^\intercal \boldsymbol{L}_i = \beta^\intercal \boldsymbol{L}_i^*$$

where $\beta = (\beta_0, \beta_1, \boldsymbol{\beta_l}^\intercal)^\intercal$. Then, the coefficient of exposure, β_1, corresponds to the difference in adjusted RMST. The estimates of β can be estimated from the generalized estimating equation (GEE) (Liang and Zeger, 1986; Zeger and Liang, 1986)

$$\sum_i \left(\frac{\partial}{\partial \beta} (\beta^\intercal \boldsymbol{L}_i^*) \right)^\intercal V_i^{-1} \left(\mu_{\tau,\text{pseudo},i} - \beta^\intercal \boldsymbol{L}_i^* \right) = 0$$

where V_i is a working covariance matrix. A sandwich variance estimator can be used to estimate the variance of $\widehat{\beta}$ (White, 1982). See details for the variance estimate for $\widehat{\beta}$ in Andersen, Klein, and Rosthøj (2003).

The pseudo-observations can be obtained using `pseudomean` function in `pseudo` package (Perme, Gerster, and Rodrigues, 2017) in R. Based on the obtained pseudo-observations, the estimate of the difference in adjusted RMST and its sandwich variance estimate can be obtained using `geeglm` function in `geepack` package (Højsgaard, Halekoh, and Yan, 2006) in R. See details for `pseudo` package in Klein et al. (2008). We provide example code in Section A.1.2.

3.2.1.2 ANCOVA-type model

Tian, Zhao, and Wei (2014) proposed the analysis of covariance (ANCOVA)-type covariate adjustment method using regression model for estimating the difference in adjusted RMST. Followed from Andersen, Hansen, and Klein (2004)'s regression model, Tian, Zhao, and Wei (2014) considered the linear model

$$E(T_{\tau,i} \mid A_i, \boldsymbol{L}_i) = \beta_0 + \beta_1 A_i + \beta_{\boldsymbol{l}}^\intercal \boldsymbol{L}_i = \beta^\intercal \boldsymbol{L}_i^*$$

where $T_{\tau,i} = \min\{T_i, \tau\}$, $\beta = (\beta_0, \beta_1, \beta_{\boldsymbol{l}}^\intercal)^\intercal$, and $\boldsymbol{L}_i^* = (1, A_i, \boldsymbol{L}_i^\intercal)^\intercal$. Then, the coefficient of exposure, β_1, corresponds to the difference in adjusted RMST.

To estimate the coefficients β while considering the censoring, Tian, Zhao, and Wei (2014) considered the inverse probability (IP) censoring weighted estimating equation

$$S(\beta) = \frac{1}{n} \sum_{i=1}^{N} \frac{\widetilde{\Delta}_i}{\widehat{G}(T_{\tau,i})} \boldsymbol{L}_i^* \left\{ T_{\tau,i} - \beta^\intercal \boldsymbol{L}_i^* \right\}$$

where $\widetilde{\Delta}_i = I(T_{\tau,i} \leq C_i)$ and $\widehat{G}(T_{\tau,i})$ is the Kaplan–Meier estimator of the censoring time C based on $\{(\widetilde{T}_i, 1-\Delta_i);\ i = 1, \ldots, N\}$. They showed that under mild regularity conditions, $n^{1/2}(\widehat{\beta} - \beta)$ is asymptotically normal, and thus the asymptotic variance of $\widehat{\beta}$ can be obtained. See details in the supplementary material of Tian, Zhao, and Wei (2014).

The estimate of the difference in adjusted RMST using ANCOVA-type model and its asymptotic variance can be obtained using `rmst2` function with "covariates" argument in `survRM2` package (Uno et al., 2022) in R. See details for `survRM2` package in Uno (2015). We provide example code in Section A.1.3.

3.2.2 Inverse probability weighting

3.2.2.1 IP weighted Cox model

The Cox proportional hazards model (hereinafter referred to as Cox model) is widely used to compare the survival experiences between the exposure groups after adjusting for prognostic factors (or risk factors) that affect the outcome. Given the confounders $\boldsymbol{L}$, Cox model (Cox, 1972, 1975) is

$$h(t; A_i, \boldsymbol{L}_i) = h_0(t) \exp(\beta_1 A_i + \beta_{\boldsymbol{l}}^\intercal \boldsymbol{L}_i) = h_0(t) \exp(\beta^\intercal \boldsymbol{L}_i^*) \tag{3.2}$$

where $h_0(t)$ is the baseline hazard function, $\beta = (\beta_1, \beta_{\boldsymbol{l}}^\intercal)^\intercal$, and $\boldsymbol{L}_i^* = (A_i, \boldsymbol{L}_i^\intercal)^\intercal$.

Cole and Hernán (2004) described the method to estimate the adjusted survival curve using Cox model weighted by IP weights. Using the PS (2.1), the IP weights are defined by

$$\mathrm{w}_i = \frac{I(A_i = 1)}{P(A_i = 1 \mid \boldsymbol{L}_i)} + \frac{I(A_i = 0)}{P(A_i = 0 \mid \boldsymbol{L}_i)}. \tag{3.3}$$

They proposed using a null Cox model with IP weights that is stratified on exposure levels. Then, the adjusted survival curve under exposure a, $\widehat{S}^a_{\mathrm{cox}}(t)$, can be estimated from this IP weighted Cox model. The difference in adjusted RMST is obtained by

$$\widehat{\mu}^1_{\tau,\mathrm{cox}} - \widehat{\mu}^0_{\tau,\mathrm{cox}} = \int_0^\tau \widehat{S}^1_{\mathrm{cox}}(t)\, \mathrm{d}t - \int_0^\tau \widehat{S}^0_{\mathrm{cox}}(t)\, \mathrm{d}t.$$

When the IP weights are highly variable, one can use the stabilized IP weights (Robins, 1998)

$$\mathrm{sw}_i = \frac{I(A_i = 1)P(A_i = 1)}{P(A_i = 1 \mid L_i)} + \frac{I(A_i = 0)P(A_i = 0)}{P(A_i = 0 \mid L_i)} \tag{3.4}$$

instead of non-stabilized IP weights (3.3). Although the robust variance estimator for the estimate of hazard ratio can be estimated by using the partial

likelihood for the IP weighted Cox model (Lin and Wei, 1989; Binder, 1992; Shu et al., 2021), to best our knowledge, there is no study on estimating the variance of adjusted RMST estimated from IP weighted Cox model. Alternatively, we can use the bootstrap method to obtain the variance of the estimate.

The estimate of the difference in adjusted RMST using IP weighted Cox model can be obtained by using `coxph` and `survfit` functions in `survival` package (Therneau, 2022) and the `rmst` function in `RISCA` package (Foucher et al., 2022) in R. Otherwise, one can obtain the estimate by using `svydesign` and `svykm` functions in `survey` package (Lumley, 2010) and `rmst` function in `RISCA` package (Foucher et al., 2022) in R. Whichever method is used, the values of the estimate are the same. We provide example code in Section A.1.4.

3.2.2.2 Adjusted Kaplan–Meier estimator

Xie and Liu (2005) proposed the adjusted Kaplan-Meier estimator using IP weighting method. They considered the weighted number of events under exposure a as $\widetilde{d}_j^a = \sum_{i:T_i=t_j} \left(\delta_i I(A_i = a)/\widehat{P}(A_i = a \mid \boldsymbol{L}_i)\right)$ and the weighted number at risk as $\widetilde{Y}_j^a = \sum_{i:T_i \geq t_j} \left(I(A_i = a)/\widehat{P}(A_i = a \mid \boldsymbol{L}_i)\right)$. Then, the adjusted Kaplan-Meier estimator at time t under exposure a is obtained by

$$\widehat{S}^a_{\text{adj}}(t) = \begin{cases} 1 & \text{if } t < t_1^a \\ \prod_{j:t_j \leq t} \left(1 - \frac{\widetilde{d}_j^a}{\widetilde{Y}_j^a}\right) & \text{if } t \geq t_1^a \end{cases} \tag{3.5}$$

if $\widetilde{Y}_j^a > 0$ and where t_1^a is the first event time for the exposure group a. Xie and Liu (2005) shown that under the identifiability conditions (**A1**)–(**A4**), the adjusted Kaplan-Meier estimator (3.5) is a consistent estimate of $S^a(t)$. For details of the theoretic results and proofs, see Section 3 and Appendix A.2 and A.3 in Xie and Liu (2005). Based on the adjusted Kaplan-Meier estimator

(3.5), the estimate of adjusted RMST up to τ under exposure a is

$$\widehat{\mu}^a_{\text{adj},\tau} = \int_0^\tau \widehat{S}^a_{\text{adj}}(t)\ \mathrm{d}t \tag{3.6}$$

and the estimate of the difference in adjusted RMST is

$$\begin{aligned}
\widehat{\mu}^1_{\text{adj},\tau} - \widehat{\mu}^0_{\text{adj},\tau} &= \int_0^\tau \left[\prod_{j:t_j \le t} \left(1 - \frac{\widetilde{d}^1_j}{\widetilde{Y}^1_j}\right) - \prod_{j:t_j \le t} \left(1 - \frac{\widetilde{d}^0_j}{\widetilde{Y}^0_j}\right) \right] \mathrm{d}t \\
&= \int_0^\tau \left[\prod_{t_j \le t} \left(1 - \frac{\sum_{i:T_i=t_j} \frac{\delta_i I(A_i=1)}{\widehat{P}(A_i=1|\boldsymbol{L}_i)}}{\sum_{i:T_i \ge t_j} \frac{I(A_i=1)}{\widehat{P}(A_i=1|\boldsymbol{L}_i)}}\right) \right. \\
&\qquad \left. - \prod_{t_j \le t} \left(1 - \frac{\sum_{i:T_i=t_j} \frac{\delta_i I(A_i=0)}{1-\widehat{P}(A_i=1|\boldsymbol{L}_i)}}{\sum_{i:T_i \ge t_j} \frac{I(A_i=0)}{1-\widehat{P}(A_i=1|\boldsymbol{L}_i)}}\right) \right] \mathrm{d}t.
\end{aligned} \tag{3.7}$$

Xie and Liu (2005) also shown that the asymptotic variance for the adjusted Kaplan-Meier estimator (3.5) can be estimated by

$$\widehat{V}\left[\widehat{S}^a_{\text{adj}}(t)\right] = \left[\widehat{S}^a_{\text{adj}}(t)\right]^2 \sum_{j:t_j \le t} \frac{\widetilde{d}^a_j}{M^a_j(\widetilde{Y}^a_j - \widetilde{d}^a_j)} \tag{3.8}$$

where t^a_j denotes the event time for the exposure group a and

$$M^a_j = \frac{\left[\sum_{i:T_i \ge t_j} I(A_i = a)/\widehat{P}(A_i = a \mid \boldsymbol{L}_i)\right]^2}{\sum_{i:T_i \ge t_j} \left[I(A_i = a)/\widehat{P}(A_i = a \mid \boldsymbol{L}_i)\right]^2}.$$

Based on the estimate of asymptotic variance for the adjusted Kaplan-Meier estimator (3.8), Conner et al. (2019) proposed that the asymptotic variance for the estimate of adjusted RMST (3.6) can be estimated by

$$\widehat{V}\left(\widehat{\mu}^a_{\text{adj},\tau}\right) = \sum_{j:t^a_j \le \tau} \left(\sum_{i=j}^{\tau} \widehat{S}^a_{\text{adj}}(t_i) \times (t_{i+1} - t_i) \right)^2 \frac{\widetilde{d}^a_j}{M^a_j(\widetilde{Y}^a_j - \widetilde{d}^a_j)}.$$

When the IP weights are considered fixed and the adjusted RMSTs are thus independent, the variance for the estimate of the difference in adjusted RMST

(3.7) is estimated by

$$\widehat{V}\big(\widehat{\mu}^1_{\text{adj},\tau} - \widehat{\mu}^0_{\text{adj},\tau}\big) = \widehat{V}\big(\widehat{\mu}^1_{\text{adj},\tau}\big) + \widehat{V}\big(\widehat{\mu}^0_{\text{adj},\tau}\big).$$

The estimate of the difference in adjusted RMST using the adjusted Kaplan-Meier estimator can be obtained by using `ipw.survival` and `rmst` functions in `RISCA` package (Foucher et al., 2022) in R. The asymptotic variance of the estimate can be obtained using function available in https://github.com/s-conner/akm-rmst/blob/master/AKM_rmst.R. We provide example code in Section A.1.5.

3.2.3 G-computation

Chatton et al. (2022) proposed the method using G-computation for estimating the difference in adjusted RMST. Let $H_0(t) = \int_0^t h_0(s)\,\mathrm{d}s$ be the cumulative baseline hazard function. To estimate the survival function from the Cox model (3.2), Breslow (1974) proposed that both $H_0(t)$ and $\beta = \big(\beta_1, \beta_{\boldsymbol{l}}^\intercal\big)^\intercal$ can be simultaneously estimated using the Breslow's likelihood. Then, the estimates of $H_0(t)$ can be obtained by

$$\widehat{H}_0(t) = \sum_{t_i \le t} \frac{1}{\sum_{j \in R(t_i)} \exp\left(\widehat{\beta}_1 A_j + \widehat{\beta}_{\boldsymbol{l}}^\intercal \boldsymbol{L}_j\right)}$$

where $R(t_i)$ is the risk set at time t_i, and $\widehat{H}_0(t)$ is referred to as the Breslow estimator of the cumulative baseline hazard function. Using $\widehat{H}_0(t)$ and $\widehat{\beta} = \big(\widehat{\beta}_1, \widehat{\beta}_{\boldsymbol{l}}^\intercal\big)^\intercal$, the estimate of adjusted survival function under exposure a given the potential confounders $\boldsymbol{L}$ is given by

$$\widehat{S}^a_{\text{gc}}(t) = \frac{1}{N}\sum_{i=1}^{N} \exp\left[-\widehat{H}_0(t) \times \exp(\widehat{\beta}_1 a + \widehat{\beta}_{\boldsymbol{l}}^\intercal \boldsymbol{L}_i)\right].$$

Then, the difference in adjusted RMST is obtained by

$$\widehat{\mu}^1_{\tau,\mathrm{gc}} - \widehat{\mu}^0_{\tau,\mathrm{gc}} = \int_0^\tau \widehat{S}^1_{\mathrm{gc}}(t)\,\mathrm{d}t - \int_0^\tau \widehat{S}^0_{\mathrm{gc}}(t)\,\mathrm{d}t.$$

To best our knowledge, there is no study on estimating the variance of adjusted RMST estimated from G-computation method. Alternatively, we can use the bootstrap method to obtain the variance of the estimate.

The estimate of the difference in adjusted RMST using G-computation with its bootstrap confidence interval can be obtained by using `gc.survival` function in `RISCA` package (Foucher et al., 2022) in R. We provide example code in Section A.1.6.

3.3 Simulation study 1

We performed the simulation studies for comparing the performance of five methods described in Section 3.2. To evaluate the estimate of the difference in adjusted RMST, we examined the bias of the estimate and the coverage rate of the confidence interval. For all methods, we conducted 1,000 simulation replications for each parameter combination and constructed the 95% percentile bootstrap confidence interval with bootstrap resampling $B = 1{,}000$ times. All analyses of this thesis were used through R (Version 4.1.0) (R Core Team, 2021) using 8 CPU cores in parallel.

3.3.1 Simulation settings

We randomly generated $N = 500$ of the confounder L_1 from normal distribution with mean 1 and standard deviation 0.25 and the confounder L_2 from Bernoulli distribution with probability 0.5. The exposure A was generated from Bernoulli

distribution with the true PS $P(A = 1 \mid L_1, L_2) = \text{expit}\{\beta_0 + 1.5L_1 + 0.8L_2\}$ where $\text{expit}\{x\} = 1/(1 + \exp\{-x\})$ and β_0 was set at different values to yield an approximately certain probability ($\beta_0 = -1.9$ for $E(A \mid L_1, L_2) \approx 0.5$ and $\beta_0 = -0.425$ for $E(A \mid L_1, L_2) \approx 0.8$).

The potential outcomes for event times T^a under exposure a were randomly generated using the inverse transform sampling method via Cox model with Weibull-distributed baseline hazard as a function of L_1 and L_2 (Bender, Augustin, and Blettner, 2005) where

$$T^a = 5 \times \left[-\frac{\log(U^a)}{0.95 \times \exp\{\log(2.5)a + \log(1.2)L_1 + \log(0.7)L_2\}} \right]^{1/1.8} \tag{3.9}$$

and U^a were generated from uniform distribution on $(0, 1)$. Then, the event times $T = I(A = 1) \times T^1 + I(A = 0) \times T^0$. And, the censoring times C were randomly generated from Weibull distribution with scale parameter λ and shape parameter $\nu = 0.6$. We set the values of scale parameter $\lambda \in \{0.052, 0.178, 0.354, 0.643, 1.365\}$ to obtain the simulated data that correspond to approximate censoring rate $c \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$, respectively. Then, the observed survival time is the minimum of the event time and the censoring time (i.e., $\widetilde{T} = \min\{T, C\}$). Two values of specific time point $\tau \in \{1, 3\}$ are considered. Then, the total number of combinations of β_0, c, and τ is 20.

3.3.2 True value of difference in RMST

To evaluate bias and coverage rate, we need to know the true value of difference in RMST. Based on the definition of difference in RMST expressed in (2.2), the true value of RMST under exposure a is written as

$$\mu_\tau^a = \int_0^\tau S^a(t)\, \mathrm{d}t$$

$$= \int_0^\tau \int_l S^a(t;l)\,\mathrm{d}l\,\mathrm{d}t$$
$$= \int_l \int_0^\tau S^a(t;l)\,\mathrm{d}t\,\mathrm{d}l = \int_l \mu_\tau^a(l)\,\mathrm{d}l.$$

Because the true event times were generated from Cox model with Weibull-distributed baseline hazard (3.9), the true value of conditional RMST is

$$\begin{aligned}
\mu_\tau^a(l) &= E[\min\{T^a, \tau\} \mid L] \\
&= \int_0^\tau S^a(t;l)\,\mathrm{d}t \\
&= \int_0^\tau \left[\exp\left(-\frac{\lambda t^\nu}{5^\nu}\right)\right]^{\exp\{\log(2.5)a+\log(1.2)l_1+\log(0.7)l_2\}} \mathrm{d}t \\
&= \frac{5 \times \gamma\left(\frac{1}{\nu}, \frac{\exp\{\log(2.5)a + \log(1.2)l_1 + \log(0.7)l_2\}\lambda\tau^\nu}{5^\nu}\right)}{\nu \times \left(\exp\{\log(2.5)a + \log(1.2)l_1 + \log(0.7)l_2\}\right)^{1/\nu} \times \lambda^{1/\nu}}
\end{aligned} \tag{3.10}$$

where $\gamma(s,x) = \int_0^x t^{s-1}\exp(-t)\,\mathrm{d}t$ is the lower incomplete gamma function. A detailed proof of (3.10) is given in Appendix A.2. Then, we can obtain the approximate value of true difference in RMST by calculating the values of standardized mean, as follows:

$$\begin{aligned}
\mu_\tau^1 - \mu_\tau^0 &= \int_l \mu_\tau^1(l)\,\mathrm{d}l - \int_l \mu_\tau^0(l)\,\mathrm{d}l \\
&\approx \frac{1}{n}\sum_{i=1}^{n} \mu_\tau^1(l_i) - \frac{1}{N-n}\sum_{i=1}^{N-n} \mu_\tau^0(l_i).
\end{aligned} \tag{3.11}$$

Note that the true value of difference in RMST is irrelevant to both β_0 and c in our simulation setting. Therefore, we first generated the confounders L_1 and L_2 from superpopulation with sample size $N = 1{,}000{,}000$, and based on them, calculated (3.11) using values of (3.10) for each exposure group. When $\lambda = 0.95$ and $\nu = 1.8$, the approximate value of true difference in RMST $\mu_1^1 - \mu_1^0 = -0.027$ for $\tau = 1$ and $\mu_3^1 - \mu_3^0 = -0.417$ for $\tau = 3$, respectively.

3.3.3 Simulation study 1 results

The simulation results for sample size $N = 500$ are reported in Tables 3.1–3.5 and Figures 3.1 and 3.2. In our simulation studies, when the all models are correctly specified, the estimates of the difference in adjusted RMST for all methods showed the unbiased estimators, except for IP weighted Cox model when the censoring rate was 0.9. In terms of the coverage rate, all methods tended to achieve the desired coverage rate (i.e., 0.95) and did not appear to be conservative, except that the true PS was 0.8 and the censoring rate was 0.9. When the true PS was 0.8 and the censoring rate was 0.9 (specifically, when $\tau = 3$), the estimate of the IP weighted Cox model was slightly biased, and the coverage rate of the confidence interval exhibited poor coverage rate.

The simulation results for sample size N = 1,000 are reported in Tables A.3.1–A.3.5 and Figures A.3.1 and A.3.2. These results were similar to

Table 3.1 Simulation study 1 results ($N = 500$): Pseudo-observation

τ	RMST (True)	$E(A\|\boldsymbol{L})$	Censoring rate	RMST (Estimate)	Bias	95% confidence interval	Coverage rate
1	-0.027	0.5	0.1	-0.027	0.000	[-0.050, -0.004]	0.947
			0.3	-0.027	0.000	[-0.050, -0.003]	0.944
			0.5	-0.027	0.000	[-0.052, -0.003]	0.943
			0.7	-0.027	0.001	[-0.054, 0.000]	0.944
			0.9	-0.026	0.001	[-0.060, 0.005]	0.938
		0.8	0.1	-0.027	0.000	[-0.052, -0.001]	0.928
			0.3	-0.029	-0.002	[-0.053, -0.003]	0.926
			0.5	-0.027	0.000	[-0.053, 0.001]	0.925
			0.7	-0.028	-0.001	[-0.056, 0.001]	0.920
			0.9	-0.027	0.000	[-0.062, 0.010]	0.913
3	-0.417	0.5	0.1	-0.416	0.001	[-0.559, -0.269]	0.957
			0.3	-0.416	0.001	[-0.564, -0.259]	0.935
			0.5	-0.415	0.001	[-0.581, -0.247]	0.944
			0.7	-0.408	0.008	[-0.600, -0.213]	0.942
			0.9	-0.414	0.002	[-0.710, -0.117]	0.947
		0.8	0.1	-0.427	-0.010	[-0.595, -0.257]	0.940
			0.3	-0.424	-0.007	[-0.603, -0.243]	0.947
			0.5	-0.421	-0.005	[-0.617, -0.219]	0.948
			0.7	-0.423	-0.007	[-0.655, -0.182]	0.946
			0.9	-0.423	-0.006	[-0.774, -0.058]	0.942

Table 3.2 Simulation study 1 results ($N = 500$): ANCOVA-type model

τ	RMST (True)	$E(A\|\boldsymbol{L})$	Censoring rate	RMST (Estimate)	Bias	95% confidence interval	Coverage rate
1	-0.027	0.5	0.1	-0.027	0.000	[-0.050, -0.004]	0.948
			0.3	-0.027	0.000	[-0.051, -0.003]	0.943
			0.5	-0.027	0.000	[-0.052, -0.003]	0.947
			0.7	-0.027	0.000	[-0.054, 0.000]	0.944
			0.9	-0.026	0.001	[-0.061, 0.006]	0.938
		0.8	0.1	-0.027	0.000	[-0.053, -0.001]	0.929
			0.3	-0.029	-0.002	[-0.054, -0.003]	0.922
			0.5	-0.027	0.000	[-0.053, 0.001]	0.923
			0.7	-0.028	-0.001	[-0.057, 0.001]	0.919
			0.9	-0.027	0.000	[-0.063, 0.011]	0.907
3	-0.417	0.5	0.1	-0.416	0.001	[-0.558, -0.270]	0.958
			0.3	-0.416	0.001	[-0.564, -0.259]	0.934
			0.5	-0.416	0.001	[-0.582, -0.246]	0.952
			0.7	-0.409	0.007	[-0.607, -0.209]	0.942
			0.9	-0.421	-0.004	[-0.773, -0.091]	0.959
		0.8	0.1	-0.427	-0.010	[-0.597, -0.259]	0.934
			0.3	-0.424	-0.008	[-0.604, -0.240]	0.949
			0.5	-0.422	-0.006	[-0.618, -0.218]	0.945
			0.7	-0.422	-0.006	[-0.657, -0.181]	0.941
			0.9	-0.432	-0.015	[-0.856, -0.022]	0.940

Table 3.3 Simulation study 1 results ($N = 500$): IP weighted Cox model

τ	RMST (True)	$E(A\|\boldsymbol{L})$	Censoring rate	RMST (Estimate)	Bias	95% confidence interval	Coverage rate
1	-0.027	0.5	0.1	-0.027	0.000	[-0.050, -0.004]	0.950
			0.3	-0.027	0.000	[-0.050, -0.004]	0.937
			0.5	-0.027	0.000	[-0.052, -0.003]	0.940
			0.7	-0.026	0.001	[-0.053, 0.000]	0.943
			0.9	-0.026	0.001	[-0.059, 0.005]	0.933
		0.8	0.1	-0.027	0.000	[-0.051, -0.002]	0.918
			0.3	-0.028	-0.001	[-0.052, -0.005]	0.928
			0.5	-0.026	0.001	[-0.051, 0.000]	0.913
			0.7	-0.027	0.000	[-0.053, 0.000]	0.913
			0.9	-0.024	0.003	[-0.058, 0.016]	0.905
3	-0.417	0.5	0.1	-0.414	0.002	[-0.557, -0.267]	0.955
			0.3	-0.414	0.002	[-0.564, -0.258]	0.938
			0.5	-0.414	0.003	[-0.580, -0.243]	0.946
			0.7	-0.406	0.010	[-0.598, -0.211]	0.944
			0.9	-0.408	0.008	[-0.720, -0.072]	0.954
		0.8	0.1	-0.420	-0.003	[-0.589, -0.248]	0.928
			0.3	-0.416	0.000	[-0.594, -0.231]	0.941
			0.5	-0.414	0.003	[-0.609, -0.210]	0.941
			0.7	-0.417	-0.001	[-0.642, -0.180]	0.943
			0.9	-0.348	0.068	[-0.717, 0.586]	0.877

Table 3.4 Simulation study 1 results ($N = 500$): Adjusted Kaplan-Meier

τ	RMST (True)	$E(A\|\boldsymbol{L})$	Censoring rate	RMST (Estimate)	Bias	95% confidence interval	Coverage rate
1	-0.027	0.5	0.1	-0.027	0.000	[-0.050, -0.004]	0.950
			0.3	-0.027	0.000	[-0.051, -0.004]	0.936
			0.5	-0.027	0.000	[-0.052, -0.003]	0.940
			0.7	-0.027	0.001	[-0.054, 0.000]	0.943
			0.9	-0.026	0.001	[-0.060, 0.005]	0.933
		0.8	0.1	-0.027	0.001	[-0.051, -0.002]	0.918
			0.3	-0.028	-0.001	[-0.052, -0.005]	0.927
			0.5	-0.026	0.001	[-0.051, 0.000]	0.914
			0.7	-0.027	0.000	[-0.054, 0.000]	0.913
			0.9	-0.026	0.001	[-0.059, 0.006]	0.907
3	-0.417	0.5	0.1	-0.415	0.001	[-0.559, -0.268]	0.956
			0.3	-0.416	0.001	[-0.566, -0.259]	0.936
			0.5	-0.415	0.001	[-0.582, -0.244]	0.946
			0.7	-0.408	0.008	[-0.601, -0.212]	0.946
			0.9	-0.415	0.002	[-0.714, -0.120]	0.951
		0.8	0.1	-0.418	-0.002	[-0.588, -0.245]	0.929
			0.3	-0.415	0.002	[-0.594, -0.228]	0.944
			0.5	-0.412	0.005	[-0.609, -0.206]	0.943
			0.7	-0.415	0.002	[-0.642, -0.174]	0.942
			0.9	-0.409	0.007	[-0.749, -0.023]	0.931

Table 3.5 Simulation study 1 results ($N = 500$): G-computation

τ	RMST (True)	$E(A\|\boldsymbol{L})$	Censoring rate	RMST (Estimate)	Bias	95% confidence interval	Coverage rate
1	-0.027	0.5	0.1	-0.027	0.000	[-0.038, -0.017]	0.946
			0.3	-0.027	0.000	[-0.039, -0.016]	0.943
			0.5	-0.027	0.000	[-0.040, -0.015]	0.958
			0.7	-0.027	0.000	[-0.043, -0.014]	0.955
			0.9	-0.027	0.000	[-0.051, -0.007]	0.944
		0.8	0.1	-0.027	0.000	[-0.038, -0.018]	0.943
			0.3	-0.027	0.000	[-0.038, -0.017]	0.941
			0.5	-0.027	0.000	[-0.039, -0.016]	0.950
			0.7	-0.027	0.000	[-0.041, -0.013]	0.935
			0.9	-0.026	0.001	[-0.049, -0.005]	0.950
3	-0.417	0.5	0.1	-0.416	0.000	[-0.518, -0.327]	0.953
			0.3	-0.415	0.002	[-0.526, -0.313]	0.954
			0.5	-0.416	0.000	[-0.542, -0.294]	0.947
			0.7	-0.412	0.004	[-0.573, -0.256]	0.949
			0.9	-0.411	0.006	[-0.686, -0.142]	0.947
		0.8	0.1	-0.418	-0.001	[-0.517, -0.319]	0.935
			0.3	-0.415	0.002	[-0.526, -0.301]	0.953
			0.5	-0.411	0.005	[-0.543, -0.277]	0.952
			0.7	-0.412	0.005	[-0.581, -0.237]	0.946
			0.9	-0.401	0.015	[-0.697, -0.086]	0.951

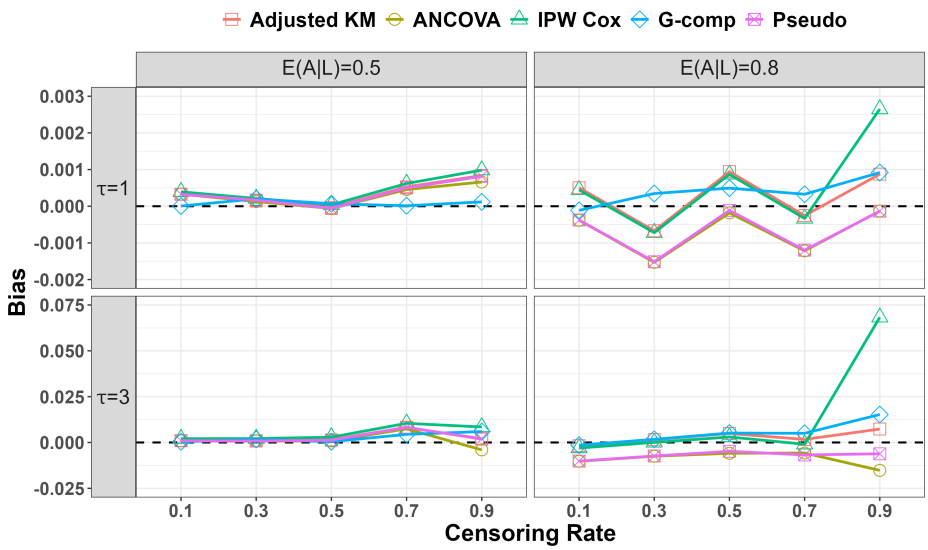


Figure 3.1 Bias for simulation study 1 ($N = 500$). Adjusted KM = adjusted Kaplan-Meier estimator; ANCOVA = ANCOVA-type model; IPW Cox = IP weighted Cox model; G-comp = G-computation; Pseudo = pseudo-observation. Dashed line in the plot represents a bias of 0.

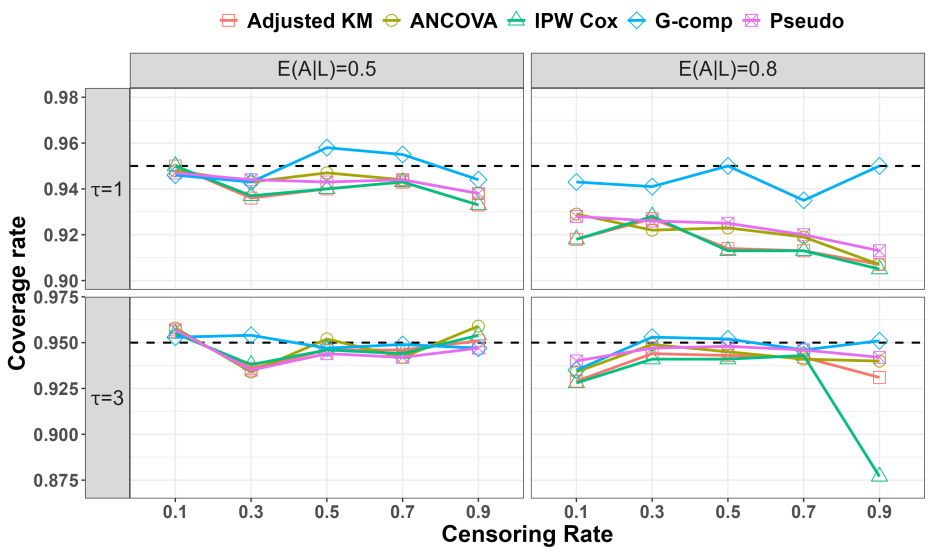


Figure 3.2 Coverage rate for simulation study 1 ($N = 500$). Adjusted KM = adjusted Kaplan-Meier estimator; ANCOVA = ANCOVA-type model; IPW Cox = IP weighted Cox model; G-comp = G-computation; Pseudo = pseudo-observation. Dashed line in the plot represents a coverage rate of 0.95.

those in sample size $N = 500$. Unlike the results for sample size $N = 500$, when the true PS was 0.8 and the censoring rate was 0.9, the bias for the estimate of the IP weighted Cox model was vanished, and the coverage rate of the confidence interval achieved the nominal coverage rate 0.95.

In summary, when there is neither a large censoring rate nor extreme PS, we confirmed that all the methods being considered provide the unbiased estimates with the percentile bootstrap confidence intervals achieving near nominal coverage probability. As the sample size increased, the bias was reduced and the coverage rate seemed to improve.

3.4 Real data analysis 1: Colon cancer data

Colon cancer data were included patients who (a) were newly diagnosed with colon adenocarcinoma from January 1, 2009 to December 31, 2015, (b) had early-stage colon cancer, which was defined as clinical stage 0/I disease on staging abdominal computed tomography (CT), and (c) did not have synchronous rectal cancer located within 15 cm of the anal verge (Lee et al., 2023). Patients were divided into either of with-staging chest CT or without-staging chest CT group according to whether they underwent staging chest CT. The with-staging chest CT group included 606 patients, and the without-staging chest CT group included 385 patients. Survival outcome was overall survival, which was defined as the time interval from the date of staging abdominal CT scan to the date of death from any cause. The date of staging abdominal CT scan was different for each patients. Patients without death records were administrative censored on December 31, 2019, and there is no censoring apart from administrative censoring at the end of follow-up. Of 991 patients, there were 111 events (76 for with-staging chest CT group versus 35 for without-staging chest CT group).

In Lee et al. (2023), the minimally sufficient adjustment set of covariates which was consisted of prognostic factors that needed to be adjusted for included age at study entry (years), sex (male vs. female), smoking behavior (never-, former-, and current-smoker), year at diagnosis (2009-2012 vs. 2013-2015), referral from another hospital (yes vs. no), Charlson comorbidity index (no, mild, moderate, and severe), family history of colonic neoplasm (yes vs. no), family history of any cancer (yes vs. no), endoscopic appearance (superficial vs. advanced), and histologic grade via endoscopic procedure (well, moderately, and poorly differentiated). The minimally sufficient adjustment set of covariates was identified from the causal diagram. Detailed information about data is given in the supplementary material of Lee et al. (2023). They measured the difference in RMST between the exposure groups (with-staging chest CT group minus without-staging chest CT group). Therefore, negative values of the difference in RMST indicate an increased risk after the use of staging chest CT. The difference in unadjusted RMST was estimated using the Kaplan–Meier estimator (3.1). The difference in adjusted RMST was estimated using the adjusted survival curve from the IP weighted pooled logistic regression model. See details for estimating the difference in adjusted RMST from the IP weighted pooled logistic regression model in Appendix A.4.

Using this data, we compared whether there was a difference between the methods for the estimates of the difference in average overall survival time up to τ. Because Lee et al. (2023) were considered a 5 years RMST in the primary analysis, we also set a specific time point τ to 5 years (12×5 months). For five methods described in Section 3.2, we estimated the differences in adjusted RMST and constructed the 95% percentile bootstrap confidence intervals with

Table 3.6 Real data analysis results for colon cancer data

Method	RMST difference (Months)	95% confidence interval (Months)	
		Lower	Upper
Unadjusted Kaplan-Meier estimator	-0.779	-1.620	-0.014
Pooled logistic regression	0.424	-0.797	2.058
Pseudo-observation	0.402	-0.981	2.450
ANCOVA-type model	0.459	-0.804	1.754
IP weighted Cox model	0.486	-1.038	2.877
Adjusted Kaplan-Meier estimator	0.492	-1.037	2.899
G-computation	0.481	-0.535	1.653

bootstrap resampling $B = 1{,}000$ times. And, we compared the results with those of Lee et al. (2023).

The results are reported in Table 3.6. The results in the first two rows of Table 3.6 (i.e., unadjusted Kaplan-Meier estimator and pooled logistic regression) were the same as in Lee et al. (2023). The point estimate of the difference in unadjusted RMST up to 5 years was statistically significant (-0.779 [95% CI: -1.620 to -0.014] months). However, after adjusting for potential confounders, the point estimates of the difference in adjusted RMST up to 5 years were not statistically significant although there were sightly different in terms of the point estimates and confidence intervals. Therefore, after adjusting for potential confounders, the use of staging chest CT did not affect the overall survival in patients with early-stage colon cancer.

Chapter 4

Sensitivity Analysis

In this Chapter, using an adjusted Kaplan-Meier curve (Xie and Liu, 2005) described in Section 3.2.2.2 and adapting a parametric sensitivity model for non-survival data (Zhao, Small, and Bhattacharya, 2019), we propose a novel sensitivity analysis method for the difference in adjusted RMST that can be performed when the residual confounding by unmeasured confounders is suspected. Alternative sensitivity analysis methods for the other estimation methods and their limitations are discussed in Chapter 5.

The proposed method will be formulated an optimization problem. Simple analytic solutions are derived for some special cases. In general cases, we explain the practical optimization methods to solve the optimization problem.

4.1 Background

Note that in order to interpret the estimate of the difference in adjusted RMST (3.7) as a causal effect, the conditional exchangeability assumption (**A3**) is critical. When there exists an unmeasured confounder, this opens the backdoor path, and thus the assumption (**A3**) is violated (Pearl, 1995). Furthermore,

$\widehat{P}(A_i = a \mid \boldsymbol{L}_i)$ may not be a consistent estimate of the true PS, which makes (3.7) a biased estimate of the causal effect. The key problem is that it is generally untestable from observed data to check whether the assumption (**A3**) is violated (Greenland and Robins, 1986). Therefore, a sensitivity analysis should be considered to investigate how sensitive the estimate of the difference in adjusted RMST (3.7) is to the degree of unmeasured confounding. Therefore, we propose a novel sensitivity analysis method which considers unmeasured confounding for evaluating the estimate of the difference in adjusted RMST.

4.2 Sensitivity model

We extend the marginal sensitivity model proposed by Tan (2006) and Zhao, Small, and Bhattacharya (2019) to survival analysis and follow the notation stated in Rosenbaum (2002) and Dorn and Guo (2022). Assume that if we adjusted for unmeasured confounders (which we denoted by $\boldsymbol{U}$) along with measured confounders $\boldsymbol{L}$, then all confounding is removed. Thus, we can define $e_0(\boldsymbol{l}, \boldsymbol{u}) = P(A = 1 \mid \boldsymbol{L} = \boldsymbol{l}, \boldsymbol{U} = \boldsymbol{u})$ as the true PS. Also, denote $e(\boldsymbol{l}) = P(A = 1 \mid \boldsymbol{L} = \boldsymbol{l})$ as the PS only based on measured confounders $\boldsymbol{L}$. The following is a key definition required to propose our sensitivity analysis method.

Definition 4.1 (**Zhao's marginal sensitivity model**) For a fixed sensitivity parameter $\Lambda \geq 1$, assume that the true PS $e_0(\boldsymbol{l}, \boldsymbol{u}) \in \mathcal{E}(\Lambda)$ where the set of marginal sensitivity models is defined by

$$\mathcal{E}(\Lambda) = \left\{ e(\boldsymbol{l}, \boldsymbol{u}) : \frac{1}{\Lambda} \leq \frac{\text{odds}\{e(\boldsymbol{l}, \boldsymbol{u})\}}{\text{odds}\{e(\boldsymbol{l})\}} \leq \Lambda, \text{ for all } \boldsymbol{l} \in \boldsymbol{L} \text{ and } \boldsymbol{u} \in \boldsymbol{U} \right\}$$

and $\text{odds}\{p\} = p/(1-p)$.

Remark 4.1 The marginal sensitivity model implies that within each stratum of the measured confounders $\boldsymbol{L}$, measuring the unmeasured confounders $\boldsymbol{U}$ can only change the odds of $e(\boldsymbol{l})$ by a factor of at most Λ. It thus means that under the specified sensitivity model $e(\boldsymbol{l}, \boldsymbol{u})$,

$$P\left(\text{odds}\{e(\boldsymbol{l}, \boldsymbol{u})\} \in \left[\frac{1}{\Lambda} \times \text{odds}\{e(\boldsymbol{l})\},\ \Lambda \times \text{odds}\{e(\boldsymbol{l})\}\right] \,\middle|\, \boldsymbol{U} = \boldsymbol{u}\right) = 1.$$

Practically, $e(\boldsymbol{l})$ is often estimated parametrically, and thus we can denote its estimate as $e_\beta(\boldsymbol{l})$ which is closest to $e(\boldsymbol{l})$ in Kullback–Leibler divergence. As in Zhao, Small, and Bhattacharya (2019), it is handy to represent $e_\beta(\boldsymbol{l})$ and $e_0(\boldsymbol{l}, \boldsymbol{u})$ as log odds scale. Denote $g_\beta(\boldsymbol{l}) = \text{logit}\{e_\beta(\boldsymbol{l})\}$ and $g_0(\boldsymbol{l}, \boldsymbol{u}) = \text{logit}\{e_0(\boldsymbol{l}, \boldsymbol{u})\}$ where $\text{logit}\{p\} = \log\{p/(1-p)\}$. And, let $h_{\beta_0}(\boldsymbol{l}, \boldsymbol{u}) = g_\beta(\boldsymbol{l}) - g_0(\boldsymbol{l}, \boldsymbol{u})$ be the log odds scale difference between $e_\beta(\boldsymbol{l})$ and $e_0(\boldsymbol{l}, \boldsymbol{u})$. Through the parameterization of PS and the newly introduced notations, we can redefine Definition 4.1 as follows:

Definition 4.2 (**Zhao's parametric sensitivity model**) For $\Lambda \geq 1$, assume that $h_{\beta_0}(\boldsymbol{l}, \boldsymbol{u}) \in \varepsilon(\Lambda)$ where the set of parametric sensitivity models is defined by

$$\varepsilon(\Lambda) = \left\{h(\boldsymbol{l}, \boldsymbol{u}) : \sup |h(\boldsymbol{l}, \boldsymbol{u})| \leq \log(\Lambda),\ \text{for all } \boldsymbol{l} \in \boldsymbol{L} \text{ and } \boldsymbol{u} \in \boldsymbol{U}\right\}.$$

The parametric sensitivity model $h(\boldsymbol{l}, \boldsymbol{u})$ implies that measuring $\boldsymbol{U}$ can only change the log odds of PS, $g_\beta(\boldsymbol{l})$, by a factor of at most $\log(\Lambda)$. Accordingly, we can define the following shifted PS.

Definition 4.3 (**Shifted propensity score**) Under any specified sensitivity model $h(\boldsymbol{l}, \boldsymbol{u}) \in \varepsilon(\Lambda)$, the shifted PS is defined by

$$e^{(h)}(\boldsymbol{l}, \boldsymbol{u}) = \left[1 + \exp\{h(\boldsymbol{l}, \boldsymbol{u}) - g_\beta(\boldsymbol{l})\}\right]^{-1}. \tag{4.1}$$

Remark 4.2 Practically, the shifted PS (4.1) is an alternative PS obtained by taking into account a user-specified sensitivity parameter Λ which reflects the degree of unmeasured confounding. Note that if $\Lambda = 1$ and thus $h(\boldsymbol{l}, \boldsymbol{u}) = h_{\beta_0}(\boldsymbol{l}, \boldsymbol{u})$, then $e^{(h)}(\boldsymbol{l}, \boldsymbol{u})$ is equal to the true PS $e_0(\boldsymbol{l}, \boldsymbol{u})$, so there is no unmeasured confounding.

The corresponding shifted PS (4.1) can be estimated by

$$\widehat{e}^{(h)}(\boldsymbol{l}, \boldsymbol{u}) = \frac{1}{1 + \exp\{h(\boldsymbol{l}, \boldsymbol{u}) - \widehat{g}_\beta(\boldsymbol{l})\}} \tag{4.2}$$

where $\widehat{g}_\beta(\boldsymbol{l}) = \text{logit}\{\widehat{P}_\beta(A = 1 \mid \boldsymbol{L} = \boldsymbol{l})\}$ which is estimated from observed data.

4.3 Estimate of difference in bias-adjusted RMST

When unmeasured confounding is suspected, by substituting $\widehat{P}(A = 1 \mid \boldsymbol{L})$ in the estimate of the difference in adjusted RMST (3.7) with the estimate of the shifted PS (4.2), we can rewrite expression (3.7) as

$$\begin{aligned}\widehat{\mu}_\tau^{(h),1} - \widehat{\mu}_\tau^{(h),0} = \int_0^\tau \Bigg[&\prod_{t_j \le t} \left(1 - \frac{\sum_{i:T_i=t_j} \frac{\delta_i I(A_i=1)}{\widehat{e}^{(h)}(\boldsymbol{l}_i, \boldsymbol{u}_i)}}{\sum_{i:T_i \ge t_j} \frac{I(A_i=1)}{\widehat{e}^{(h)}(\boldsymbol{l}_i, \boldsymbol{u}_i)}}\right) \\ &- \prod_{t_j \le t} \left(1 - \frac{\sum_{i:T_i=t_j} \frac{\delta_i I(A_i=0)}{1-\widehat{e}^{(h)}(\boldsymbol{l}_i, \boldsymbol{u}_i)}}{\sum_{i:T_i \ge t_j} \frac{I(A_i=0)}{1-\widehat{e}^{(h)}(\boldsymbol{l}_i, \boldsymbol{u}_i)}}\right) \Bigg] \, \mathrm{d}t \end{aligned} \tag{4.3}$$

and denote the set of (4.3)

$$\{\widehat{\mu}_\tau^{(h),1} - \widehat{\mu}_\tau^{(h),0} : h(\boldsymbol{l}, \boldsymbol{u}) \in \varepsilon(\Lambda)\} \tag{4.4}$$

as a partially identified region of (4.3). We will refer to (4.3) simply as "the estimate of the difference in bias-adjusted RMST" in the rest of this thesis.

Note that because the true value of $e^{(h)}(\boldsymbol{l}, \boldsymbol{u})$ is generally not identifiable from data, the value of (4.3) cannot directly estimated. However, we can estimate the sensitivity range (i.e., minimum and maximum) of the point estimates (4.3) and confidence interval for partially identified region (4.4), given the specified sensitivity parameter Λ. We defer the construction of the confidence interval until Section 4.7. For now, we will focus on constructing the sensitivity range of (4.3).

4.4 Sensitivity range

In Section 3.1, we assumed that the first n of N subjects are exposed and the rest $N - n$ are unexposed. We can simplify expression for the estimate of the difference in bias-adjusted RMST (4.3) by substituting $\widehat{e}^{(h)}(\boldsymbol{l}, \boldsymbol{u})$ in (4.3) with the right-hand side of (4.2) and by introducing the variables w_i and z_i for each exposure group separately. Let $w_i = \exp\{-\widehat{g}_\beta(\boldsymbol{L}_i)\}$ for the exposed group $(i = 1, \ldots, n)$ and $w_i = \exp\{\widehat{g}_\beta(\boldsymbol{L}_i)\}$ for the unexposed group $(i = n+1, \ldots, N)$, respectively. Also, let $z_i = \exp\{h(\boldsymbol{L}_i, \boldsymbol{U}_i)\}$ for the exposed group and $z_i = \exp\{-h(\boldsymbol{L}_i, \boldsymbol{U}_i)\}$ for the unexposed group, respectively. Because the postulated sensitivity model $h(\boldsymbol{L}_i, \boldsymbol{U}_i) \in [-\log(\Lambda), \log(\Lambda)]$, it is clear that $z_i \in [1/\Lambda, \Lambda]$. Therefore, the sensitivity range of (4.3) can be evaluated by solving the optimization problem as follows:

$$\begin{aligned}
\min \text{or} \max \; & \int_0^\tau \left[\prod_{t_j \le t} \left(1 - \frac{\sum_{i:T_i=t_j} \delta_i I(A_i = 1)[1 + z_i w_i]}{\sum_{i:T_i \ge t_j} I(A_i = 1)[1 + z_i w_i]} \right) \right. \\
& \left. - \prod_{t_j \le t} \left(1 - \frac{\sum_{i:T_i=t_j} \delta_i I(A_i = 0)[1 + z_i w_i]}{\sum_{i:T_i \ge t_j} I(A_i = 0)[1 + z_i w_i]} \right) \right] \mathrm{d}t \\
\text{subject to} \quad & \frac{1}{\Lambda} \le z_i \le \Lambda, \text{ for } i = 1, \ldots, N. \qquad (4.5)
\end{aligned}$$

The lower and upper bounds of sensitivity range correspond to the solution of minimization and maximization problems in (4.5), respectively. Note that solving the optimization problem (4.5) only depend on the optimization parameter z_i. Therefore, when performing our sensitivity analysis method, it does not matter which model is used to estimate $g_\beta(\boldsymbol{L}_i)$.

Because the optimization problem (4.5) can be separated into two parts related to the exposed and unexposed groups, we can obtain the solution of minimization or maximization problem in (4.5) by 1) solving both minimization or maximization problem for the exposed group

$$\begin{aligned} &\min \text{or} \max \int_0^\tau \prod_{t_j \le t} \left(1 - \frac{\sum_{i:T_i=t_j} \delta_i[1 + z_i w_i]}{\sum_{i:T_i \ge t_j}[1 + z_i w_i]}\right) \mathrm{d}t \\ &\text{subject to} \quad \frac{1}{\Lambda} \le z_i \le \Lambda, \text{ for } i = 1, \ldots, n, \end{aligned} \tag{4.6}$$

where $w_i = \exp\{-\widehat{g}_\beta(\boldsymbol{L}_i)\}$ and $z_i = \exp\{h(\boldsymbol{L}_i, \boldsymbol{U}_i)\}$ and maximization or minimization problem for the unexposed group

$$\begin{aligned} &\max \text{or} \min \int_0^\tau \prod_{t_j \le t} \left(1 - \frac{\sum_{i:T_i=t_j} \delta_i[1 + z_i w_i]}{\sum_{i:T_i \ge t_j}[1 + z_i w_i]}\right) \mathrm{d}t \\ &\text{subject to} \quad \frac{1}{\Lambda} \le z_i \le \Lambda, \text{ for } i = n+1, \ldots, N, \end{aligned} \tag{4.7}$$

where $w_i = \exp\{\widehat{g}_\beta(\boldsymbol{L}_i)\}$ and $z_i = \exp\{-h(\boldsymbol{L}_i, \boldsymbol{U}_i)\}$ and 2) computing the difference between solutions of (4.6) and (4.7).

Note that to minimize the objective function in (4.6) or (4.7) subject to the optimization parameter $z_i \in [1/\Lambda, \Lambda]$, z_i's for subjects who are censored should be equal to $1/\Lambda$ because z_i's for censored subjects are only included in the denominator. Similarly, to maximize the objective function in (4.6) or (4.7), z_i's for censored subjects should be equal to Λ. Therefore, by using the event indicator δ_i, the optimization problems (4.6) and (4.7) can be divided

according to whether or not subjects experience the event of interest. Consider the minimization or maximization problem in (4.6) first (i.e., optimization problems for the exposed group). Each problem in (4.6) can be simplified as follows:

$$\begin{aligned} &\min \int_0^\tau \prod_{t_j \le t} \left(1 - \frac{\sum_{i:\{T_i = t_j, \delta_i = 1\}} (1 + z_i w_i)}{\sum_{i:\{T_i \ge t_j, \delta_i = 1\}} (1 + z_i w_i) + \sum_{i:\{T_i \ge t_j, \delta_i = 0\}} (1 + \frac{1}{\Lambda} w_i)} \right) \mathrm{d}t \\ &\text{subject to} \quad \frac{1}{\Lambda} \le z_i \le \Lambda, \text{ for } i = 1, \ldots, n, \end{aligned} \tag{4.8}$$

or

$$\begin{aligned} &\max \int_0^\tau \prod_{t_j \le t} \left(1 - \frac{\sum_{i:\{T_i = t_j, \delta_i = 1\}} (1 + z_i w_i)}{\sum_{i:\{T_i \ge t_j, \delta_i = 1\}} (1 + z_i w_i) + \sum_{i:\{T_i \ge t_j, \delta_i = 0\}} (1 + \Lambda w_i)} \right) \mathrm{d}t \\ &\text{subject to} \quad \frac{1}{\Lambda} \le z_i \le \Lambda, \text{ for } i = 1, \ldots, n, \end{aligned} \tag{4.9}$$

where $w_i = \exp\{-\widehat{g}_\beta(\boldsymbol{L}_i)\}$ and $z_i = \exp\{h(\boldsymbol{L}_i, \boldsymbol{U}_i)\}$. As can be seen in optimization problems (4.8) and (4.9), z_i's for censored subjects are already determined as $1/\Lambda$ or Λ, so that we only need to solve the problems for m (the number of events in the exposed group) out of n optimization parameters $\boldsymbol{z} = (z_1, \ldots, z_n)$. It implies that as the censoring rate decreases, the more $\boldsymbol{z}$ need to be determined, and thus the more computational time is required to obtain the solution. Ultimately, the censoring rate is a key factor for computational time in our optimization problem. The maximization or minimization problem in (4.7) can be simplified in a similar way (i.e., optimization problem for the unexposed group):

$$\begin{aligned} &\max \int_0^\tau \prod_{t_j \le t} \left(1 - \frac{\sum_{i:\{T_i = t_j, \delta_i = 1\}} (1 + z_i w_i)}{\sum_{i:\{T_i \ge t_j, \delta_i = 1\}} (1 + z_i w_i) + \sum_{i:\{T_i \ge t_j, \delta_i = 0\}} (1 + \Lambda w_i)} \right) \mathrm{d}t \\ &\text{subject to} \quad \frac{1}{\Lambda} \le z_i \le \Lambda, \text{ for } i = n+1, \ldots, N, \end{aligned} \tag{4.10}$$

or

$$\min \int_0^\tau \prod_{t_j \le t} \left(1 - \frac{\sum_{i:\{T_i=t_j,\delta_i=1\}}(1+z_i w_i)}{\sum_{i:\{T_i\ge t_j,\delta_i=1\}}(1+z_i w_i) + \sum_{i:\{T_i\ge t_j,\delta_i=0\}}(1+\frac{1}{\Lambda} w_i)} \right) \mathrm{d}t$$

$$\text{subject to} \quad \frac{1}{\Lambda} \le z_i \le \Lambda, \text{ for } i = n+1, \ldots, N, \tag{4.11}$$

where $w_i = \exp\{\widehat{g}_\beta(\boldsymbol{L}_i)\}$ and $z_i = \exp\{-h(\boldsymbol{L}_i, \boldsymbol{U}_i)\}$. Similarly in this case, we only need to solve the optimization problem for $D-m$ (the number of events in the unexposed group) out of $N-n$ optimization parameters $\boldsymbol{z} = (z_{n+1}, \ldots, z_N)$.

4.5 Analytic solution to bias-adjusted RMST in special case

There are special settings in which optimization problems (4.8–4.11) can be solved analytically. Consider a closed cohort that the study entry times $t_0 = 0$ are the same for all subjects and there is no censoring apart from administrative censoring at the end of follow-up. Without loss of generality, let the first m subjects experience the event of interest in the exposed group (i.e., $\delta_1 = \cdots = \delta_m = 1$ and $\delta_{m+1} = \cdots = \delta_n = 0$) and the first $D-m$ subjects experience the event in the unexposed group (i.e., $\delta_{n+1} = \cdots = \delta_{n-m+D} = 1$ and $\delta_{n-m+D+1} = \cdots = \delta_N = 0$). Also, assume that the event times are continuous and ordered increasingly (i.e., $t_1 < \cdots < t_m$ in the exposed group and $t_{n+1} < \cdots < t_{n-m+D}$ in the unexposed group). Since we assume that the only censoring is due to administrative censoring, the last event time in each group (t_m or t_{n-m+D}) is less than or equal to the administrative censoring time.

When the pre-specified time point $\tau \in (t_{k-1}, t_k]$ for any $k \in \{2, \ldots, m\}$, the objective function in (4.8) can be reduced to linear fractional programming, as

follows (a detailed proof is given in Appendix B.1):

$$\begin{aligned}
&\int_0^\tau \prod_{t_j \le t} \left(1 - \frac{\sum_{i:\{T_i=t_j,\delta_i=1\}}(1+z_i w_i)}{\sum_{i:\{T_i\ge t_j,\delta_i=1\}}(1+z_i w_i) + \sum_{i:\{T_i\ge t_j,\delta_i=0\}}(1+\frac{1}{\Lambda}w_i)}\right) \mathrm{d}t \\
=&\sum_{l=0}^{k-1} \prod_{t_j \le t_l} \left(\frac{\sum_{i:\{T_i>t_j,\delta_i=1\}}(1+z_i w_i)+\sum_{i:\{T_i\ge t_j,\delta_i=0\}}(1+\frac{1}{\Lambda}w_i)}{\sum_{i:\{T_i\ge t_j,\delta_i=1\}}(1+z_i w_i)+\sum_{i:\{T_i\ge t_j,\delta_i=0\}}(1+\frac{1}{\Lambda}w_i)}\right)(t_{l+1}-t_l) \\
=&\frac{\sum_{i=1}^n (1+z_i w_i)t_i}{\sum_{i=1}^n (1+z_i w_i)}
\end{aligned}$$

where $\boldsymbol{t} = (t_1, \ldots, t_n) = (t_1, \ldots, t_{k-1}, \tau, \ldots, \tau)$, $w_i = \exp\{-\widehat{g}_\beta(\boldsymbol{L}_i)\}$, and $\boldsymbol{z} = (z_1, \ldots, z_n) = (z_1, \ldots, z_m, 1/\Lambda, \ldots, 1/\Lambda)$, for $i = 1, \ldots, n$. Therefore, the optimization problems (4.8) and (4.9) can be computed by using linear fractional programming

$$\begin{aligned}
&\min \text{or} \max \ \frac{\sum_{i=1}^n (1+z_i w_i)t_i}{\sum_{i=1}^n (1+z_i w_i)} \\
&\text{subject to} \quad \frac{1}{\Lambda} \le z_i \le \Lambda, \ \text{for } i = 1, \ldots, n
\end{aligned} \tag{4.12}$$

where $\boldsymbol{t} = (t_1, \ldots, t_{k-1}, \tau, \ldots, \tau)$, $w_i = \exp\{-\widehat{g}_\beta(\boldsymbol{L}_i)\}$, and $\boldsymbol{z} = (z_1, \ldots, z_n) = (z_1, \ldots, z_m, 1/\Lambda, \ldots, 1/\Lambda)$ for the minimization problem and $\boldsymbol{z} = (z_1, \ldots, z_m, \Lambda, \ldots, \Lambda)$ for the maximization problem. Also, when $\tau \in (t_{k-1}, t_k]$ for any $k \in \{n+2, \ldots, n-m+D\}$, the optimization problems (4.10) and (4.11) can be computed by using linear fractional programming

$$\begin{aligned}
&\max \text{or} \min \ \frac{\sum_{i=n+1}^N (1+z_i w_i)t_i}{\sum_{i=n+1}^N (1+z_i w_i)} \\
&\text{subject to} \quad \frac{1}{\Lambda} \le z_i \le \Lambda, \ \text{for } i = n+1, \ldots, N
\end{aligned} \tag{4.13}$$

where $\boldsymbol{t} = (t_{n+1}, \ldots, t_N) = (t_{n+1}, \ldots, t_{k-1}, \tau, \ldots, \tau)$, $w_i = \exp\{\widehat{g}_\beta(\boldsymbol{L}_i)\}$, and $\boldsymbol{z} = (z_{n+1}, \ldots, z_N) = (z_{n+1}, \ldots, z_{n-m+D}, \Lambda, \ldots, \Lambda)$ for the maximization problem and $\boldsymbol{z} = (z_{n+1}, \ldots, z_{n-m+D}, 1/\Lambda, \ldots, 1/\Lambda)$ for the minimization problem.

Because the problems (4.12) and (4.13) can be transformed to linear programming by the Charnes-Cooper transformation (Charnes and Cooper, 1962), it can be computed effectively by solving linear programming based on Proposition 2 in Zhao, Small, and Bhattacharya (2019). It implies that to obtain the solution of optimization problems (4.8) and (4.9), we have only to 1) compute the objective functions of minimization and maximization problems in (4.12) for at most m candidates of $\boldsymbol{z}$ where

$$z_i = \begin{cases} \Lambda & \text{if } 1 \le i \le v \\ \frac{1}{\Lambda} & \text{if } v+1 \le i \le m \end{cases} \quad \& \quad z_i = \begin{cases} \frac{1}{\Lambda} & \text{if } 1 \le i \le v \\ \Lambda & \text{if } v+1 \le i \le m \end{cases} \tag{4.14}$$

for $v \in \{1, \ldots, m\}$, respectively, and 2) choose the minimum and maximum value of them as the solutions of minimization and maximization problems, respectively. Similarly, to obtain the solutions of the maximization and minimization problems in (4.13), we only need to consider at most $D-m$ candidates of $\boldsymbol{z}$ where

$$z_i = \begin{cases} \frac{1}{\Lambda} & \text{if } n+1 \le i \le v \\ \Lambda & \text{if } v+1 \le i \le n-m+D \end{cases} \quad \& \quad z_i = \begin{cases} \Lambda & \text{if } n+1 \le i \le v \\ \frac{1}{\Lambda} & \text{if } v+1 \le i \le n-m+D \end{cases} \tag{4.15}$$

for $v \in \{n+1, \ldots, n-m+D\}$, respectively.

There is an alternative setting which translates (4.8–4.11) into linear programming. Consider that a closed cohort where the study entry times are the same for all subjects. Without loss of generality, let the event times be continuous and ordered increasingly (i.e., $t_1 < \cdots < t_m$ in the exposed group and $t_{n+1} < \cdots < t_{n-m+D}$ in the unexposed group). Also, the censoring times are ordered non-decreasingly in each exposure group, respectively (i.e.,

$t_{m+1} \leq \ldots \leq t_n$ in the exposed group and $t_{n-m+D+1} \leq \ldots \leq t_N$ in the unexposed group). In the exposed group, if $\tau \in (t_{k-1}, t_k]$ and the minimum censoring time $t_{m+1} \geq t_{k-1}$ for certain $k \in \{2, \ldots, m\}$, then the problems (4.8) and (4.9) can be reduced to linear programming (4.12). A proof is included in Appendix B.2. Also, it can be considered similarly to the problems (4.10) and (4.11) in the unexposed group.

In real applications, if the both minimum censoring times in each exposure group are longer than or equal to τ, we solve the optimization problem by reducing it to linear programming problem. Otherwise, use the optimization methods explained in the next Section.

4.6 Methods for solution of optimization problem in general case

In many survival analysis, the study entry times are different for each subject, or censoring is not necessarily just administrative censoring. In this case, the optimization problems (4.8–4.11) cannot be translated into linear fractional programming, so some optimization parameters $\boldsymbol{z}$ may not converge to boundary value of $1/\Lambda$ or Λ but converge to value between $1/\Lambda$ and Λ. We prove this by constructing a counter-example (See details in Appendix B.3). In our counter-example, for simplicity, we consider only four subjects, all in the exposed group. Also, the study entry times are the same for all subjects. Additionally, let the first, third, and fourth subjects experience the event of interest, but the second subject be censored (i.e., $\delta_1 = \delta_3 = \delta_4 = 1$ and $\delta_2 = 0$). And, let the survival times be ordered increasingly (i.e., $t_1 < t_2 < t_3 < t_4$). Taking $\tau = t_4$, the objective function in (4.8) is represented as follows (a detailed proof is given

in Appendix B.3):

$$\int_0^\tau \prod_{t_j \le t} \left(1 - \frac{\sum_{i:\{T_i=t_j,\delta_i=1\}}(1+z_i w_i)}{\sum_{i:\{T_i\ge t_j,\delta_i=1\}}(1+z_i w_i) + \sum_{i:\{T_i\ge t_j,\delta_i=0\}}(1+\frac{1}{\Lambda}w_i)}\right) \mathrm{d}t$$
$$= t_1 + \left(\frac{(1+\frac{1}{\Lambda}w_2) + (1+z_3 w_3) + (1+\frac{1}{\Lambda}w_4)}{(1+\Lambda w_1) + (1+\frac{1}{\Lambda}w_2) + (1+z_3 w_3) + (1+\frac{1}{\Lambda}w_4)}\right)$$
$$\times \left(\frac{(1+z_3 w_3)(t_3 - t_1) + (1+\frac{1}{\Lambda}w_4)(\tau - t_1)}{(1+z_3 w_3) + (1+\frac{1}{\Lambda}w_4)}\right). \tag{4.16}$$

Note that because the numerator and denominator of (4.16) are consisted by a quadratic function of z_3, this objective function can be locally convex or concave for z_3 between $1/\Lambda$ and Λ depending on the situation. In Appendix B.3, we showed that z_3 converges to values between $1/\Lambda$ and Λ with simple numerical data. It implies that when solving the optimization problems (4.8–4.11), some of z_i may not converge to boundary value $1/\Lambda$ or Λ.

Because there is no closed-form solution for almost $\boldsymbol{z}$ in general setting and each of $\boldsymbol{z}$ is ranged between $1/\Lambda$ and Λ, the optimum values of $\boldsymbol{z}$ can be determined by using an optimization algorithm such as the L-BFGS-B optimization method (Byrd et al., 1995) in the `optim` function of `stats` package in R. We refer to this method as "the direct optimization method". We used the `optimParallel` function in `optimParallel` package (Gerber and Furrer, 2019) which provides a parallel extension of the L-BFGS-B optimization method in the `optim` function. Although parallel processing reduces the computing time, as the censoring rate decreases, the more optimum values of $\boldsymbol{z}$ need to be determined, and thus the more computational time is taken to solve the optimization problem. Therefore, we propose alternative method with reasonable computational time to find the solution.

In minimizing the problem (4.8), it can be easily known that z_1 should be equal to Λ and z_m should be equal to $1/\Lambda$. When the problem (4.8) is solved

using the direct optimization method and the index i increases sequentially, it is empirically confirmed that z_i converges to Λ from $i = 1$ to one particular $i = u$ and converges to $1/\Lambda$ from another particular $i = u+v$ to $i = m$, for some unknown constants u and v. For $i \in (u, u + v)$, z_i can converge to any value between $1/\Lambda$ and Λ. Although there are some z_i's which may not converge to boundary value, we empirically founded that when there is neither a large censoring rate nor extreme PS, almost all of $\boldsymbol{z}$ converge to boundary value $1/\Lambda$ or Λ and there is only one changing point from Λ to $1/\Lambda$. This pattern was similarly founded in the other problems (4.9–4.11) as well. It suggests that one of the candidates of $\boldsymbol{z}$ as in (4.14) or (4.15) may be a practical solution to our optimization problem. Therefore, the objective function in minimization problem (4.8) is computed for m candidates of $\boldsymbol{z}$ as in (4.14), and then, one choose the minimum value of them as the solution of minimization problem. Also, the other problems (4.9–4.11) can be solved in similar way. We refer to this method as "the approximate optimization method". To further reduce the computing time, we used the `parSapply` function in `parallel` package (R Core Team, 2021) in R. Since there is no simple theoretical solution, we will resort the performance of our approximate optimization method to Monte Carlo simulations, as described in Section 4.8.1.

4.7 Confidence interval for partially identified region

In addition to the sensitivity range of the point estimates (4.3), Zhao, Small, and Bhattacharya (2019) proposed the percentile bootstrap confidence interval with at least $1 - \alpha$ coverage rate for partially identified region (4.4) which is

asymptotically valid and computationally tractable. For details of the theorems and lemma, see Section 4.3.2 in Zhao, Small, and Bhattacharya (2019). In this thesis, we briefly explain how to construct the $1-\alpha$ confidence interval for partially identified region (4.4) as follows:

(a) Obtain the number B bootstrap samples by taking samples of size N from the original data $(\widetilde{T}_i, \delta_i, A_i, \boldsymbol{L}_i)$, $i = 1, \ldots, N$, using random sampling with replacement.

(b) In each bootstrap sample b, for $b = 1, \ldots, B$, re-estimate the PS and then calculate the sensitivity range by solving the optimization problem (4.5). i.e., lower bound of range: $\min[\widehat{\mu}_{\tau,b}^{(h),1} - \widehat{\mu}_{\tau,b}^{(h),0}]$ and upper bound of range: $\max[\widehat{\mu}_{\tau,b}^{(h),1} - \widehat{\mu}_{\tau,b}^{(h),0}]$, respectively.

(c) Finally, use the $\alpha/2$ percentile among lower bounds of the bootstrap sensitivity ranges, $Q_{\alpha/2}\left\{\left(\min[\widehat{\mu}_{\tau,b}^{(h),1} - \widehat{\mu}_{\tau,b}^{(h),0}]\right)_{b\in[B]}\right\}$, as a lower limit of confidence interval and the $1-\alpha/2$ percentile among upper bounds of the bootstrap sensitivity ranges, $Q_{1-\alpha/2}\left\{\left(\max[\widehat{\mu}_{\tau,b}^{(h),1} - \widehat{\mu}_{\tau,b}^{(h),0}]\right)_{b\in[B]}\right\}$, as a upper limit of confidence interval for partially identified region.

Remark 4.3 Note that $\Lambda = 1.0$ means that there is no unmeasured confounding. In this case, the $1-\alpha$ percentile bootstrap confidence interval for partially identified region is the same as the $1-\alpha$ percentile bootstrap confidence interval of point estimate for the difference in adjusted RMST (3.7) obtained from the adjusted Kaplan-Meier estimator (3.5).

4.8 Simulation study 2

In this Section, we perform two simulation studies to evaluate the performance of 1) approximate optimization method and 2) sensitivity range and percentile bootstrap confidence interval.

4.8.1 Simulation study 2.1: Bias and computational time

First, assume that the sensitivity range of the point estimates for the difference in bias-adjusted RMST obtained from the direct optimization method is considered as the correctly calculated sensitivity range. Then, we illustrate that the approximate optimization method is valid and practical by comparing the bias and computational time between the two optimization methods. We simulated the data as described in Section 3.3.1 with the following modifications: (i) let L_2 be the unmeasured confounder, (ii) consider the four values of the sensitivity parameter $\Lambda \in \{1.1, 1.3, 1.5, 2.0\}$. Then, the total number of combinations of β_0, c, Λ, and τ is 80.

The parametric approximation of the PS, $P_\beta(A = 1 \mid L_1)$, was estimated from logistic regression model. When estimating $P_\beta(A = 1 \mid L_1)$ from superpopulation with sample size N = 1,000,000, $\widehat{P}_\beta(A = 1 \mid L_1)$ was range from 0.1321 to 0.8609 for $\beta_0 = -1.9$ (i.e., $P_\beta(A = 1 \mid L_1, L_2) \approx 0.5$) and from 0.3780 to 0.9640 for $\beta_0 = -0.425$ (i.e., $P_\beta(A = 1 \mid L_1, L_2) \approx 0.8$). Also, the approximately true values of $\exp\{h_{\beta_0}(\boldsymbol{l}, \boldsymbol{u})\}$ were $\exp\{\widehat{h}_{\beta_0}(\boldsymbol{l}, \boldsymbol{u})\} = \widehat{g}_{\beta_0}(\boldsymbol{l}) - \widehat{g}_0(\boldsymbol{l}, \boldsymbol{u}) \in (1/1.602, 1.606)$ for $\beta_0 = -1.9$ and $\exp\{\widehat{h}_{\beta_0}(\boldsymbol{l}, \boldsymbol{u})\} \in (1/1.637, 1.489)$ for $\beta_0 = -0.425$ where $\widehat{g}_{\beta_0}(\boldsymbol{l})$ and $\widehat{g}_0(\boldsymbol{l}, \boldsymbol{u})$ were estimated from superpopulation with sample size N = 1,000,000.

For each combination of coefficient for intercept term β_0 in the exposure generating model and censoring rate c, we simulated the data 200 times to obtain the mean and range (minimum and maximum) of (i) bias and (ii) computational time difference and ratio. When calculating the sensitivity range of the estimates for the difference in bias-adjusted RMST, the bias was

measured by calculating the difference between the lower (or upper) bound of the estimates of the difference in bias-adjusted RMST obtained by the approximate optimization method and that obtained by the direct optimization method. Also, the computational time difference and ratio between the direct optimization method and the approximate optimization method per each Monte Carlo data was measured.

First, consider the case when $\beta_0 = -1.9$. The simulation results are reported in Figures 4.1 and 4.2 and Tables B.4.1–B.4.4. Regardless of Λ and τ, when the censoring rate was less than 0.7, the bias was hardly different from 0, as seen in Figure 4.1 and Tables B.4.1 and B.4.2 (maximum mean bias for lower bound: 0.000000 [range: 0.000000 to 0.000004]; minimum mean bias for upper bound: 0.000000 [range: -0.000003 to 0.000000]). However, when the censoring rate was greater than or equal to 0.7, the bias was slightly different from 0 (maximum mean bias for lower bound: 0.000180 [range: 0.000000 to 0.017455]; minimum mean bias for upper bound: -0.000218 [range: -0.024271 to 0.000022]). Note that, in contrast to the increase in bias when the censoring rate was greater than or equal to 0.7, the computational time difference and ratio between two methods was not significantly different, as seen in Figure 4.2 and Tables B.4.3 and B.4.4 (maximum mean computational time difference: 4.07 [range: 2.18 to 6.44] seconds; maximum mean computational time ratio: 2.46 [range: 1.80 to 3.07] times). The results when $\beta_0 = -0.425$ were similar to those in $\beta_0 = -1.9$, as can be seen from Figures B.5.1 and B.5.2 and Tables B.5.1–B.5.4 in Appendix B.5.

These simulation results suggest that if the censoring rate is less than 0.7, the approximate optimization method is not inferior to the direct optimization method in terms of bias, but is superior to in terms of computational time.

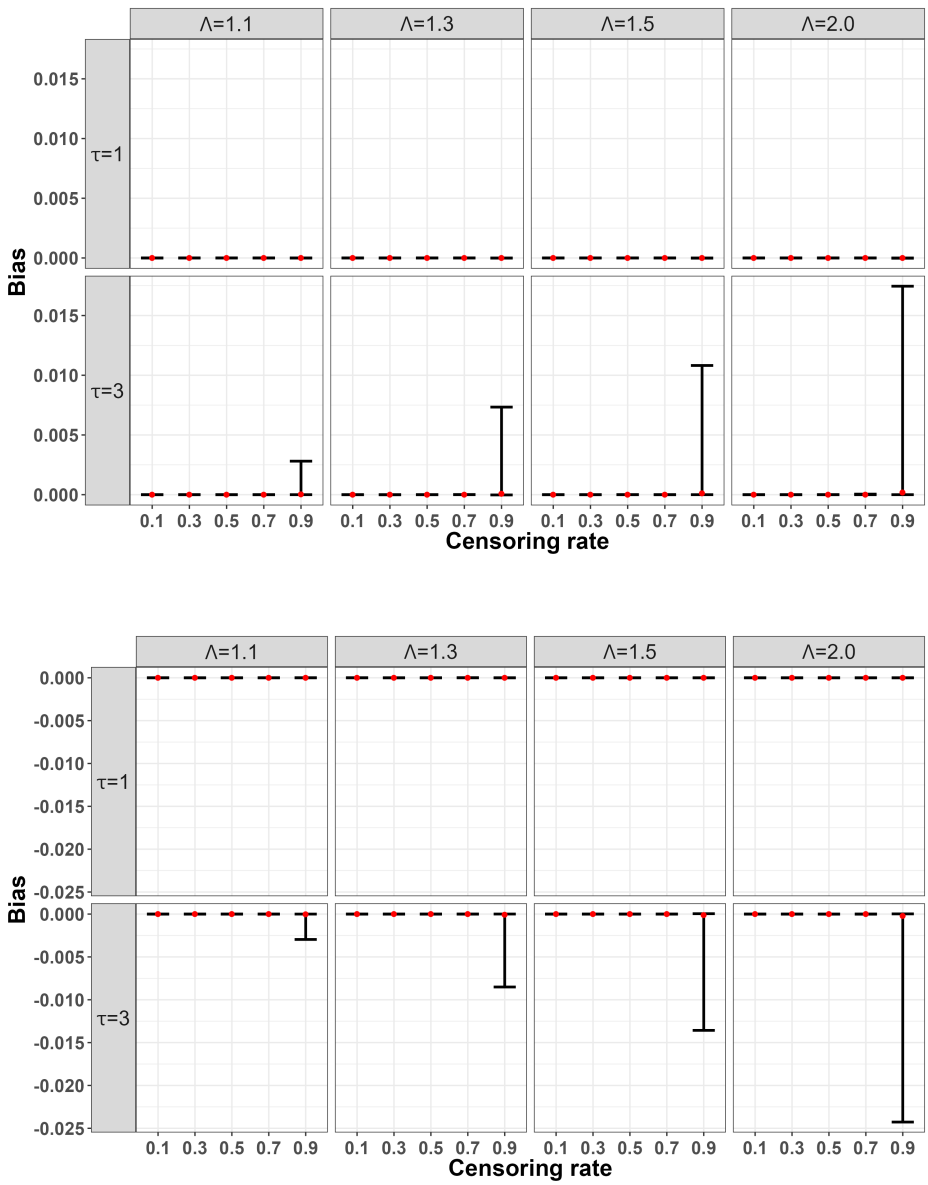


Figure 4.1 (Top): Bias for lower bound of sensitivity range. (Bottom): Bias for upper bound of sensitivity range. The mean is represented by red dot. The range of bias is represented by lower and upper horizontal bar.

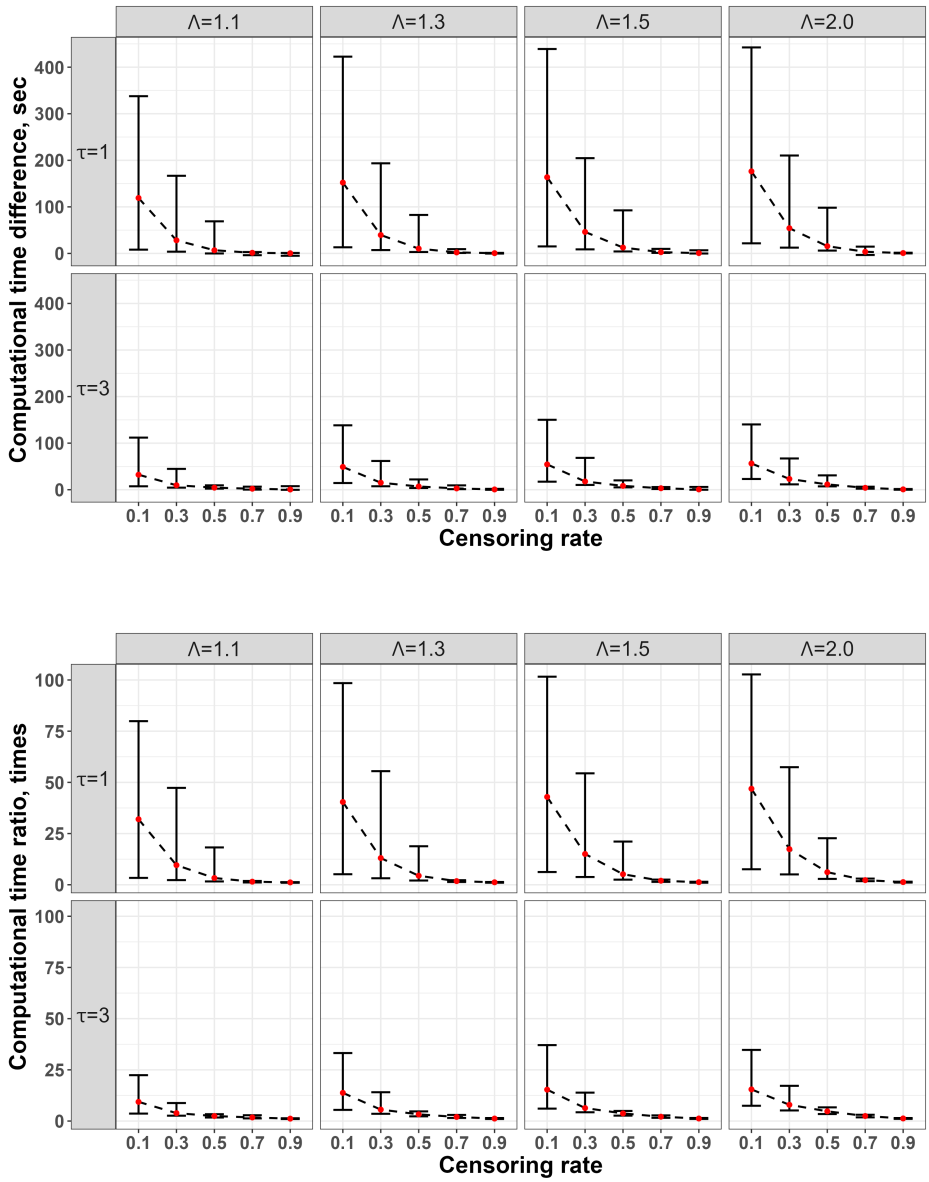


Figure 4.2 (Top): Difference in computational time between two methods per each Monte Carlo data. (Bottom): Ratio of computational time between two methods per each Monte Carlo data. The mean is represented by red dot. The range of computational time difference is represented by lower and upper horizontal bar.

Therefore, we can use the approximate optimization method in case that the censoring rate is less than 0.7. Otherwise, one can use the direct optimization method to perform our sensitivity analysis, although it takes slightly more computational time than the approximate optimization method.

4.8.2 Simulation study 2.2: Sensitivity range and coverage rate

To evaluate the coverage rate of the percentile bootstrap confidence interval for partially identified region (4.4) described in Section 4.7, we conducted a second simulation study. As a result of the simulation results in Section 4.8.1, we constructed the percentile bootstrap confidence interval by using the approximate optimization method.

We simulated the data as described in Section 3.3.1 with the following modifications: (i) let L_2 be the unmeasured confounder, (ii) consider $\beta_0 = -1.9$ only, (iii) administrative censoring time was set to t_c for all subjects where t_c was adjusted to obtain the simulated data that correspond to approximate censoring rate c, and (iv) consider the three values of censoring rate $c \in \{0.0, 0.3, 0.5\}$ and the five values of the sensitivity parameter $\Lambda \in \{1.0, 1.1, 1.3, 1.5, 2.0\}$. Then, the total number of combinations of β_0, c, Λ, τ is 30. We simulated 1,000 replications with sample size $N = 500$ for each censoring rate and constructed the 90% and 95% percentile bootstrap confidence interval for partially identified region with bootstrap resampling B = 1,000 times, respectively.

Considering that the PS was estimated using the measured confounder only (i.e., L_2 is unmeasured) and that the degree of unmeasured confounding would be expected to be on the order of Λ, the true partially identified regions were approximately calculated by using superpopulation with sample size

$N = 1{,}000{,}000$ and true event times T^a (i.e., survival times when all subjects experience the event of interest), for each combination of τ and β_0. In this case, because there is no censored subject, the approximately true partially identified region can be obtained by using analytic solution described in Section 4.5. Based on 1,000 bootstrap replicates, both the median values for lower and upper bounds of sensitivity range and the respective 90% or 95% percentile bootstrap confidence interval were calculated. Also, the coverage rate was calculated as the proportion of replications that the percentile bootstrap confidence interval covered the true partially identified region.

The simulation results are reported in Table 4.1. The median sensitivity range did not seem to differ from the approximately true partially identified region although there was a slight difference when the censoring rate was 0.5 and Λ increased gradually. The percentile bootstrap confidence intervals for partially identified region had desired coverage rate (90% or 95%) and did not appear to be conservative, although there was a slight difference from the desired coverage rate as the censoring rate increased.

4.9 Real data analysis 2

We applied the proposed sensitivity analysis methods described in this Section to 1) German breast cancer study group (GBSG) data, available in the `survival` package (Therneau, 2022), and 2) non–small cell lung cancer (NSCLC) data, recently published medical research in Song et al. (2021). In both data, the sensitivity range of the point estimates for the difference in bias-adjusted RMST at the pre-specified time point and the 95% confidence interval for theirs partially identified region that is constructed by the percentile bootstrap with 1,000 replications were calculated.

Table 4.1 Sensitivity range and confidence interval with coverage rate ($\beta_0 = -1.9$)

τ	Censoring rate	Λ	Partially identified region (Approximately true)	Sensitivity range (Median)	Confidence level $(1-\alpha)$: 0.90		Confidence level $(1-\alpha)$: 0.95	
					Confidence interval	Coverage	Confidence interval	Coverage
1	0.0	1.0	[-0.0250, -0.0250]	[-0.0248, -0.0248]	[-0.0436, -0.0065]	0.895	[-0.0472, -0.0034]	0.941
		1.1	[-0.0305, -0.0196]	[-0.0303, -0.0195]	[-0.0497, -0.0016]	0.886	[-0.0534, 0.0018]	0.937
		1.3	[-0.0406, -0.0103]	[-0.0399, -0.0106]	[-0.0611, 0.0076]	0.861	[-0.0653, 0.0111]	0.926
		1.5	[-0.0497, -0.0024]	[-0.0491, -0.0028]	[-0.0719, 0.0158]	0.847	[-0.0767, 0.0196]	0.908
		2.0	[-0.0696, 0.0133]	[-0.0684, 0.0127]	[-0.0957, 0.0342]	0.820	[-0.1016, 0.0387]	0.891
	0.3	1.0	[-0.0250, -0.0250]	[-0.0248, -0.0248]	[-0.0436, -0.0067]	0.896	[-0.0472, -0.0032]	0.948
		1.1	[-0.0305, -0.0196]	[-0.0303, -0.0195]	[-0.0496, -0.0017]	0.887	[-0.0535, 0.0018]	0.935
		1.3	[-0.0406, -0.0103]	[-0.0399, -0.0106]	[-0.0610, 0.0076]	0.865	[-0.0655, 0.0112]	0.919
		1.5	[-0.0497, -0.0024]	[-0.0491, -0.0028]	[-0.0715, 0.0161]	0.843	[-0.0763, 0.0198]	0.906
		2.0	[-0.0696, 0.0133]	[-0.0684, 0.0127]	[-0.0957, 0.0343]	0.825	[-0.1012, 0.0388]	0.889
	0.5	1.0	[-0.0250, -0.0250]	[-0.0248, -0.0248]	[-0.0434, -0.0067]	0.893	[-0.0470, -0.0032]	0.944
		1.1	[-0.0305, -0.0196]	[-0.0303, -0.0195]	[-0.0495, -0.0017]	0.888	[-0.0534, 0.0017]	0.937
		1.3	[-0.0406, -0.0103]	[-0.0399, -0.0106]	[-0.0611, 0.0076]	0.868	[-0.0650, 0.0111]	0.918
		1.5	[-0.0497, -0.0024]	[-0.0491, -0.0028]	[-0.0717, 0.0162]	0.839	[-0.0763, 0.0198]	0.904
		2.0	[-0.0696, 0.0133]	[-0.0684, 0.0127]	[-0.0956, 0.0341]	0.826	[-0.1010, 0.0386]	0.885
3	0.0	1.0	[-0.3846, -0.3846]	[-0.3866, -0.3866]	[-0.5033, -0.2726]	0.909	[-0.5235, -0.2498]	0.954
		1.1	[-0.4453, -0.3234]	[-0.4470, -0.3268]	[-0.5624, -0.2118]	0.907	[-0.5833, -0.1886]	0.954
		1.3	[-0.5499, -0.2149]	[-0.5506, -0.2195]	[-0.6639, -0.1015]	0.910	[-0.6845, -0.0793]	0.949
		1.5	[-0.6377, -0.1211]	[-0.6377, -0.1260]	[-0.7493, -0.0075]	0.908	[-0.7702, 0.0159]	0.947
		2.0	[-0.8084, 0.0691]	[-0.8079, 0.0635]	[-0.9182, 0.1863]	0.901	[-0.9398, 0.2086]	0.944
	0.3	1.0	[-0.3846, -0.3846]	[-0.3866, -0.3866]	[-0.5029, -0.2715]	0.910	[-0.5250, -0.2514]	0.954
		1.1	[-0.4453, -0.3234]	[-0.4470, -0.3268]	[-0.5622, -0.2110]	0.909	[-0.5843, -0.1891]	0.954
		1.3	[-0.5499, -0.2149]	[-0.5506, -0.2195]	[-0.6637, -0.1020]	0.909	[-0.6858, -0.0804]	0.950
		1.5	[-0.6377, -0.1211]	[-0.6377, -0.1260]	[-0.7498, -0.0069]	0.907	[-0.7714, 0.0150]	0.947
		2.0	[-0.8084, 0.0691]	[-0.8079, 0.0635]	[-0.9186, 0.1851]	0.902	[-0.9399, 0.2070]	0.946
	0.5	1.0	[-0.3846, -0.3846]	[-0.3865, -0.3865]	[-0.5024, -0.2717]	0.911	[-0.5232, -0.2493]	0.952
		1.1	[-0.4453, -0.3234]	[-0.4470, -0.3268]	[-0.5620, -0.2108]	0.910	[-0.5821, -0.1881]	0.950
		1.3	[-0.5499, -0.2149]	[-0.5506, -0.2195]	[-0.6638, -0.1007]	0.907	[-0.6843, -0.0793]	0.947
		1.5	[-0.6377, -0.1211]	[-0.6375, -0.1257]	[-0.7490, -0.0069]	0.908	[-0.7702, 0.0161]	0.946
		2.0	[-0.8084, 0.0691]	[-0.8079, 0.0636]	[-0.9186, 0.1858]	0.903	[-0.9403, 0.2107]	0.940

4.9.1 Real data analysis 2.1: GBSG data

GBSG data were included for patients with histologically verified primary breast cancer (positive regional lymph nodes but no distant metastases) to investigate the effects of chemotherapy and hormone therapy on the recurrence-free survival time (Schumacher et al., 1994). This data contained 686 patients with complete data for the prognostic factors. Survival outcome was recurrence-free survival time (days). Exposure was 2 years of hormonal therapy with tamoxifen. Event indicator was dichotomized into alive without recurrence and recurrence or death (censoring rate: 61.7% [235 of 440] for hormone-treated group and 53.4% [152 of 246] for hormone-untreated group). And, prognostic factors which were considered in Model Section of Royston and Altman (2013) were age at primary surgery (years), menopausal status (premenopausal vs. postmenopausal), tumor size (mm), the number of positive lymph nodes (n), and estrogen receptors (fmol/l). We preprocessed this data as delineated in Royston and Altman (2013). Detailed information about GBSG data is shown in Schumacher et al. (1994) and Royston and Altman (2013).

Using this data, we validated whether the mean recurrence-free survival time between hormonal therapy groups differs when patients are followed up to τ. Because Royston and Altman (2013) were considered 2 years and 5 years recurrence-free survival probabilities, we considered here two values of specific time point τ (2 years [365.25×2 days] and 5 years [365.25×5 days]). And, we performed a sensitivity analysis for unmeasured confounding via our proposed method. First, the PS was estimated via a logistic regression model conditioning on prognostic factors. In the PS model, age was transformed into both age^3 and age$^3 \times \log(\text{age})$, tumor size was categorized into three groups

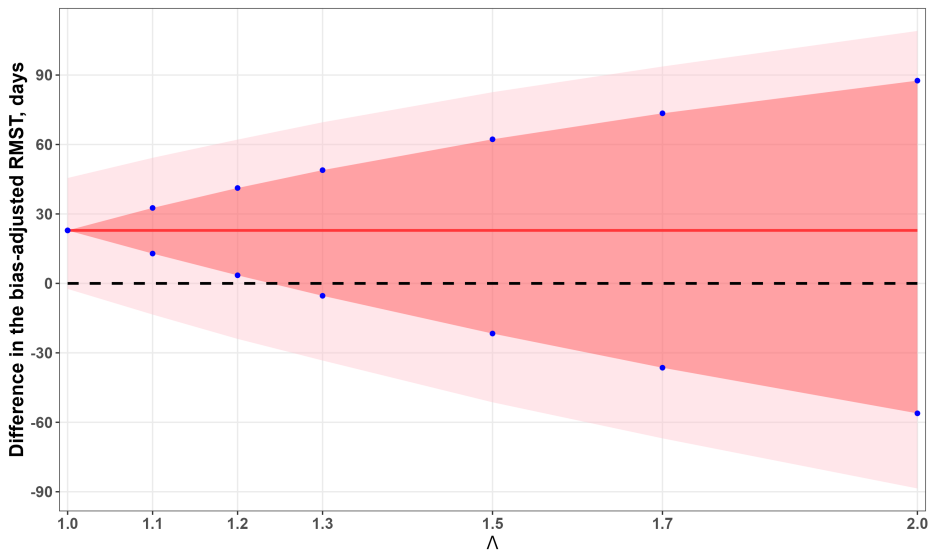
($\leq$ 20 mm, 20 to 50 mm, and $>$ 50 mm), and the number of positive lymph nodes was transformed into square root value. We considered the seven values of the sensitivity parameter $\Lambda \in \{1.0, 1.1, 1.2, 1.3, 1.5, 1.7, 2.0\}$.

In GBSG data, the bias of lower and upper bound of sensitivity range between the approximate and direct optimization methods was hardly different from 0 (maximum bias for lower bound of sensitivity range $<$ 0.000001 and minimum bias for upper bound of sensitivity range $>$ -0.000001, for $\tau = 2$ years; maximum bias for lower bound of sensitivity range $<$ 0.000001 and minimum bias for upper bound of sensitivity range -0.000102, for $\tau = 5$ years). Also, the bias of lower and upper bound of 95% confidence interval was less than 0.000001 and greater than -0.000001 for both values of τ. Therefore, we only report the results obtained by the approximate optimization method.

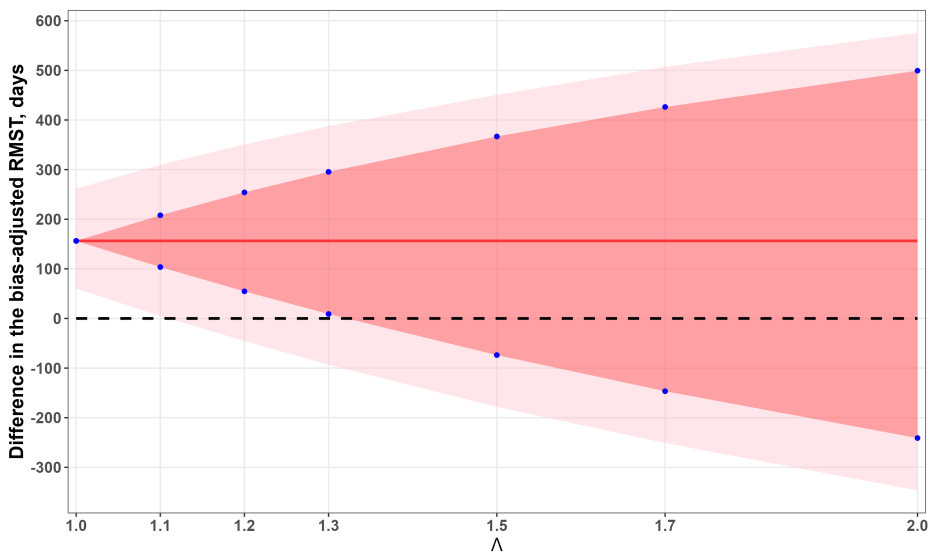
In Table 4.2 and Figure 4.3, the point estimate of the difference in adjusted RMST up to 5 years (i.e., $\Lambda = 1.0$) was statistically significant (156.32 [95%

Table 4.2 Sensitivity range and 95% confidence interval for difference in bias-adjusted RMST up to 2 and 5 years

τ	Λ	Sensitivity range (days)		95% confidence interval (days)	
		Lower bound	Upper bound	Lower bound	Upper bound
2y	1.0	22.89	22.89	-2.37	45.52
	1.1	12.88	32.59	-13.44	54.23
	1.2	3.50	41.20	-23.97	62.11
	1.3	-5.35	48.92	-33.30	69.58
	1.5	-21.66	62.24	-51.35	82.59
	1.7	-36.42	73.47	-66.92	93.68
	2.0	-56.13	87.59	-88.57	109.04
5y	1.0	156.32	156.32	60.22	261.67
	1.1	103.62	207.99	4.82	309.70
	1.2	54.67	254.07	-45.45	350.81
	1.3	8.96	295.44	-93.26	387.46
	1.5	-73.75	366.76	-178.17	450.61
	1.7	-146.62	426.23	-251.04	507.01
	2.0	-241.02	499.29	-346.60	575.41



$\tau = 2$ years



$\tau = 5$ years

Figure 4.3 Top: Difference in bias-adjusted RMST up to 2 years. Bottom: Difference in bias-adjusted RMST up to 5 years. Red solid line represents the difference in adjusted RMST. Black dashed line represents 0. Dark pink region represents the interval of point estimate. Light pink region represents the 95% confidence interval of point estimate. Blue points represent the lower and upper bound of sensitivity range using the sensitivity parameter $\Lambda \in \{1.0, 1.1, 1.2, 1.3, 1.5, 1.7, 2.0\}$.

CI: 60.22 to 261.67] days), although during 2 years of follow-up, it was not significant (22.89 [95% CI: -2.37 to 45.52] days). As a result of the sensitivity analysis, results drawn only from measured prognostic factors were quite sensitive to unmeasured confounding because the confidence interval contained 0 for sensitivity parameter Λ near 1.1. In the previous study (Schumacher et al., 1994), the effect of hormonal therapy on recurrence-free survival was not statistically significant (hazard ratio: 0.75 [95% CI: 0.54 to 1.04]; $P = .084$), which was evaluated by means of a multivariate Cox model. In our sensitivity analysis results, by using the difference in RMST (i.e., absolute risk) as the effect measure of exposure, we showed that even in the presence of weak unmeasured confounding, the 95% confidence interval contained 0. It suggests that great caution is required in interpreting the results and additional studies are needed to determine the effect of hormonal therapy with tamoxifen on recurrence-free survival.

4.9.2 Real data analysis 2.2: NSCLC data

NSCLC data were included patients with initial diagnosis of NSCLC to evaluate the prognostic performance of the proposed N descriptors for clinical staging (Song et al., 2021). This data contained 1,271 patients who divided by four clinical N stages (cN0, cN1, cN2, and cN3). For the illustration of our method, we only considered 248 patients in cN1 or cN2 group. Survival outcome was overall survival, which was measured from the date of staging chest CT to the date of any cause of death. The study entry times were different for each patients, but patients without death records were administrative censored on May 15, 2020 (censoring rate: 48.8% [42 of 86] for cN1 group and 25.3% [41 of 162] for cN2 group). Prognostic factors which were considered in Cox model of

Song et al. (2021) were age ($\leq$ 60 years vs. $>$ 60 years), sex (male vs. female), smoking history (never-, former-, and current-smoker), family history of lung cancer (yes vs. no), tumor type (solid vs. subsolid), tumor location (upper or middle vs. lower), histologic type (adenocarcinoma vs. others), and clinical T stage (cTis/cT1, cT2, cT3, and cT4). Detailed information about data is given in Song et al. (2021).

Using this data, we compared how the mean overall survival time between risk groups (cN1 vs. cN2) differs when patients are followed up to τ. Because Song et al. (2021) were considered a 5 years overall survival probabilities, we also set a specific time point τ to 5 years (365.25×5 days). And, we performed a sensitivity analysis for unmeasured confounding via our proposed method. First, to balance the probability of being in each risk group, the probability belonging to cN2 conditioning on prognostic factors (alike PS) was estimated via a logistic regression model. Among those prognostic factors considered in Song et al. (2021), tumor type was excluded from the model because patients with subsolid tumor were only 3 and 1 in cN1 group and cN2 group, respectively. Also, clinical T stage was excluded from the model because it was simultaneously measured with the clinical N stage. We considered the seven values of the sensitivity parameter $\Lambda \in \{1.0, 1.1, 1.2, 1.3, 1.5, 1.7, 2.0\}$.

In NSCLC data, the mean bias of lower and upper bound of sensitivity range between methods was hardly different from 0 (maximum bias for lower bound of sensitivity range: < 0.000001 and minimum bias for upper bound of sensitivity range: > -0.000001). Also, the mean bias of lower and upper bound of 95% confidence interval was less than 0.000005 and greater than -0.000001. Therefore, we only report the results obtained by the approximate optimization method.

Table 4.3 Sensitivity range and 95% confidence interval for difference in bias-adjusted RMST up to 5 years

Λ	Sensitivity range (days)		95% confidence interval (days)	
	Lower bound	Upper bound	Lower bound	Upper bound
1.0	-379.61	-379.61	-546.02	-208.97
1.1	-433.79	-324.37	-594.28	-150.17
1.2	-482.38	-273.13	-637.45	-96.95
1.3	-526.35	-225.45	-677.52	-45.74
1.5	-602.87	-139.11	-744.85	43.32
1.7	-667.10	-63.09	-802.77	118.08
2.0	-746.28	35.00	-874.91	219.37

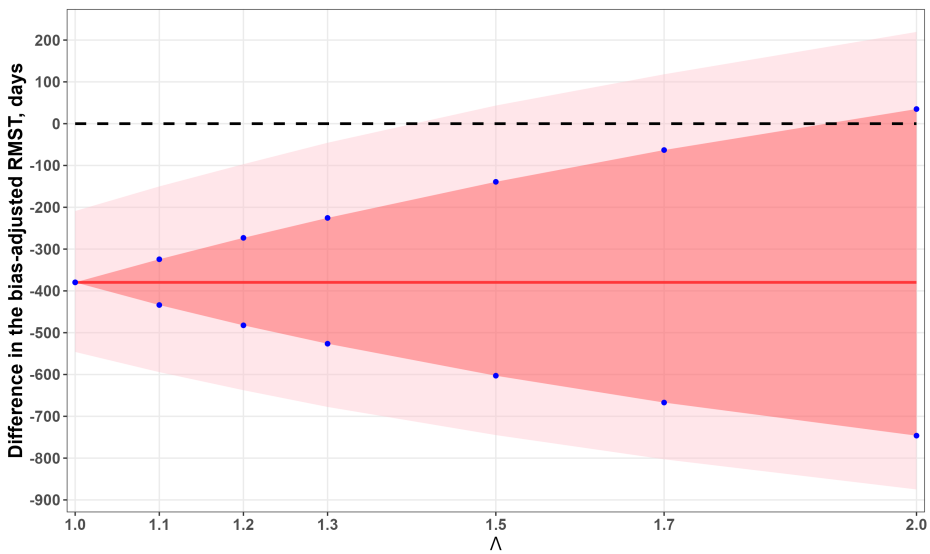


Figure 4.4 Difference in bias-adjusted RMST up to 5 years. Red solid line represents the difference in adjusted RMST. Black dashed line represents 0. Dark pink region represents the interval of point estimate. Light pink region represents the 95% confidence interval of point estimate. Blue points represent the lower and upper bound of sensitivity range using the sensitivity parameter $\Lambda \in \{1.0, 1.1, 1.2, 1.3, 1.5, 1.7, 2.0\}$.

In Table 4.3 and Figure 4.4, the point estimate of the difference in bias-adjusted RMST up to 5 years for $\Lambda = 0$ was statistically significant (-379.61 [95% CI: -546.02 to -208.97] days). To the best our knowledge, it is thought that

there is no unmeasured confounder that would change the probability of being in each risk group as strong as age. When the other prognostic factors (i.e., sex, smoking history, family history of lung cancer, tumor location, and histologic type) are considered as common confounders, the odds ratio of the probabilities of being in each risk group with and without the age variable changes at most 1.2629 (range: [1/1.1426, 1.2629]). Because the confidence interval did not contain 0 for sensitivity parameter Λ near 1.4 in our sensitivity analysis, the result drawn only from measured prognostic factors was considerably robust to unmeasured confounding. According to the American Joint Committee on Cancer (AJCC) (Edition et al., 2017), the 5 years overall survival probabilities for clinical N1 and N2 stage were 37% and 23%, respectively. Also, the hazard ratio between risk groups estimated by a Cox model adjusting for histology grade, sex, age, and geographical region were 1.42 ($P < 0.0001$). Along with AJCC, our sensitivity analysis further strengthens the evidence that there is the difference in mean life expectancy between two risk groups.

Chapter 5

Discussion

The difference in RMST has recently been used frequently as an alternative measure to hazard ratio for survival analysis in medical fields (Royston and Parmar, 2013; Uno et al., 2014; Trinquart et al., 2016; Kim, Uno, and Wei, 2017; Pak et al., 2017; Kloecker et al., 2020; Han and Jung, 2022). To reduce bias of confounding, many statistical methods that make it possible to estimate the adjusted RMST have been proposed (Andersen, Hansen, and Klein, 2004; Cole and Hernán, 2004; Xie and Liu, 2005; Tian, Zhao, and Wei, 2014; Chatton et al., 2022). Based on simulation studies, we confirmed that all the methods being considered provide the unbiased estimates with the percentile bootstrap confidence intervals achieving near nominal coverage probability when there is neither a large censoring rate nor extreme PS. Also, the bias was reduced and the coverage rate seemed to improve, as the sample size increased.

Despite increasing usage of the difference in RMST as an effect measure, to our knowledge, there were few available sensitivity analysis methods for unmeasured confounding when evaluating the estimate of the difference in adjusted RMST. In this thesis, we proposed a novel propensity score-based

sensitivity analysis method for the estimate of bias-adjusted RMST to assess the impact of probably possible unmeasured confounding in observational survival studies. The proposed method was a direct extension of existing sensitivity model (Zhao, Small, and Bhattacharya, 2019) which quantifies the degree of unmeasured confounding where subjects who might appear similar in terms of measured prognostic factors $\boldsymbol{L}$ may be different in their odds of receiving the exposure by at most Λ. Given a user-specified sensitivity parameter Λ, one can obtain a sensitivity range of the estimates for the difference in bias-adjusted RMST up to pre-specified time point τ along with a confidence interval with asymptotically at least $1 - \alpha$ coverage probability.

To obtain the sensitivity range of the estimates for the difference in bias-adjusted RMST, we should solve the optimization problem (4.5). However, as seen in Section 4.4, the optimization problem (4.5) could not be transformed into linear (fractional) programming problem in general survival analysis, and thus there is no analytic solution. Although it can be directly solved by using an optimization algorithm such as L-BFGS-B, the computational time was non-negligible. Therefore, we proposed an approximate optimization method and showed that by resorting to intensive Monte Carlo simulation studies, it can be an alternative method that is not inferior to the direct optimization method in terms of bias but superior in terms of computational time. In performing our sensitivity analysis, we recommend using the approximate optimization method in case that the censoring rate is less than 0.7. Otherwise, one may use the direct optimization method although it takes slightly more computational time than the approximate optimization method.

Compared with other sensitivity analysis methods, the proposed method has the advantage that one can perform the sensitivity analysis regardless

of model used for estimating the PS. Also, no assumptions for the relation between and the distribution of measured and unmeasured confounders are required. Furthermore, our method needs only one sensitivity parameter Λ. However, the challenge lies in specifying a sensible sensitivity parameter Λ which quantifies the degree of unmeasured confounding. Therefore, considerable domain knowledge about the assignment of exposure may be needed to limit a controversy about potential unmeasured confounding. In other words, we need to know how much unmeasured confounding would change the PS. Taking our second real data analysis (NSCLC) in Section 4.9.2 as an example, when considering the overall survival time in non–small cell lung cancer and the nature of age, one might argue that unmeasured confounding will not be as strongly associated with a probability of being in each risk group (cN1 vs. cN2) as is the age variable.

Our sensitivity analysis method was based on the IP weighted Kaplan-Meier curve via Xie and Liu (2005). Alternatively, when the unmeasured confounding is suspected and the true effect of exposure is represented by a regression model that includes the exposure as well as both the measured and unmeasured confounders (such as pseudo-observation method described in Section 3.2.1.1 or ANCOVA-type model described in Section 3.2.1.2), Lin, Psaty, and Kronmal (1998)'s sensitivity analysis method can be used. However, it has many limitations and is not easy to apply in practice. We simply described their sensitivity analysis method for unmeasured confounding and its limitations in Appendix C. Also, one might want to perform the sensitivity analysis on the results of Cox model adjusted for potential confounders using IP weighting proposed by Cole and Hernán (2004) described in Section 3.2.2.1 or using G-computation proposed by Chatton et al. (2022) described in Section 3.2.3.

However, the parametric sensitivity model (Definition 4.2) may not be directly applicable to methods of Cole and Hernán (2004) and Chatton et al. (2022) since the equation for the estimate of exposure effect expressed as the shifted PS (such as the estimate of the difference in bias-adjusted RMST) cannot be readily derived from Cox model. Further studies are needed to develop the sensitivity analysis method for unmeasured confounding based on Cole and Hernán (2004)'s adjusted survival curve or Chatton et al. (2022)'s G-computation method.

Our study has several limitations. First, our proposed method depends on the assumptions such as the positivity and consistency which are in fact not unique to our method but necessary in any propensity score-based methods. Similarly, the proposed sensitivity analysis methods can be unstable when the PSs are highly variable. In this case, one can stabilize the PSs via truncation (Potter, 1993; Cole and Hernán, 2008). Second, we only considered a single binary exposure. It can be extended to multi-valued categorical, ordinal, or continuous exposures although the underlying sensitivity model may become complex. Finally, the performance of our sensitivity analysis methods inevitably resorted to simulation studies, so that more evidence will be needed to use widely in real applications. However, based on our intensive simulation studies, the median sensitivity range did not seem to differ from the approximately true partially identified region, and the percentile bootstrap confidence intervals for partially identified region had desired coverage rate and did not appear to be conservative, although there was a slight degradation in performance as the censoring rate increased.

In summary, we never know whether the assumption of conditional exchangeability is satisfied in observational study. Consequently, it is of interest

to conduct a sensitivity analysis to quantify how much the analysis results are varied by unmeasured confounding. In accordance with importance of sensitivity analysis, we proposed a propensity score-based sensitivity analysis method for unmeasured confounding of the difference in adjusted RMST in observational survival studies.

Bibliography

[1] Abramowitz, M. and Stegun, I. A. (1964). *Handbook of mathematical functions with formulas, graphs, and mathematical tables* (Volume 55). US Government printing office.

[2] Andersen, P. K., Hansen, M. G., and Klein, J. P. (2004). Regression analysis of restricted mean survival time based on pseudo-observations. *Lifetime data analysis*, **10**(4), 335–350.

[3] Andersen, P. K., Klein, J. P., and Rosthøj, S. (2003). Generalised linear models for correlated pseudo-observations, with applications to multi-state models. *Biometrika*, **90**(1), 15–27.

[4] Andersen, P. K. and Perme, M. P. (2010). Pseudo-observations in survival analysis. *Statistical methods in medical research*, **19**(1), 71–99.

[5] Bender, R., Augustin, T., and Blettner, M. (2005). Generating survival times to simulate Cox proportional hazards models. *Statistics in medicine*, **24**(11), 1713–1723.

[6] Binder, D. A. (1992). Fitting Cox's proportional hazards models from survey data. *Biometrika*, **79**(1), 139–147.

[7] Breslow, N. E. (1974). Covariance analysis of censored survival data. *Biometrics*, **30**(1), 89–99.

[8] Brumback, B. A., Hernán, M. A., Haneuse, S. J., and Robins, J. M. (2004). Sensitivity analyses for unmeasured confounding assuming a marginal structural model for repeated measures. *Statistics in medicine*, **23**(5), 749–767.

[9] Byrd, R. H., Lu, P., Nocedal, J., and Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on scientific computing*, **16**(5), 1190–1208.

[10] Charnes, A. and Cooper, W. W. (1962). Programming with linear fractional functionals. *Naval Research logistics quarterly*, **9**(3-4), 181–186.

[11] Chatton, A., Borgne, F. L., Leyrat, C., and Foucher, Y. (2022). G-computation and doubly robust standardisation for continuous-time data: A comparison with inverse probability weighting. *Statistical Methods in Medical Research*, **31**(4), 706–718.

[12] Chen, P. Y. and Tsiatis, A. A. (2001). Causal inference on the difference of the restricted mean lifetime between two groups. *Biometrics*, **57**(4), 1030–1038.

[13] Cole, S. R. and Hernán, M. A. (2004). Adjusted survival curves with inverse probability weights. *Computer methods and programs in biomedicine*, **75**(1), 45–49.

[14] Cole, S. R. and Hernán, M. A. (2008). Constructing inverse probability weights for marginal structural models. *American journal of epidemiology*, **168**(6), 656–664.

[15] Conner, S. C., Sullivan, L. M., Benjamin, E. J., LaValley, M. P., Galea, S., and Trinquart, L. (2019). Adjusted restricted mean survival times in observational studies. *Statistics in medicine*, **38**(20), 3832–3860.

[16] Cornfield, J., Haenszel, W., Hammond, E. C., Lilienfeld, A. M., Shimkin, M. B., and Wynder, E. L. (1959). Smoking and lung cancer: Recent evidence and a discussion of some questions. *Journal of the National Cancer institute*, **22**(1), 173–203.

[17] Cox, D. R. (1958). Planning of experiments.

[18] Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, **34**(2), 187–202.

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

[20] Cox, D. R. and Oakes, D. (2018). *Analysis of survival data*. Chapman; Hall/CRC.

[21] D'Agostino, R. B., Lee, M. L., Belanger, A. J., Cupples, L. A., Anderson, K., and Kannel, W. B. (1990). Relation of pooled logistic regression to time dependent Cox regression analysis: The Framingham Heart Study. *Statistics in medicine*, **9**(12), 1501–1515.

[22] De Finetti, B. (1974). Theory of probability: A critical introductory treatment.

[23] Dorn, J. and Guo, K. (2022). Sharp sensitivity analysis for inverse propensity weighting via quantile balancing. *Journal of the American Statistical Association*, **0**(0), 1–13.

[24] Edition, S., Edge, S., Byrd, D., and American Joint Committee on Cancer and others. (2017). *AJCC cancer staging manual 8th ed*. Springer New York, NY.

[25] Foucher, Y., Le Borgne, F., Dantan, E., Gillaizeau, F., Chatton, A., Combescure, C., and Sabathe, C. (2022). *RISCA: Causal inference and prediction in cohort-based analyses* [R package version 1.0.3]. https://CRAN.R-project.org/package=RISCA

[26] Gerber, F. and Furrer, R. (2019). optimParallel: An R package providing a parallel version of the L-BFGS-B optimization method. *R Journal*, **11**(1), 352–358.

[27] Greenland, S. and Robins, J. M. (1986). Identifiability, exchangeability, and epidemiological confounding. *International journal of epidemiology*, **15**(3), 413–419.

[28] Greenwood, M. (1926). The "errors of sampling" of the survivorship tables. *Reports on public health and medical subjects*.

[29] Han, K. and Jung, I. (2022). Restricted mean survival time for survival analysis: A quick guide for clinical researchers. *Korean Journal of Radiology*, **23**(5), 495–499.

[30] Hernán, M. A. (2010). The hazards of hazard ratios. *Epidemiology (Cambridge, Mass.)*, **21**(1), 13.

[31] Hernán, M. A., Brumback, B., and Robins, J. M. (2000). Marginal structural models to estimate the causal effect of zidovudine on the survival of hiv-positive men. *Epidemiology*, **11**(5), 561–570.

[32] Hernán, M. A. and Robins, J. M. (2022). *Causal inference: What if*. Boca Raton: Chapman & Hall/CRC.

[33] Højsgaard, S., Halekoh, U., and Yan, J. (2006). The R package geepack for generalized estimating equations. *Journal of statistical software*, **15**(2), 1–11.

[34] Holland, P. W. (1986). Statistics and causal inference. *Journal of the American statistical Association*, **81**(396), 945–960.

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

[36] Huang, R., Xu, R., and Dulai, P. S. (2020). Sensitivity analysis of treatment effect to unmeasured confounding in observational studies with survival and competing risks outcomes. *Statistics in Medicine*, **39**(24), 3397–3411.

[37] Irwin, J. (1949). The standard error of an estimate of expectation of life, with special reference to expectation of tumourless life in experiments with mice. *Epidemiology & Infection*, **47**(2), 188–189.

[38] Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, **53**(282), 457–481.

[39] Kim, D. H., Uno, H., and Wei, L. J. (2017). Restricted mean survival time as a measure to interpret clinical trial results. *JAMA cardiology*, **2**(11), 1179–1180.

[40] Klein, J. P., Gerster, M., Andersen, P. K., Tarima, S., and Perme, M. P. (2008). SAS and R functions to compute pseudo-values for censored data regression. *Computer methods and programs in biomedicine*, **89**(3), 289–300.

[41] Klein, J. P. and Moeschberger, M. L. (2003). *Survival analysis: Techniques for censored and truncated data* (Volume 1230). Springer.

[42] Kloecker, D. E., Davies, M. J., Khunti, K., and Zaccardi, F. (2020). Uses and limitations of the restricted mean survival time: Illustrative examples from cardiovascular outcomes and mortality trials in type 2 diabetes. *Annals of internal medicine*, **172**(8), 541–552.

[43] Klungsøyr, O., Sexton, J., Sandanger, I., and Nygård, J. F. (2009). Sensitivity analysis for unmeasured confounding in a marginal structural Cox proportional hazards model. *Lifetime data analysis*, **15**(2), 278–294.

[44] Lee, S., Lee, K. H., Park, J. H., Kim, H. Y., Choi, Y., and Lee, K. H. (2023). Staging chest CT in patients with early colon cancer: Analysis of its prognostic impact using a causal diagram. Manuscript submitted to American Journal of Roentgenology.

[45] Liang, K. Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, **73**(1), 13–22.

[46] Lin, D. Y., Psaty, B. M., and Kronmal, R. A. (1998). Assessing the sensitivity of regression results to unmeasured confounders in observational studies. *Biometrics*, **54**(3), 948–963.

[47] Lin, D. Y. and Wei, L. J. (1989). The robust inference for the Cox proportional hazards model. *Journal of the American statistical Association*, **84**(408), 1074–1078.

[48] Lin, N. X., Logan, S., and Henley, W. E. (2013). Bias and sensitivity analysis when estimating treatment effects from the Cox model with omitted covariates. *Biometrics*, **69**(4), 850–860.

[49] Lumley, T. (2010). *Complex surveys: A guide to analysis using R: A guide to analysis using R*. John Wiley; Sons.

[50] Pak, K., Uno, H., Kim, D. H., Tian, L., Kane, R. C., Takeuchi, M., Fu, H., Claggett, B., and Wei, L. J. (2017). Interpretability of cancer clinical trial results using restricted mean survival time as an alternative to the hazard ratio. *JAMA oncology*, **3**(12), 1692–1696.

[51] Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, **82**(4), 669–688.

[52] Perme, M. P., Gerster, M., and Rodrigues, K. (2017). *Pseudo: Computes pseudo-observations for modeling* [R package version 1.4.3]. https://CRAN.R-project.org/package=pseudo

[53] Potter, F. J. (1993). The effect of weight trimming on nonlinear survey estimates. *Proceedings of the American Statistical Association, Section on Survey Research Methods*, **758763**.

[54] R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

[55] Robins, J. (1986). A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Mathematical modelling*, **7**(9-12), 1393–1512.

[56] Robins, J. M. (1989). The analysis of randomized and non-randomized AIDS treatment trials using a new approach to causal inference in longitudinal studies. *Health service research methodology: a focus on AIDS*, 113–159.

[57] Robins, J. M. (1998). Marginal structural models. in: 1997 proceedings of the section on bayesian statistical science. *American Statistical Association*, 1–10.

[58] Robins, J. M. (1999). Association, causation, and marginal structural models. *Synthese*, **1**(121), 151–179.

[59] Robins, J. M., Rotnitzky, A., and Scharfstein, D. O. (2000). Sensitivity analysis for selection bias and unmeasured confounding in missing data and causal inference models. *Statistical models in epidemiology, the environment, and clinical trials* (pp. 1–94). Springer.

[60] Rosenbaum, P. R. (2002). Overt bias in observational studies. *Observational studies* (pp. 71–104). Springer.

[61] Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, **70**(1), 41–55.

[62] Royston, P. and Altman, D. G. (2013). External validation of a Cox prognostic model: Principles and methods. *BMC medical research methodology*, **13**(1), 1–15.

[63] Royston, P. and Parmar, M. K. (2011). The use of restricted mean survival time to estimate the treatment effect in randomized clinical trials when the proportional hazards assumption is in doubt. *Statistics in medicine*, **30**(19), 2409–2421.

[64] Royston, P. and Parmar, M. K. (2013). Restricted mean survival time: An alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. *BMC medical research methodology*, **13**(1), 1–15.

[65] Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, **66**(5), 688.

[66] Rubin, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *The Annals of statistics*, **6**(1), 34–58.

[67] Rubin, D. B. (1980). Randomization analysis of experimental data: The Fisher randomization test comment. *Journal of the American statistical association*, **75**(371), 591–593.

[68] Scharfstein, D. O., Rotnitzky, A., and Robins, J. M. (1999). Adjusting for nonignorable drop-out using semiparametric nonresponse models. *Journal of the American Statistical Association*, **94**(448), 1096–1120.

[69] Schumacher, M., Bastert, G., Bojar, H., Hübner, K., Olschewski, M., Sauerbrei, W., Schmoor, C., Beyerle, C., Neumann, R., and Rauschecker, H. (1994). Randomized 2 x 2 trial evaluating hormonal treatment and the duration of chemotherapy in node-positive breast cancer patients. German Breast Cancer Study Group. *Journal of Clinical Oncology*, **12**(10), 2086–2093.

[70] Shu, D., Young, J. G., Toh, S., and Wang, R. (2021). Variance estimation in inverse probability weighted Cox models. *Biometrics*, **77**(3), 1101–1117.

[71] Song, H., Yoon, S. H., Kim, J., Kim, J., Lee, K. W., Lee, W., Lee, S., Kim, K., Lee, C.-T., Chung, J.-H. et al. (2021). Application of N descriptors proposed by the international association for the study of lung cancer in clinical staging. *Radiology*, **300**(2), 450–457.

[72] Stensrud, M. J., Aalen, J. M., Aalen, O. O., and Valberg, M. (2019). Limitations of hazard ratios in clinical trials. *European heart journal*, **40**(17), 1378–1383.

[73] Stensrud, M. J. and Hernán, M. A. (2020). Why test for proportional hazards? *Jama*, **323**(14), 1401–1402.

[74] Tan, Z. (2006). A distributional approach for causal inference using propensity scores. *Journal of the American Statistical Association*, **101**(476), 1619–1637.

[75] Therneau, T. M. (2022). *A package for survival analysis in R* [R package version 3.4-0]. https://CRAN.R-project.org/package=survival

[76] Thompson, W. (1977). On the treatment of grouped observations in life studies. *Biometrics*, **33**(3), 463–470.

[77] Tian, L., Zhao, L., and Wei, L. J. (2014). Predicting the restricted mean event time with the subject's baseline covariates in survival analysis. *Biostatistics*, **15**(2), 222–233.

[78] Trinquart, L., Jacot, J., Conner, S. C., and Porcher, R. (2016). Comparison of treatment effects measured by the hazard ratio and by the ratio of restricted mean survival times in oncology randomized controlled trials. *Journal of Clinical Oncology*, **34**(15), 1813–1819.

[79] Uno, H. (2015). Vignette for survRM2 package: Comparing two survival curves using the restricted mean survival time. https://cran.r-project.org/web/packages/survRM2/vignettes/survRM2-vignette3-2.html

[80] Uno, H., Claggett, B., Tian, L., Inoue, E., Gallo, P., Miyata, T., Schrag, D., Takeuchi, M., Uyama, Y., Zhao, L. et al. (2014). Moving beyond

the hazard ratio in quantifying the between-group difference in survival analysis. *Journal of clinical Oncology*, **32**(22), 2380–2385.

[81] Uno, H., Tian, L., Horiguchi, M., Cronin, A., Battioui, C., and Bell, J. (2022). *survRM2: Comparing restricted mean survival time* [R package version 1.0-4]. https://CRAN.R-project.org/package=survRM2

[82] Westreich, D. and Cole, S. R. (2010). Invited commentary: Positivity in practice. *American journal of epidemiology*, **171**(6), 674–677.

[83] White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica: Journal of the econometric society*, **50**(1), 1–25.

[84] Xie, J. and Liu, C. (2005). Adjusted Kaplan–Meier estimator and log-rank test with inverse probability of treatment weighting for survival data. *Statistics in medicine*, **24**(20), 3089–3110.

[85] Zeger, S. L. and Liang, K. Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*, **42**(1), 121–130.

[86] Zhao, Q., Small, D. S., and Bhattacharya, B. B. (2019). Sensitivity analysis for inverse probability weighting estimators via the percentile bootstrap. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **81**(4), 735–761.

Appendix A

Appendix for Chapter 3

A.1 Example R codes

Here, we consider simulation setting described in Section 3.3.1 as an example with the following modifications: (i) sample size N is 10,000 (ii) consider censoring rate to be 0.5 (i.e., the value of scale parameter $\lambda = 0.354$) (iii) consider the true PS to be approximately 0.5 only (i.e., $\beta_0 = -1.9$).

```
## Function for making simulation dataset
simulate_data <- function(dataset, baseline="Weibull", N=10000,
                          params=list(lambda=0.95, nu=1.8),
                          params.cen=list(lambda=0.354, nu=0.6),
                          cen.rate=0.5, b0=-1.9) {
  ## Coef of the time-to-event model (Exponential or Weibull)
  bAY <- log(2.5)
  bL1Y <- log(0.7)
  bL2Y <- log(1.2)

  ## Coef of exposure model (logistic)
  b0 <- b0      # -1.9 ==> P(A=1|L)=0.5 // -0.425 ==>P(A=1|L)=0.8
  bL1A <- 0.8
  bL2A <- 1.5

  ## Generate confounders (L1~L2)
  for (i in 1:1) {
    assign(paste("L", i, sep=""), rbinom(N, 1, prob=0.5))
  }
  for (i in 2:2) {
    assign(paste("L", i, sep=""), rnorm(N, 1, 0.5^2))
```

```
}

## Generate exposure A
probA <- plogis(b0 + bL1A * L1 + bL2A * L2)
A <- rbinom(N, 1, prob=probA)

## Coefficients
beta.all <- c(bAY, bL1Y, bL2Y)

## Draw from a U(0,1) random variable
u <- runif(N)

## Simulate survival times depending on the baseline hazard (exposed)
Covariate.ex <- cbind(rep(1, N), L1, L2)
if (baseline == "Exponential") {
  t.true.ex <- 5*(-log(u))/(params$lambda * exp(Covariate.ex %*% beta.all))
} else if (baseline == "Weibull") {
  t.true.ex <- 5*(-log(u)/(params$lambda *
                             exp(Covariate.ex %*% beta.all)))^(1/params$nu)
}

## Draw from a U(0,1) random variable
ustar <- runif(N)
## Simulate survival times depending on the baseline hazard (unexposed)
Covariate.unex <- cbind(rep(0, N), L1, L2)
if (baseline == "Exponential") {
  t.true.unex <- 5*(-log(ustar))/(params$lambda *
                   exp(Covariate.unex %*% beta.all))
} else if (baseline == "Weibull") {
  t.true.unex <- 5*(-log(ustar)/(params$lambda *
                   exp(Covariate.unex %*% beta.all)))^(1/params$nu)
}
# plot(density(t.true))
t.true <- ifelse(A == 1, t.true.ex, t.true.unex)

## Simulate cenosring times
t.cen <- rweibull(N, shape=params.cen$nu,
                  scale=1/(params.cen$lambda^(1/params.cen$nu)))

## Make observed survival time
t.obs <- ifelse(t.true <= t.cen, t.true, t.cen)
status <- ifelse(t.true <= t.cen, 1, 0)

## True survival probability
st.ex <- exp(-params$lambda * (t.true.ex/5)^params$nu *
               exp(Covariate.ex %*% beta.all))
st.unex <- exp(-params$lambda * (t.true.unex/5)^params$nu *
                 exp(Covariate.unex %*% beta.all))

## Return data frame
dat.temp <- data.frame(dataset=dataset, baseline=baseline,
                       id=c(1:N), N=N, cen.rate=cen.rate,
                       A=A, L1=L1, L2=L2,
```

```
                          t.true.ex=t.true.ex, t.true.unex=t.true.unex,
                          t.true=t.true, t.cen=t.cen, t.obs=t.obs,
                          status=status, st.ex=st.ex, st.unex=st.unex,
                          stringsAsFactors=FALSE, row.names=NULL)
  return(dat.temp)
}

## Make simulation dataset
set.seed(221205)
sim.dat <- simulate_data(dataset=1, baseline="Weibull", N=10000,
                          params=list(lambda=0.95, nu=1.8),
                          params.cen=list(lambda=0.354, nu=0.6),
                          cen.rate=0.5, b0=-1.9)
```

A.1.1 R code for Kaplan-Meier estimator

```
### Estimate of RMST difference in RCT
## Method 1
library(survival)
library(RISCA)
kaplan.fit <- survfit(Surv(t.obs,status) ~ A, data=sim.dat)
dat.surv <- data.frame(strata=summary(kaplan.fit)$strata,
                       time=summary(kaplan.fit)$time,
                       surv=summary(kaplan.fit)$surv)
rmst.0.0 <- rmst(times=dat.surv[dat.surv$strata == "A=0",]$time,
                 surv.rates=dat.surv[dat.surv$strata == "A=0",]$surv,
                 max.time=5, type='s')
rmst.1.0 <- rmst(times=dat.surv[dat.surv$strata == "A=1",]$time,
                 surv.rates=dat.surv[dat.surv$strata == "A=1",]$surv,
                 max.time=5, type='s')
# Estimate of RMST difference
print(rmst.1.0 - rmst.0.0, 5)
```

```
## [1] -0.96697
```

```
### Method 2
library(survRM2)
# Using rmst2 function without "covariates" argument
rmst2.without.cov <- rmst2(time=sim.dat$t.obs,
                           status=sim.dat$status, arm=sim.dat$A, tau=5)
print(rmst2.without.cov)
```

```
## The truncation time: tau = 5  was specified.
##
## Restricted Mean Survival Time (RMST) by arm
##                Est.    se lower .95 upper .95
## RMST (arm=1) 2.693 0.025     2.645     2.741
## RMST (arm=0) 3.660 0.026     3.609     3.711
##
## Restricted Mean Time Lost (RMTL) by arm
```

```
##                Est.    se lower .95 upper .95
## RMTL (arm=1) 2.307 0.025     2.259     2.355
## RMTL (arm=0) 1.340 0.026     1.289     1.391
##
## Between-group contrast
##                          Est. lower .95 upper .95 p
## RMST (arm=1)-(arm=0) -0.967    -1.037    -0.897 0
## RMST (arm=1)/(arm=0)  0.736     0.719     0.753 0
## RMTL (arm=1)/(arm=0)  1.722     1.648     1.798 0
```

```
# Estimate of RMST difference
print(rmst2.without.cov$unadjusted.result[1,1], 5)
```

```
## [1] -0.96697
```

A.1.2 R code for pseudo-observation

```
## Andersen et al.'s RMST regression-based method (Pseudo-observation)
library(pseudo)
library(geepack)
rmst.pseudo <- pseudomean(time=sim.dat$t.obs, event=sim.dat$status, tmax=5)
fit_pseudo <- geeglm(rmst.pseudo ~ A + L1 + L2, data=sim.dat, id=id,
                     family="gaussian", corstr="independence", scale.fix=F)
summary(fit_pseudo)
```

```
## Call:
## geeglm(formula = rmst.pseudo ~ A + L1 + L2, family = "gaussian",
##     data = sim.dat, id = id, corstr = "independence", scale.fix = F)
##
##  Coefficients:
##             Estimate  Std.err     Wald Pr(>|W|)
## (Intercept)  3.62491  0.07770 2176.350   <2e-16 ***
## A           -1.02840  0.03765  746.123   <2e-16 ***
## L1           0.36292  0.03704   96.025   <2e-16 ***
## L2          -0.11228  0.07471    2.259    0.133
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation structure = independence
## Estimated Scale Parameters:
##
##             Estimate Std.err
## (Intercept)    3.334 0.03988
## Number of clusters:   10000  Maximum cluster size: 1
```

```
# Estimate of RMST difference with Wald confidence interval
print(c(coef(fit_pseudo)[2], confint.default(fit_pseudo)[2,]), 5)
```

```
##        A     2.5 %    97.5 %
## -1.02840  -1.10219  -0.95461
```

A.1.3 R code for ANCOVA-type model

```
## Tian et al.'s RMST regression-based method (ANCOVA-type adjusted analyses)
library(survRM2)
# Using rmst2 function with "coavariates" argument
rmst2.with.cov <- rmst2(time=sim.dat$t.obs, status=sim.dat$status,
                    arm=sim.dat$A, tau=5, covariates=sim.dat[,c("L1","L2")])
print(rmst2.with.cov)
```

```
## The truncation time: tau = 5  was specified.
##
## Summary of between-group contrast (adjusted for the covariates)
##                         Est. lower .95 upper .95 p
## RMST (arm=1)-(arm=0) -1.026    -1.098    -0.954 0
## RMST (arm=1)/(arm=0)  0.722     0.706     0.739 0
## RMTL (arm=1)/(arm=0)  1.778     1.701     1.859 0
##
## Model summary (difference of RMST)
##             coef se(coef)       z     p lower .95 upper .95
## intercept  3.648    0.081  44.836 0.000     3.489     3.808
## arm       -1.026    0.037 -27.966 0.000    -1.098    -0.954
## L1         0.363    0.039   9.336 0.000     0.286     0.439
## L2        -0.137    0.079  -1.739 0.082    -0.292     0.017
##
## Model summary (ratio of RMST)
##             coef se(coef)       z    p exp(coef) lower .95 upper .95
## intercept  1.292    0.025  51.054 0.00     3.641     3.465     3.827
## arm       -0.325    0.012 -27.543 0.00     0.722     0.706     0.739
## L1         0.114    0.012   9.355 0.00     1.121     1.094     1.148
## L2        -0.043    0.025  -1.750 0.08     0.957     0.912     1.005
##
## Model summary (ratio of time-lost)
##             coef se(coef)      z     p exp(coef) lower .95 upper .95
## intercept  0.295    0.046  6.345 0.000     1.343     1.226     1.471
## arm        0.576    0.023 25.306 0.000     1.778     1.701     1.859
## L1        -0.199    0.022 -9.225 0.000     0.820     0.786     0.855
## L2         0.075    0.043  1.717 0.086     1.077     0.990     1.173
```

```
# Estimate of RMST difference
print(rmst2.with.cov$adjusted.result[1,1], 5)
```

```
## [1] -1.02644
```

A.1.4 R code for IP weighted Cox model

```
library(survival)
library(RISCA)
## Calculate the stabilized IP weights
denom.fit <- glm(A ~ L1 + L2, family=binomial, data=sim.dat)
pd.A <- predict(denom.fit, type="response")
numer.fit <- glm(A~1, family=binomial(), data=sim.dat)
pn.A <- predict(numer.fit, type="response")
sim.dat$sw <- ifelse(sim.dat$A == 0, ((1-pn.A)/(1-pd.A)), (pn.A/pd.A))

#### Cole and Hernan's adjusted survival curve with IP Weights
### Method 1
## Fitting a weighted Cox model with robust standard errors
cox.fit <- coxph(Surv(t.obs,status) ~ strata(A) + cluster(id),
                 weights=sw, data=sim.dat)
adsurv.fit <- survfit(cox.fit)
dat.surv <- data.frame(strata=summary(adsurv.fit)$strata,
                       time=summary(adsurv.fit)$time,
                       surv=summary(adsurv.fit)$surv)
rmst.0.0 <- rmst(times=dat.surv[dat.surv$strata == "A=0",]$time,
                 surv.rates=dat.surv[dat.surv$strata == "A=0",]$surv,
                 max.time=5, type='s')
rmst.1.0 <- rmst(times=dat.surv[dat.surv$strata == "A=1",]$time,
                 surv.rates=dat.surv[dat.surv$strata == "A=1",]$surv,
                 max.time=5, type='s')
# Estimate of RMST difference
print(rmst.1.0 - rmst.0.0, 5)
```

```
## [1] -1.02274
```

```
### Method 2
## Using survey package
library(survey)
library(RISCA)
sdes <- svydesign(id=~ 0, weights=~sw, data=sim.dat)
dfit <- svykm(Surv(t.obs, status) ~ A, design=sdes, se=TRUE)
rmst.0.2 <- rmst(times=dfit$`0`$time, surv.rates=dfit$`0`$surv,
                 max.time=5, type='s')
rmst.1.2 <- rmst(times=dfit$`1`$time, surv.rates=dfit$`1`$surv,
                 max.time=5, type='s')
# Estimate of RMST difference
print(rmst.1.2 - rmst.0.2, 5)
```

```
## [1] -1.02274
```

A.1.5 R code for adjusted Kaplan-Meier estimator

```
### Xie and Liu's adjusted Kaplan-Meier estimator
library(RISCA)
coxipw <- ipw.survival(times=sim.dat$t.obs, failures=sim.dat$status,
                       variable=sim.dat$A, weights=sim.dat$sw)
rmst.0.3 <- rmst(times=coxipw$table.surv$times[coxipw$table.surv$variable == 0],
        surv.rates=coxipw$table.surv$survival[coxipw$table.surv$variable == 0],
        max.time=5, type='s')
rmst.1.3 <- rmst(times=coxipw$table.surv$times[coxipw$table.surv$variable == 1],
        surv.rates=coxipw$table.surv$survival[coxipw$table.surv$variable == 1],
        max.time=5, type='s')
# Estimate of RMST difference
print(rmst.1.3 - rmst.0.3, 5)
```

```
## [1] -1.02297
```

```
## Use Conner et al's "akm_rsmt" R function from GitHub
## to estimate the variance for RMST difference
akm_rsmt <- devtools::source_url(
  "https://github.com/s-conner/akm-rmst/blob/master/AKM_rmst.R?raw=TRUE"
)$value
akm.rsmt.with.sw <- akm_rsmt(time=sim.dat$t.obs, status=sim.dat$status,
                             group=factor(sim.dat$A), weight=sim.dat$sw, tau=5,
                             alpha=.05, xaxismin=0, xaxismax=max(sim.dat$t.obs))
```

```
## RMST calculated up to tau = 5
##
## Restricted Mean Survival Time (RMST) per Group
##
##            RMST    SE
## Group 0   3.679 0.027
## Group 1   2.656 0.025
##
## Restricted Mean Survival Time (RMST) Differences
##
##                     Est.    SE    CIL    CIU p
## Groups 1 vs. 0   -1.023 0.037 -1.095 -0.951 0
##
## Restricted Mean Survival Time (RMST) Ratios
##
##                  Log Est.    SE  Est.   CIL   CIU p
## Groups 1 vs. 0     -0.326 0.012 0.722 0.705 0.739 0
```

A.1.6 R code for G-computation

```
### Chatton et al's G-computation with a Cox model
library(RISCA)
## Fit Cox model with confounders
cox.cdt <- coxph(Surv(t.obs,status) ~ A + L1 + L2, data=sim.dat, x=TRUE)
## Set the iterations to 10 because of computational time
gc.ate <- gc.survival(object=cox.cdt, data=sim.dat, group="A", times="t.obs",
                      failures="status", max.time=5, iterations=10, effect="ATE",
                      n.cluster=1)
# Estimate of RMST difference with bootstrap confidence interval
print(gc.ate$delta[1,3,4], 5)
```

```
##      estimate   ci.lower   ci.upper
## 1   -1.03037   -1.12190   -0.93884
```

A.2 Proof of true value for difference in RMST

In the equation for the true value of conditional RMST (3.10), the first and second equalities follow from the definition of RMST (2.2). The third equality holds because the true event times were generated from Cox model with Weibull-distributed baseline hazard (3.9). Now, we show in detail how the last equality holds.

Proof. To begin with, let $k = \exp\{\log(2.5)a + \log(1.2)l_1 + \log(0.7)l_2\}$. Recall from (3.10) that the true value of conditional RMST is

$$\int_0^\tau \left[\exp\left(-\frac{\lambda t^\nu}{5^\nu}\right)\right]^{\exp\{\log(2.5)a+\log(1.2)l_1+\log(0.7)l_2\}} \mathrm{d}t$$
$$= \int_0^\tau \left[\exp\left(-\frac{\lambda t^\nu}{5^\nu}\right)\right]^k \mathrm{d}t = \int_0^\tau \left[\exp\left(-\frac{k\lambda t^\nu}{5^\nu}\right)\right] \mathrm{d}t.$$

When substituting t for $\frac{5}{k^{\frac{1}{\nu}}\lambda^{\frac{1}{\nu}}}u$, $u = \frac{k^{\frac{1}{\nu}}\lambda^{\frac{1}{\nu}}t}{5}$ and $\mathrm{d}t = \frac{5}{k^{\frac{1}{\nu}}\lambda^{\frac{1}{\nu}}}\,\mathrm{d}u$. Then,

$$\int_0^{\tau}\left[\exp\left(-\frac{k\lambda t^{\nu}}{5^{\nu}}\right)\right]\mathrm{d}t = \frac{5}{k^{\frac{1}{\nu}}\lambda^{\frac{1}{\nu}}}\int_0^{\frac{k^{\frac{1}{\nu}}\lambda^{\frac{1}{\nu}}\tau}{5}}\exp(-u^{\nu})\,\mathrm{d}u. \tag{A.2.1}$$

Note that $\int \exp(-u^{\nu})\,\mathrm{d}u$ is a special integral of incomplete gamma function (Abramowitz and Stegun, 1964). That is,

$$\int \exp(-u^{\nu})\,\mathrm{d}u = -\frac{\Gamma(\frac{1}{\nu}, u^{\nu})}{\nu}. \tag{A.2.2}$$

where $\Gamma(s,x) = \int_x^{\infty} t^{s-1}\exp(-t)\,\mathrm{d}t$ is the upper incomplete gamma function. Using the fact that (A.2.2), (A.2.1) is written by

$$\begin{aligned}
&\frac{5}{k^{\frac{1}{\nu}}\lambda^{\frac{1}{\nu}}}\int_0^{\frac{k^{\frac{1}{\nu}}\lambda^{\frac{1}{\nu}}\tau}{5}}\exp(-u^{\nu})\,\mathrm{d}u \\
&= \frac{5}{k^{\frac{1}{\nu}}\lambda^{\frac{1}{\nu}}}\left[-\frac{\Gamma(\frac{1}{\nu}, u^{\nu})}{\nu}\right]_0^{\frac{k^{\frac{1}{\nu}}\lambda^{\frac{1}{\nu}}\tau}{5}} \\
&= \frac{5}{\nu k^{\frac{1}{\nu}}\lambda^{\frac{1}{\nu}}}\left[-\Gamma\left(\frac{1}{\nu}, \frac{k\lambda\tau^{\nu}}{5^{\nu}}\right) + \Gamma\left(\frac{1}{\nu}, 0\right)\right].
\end{aligned} \tag{A.2.3}$$

Because the ordinary gamma function is defined as

$$\Gamma(s) = \int_0^{\infty} t^{s-1}\exp(-t)\,\mathrm{d}t$$

and the lower incomplete gamma function is defined as

$$\gamma(s,x) = \int_0^{x} t^{s-1}exp(-t)\,\mathrm{d}t,$$

we have

$$\Gamma(s) = \Gamma(s,0) \quad \text{and} \quad \gamma(s,x) + \Gamma(s,x) = \Gamma(s). \tag{A.2.4}$$

Then, using properties in (A.2.4), (A.2.3) is reduced by

$$\frac{5}{\nu k^{\frac{1}{\nu}}\lambda^{\frac{1}{\nu}}}\left[-\Gamma\left(\frac{1}{\nu},\frac{k\lambda\tau^{\nu}}{5^{\nu}}\right)+\Gamma\left(\frac{1}{\nu},0\right)\right]=\frac{5}{\nu k^{\frac{1}{\nu}}\lambda^{\frac{1}{\nu}}}\left[\gamma\left(\frac{1}{\nu},\frac{k\lambda\tau^{\nu}}{5^{\nu}}\right)\right].$$

Therefore, the true value of conditional RMST is

$$\mu_{\tau}^{a}(l)=\frac{5\times\gamma\left(\frac{1}{\nu},\frac{\exp\{\log(2.5)a+\log(1.2)l_1+\log(0.7)l_2\}\lambda\tau^{\nu}}{5^{\nu}}\right)}{\nu\times\left(\exp\{\log(2.5)a+\log(1.2)l_1+\log(0.7)l_2\}\right)^{1/\nu}\times\lambda^{1/\nu}}. \qquad \square$$

A.3 Simulation study 1 results for sample size 1,000

Table A.3.1 Simulation study 1 results ($N = 1{,}000$): Pseudo-observation

τ	True RMST	$E(A\|\boldsymbol{L})$	Censoring rate	RMST	Bias	95% CI	Coverage
1	-0.027	0.5	0.1	-0.027	0.000	[-0.043, -0.011]	0.945
			0.3	-0.027	0.000	[-0.044, -0.011]	0.949
			0.5	-0.027	0.000	[-0.045, -0.009]	0.938
			0.7	-0.027	0.000	[-0.046, -0.009]	0.937
			0.9	-0.027	0.001	[-0.050, -0.004]	0.947
		0.8	0.1	-0.028	-0.001	[-0.045, -0.010]	0.935
			0.3	-0.028	-0.001	[-0.046, -0.009]	0.924
			0.5	-0.028	-0.001	[-0.047, -0.008]	0.945
			0.7	-0.028	-0.001	[-0.049, -0.007]	0.933
			0.9	-0.027	0.000	[-0.052, -0.001]	0.943
3	-0.417	0.5	0.1	-0.416	0.001	[-0.517, -0.314]	0.945
			0.3	-0.418	-0.001	[-0.527, -0.310]	0.950
			0.5	-0.417	-0.001	[-0.534, -0.297]	0.939
			0.7	-0.414	0.002	[-0.551, -0.274]	0.934
			0.9	-0.411	0.006	[-0.617, -0.197]	0.950
		0.8	0.1	-0.426	-0.010	[-0.542, -0.305]	0.930
			0.3	-0.427	-0.011	[-0.553, -0.295]	0.953
			0.5	-0.425	-0.009	[-0.564, -0.281]	0.955
			0.7	-0.426	-0.009	[-0.587, -0.259]	0.941
			0.9	-0.419	-0.002	[-0.664, -0.158]	0.948

Table A.3.2 Simulation study 1 results ($N = 1{,}000$): ANCOVA-type model

τ	True RMST	$E(A\|\boldsymbol{L})$	Censoring rate	RMST	Bias	95% CI	Coverage
1	-0.027	0.5	0.1	-0.027	0.000	[-0.043, -0.011]	0.942
			0.3	-0.027	0.000	[-0.044, -0.011]	0.950
			0.5	-0.027	0.000	[-0.045, -0.009]	0.936
			0.7	-0.027	0.000	[-0.046, -0.008]	0.938
			0.9	-0.027	0.000	[-0.051, -0.003]	0.952
		0.8	0.1	-0.028	-0.001	[-0.046, -0.010]	0.936
			0.3	-0.028	-0.001	[-0.046, -0.009]	0.926
			0.5	-0.028	-0.001	[-0.047, -0.008]	0.949
			0.7	-0.028	-0.001	[-0.049, -0.007]	0.928
			0.9	-0.027	0.000	[-0.052, -0.001]	0.942
3	-0.417	0.5	0.1	-0.416	0.001	[-0.517, -0.315]	0.945
			0.3	-0.418	-0.001	[-0.527, -0.311]	0.950
			0.5	-0.418	-0.001	[-0.535, -0.298]	0.938
			0.7	-0.414	0.002	[-0.555, -0.274]	0.944
			0.9	-0.413	0.003	[-0.641, -0.189]	0.958
		0.8	0.1	-0.426	-0.010	[-0.542, -0.306]	0.929
			0.3	-0.427	-0.011	[-0.553, -0.296]	0.945
			0.5	-0.426	-0.009	[-0.564, -0.282]	0.942
			0.7	-0.428	-0.011	[-0.590, -0.262]	0.940
			0.9	-0.424	-0.008	[-0.693, -0.145]	0.941

Table A.3.3 Simulation study 1 results ($N = 1{,}000$): IP weighted Cox model

τ	True RMST	$E(A\|\boldsymbol{L})$	Censoring rate	RMST	Bias	95% CI	Coverage
1	-0.027	0.5	0.1	-0.027	0.000	[-0.043, -0.011]	0.947
			0.3	-0.027	0.000	[-0.044, -0.011]	0.950
			0.5	-0.027	0.000	[-0.045, -0.009]	0.940
			0.7	-0.027	0.000	[-0.046, -0.009]	0.937
			0.9	-0.026	0.001	[-0.050, -0.004]	0.942
		0.8	0.1	-0.027	0.000	[-0.044, -0.009]	0.938
			0.3	-0.027	0.000	[-0.045, -0.009]	0.927
			0.5	-0.027	0.000	[-0.045, -0.007]	0.935
			0.7	-0.027	0.000	[-0.047, -0.007]	0.919
			0.9	-0.026	0.001	[-0.050, -0.001]	0.928
3	-0.417	0.5	0.1	-0.416	0.001	[-0.517, -0.314]	0.945
			0.3	-0.417	0.000	[-0.526, -0.309]	0.954
			0.5	-0.417	0.000	[-0.534, -0.296]	0.937
			0.7	-0.413	0.003	[-0.549, -0.273]	0.936
			0.9	-0.408	0.009	[-0.611, -0.193]	0.948
		0.8	0.1	-0.418	-0.001	[-0.539, -0.298]	0.940
			0.3	-0.419	-0.003	[-0.544, -0.283]	0.946
			0.5	-0.417	0.000	[-0.556, -0.270]	0.953
			0.7	-0.418	-0.002	[-0.578, -0.250]	0.944
			0.9	-0.406	0.011	[-0.645, -0.054]	0.950

Table A.3.4 Simulation study 1 results ($N = 1{,}000$): Adjusted Kaplan-Meier

τ	True RMST	$E(A\|\boldsymbol{L})$	Censoring rate	RMST	Bias	95% CI	Coverage
1	-0.027	0.5	0.1	-0.027	0.000	[-0.043, -0.011]	0.946
			0.3	-0.027	0.000	[-0.044, -0.011]	0.950
			0.5	-0.027	0.000	[-0.045, -0.009]	0.941
			0.7	-0.027	0.000	[-0.046, -0.009]	0.936
			0.9	-0.026	0.001	[-0.050, -0.004]	0.942
		0.8	0.1	-0.027	0.000	[-0.044, -0.009]	0.938
			0.3	-0.027	0.000	[-0.045, -0.009]	0.927
			0.5	-0.027	0.000	[-0.045, -0.007]	0.934
			0.7	-0.027	0.000	[-0.047, -0.007]	0.919
			0.9	-0.026	0.001	[-0.050, -0.001]	0.928
3	-0.417	0.5	0.1	-0.416	0.000	[-0.518, -0.314]	0.945
			0.3	-0.418	-0.001	[-0.527, -0.310]	0.953
			0.5	-0.417	-0.001	[-0.535, -0.297]	0.936
			0.7	-0.414	0.002	[-0.550, -0.274]	0.937
			0.9	-0.411	0.006	[-0.615, -0.196]	0.949
		0.8	0.1	-0.417	-0.001	[-0.538, -0.297]	0.938
			0.3	-0.418	-0.002	[-0.544, -0.282]	0.946
			0.5	-0.416	0.001	[-0.556, -0.269]	0.954
			0.7	-0.417	0.000	[-0.578, -0.247]	0.944
			0.9	-0.408	0.009	[-0.652, -0.148]	0.940

Table A.3.5 Simulation study 1 results ($N = 1{,}000$): G-computation

τ	True RMST	$E(A\|\boldsymbol{L})$	Censoring rate	RMST	Bias	95% CI	Coverage
1	-0.027	0.5	0.1	-0.027	0.000	[-0.035, -0.020]	0.939
			0.3	-0.027	0.000	[-0.036, -0.020]	0.944
			0.5	-0.027	0.000	[-0.036, -0.019]	0.940
			0.7	-0.027	0.000	[-0.037, -0.017]	0.933
			0.9	-0.027	0.000	[-0.043, -0.013]	0.953
		0.8	0.1	-0.027	0.000	[-0.034, -0.020]	0.941
			0.3	-0.027	0.000	[-0.035, -0.020]	0.952
			0.5	-0.027	0.000	[-0.036, -0.019]	0.942
			0.7	-0.027	0.000	[-0.037, -0.017]	0.942
			0.9	-0.026	0.001	[-0.042, -0.012]	0.950
3	-0.417	0.5	0.1	-0.416	0.000	[-0.485, -0.352]	0.946
			0.3	-0.418	-0.002	[-0.497, -0.347]	0.949
			0.5	-0.418	-0.001	[-0.507, -0.334]	0.928
			0.7	-0.414	0.002	[-0.528, -0.309]	0.933
			0.9	-0.412	0.004	[-0.604, -0.225]	0.942
		0.8	0.1	-0.417	0.000	[-0.488, -0.348]	0.944
			0.3	-0.418	-0.001	[-0.497, -0.339]	0.960
			0.5	-0.415	0.001	[-0.508, -0.324]	0.948
			0.7	-0.415	0.002	[-0.532, -0.295]	0.944
			0.9	-0.408	0.009	[-0.616, -0.193]	0.946

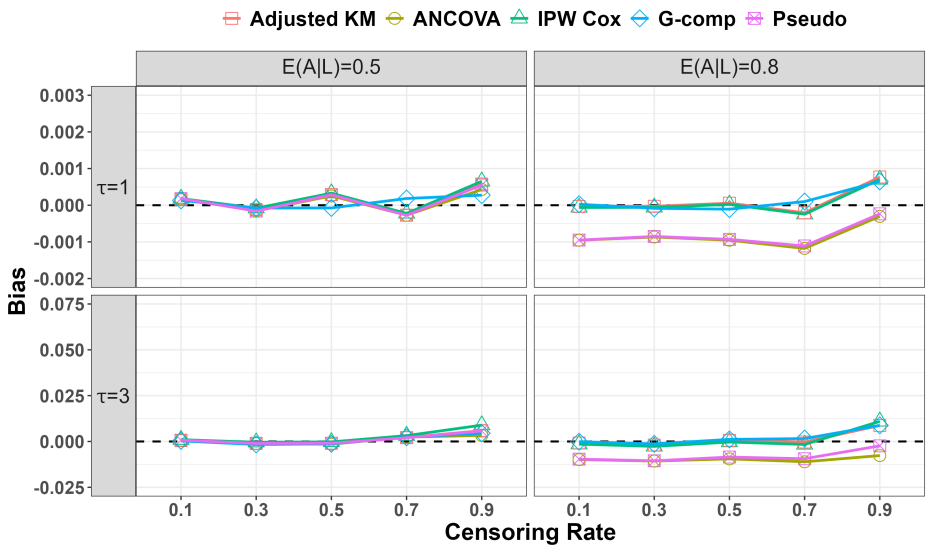


Figure A.3.1 Bias for simulation study 1 (N = 1,000). Adjusted KM = adjusted Kaplan-Meier estimator; ANCOVA = ANCOVA-type model; IPW Cox = IP weighted Cox model; G-comp = G-computation; Pseudo = pseudo-observation. Dashed line in the plot represents a bias of 0.

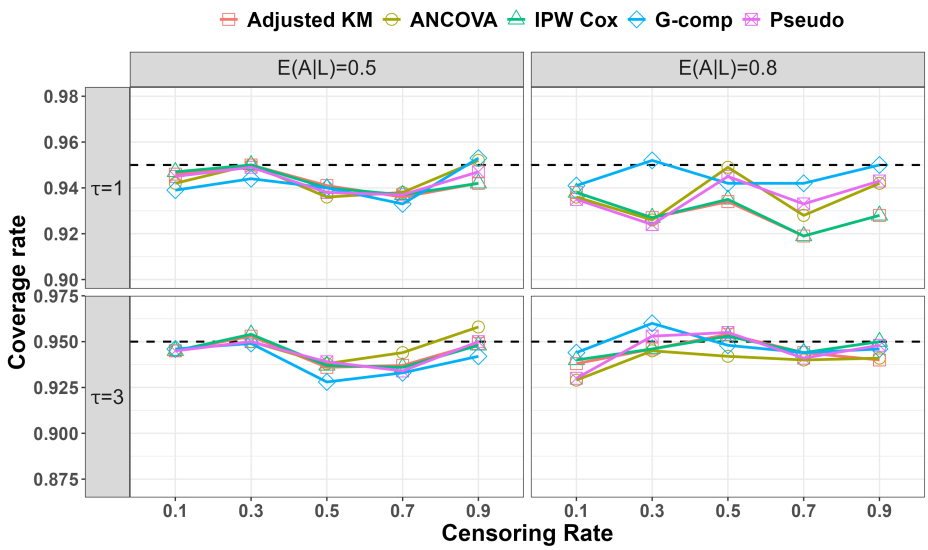


Figure A.3.2 Coverage rate for simulation study 1 (N = 1,000). Adjusted KM = adjusted Kaplan-Meier estimator; ANCOVA = ANCOVA-type model; IPW Cox = IP weighted Cox model; G-comp = G-computation; Pseudo = pseudo-observation. Dashed line in the plot represents a coverage rate of 0.95.

A.4 Pooled logistic regression model

Besides the Cox model, there is a simple method to parametrically estimate the (discrete-time) hazards which use a logistic regression model from data that transform individual data (with one row per each subject) into person-time format (with one row per person-time) (D'Agostino et al., 1990). We refer to this regression model as a pooled logistic regression model. To adjust for confounding, Hernán, Brumback, and Robins (2000) described a pooled logistic regression using the IP of treatment and censoring weights when there exist time-dependent confounders and selection bias due to loss to follow-up. Under assumptions (**A1**)–(**A4**) and without time-dependent confounders, we exploit and adapt the procedure described by Hernán (2010) to obtain the adjusted survival curve from the IP weighted pooled logistic regression, and thus we can estimate the difference in adjusted RMST as follows.

First, restructure the individual data, which has the (stabilized) IP weights calculated in advance as a separate variable, into the person-time format data. For example, the first row contains information about first subject at time 0, the second row contains information about first subject at time 1, and so on. This process continues until the follow up of first subject is end. In this way, each subject has multiple row per person-time in the person-time format data. Additionally, if the first subject experienced the event of interest, the event indicator variable of the last row is set to 1 and the remaining rows are set to 0. Otherwise (i.e., if the first subject is censored), the event indicator variables of all rows are set to 0. The other subjects also expand data in the same way.

Second, using the person-time data format, fit a pooled logistic regression model weighted by the (stabilized) IP weights (3.3) or (3.4)

$$\text{logit } P(D_{k+1} = 1 \mid D_k = 0, A) = \theta_{0,k} + \theta_1 A \tag{A.4.1}$$

where $\text{logit}\{p\} = \log\{p/(1-p)\}$, D_{k+1} is the event indicator between time k to time $k+1$, and $\theta_{0,k}$ is the time-varying intercept. Assume that all subjects had to survive at time 0 (i.e. $D_0 = 0$). We can estimate $\theta_{0,k}$ based on some flexible function of time (e.g., polynomial splines or cubic splines). Note that when $P(D_{k+1} = 1 \mid D_k = 0, A)$ is close to zero and $P(D_{k+1} = 0 \mid D_k = 0, A)$ is thus close to one, the IP weighted pooled logistic regression model (A.4.1) approximates the IP weighted Cox model, because $P(D_{k+1} = 1 \mid D_k = 0, A)$ is approximately equal to the hazard $P(T = k+1 \mid T > k, A)$ and the log odds of hazard is

$$\begin{aligned}
\text{logit } P(D_{k+1} = 1 \mid D_k = 0, A) &= \log\left(\frac{P(D_{k+1} = 1 \mid D_k = 0, A)}{1 - P(D_{k+1} = 1 \mid D_k = 0, A)}\right) \\
&\approx \log P(D_{k+1} = 1 \mid D_k = 0, A) \\
&\approx \log P(T = k+1 \mid T > k, A) \\
&= \log\{h_0(k)\exp(\beta_1 A)\} \\
&= \log\{h_0(k)\} + \beta_1 A = \beta_{0,k} + \beta_1 A. \qquad \text{(A.4.2)}
\end{aligned}$$

where $\text{logit}\{p\} = \log\{p/(1-p)\}$. Therefore, if $P(D_{k+1} = 1 \mid D_k = 0, A)$ is close to zero, the log odds ratio θ_1 in (A.4.1) approximately equals to the log hazard ratio β_1 in (A.4.2) (Thompson, 1977). The condition that $P(D_{k+1} = 1 \mid D_k = 0, A) \approx 0$ can almost always be ensured to hold because if one set the time interval short enough (e.g., change time interval from days to hours or minutes), $P(D_{k+1} = 1 \mid D_k = 0, A)$ will be close to zero. In other words, we need to set

the number of event of interest to be rare in each time interval (as a rule of thumb, $< 10\%$) (Hernán and Robins, 2022). Technically, under assumptions (**A1**)–(**A4**), the IP weighted pooled logistic regression model (A.4.1) estimates the parameters of the marginal structural logistic model

$$\text{logit } P(D^a_{k+1} = 1 \mid D^a_k = 0) = \theta_{0,k} + \theta_1 a.$$

That is, the IP weighted pooled logistic regression model estimates the ratio of the hazards (i.e., $\exp\{\theta_1\}$) that would have been observed if all subjects had been exposed ($a = 1$) and if all subjects had been unexposed ($a = 0$).

Finally, the survival probability at time $k+1$ under exposure a, $P(D^a_{k+1} = 0)$, can obtained by multiplying one minus the hazard $P(D^a_m = 0 \mid D^a_{m-1} = 0)$ over all previous times $m = 1, \ldots, k+1$, as follows:

$$\begin{aligned} P(D^a_{k+1} = 0) &= \prod_{m=1}^{k+1} \left[P(D^a_m = 0 \mid D^a_{m-1} = 0)\right] \\ &= \prod_{m=1}^{k+1} \left[1 - P(D^a_m = 1 \mid D^a_{m-1} = 0)\right]. \end{aligned}$$

Then, if we substitute the estimates of $P(D^a_m = 1 \mid D^a_{m-1} = 0)$ for $m = 1, \ldots, k+1$ from the IP weighted pooled logistic regression model (A.4.1) into the above formula, the estimate of the survival probability $P(D^a_{k+1} = 0)$ can be obtained. Therefore, using the estimate of the survival probability, we can easily estimate the adjusted survival curve, and also obtain the estimates of adjusted RMST for each exposure group and their difference by integrating the area under the adjusted survival curve up to a specific time point.

To best our knowledge, there is no study on estimating the variance of the adjusted RMST estimated from G-computation method. Alternatively, we can use the bootstrap method to obtain the variance of the estimate.

Appendix B

Appendix for Chapter 4

B.1 Proof of reducing (4.8) to linear fractional programming in special case

In Section 4.5, we considered a closed cohort that the study entry times $t_0 = 0$ are the same for all subjects and there is no censoring apart from administrative censoring at the end of follow-up. Without loss of generality, let the first m subjects experience the events of interest (i.e., $\delta_1 = \cdots = \delta_m = 1$ and $\delta_{m+1} = \cdots = \delta_n = 0$) in the exposed group and the first $D - m$ subjects experience the events of interest (i.e., $\delta_{n+1} = \cdots = \delta_{n-m+D} = 1$ and $\delta_{n-m+D+1} = \cdots = \delta_N = 0$) in the unexposed group. Also, assume that the event times are continuous and ordered increasingly (i.e., $t_1 < \cdots < t_m$ in the exposed group and $t_{n+1} < \cdots < t_{n-m+D}$ in the unexposed group). Since we assume that the only censoring is due to administrative censoring, the last event time in each group (t_m or t_{n-m+D}) is less than or equal to the administrative censoring time. Here, we prove only for the exposed group. Proof for the unexposed group is similar.

Let the pre-specified time point $\tau \in (t_{k-1}, t_k]$ for any $k \in \{2, \ldots, m\}$. In this setting, we can rewrite the objective function in (4.8) as:

$$
\begin{aligned}
&\widehat{\mu}_{\tau}^{(h),1} \\
&= \int_0^{\tau} \prod_{t_j \le t} \left(1 - \frac{\sum_{i:\{T_i = t_j, \delta_i = 1\}} (1 + z_i w_i)}{\sum_{i:\{T_i \ge t_j, \delta_i = 1\}} (1 + z_i w_i) + \sum_{i:\{T_i \ge t_j, \delta_i = 0\}} (1 + \frac{1}{\Lambda} w_i)} \right) \mathrm{d}t \\
&= \sum_{l=0}^{k-1} \prod_{t_j \le t_l} \left(1 - \frac{\sum_{i:\{T_i = t_j, \delta_i = 1\}} (1 + z_i w_i)}{\sum_{i:\{T_i \ge t_j, \delta_i = 1\}} (1 + z_i w_i) + \sum_{i:\{T_i \ge t_j, \delta_i = 0\}} (1 + \frac{1}{\Lambda} w_i)} \right) (t_{l+1} - t_l) \\
&= \sum_{l=0}^{k-1} \prod_{t_j \le t_l} \left(\frac{\sum_{i:\{T_i > t_j, \delta_i = 1\}} (1 + z_i w_i) + \sum_{i:\{T_i \ge t_j, \delta_i = 0\}} (1 + \frac{1}{\Lambda} w_i)}{\sum_{i:\{T_i \ge t_j, \delta_i = 1\}} (1 + z_i w_i) + \sum_{i:\{T_i \ge t_j, \delta_i = 0\}} (1 + \frac{1}{\Lambda} w_i)} \right) (t_{l+1} - t_l)
\end{aligned}
\tag{B.1.1}
$$

where $z_i = \exp\{h(\boldsymbol{L}_i, \boldsymbol{U}_i)\}$ and $w_i = \exp\{-\widehat{g}_{\beta}(\boldsymbol{L}_i)\}$. Note that the survival probability at and after t_k does not affect the estimate of the bias-adjusted RMST up to $\tau \in (t_{k-1}, t_k]$. Because we assume that the event times are continuous and ordered increasingly, for $j = 0, \ldots, k-1$, we can rewrite the numerator of fraction term in (B.1.1) as $(1 + z_{j+1} w_{j+1}) + \cdots + (1 + z_m w_m) + (1 + \frac{1}{\Lambda} w_{m+1}) + \cdots + (1 + \frac{1}{\Lambda} w_n)$ and denote it as f_j^n. Similarly, for $j = 1, \ldots, k$, we can rewrite the denominator as $(1 + z_j w_j) + \cdots + (1 + z_m w_m) + (1 + \frac{1}{\Lambda} w_{m+1}) + \cdots + (1 + \frac{1}{\Lambda} w_n)$ and denote it as f_j^d. Since there is no event of interest at t_0, $f_0^d = (1 + z_1 w_1) + \cdots + (1 + z_m w_m) + (1 + \frac{1}{\Lambda} w_{m+1}) + \cdots + (1 + \frac{1}{\Lambda} w_n)$, especially. Then, the right-hand side in (B.1.1) can be expressed as follows:

$$
\begin{aligned}
\widehat{\mu}_{\tau}^{(h),1} = \left(\frac{f_0^n}{f_0^d} \right) (t_1 - t_0) &+ \left(\frac{f_0^n}{f_0^d} \right) \left(\frac{f_1^n}{f_1^d} \right) (t_2 - t_1) + \\
&+ \left(\frac{f_0^n}{f_0^d} \right) \left(\frac{f_1^n}{f_1^d} \right) \left(\frac{f_2^n}{f_2^d} \right) (t_3 - t_2) + \\
&\vdots \\
&+ \left(\frac{f_0^n}{f_0^d} \right) \left(\frac{f_1^n}{f_1^d} \right) \left(\frac{f_2^n}{f_2^d} \right) \times \cdots \times \left(\frac{f_{k-2}^n}{f_{k-2}^d} \right) (t_{k-1} - t_{k-2}) \\
&+ \left(\frac{f_0^n}{f_0^d} \right) \left(\frac{f_1^n}{f_1^d} \right) \left(\frac{f_2^n}{f_2^d} \right) \times \cdots \times \left(\frac{f_{k-1}^n}{f_{k-1}^d} \right) (\tau - t_{k-1}).
\end{aligned}
$$

Because $f_0^n = f_0^d$ and $f_j^n = f_{j+1}^d$, then,

$$\begin{aligned}
\widehat{\mu}_\tau^{(h),1} = t_1 &+ \left(\frac{f_2^d}{f_1^d}\right)(t_2 - t_1) \\
&+ \left(\frac{f_2^d}{f_1^d}\right)\left(\frac{f_3^d}{f_2^d}\right)(t_3 - t_2) + \\
&\vdots \\
&+ \left(\frac{f_2^d}{f_1^d}\right)\left(\frac{f_3^d}{f_2^d}\right) \times \cdots \times \left(\frac{f_{k-1}^d}{f_{k-2}^d}\right)(t_{k-1} - t_{k-2}) \\
&+ \left(\frac{f_2^d}{f_1^d}\right)\left(\frac{f_3^d}{f_2^d}\right) \times \cdots \times \left(\frac{f_k^d}{f_{k-1}^d}\right)(\tau - t_{k-1}).
\end{aligned}$$

When cancelling common terms in the numerator and denominator, inserting the values of f_j^d, and rearranging the resultant equation, we have

$$\begin{aligned}
&\widehat{\mu}_\tau^{(h),1} \\
&= t_1 + \left(\frac{(1 + z_2 w_2) + \cdots + (1 + z_m w_m) + (1 + \frac{1}{\Lambda} w_{m+1}) + \cdots + (1 + \frac{1}{\Lambda} w_n)}{(1 + z_1 w_1) + \cdots + (1 + \frac{1}{\Lambda} w_n)}\right)(t_2 - t_1) \\
&\quad + \left(\frac{(1 + z_3 w_3) + \cdots + (1 + z_m w_m) + (1 + \frac{1}{\Lambda} w_{m+1}) + \cdots + (1 + \frac{1}{\Lambda} w_n)}{(1 + z_1 w_1) + \cdots + (1 + \frac{1}{\Lambda} w_n)}\right)(t_3 - t_2) + \\
&\quad \vdots \\
&\quad + \left(\frac{(1 + z_{k-1} w_{k-1}) + \cdots + (1 + z_m w_m) + (1 + \frac{1}{\Lambda} w_{m+1}) + \cdots + (1 + \frac{1}{\Lambda} w_n)}{(1 + z_1 w_1) + \cdots + (1 + \frac{1}{\Lambda} w_n)}\right)(t_{k-1} - t_{k-2}) \\
&\quad + \left(\frac{(1 + z_k w_k) + \cdots + (1 + z_m w_m) + (1 + \frac{1}{\Lambda} w_{m+1}) + \cdots + (1 + \frac{1}{\Lambda} w_n)}{(1 + z_1 w_1) + \cdots + (1 + \frac{1}{\Lambda} w_n)}\right)(\tau - t_{k-1}) \\
&= t_1 + \frac{(1 + z_2 w_2)(t_2 - t_1)}{(1 + z_1 w_1) + \cdots + (1 + \frac{1}{\Lambda} w_n)} \\
&\quad + \frac{(1 + z_3 w_3)\{(t_2 - t_1) + (t_3 - t_2)\}}{(1 + z_1 w_1) + \cdots + (1 + \frac{1}{\Lambda} w_n)} + \\
&\quad \vdots \\
&\quad + \frac{(1 + z_{k-1} w_{k-1})\{(t_2 - t_1) + (t_3 - t_2) + \cdots + (t_{k-1} - t_{k-2})\}}{(1 + z_1 w_1) + \cdots + (1 + \frac{1}{\Lambda} w_n)} \\
&\quad + \frac{\{(1 + z_k w_k) + \cdots + (1 + z_m w_m)\}\{(t_2 - t_1) + (t_3 - t_2) + \cdots + (\tau - t_{k-1})\}}{(1 + z_1 w_1) + \cdots + (1 + \frac{1}{\Lambda} w_n)} \\
&\quad + \frac{\{(1 + \frac{1}{\Lambda} w_{m+1}) + \cdots + (1 + \frac{1}{\Lambda} w_n)\}\{(t_2 - t_1) + (t_3 - t_2) + \cdots + (\tau - t_{k-1})\}}{(1 + z_1 w_1) + \cdots + (1 + \frac{1}{\Lambda} w_n)}
\end{aligned}$$

$$
\begin{aligned}
&= t_1 + \frac{(1+z_2w_2)(t_2-t_1)}{(1+z_1w_1)+\cdots+(1+\frac{1}{\Lambda}w_n)} \\
&\quad + \frac{(1+z_3w_3)(t_3-t_1)}{(1+z_1w_1)+\cdots+(1+\frac{1}{\Lambda}w_n)} + \\
&\quad \vdots \\
&\quad + \frac{(1+z_{k-1}w_{k-1})(t_{k-1}-t_1)}{(1+z_1w_1)+\cdots+(1+\frac{1}{\Lambda}w_n)} \\
&\quad + \frac{\{(1+z_kw_k)+\cdots+(1+z_mw_m)\}(\tau-t_1)}{(1+z_1w_1)+\cdots+(1+\frac{1}{\Lambda}w_n)} \\
&\quad + \frac{\{(1+\frac{1}{\Lambda}w_{m+1})+\cdots+(1+\frac{1}{\Lambda}w_n)\}(\tau-t_1)}{(1+z_1w_1)+\cdots+(1+\frac{1}{\Lambda}w_n)}
\end{aligned}
$$

$$
\begin{aligned}
&= t_1 - \frac{\{(1+z_2w_2)+\cdots+(1+\frac{1}{\Lambda}w_n)\}t_1}{(1+z_1w_1)+\cdots+(1+\frac{1}{\Lambda}w_n)} \\
&\quad + \frac{(1+z_2w_2)t_2+\cdots+(1+z_{k-1}w_{k-1})t_{k-1}+(1+z_kw_k)\tau+\cdots+(1+\frac{1}{\Lambda}w_n)\tau}{(1+z_1w_1)+\cdots+(1+\frac{1}{\Lambda}w_n)}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{(1+z_1w_1)t_1+(1+z_2w_2)t_2+\cdots+(1+z_{k-1}w_{k-1})t_{k-1}}{(1+z_1w_1)+\cdots+(1+\frac{1}{\Lambda}w_n)} \\
&\quad + \frac{(1+z_kw_k)\tau+\cdots+(1+z_mw_m)\tau+(1+\frac{1}{\Lambda}w_{m+1})\tau+\cdots+(1+\frac{1}{\Lambda}w_n)\tau}{(1+z_1w_1)+\cdots+(1+\frac{1}{\Lambda}w_n)}
\end{aligned}
$$

$$
= \frac{\sum_{i=1}^{n}(1+z_iw_i)t_i}{\sum_{i=1}^{n}(1+z_iw_i)}.
$$

where $(z_1,\ldots,z_n) = (z_1,\ldots,z_{m-1},z_m,1/\Lambda,\ldots,1/\Lambda)$ and $(t_1,\ldots,t_n) = (t_1,$ $\ldots,$ t_{k-1}, τ, $\ldots$, τ). $\square$

B.2 Proof of reducing (4.8) to linear fractional programming in alternative setting

Consider that a closed cohort where the study entry times are the same for all subjects and that the minimum censoring time is longer than or equal to the pre-specified time point τ. Also, the censoring times are ordered non-decreasingly in each exposure group, respectively (i.e., $t_{m+1} \leq \ldots \leq t_n$ in the exposed group and $t_{n-m+D+1} \leq \ldots \leq t_N$ in the unexposed group). Thus, the

minimum censoring time is t_{m+1} in the exposed group and $t_{n-m+D+1}$ in the unexposed group.

Note that even in this alternative setting, the objective function in (4.8) can be written the same as (B.1.1). Also, the survival probabilities at and after t_k, for any $k \in \{1, \ldots, m\}$, do not affect the estimate of the bias-adjusted RMST up to $\tau \in (t_{k-1}, t_k]$. Therefore, we can apply results in Appendix B.1 without further proof. In other words, we set the optimization parameters $(z_1, \ldots, z_n)$ as $(z_1, \ldots, z_{m-1}, z_m, 1/\Lambda, \ldots, 1/\Lambda)$ and the survival times $(t_1, \ldots, t_n)$ as $(t_1, \ldots, t_{k-1}, \tau, \ldots, \tau)$, and solve the optimization problem using the method described in Section 4.5.

B.3 Proof of non-convergence to boundary values

Objective function of (4.8) in a simple setting

Consider a simple setting that there are only four subjects, all in the exposed group, and the study entry times are the same for all four subjects. Additionally, let the first, third, and fourth subjects experience the event of interest and second subject be censored (i.e., $\delta_1 = \delta_3 = \delta_4 = 1$ and $\delta_2 = 0$). And, let the survival times be ordered increasingly (i.e., $t_1 < t_2 < t_3 < t_4$). In this setting, data are shown in the below Table B.3.1.

Table B.3.1 Data structure for simple setting

t_i	δ_i	w_i	A_i	z_i
$t_0 = 0$	–	–	–	–
t_1	$\delta_1 = 1$	w_1	$A_1 = 1$	z_1
t_2	$\delta_2 = 0$	w_2	$A_2 = 1$	$z_2 = \frac{1}{\Lambda}$
t_3	$\delta_3 = 1$	w_3	$A_3 = 1$	z_3
t_4	$\delta_4 = 1$	w_4	$A_4 = 1$	z_4

Note: t_i: the observed survival time, δ: the event indicator, $w_i = \exp\{-\widehat{g}_\beta(\boldsymbol{L}_i)\}$, A_i: the exposure indicator, z_i: the optimization parameter.

As in case of Appendix B.1, if $\tau = t_4$, then the objective function in (4.8) is represented as follows:

$$\begin{aligned}
\widehat{\mu}_\tau^{(h),1} &= \int_0^\tau \prod_{t_j \le t} \left[1 - \frac{\sum_{i:\{T_i = t_j, \delta_i = 1\}} (1 + z_i w_i)}{\sum_{i:\{T_i \ge t_j, \delta_i = 1\}} (1 + z_i w_i) + \sum_{i:\{T_i \ge t_j, \delta_i = 0\}} (1 + \frac{1}{\Lambda} w_i)} \right] \mathrm{d}t \\
&= t_1 + \left(\frac{(1 + z_2 w_2) + (1 + z_3 w_3) + (1 + z_4 w_4)}{(1 + z_1 w_1) + (1 + z_2 w_2) + (1 + z_3 w_3) + (1 + z_4 w_4)} \right) (t_3 - t_1) \\
&\quad + \left(\frac{(1 + z_2 w_2) + (1 + z_3 w_3) + (1 + z_4 w_4)}{(1 + z_1 w_1) + (1 + z_2 w_2) + (1 + z_3 w_3) + (1 + z_4 w_4)} \right) \\
&\quad \times \left(\frac{(1 + z_4 w_4)}{(1 + z_3 w_3) + (1 + z_4 w_4)} \right) (\tau - t_3).
\end{aligned}$$

Minimizing the above equation, the optimization parameter for the first subject (i.e., z_1) have to be Λ and that for the last subject (i.e., z_4) have to be $1/\Lambda$. Because the second subject is censored, z_2 should be equal to $1/\Lambda$. Therefore,

$$\begin{aligned}
\widehat{\mu}_\tau^{(h),1} &= t_1 + \left(\frac{(1 + z_2 w_2) + (1 + z_3 w_3) + (1 + z_4 w_4)}{(1 + z_1 w_1) + (1 + z_2 w_2) + (1 + z_3 w_3) + (1 + z_4 w_4)} \right) (t_3 - t_1) \\
&\quad + \left(\frac{(1 + z_2 w_2) + (1 + z_3 w_3) + (1 + z_4 w_4)}{(1 + z_1 w_1) + (1 + z_2 w_2) + (1 + z_3 w_3) + (1 + z_4 w_4)} \right) \\
&\quad \times \left(\frac{(1 + z_4 w_4)}{(1 + z_3 w_3) + (1 + z_4 w_4)} \right) (\tau - t_3) \\
&= t_1 + \left(\frac{(1 + \frac{1}{\Lambda} w_2) + (1 + z_3 w_3) + (1 + \frac{1}{\Lambda} w_4)}{(1 + \Lambda w_1) + (1 + \frac{1}{\Lambda} w_2) + (1 + z_3 w_3) + (1 + \frac{1}{\Lambda} w_4)} \right) (t_3 - t_1) \\
&\quad + \left(\frac{(1 + \frac{1}{\Lambda} w_2) + (1 + z_3 w_3) + (1 + \frac{1}{\Lambda} w_4)}{(1 + \Lambda w_1) + (1 + \frac{1}{\Lambda} w_2) + (1 + z_3 w_3) + (1 + \frac{1}{\Lambda} w_4)} \right) \\
&\quad \times \left(\frac{(1 + \frac{1}{\Lambda} w_4)}{(1 + z_3 w_3) + (1 + \frac{1}{\Lambda} w_4)} \right) (\tau - t_3)
\end{aligned}$$

$$
\begin{aligned}
&= t_1 - \left(\frac{(1+\frac{1}{\Lambda}w_2)+(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)}{(1+\Lambda w_1)+(1+\frac{1}{\Lambda}w_2)+(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)} \right) t_1 \\
&\quad + \left[\left(\frac{(1+\frac{1}{\Lambda}w_2)+(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)}{(1+\Lambda w_1)+(1+\frac{1}{\Lambda}w_2)+(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)} \right) \right. \\
&\qquad - \left(\frac{(1+\frac{1}{\Lambda}w_2)+(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)}{(1+\Lambda w_1)+(1+\frac{1}{\Lambda}w_2)+(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)} \right) \\
&\qquad \left. \times \left(\frac{(1+\frac{1}{\Lambda}w_4)}{(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)} \right) \right] t_3 \\
&\quad + \left(\frac{(1+\frac{1}{\Lambda}w_2)+(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)}{(1+\Lambda w_1)+(1+\frac{1}{\Lambda}w_2)+(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)} \right) \\
&\qquad \times \left(\frac{(1+\frac{1}{\Lambda}w_4)}{(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)} \right) \tau \\
&= t_1 - \left(\frac{(1+\frac{1}{\Lambda}w_2)+(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)}{(1+\Lambda w_1)+(1+\frac{1}{\Lambda}w_2)+(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)} \right) t_1 \\
&\quad + \left(\frac{(1+\frac{1}{\Lambda}w_2)+(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)}{(1+\Lambda w_1)+(1+\frac{1}{\Lambda}w_2)+(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)} \right) \\
&\qquad \times \left(\frac{(1+z_3w_3)}{(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)} \right) t_3 \\
&\quad + \left(\frac{(1+\frac{1}{\Lambda}w_2)+(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)}{(1+\Lambda w_1)+(1+\frac{1}{\Lambda}w_2)+(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)} \right) \\
&\qquad \times \left(\frac{(1+\frac{1}{\Lambda}w_4)}{(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)} \right) \tau \\
&= t_1 - \left(\frac{(1+\frac{1}{\Lambda}w_2)+(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)}{(1+\Lambda w_1)+(1+\frac{1}{\Lambda}w_2)+(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)} \right) \\
&\qquad \times \left[t_1 - \left(\frac{(1+z_3w_3)}{(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)} \right) t_3 \right. \\
&\qquad\qquad \left. - \left(\frac{(1+\frac{1}{\Lambda}w_4)}{(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)} \right) \tau \right] \\
&= t_1 - \left(\frac{(1+\frac{1}{\Lambda}w_2)+(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)}{(1+\Lambda w_1)+(1+\frac{1}{\Lambda}w_2)+(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)} \right) \\
&\qquad \times \left(\frac{(1+z_3w_3)(t_1-t_3)+(1+\frac{1}{\Lambda}w_4)(t_1-\tau)}{(1+z_3w_3)+(1+\frac{1}{\Lambda}w_4)} \right)
\end{aligned}
$$

$$= t_1 + \left(\frac{(1+\frac{1}{\Lambda}w_2) + (1+z_3w_3) + (1+\frac{1}{\Lambda}w_4)}{(1+\Lambda w_1) + (1+\frac{1}{\Lambda}w_2) + (1+z_3w_3) + (1+\frac{1}{\Lambda}w_4)} \right) \times \left(\frac{(1+z_3w_3)(t_3-t_1) + (1+\frac{1}{\Lambda}w_4)(\tau - t_1)}{(1+z_3w_3) + (1+\frac{1}{\Lambda}w_4)} \right). \tag{B.3.1}$$

Counter-example of simple numerical data

To show that some optimization parameters z may not converge to boundary value of $1/\Lambda$ or Λ but converge to value between $1/\Lambda$ and Λ, we construct a simple counter-example following the setting given in Table B.3.1. Let $\Lambda = 2$, and survival data is given as seen in Table B.3.2.

Table B.3.2 Data for counter-example

t_i	δ_i	w_i	A_i	z_i
$t_0 = 0$	–	–	–	–
$t_1 = 1$	$\delta_1 = 1$	$w_1 = 9.0$	$A_1 = 1$	$z_1 = 2$
$t_2 = 2$	$\delta_2 = 0$	$w_2 = 2.3$	$A_2 = 1$	$z_2 = \frac{1}{2}$
$t_3 = 4$	$\delta_3 = 1$	$w_3 = 1.5$	$A_3 = 1$	z_3
$t_4 = 14$	$\delta_4 = 1$	$w_4 = 0.2$	$A_4 = 1$	$z_4 = \frac{1}{2}$

In this case, (B.3.1) is written as

$$\begin{aligned} \widehat{\mu}_\tau^{(h),1} &= 1 + \left(\frac{1.5z_3 + 4.25}{1.5z_3 + 23.25} \right) \left(\frac{4.5z_3 + 17.3}{1.5z_3 + 2.1} \right) \\ &= 4 - \frac{1}{2.25} \left(\frac{69z_3 + 72.95}{(z_3 + 1.4)(z_3 + 15.5)} \right). \end{aligned}$$

Minimizing above $\widehat{\mu}_\tau^{(h),1}$ subject to $1/2 \leq z_3 \leq 2$ is equal to maximizing

$$\frac{69z_3 + 72.95}{(z_3 + 1.4)(z_3 + 15.5)}. \tag{B.3.2}$$

For $z_3 \in [1/2, 2]$, plot for equation (B.3.2) is shown in Figure B.3.1, and equation (B.3.2) is maximized at $z_3 \approx 1.168$. Therefore, this counter-example

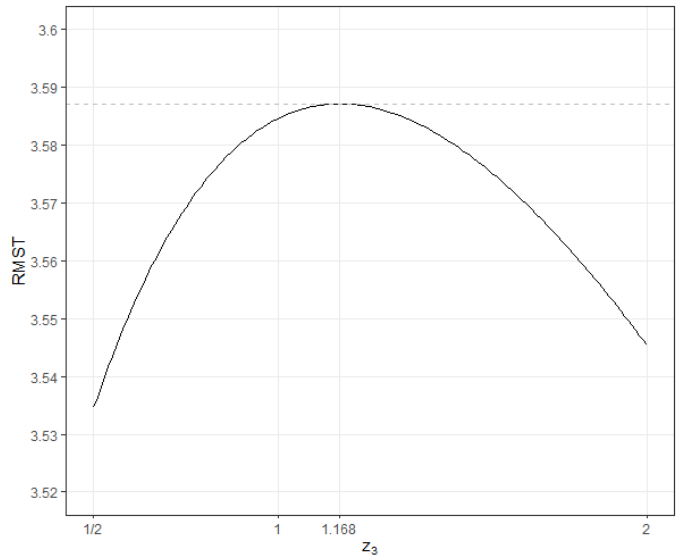


Figure B.3.1 Plot for counter-example of simple numerical data

shows that in some situations, the optimization parameters may not converge to boundary value of $1/\Lambda$ or Λ but converge to value between $1/\Lambda$ and Λ. □

B.4 Details for simulation study 2.1 ($\beta_0 = -1.9$)

Simulation 2.1 study tables ($\beta_0 = -1.9$)

Table B.4.1 Bias for lower bound of sensitivity range ($\beta_0 = -1.9$)

Λ	τ	**Censoring rate**	**Mean**	**Lower**	**Upper**
		0.1	0.000000	-0.000002	0.000000
		0.3	0.000000	0.000000	0.000000
	1	0.5	0.000000	0.000000	0.000000
		0.7	0.000000	0.000000	0.000000
		0.9	0.000000	-0.000005	0.000000
1.1		0.1	0.000000	0.000000	0.000000
		0.3	0.000000	0.000000	0.000000
	3	0.5	0.000000	0.000000	0.000000
		0.7	0.000000	0.000000	0.000000
		0.9	0.000027	0.000000	0.002806
		0.1	0.000000	0.000000	0.000000
		0.3	0.000000	0.000000	0.000000
	1	0.5	0.000000	0.000000	0.000000
		0.7	0.000000	0.000000	0.000000
		0.9	0.000000	-0.000012	0.000000
1.3		0.1	0.000000	0.000000	0.000000
		0.3	0.000000	-0.000005	0.000000
	3	0.5	0.000000	0.000000	0.000001
		0.7	0.000000	0.000000	0.000012
		0.9	0.000075	-0.000022	0.007333
		0.1	0.000000	0.000000	0.000000
		0.3	0.000000	0.000000	0.000000
	1	0.5	0.000000	0.000000	0.000000
		0.7	0.000000	0.000000	0.000000
		0.9	0.000000	-0.000017	0.000000
1.5		0.1	0.000000	0.000000	0.000000
		0.3	0.000000	0.000000	0.000000
	3	0.5	0.000000	0.000000	0.000004
		0.7	0.000000	0.000000	0.000001
		0.9	0.000111	0.000000	0.010818
		0.1	0.000000	0.000000	0.000000
		0.3	0.000000	0.000000	0.000000
	1	0.5	0.000000	0.000000	0.000000
		0.7	0.000000	0.000000	0.000000
		0.9	0.000000	-0.000022	0.000000
2		0.1	0.000000	0.000000	0.000000
		0.3	0.000000	0.000000	0.000000
	3	0.5	0.000000	0.000000	0.000002
		0.7	0.000000	-0.000001	0.000046
		0.9	0.000180	0.000000	0.017455

Table B.4.2 Bias for upper bound of sensitivity range ($\beta_0 = -1.9$)

Λ	τ	**Censoring rate**	**Mean**	**Lower**	**Upper**
		0.1	0.000000	0.000000	0.000003
		0.3	0.000000	0.000000	0.000000
	1	0.5	0.000000	0.000000	0.000000
		0.7	0.000000	0.000000	0.000004
		0.9	0.000000	0.000000	0.000006
1.1		0.1	0.000000	0.000000	0.000000
		0.3	0.000000	0.000000	0.000000
	3	0.5	0.000000	0.000000	0.000000
		0.7	0.000000	0.000000	0.000000
		0.9	-0.000028	-0.002963	0.000000
		0.1	0.000000	0.000000	0.000006
		0.3	0.000000	0.000000	0.000000
	1	0.5	0.000000	0.000000	0.000000
		0.7	0.000000	0.000000	0.000011
		0.9	0.000000	0.000000	0.000000
1.3		0.1	0.000000	0.000000	0.000000
		0.3	0.000000	0.000000	0.000000
	3	0.5	0.000000	0.000000	0.000000
		0.7	0.000000	-0.000001	0.000002
		0.9	-0.000077	-0.008510	0.000005
		0.1	0.000000	0.000000	0.000000
		0.3	0.000000	0.000000	0.000000
	1	0.5	0.000000	0.000000	0.000000
		0.7	0.000000	0.000000	0.000000
		0.9	0.000000	0.000000	0.000000
1.5		0.1	0.000000	0.000000	0.000000
		0.3	0.000000	0.000000	0.000000
	3	0.5	0.000000	0.000000	0.000000
		0.7	0.000000	-0.000001	0.000003
		0.9	-0.000119	-0.013571	0.000044
		0.1	0.000000	0.000000	0.000000
		0.3	0.000000	0.000000	0.000000
	1	0.5	0.000000	0.000000	0.000000
		0.7	0.000000	0.000000	0.000000
		0.9	0.000000	0.000000	0.000000
2		0.1	0.000000	0.000000	0.000000
		0.3	0.000000	0.000000	0.000000
	3	0.5	0.000000	-0.000003	0.000000
		0.7	0.000000	-0.000010	0.000008
		0.9	-0.000218	-0.024271	0.000022

Table B.4.3 Computational time difference ($\beta_0 = -1.9$)

Λ	τ	**Censoring rate**	**Mean**	**Lower**	**Upper**
1.1	1	0.1	119.11	8.11	337.70
		0.3	28.17	3.77	166.74
		0.5	6.85	-0.05	68.94
		0.7	1.38	-3.78	3.00
		0.9	0.39	-5.03	0.97
	3	0.1	32.20	7.34	111.89
		0.3	9.57	4.43	44.74
		0.5	4.38	2.29	9.40
		0.7	2.14	0.21	6.51
		0.9	0.46	-0.21	7.56
1.3	1	0.1	152.15	13.19	422.61
		0.3	39.51	7.22	193.54
		0.5	10.32	3.07	82.66
		0.7	2.16	0.98	9.22
		0.9	0.55	-0.27	1.14
	3	0.1	48.94	14.35	138.29
		0.3	15.03	7.40	61.71
		0.5	6.91	3.65	22.07
		0.7	2.79	1.20	9.33
		0.9	0.57	-0.19	1.84
1.5	1	0.1	163.66	15.06	439.03
		0.3	46.20	8.73	204.62
		0.5	12.63	4.24	92.52
		0.7	2.70	1.18	9.75
		0.9	0.68	-0.16	6.80
	3	0.1	54.38	17.25	150.15
		0.3	17.70	10.24	68.36
		0.5	8.22	5.01	20.02
		0.7	3.16	1.48	5.71
		0.9	0.65	-0.09	5.94
2	1	0.1	176.31	21.70	442.43
		0.3	54.20	12.44	210.29
		0.5	15.58	6.10	98.16
		0.7	3.55	-3.32	14.59
		0.9	0.80	0.12	1.51
	3	0.1	56.35	23.02	140.20
		0.3	23.13	11.49	67.10
		0.5	11.40	7.19	30.80
		0.7	4.07	2.18	6.44
		0.9	0.72	-0.05	1.48

Table B.4.4 Computational time ratio ($\beta_0 = -1.9$)

Λ	τ	**Censoring rate**	**Mean**	**Lower**	**Upper**
1.1	1	0.1	32.00	3.36	79.89
		0.3	9.57	2.28	47.31
		0.5	3.30	1.63	18.23
		0.7	1.50	1.12	1.85
		0.9	1.16	0.97	1.29
	3	0.1	9.37	3.64	22.40
		0.3	3.87	2.61	8.80
		0.5	2.44	1.73	3.35
		0.7	1.76	1.30	2.84
		0.9	1.18	0.99	1.36
1.3	1	0.1	40.42	5.15	98.46
		0.3	13.05	3.18	55.48
		0.5	4.41	2.07	18.81
		0.7	1.78	1.35	2.23
		0.9	1.21	1.00	1.37
	3	0.1	13.77	5.45	33.20
		0.3	5.59	3.54	14.07
		0.5	3.27	2.30	4.72
		0.7	2.00	1.51	2.99
		0.9	1.22	1.01	1.47
1.5	1	0.1	42.90	6.23	101.60
		0.3	15.03	3.79	54.42
		0.5	5.15	2.50	21.10
		0.7	1.97	1.40	2.52
		0.9	1.26	1.00	1.47
	3	0.1	15.33	6.07	37.08
		0.3	6.35	4.28	13.87
		0.5	3.72	2.65	4.93
		0.7	2.13	1.56	2.77
		0.9	1.24	1.02	1.51
2	1	0.1	46.91	7.57	102.72
		0.3	17.41	5.06	57.41
		0.5	6.14	2.87	22.73
		0.7	2.26	1.76	3.01
		0.9	1.30	1.07	1.50
	3	0.1	15.46	7.46	34.74
		0.3	7.94	5.18	17.20
		0.5	4.76	3.42	6.69
		0.7	2.46	1.80	3.07
		0.9	1.27	1.06	1.48

B.5 Details for simulation study 2.1 ($\beta_0 = -0.425$)

Simulation study 2.1 figures ($\beta_0 = -0.425$)

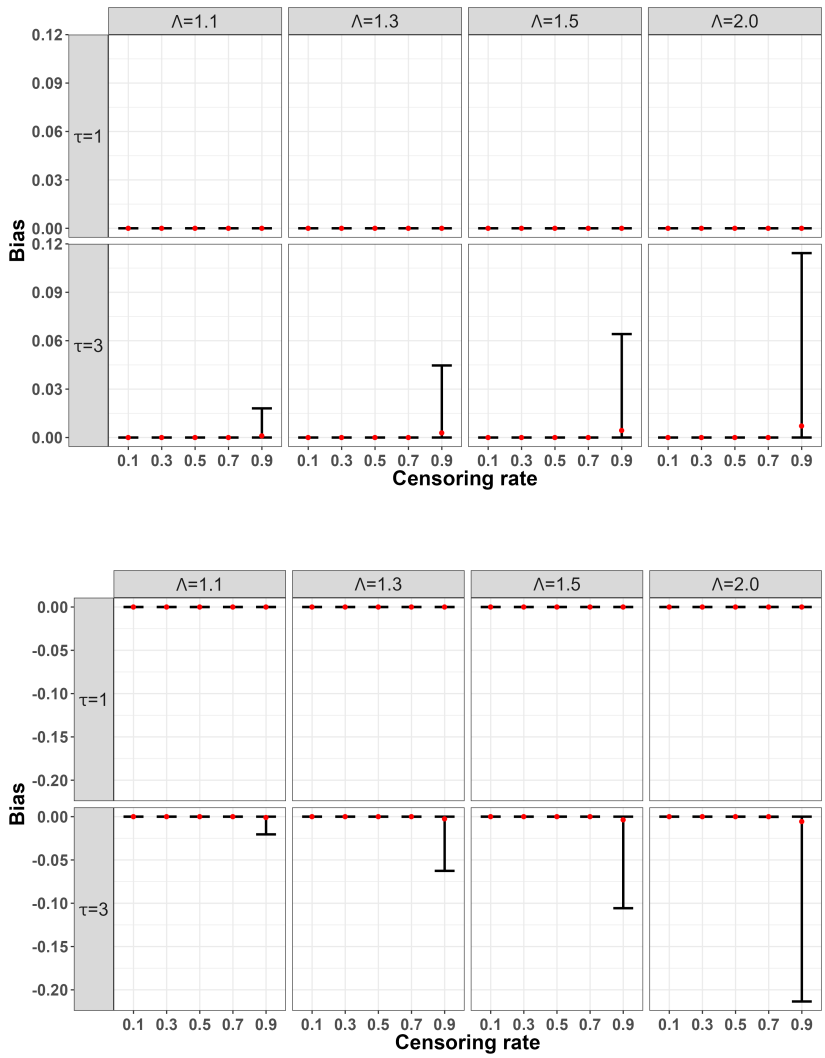


Figure B.5.1 Left: Bias for lower bound of sensitivity range. Right: Bias for upper bound of sensitivity range. The mean is represented by red dot. The range of bias is represented by lower and upper horizontal bar.

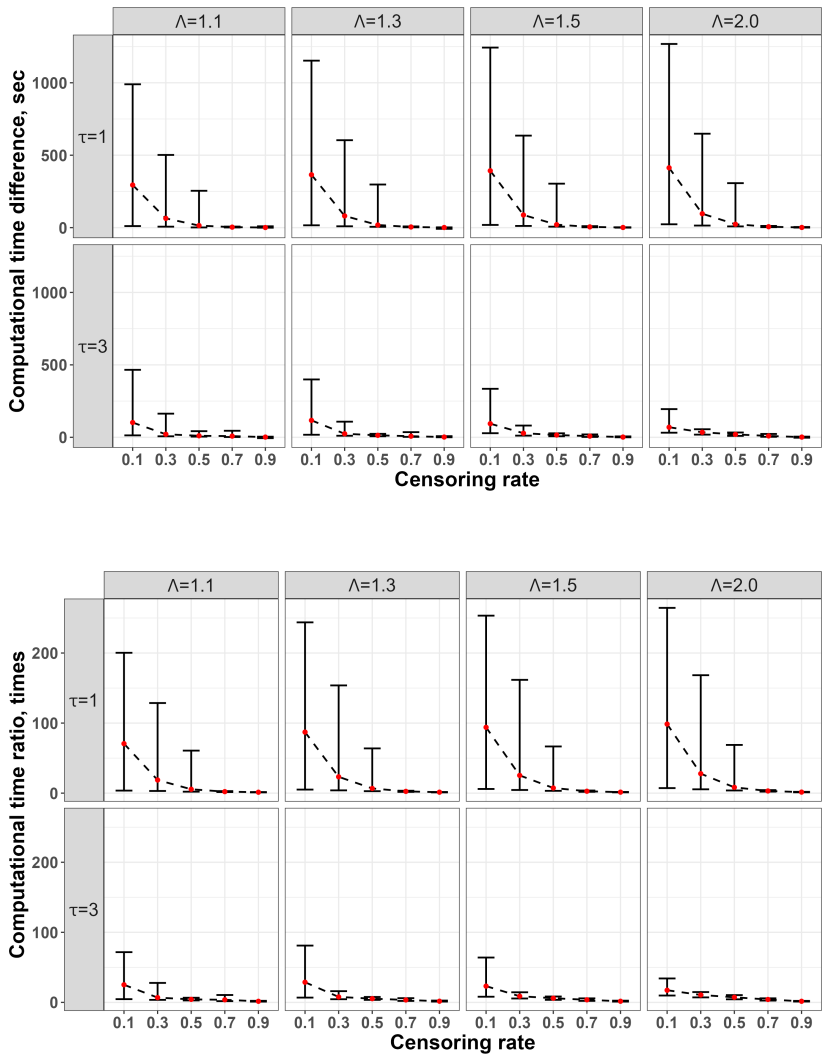


Figure B.5.2 (Top): Difference in computational time between two methods per each Monte Carlo data. (Bottom): Ratio of computational time between two methods per each Monte Carlo data. The mean is represented by red dot. The range of computational time difference is represented by lower and upper horizontal bar.

Simulation study 2.1 tables ($\beta_0 = -0.425$)

Table B.5.1 Bias for lower bound of sensitivity range ($\beta_0 = -0.425$)

Λ	τ	**Censoring rate**	**Mean**	**Lower**	**Upper**
		0.1	0.000000	-0.000004	0.000000
		0.3	0.000000	-0.000003	0.000000
	3	0.5	0.000000	-0.000003	0.000000
		0.7	0.000000	-0.000003	0.000000
		0.9	0.000000	-0.000002	0.000000
1.1		0.1	0.000000	-0.000003	0.000000
		0.3	0.000000	0.000000	0.000000
	5	0.5	0.000000	0.000000	0.000000
		0.7	0.000000	0.000000	0.000000
		0.9	0.001024	0.000000	0.018074
		0.1	0.000000	-0.000009	0.000000
		0.3	0.000000	-0.000007	0.000000
	3	0.5	0.000000	-0.000007	0.000000
		0.7	0.000000	-0.000008	0.000000
		0.9	0.000000	-0.000004	0.000000
1.3		0.1	0.000000	0.000000	0.000000
		0.3	0.000000	0.000000	0.000000
	5	0.5	0.000000	0.000000	0.000000
		0.7	0.000000	0.000000	0.000002
		0.9	0.002850	0.000000	0.044659
		0.1	0.000000	-0.000015	0.000000
		0.3	0.000000	-0.000011	0.000000
	3	0.5	0.000000	-0.000048	0.000000
		0.7	0.000000	-0.000010	0.000000
		0.9	0.000000	-0.000006	0.000000
1.5		0.1	0.000000	0.000000	0.000000
		0.3	0.000000	0.000000	0.000000
	5	0.5	0.000000	-0.000003	0.000000
		0.7	0.000000	-0.000002	0.000000
		0.9	0.004370	0.000000	0.064123
		0.1	0.000000	0.000000	0.000000
		0.3	0.000000	0.000000	0.000000
	3	0.5	0.000000	-0.000011	0.000000
		0.7	0.000000	-0.000012	0.000000
		0.9	0.000000	-0.000007	0.000000
2.0		0.1	0.000000	0.000000	0.000000
		0.3	0.000000	-0.000004	0.000000
	5	0.5	0.000000	0.000000	0.000000
		0.7	0.000000	0.000000	0.000000
		0.9	0.007139	0.000000	0.114290

Table B.5.2 Bias for upper bound of sensitivity range ($\beta_0 = -0.425$)

Λ	τ	**Censoring rate**	**Mean**	**Lower**	**Upper**
		0.1	0.000000	0.000000	0.000006
		0.3	0.000000	0.000000	0.000005
	3	0.5	0.000000	0.000000	0.000004
		0.7	0.000000	0.000000	0.000004
		0.9	0.000000	0.000000	0.000002
1.1		0.1	0.000000	0.000000	0.000000
		0.3	0.000000	0.000000	0.000000
	5	0.5	0.000000	0.000000	0.000000
		0.7	0.000000	0.000000	0.000000
		0.9	-0.000981	-0.020446	0.000000
		0.1	0.000000	0.000000	0.000000
		0.3	0.000000	0.000000	0.000000
	3	0.5	0.000000	0.000000	0.000011
		0.7	0.000000	0.000000	0.000014
		0.9	0.000000	0.000000	0.000008
1.3		0.1	0.000000	0.000000	0.000001
		0.3	0.000000	0.000000	0.000000
	5	0.5	0.000000	-0.000002	0.000000
		0.7	0.000000	-0.000050	0.000000
		0.9	-0.002541	-0.062602	0.000000
		0.1	0.000000	0.000000	0.000000
		0.3	0.000000	0.000000	0.000000
	3	0.5	0.000000	0.000000	0.000020
		0.7	0.000000	0.000000	0.000024
		0.9	0.000000	0.000000	0.000015
1.5		0.1	0.000000	0.000000	0.000011
		0.3	0.000000	-0.000001	0.000000
	5	0.5	0.000000	-0.000002	0.000000
		0.7	0.000000	-0.000022	0.000001
		0.9	-0.003670	-0.105715	0.000000
		0.1	0.000000	0.000000	0.000000
		0.3	0.000000	0.000000	0.000000
	3	0.5	0.000000	0.000000	0.000046
		0.7	0.000000	0.000000	0.000056
		0.9	0.000000	-0.000001	0.000034
2.0		0.1	0.000000	0.000000	0.000000
		0.3	0.000000	-0.000012	0.000000
	5	0.5	-0.000001	-0.000032	0.000000
		0.7	-0.000007	-0.000309	0.000000
		0.9	-0.005636	-0.213447	0.000007

Table B.5.3 Computational time difference ($\beta_0 = -0.425$)

Λ	τ	**Censoring rate**	**Mean**	**Lower**	**Upper**
1.1	3	0.1	294.35	10.89	989.53
		0.3	63.93	6.84	501.90
		0.5	14.38	1.43	254.41
		0.7	3.50	1.21	7.41
		0.9	0.88	-0.31	9.37
	5	0.1	101.92	13.32	465.69
		0.3	20.92	6.97	163.21
		0.5	10.47	4.82	42.29
		0.7	8.21	1.97	45.06
		0.9	1.43	-5.82	4.28
1.3	3	0.1	365.07	16.20	1152.71
		0.3	80.88	9.42	603.62
		0.5	18.28	5.82	297.71
		0.7	4.60	2.26	8.49
		0.9	0.98	-8.19	1.90
	5	0.1	116.92	17.60	399.34
		0.3	24.43	10.76	107.86
		0.5	13.68	7.24	23.80
		0.7	7.66	3.09	35.59
		0.9	1.78	0.19	7.85
1.5	3	0.1	392.73	18.78	1242.86
		0.3	87.43	11.80	635.26
		0.5	20.57	6.76	303.59
		0.7	5.26	2.53	10.36
		0.9	1.14	0.02	2.51
	5	0.1	93.95	28.27	334.75
		0.3	28.13	11.53	81.08
		0.5	15.72	7.21	27.54
		0.7	8.10	2.61	19.75
		0.9	1.64	-0.01	6.16
2.0	3	0.1	413.14	23.23	1267.89
		0.3	96.11	14.45	648.45
		0.5	23.56	8.66	307.20
		0.7	6.37	3.31	11.84
		0.9	1.34	0.47	4.45
	5	0.1	69.48	31.70	194.61
		0.3	34.67	18.49	55.63
		0.5	20.02	9.87	33.15
		0.7	9.41	3.58	23.12
		0.9	1.49	-2.88	4.60

Table B.5.4 Computational time ratio ($\beta_0 = -0.425$)

Λ	τ	**Censoring rate**	**Mean**	**Lower**	**Upper**
1.1	3	0.1	70.63	3.64	200.34
		0.3	18.74	3.11	128.64
		0.5	5.54	2.11	60.71
		0.7	2.16	1.58	2.91
		0.9	1.33	1.10	1.55
	5	0.1	25.13	4.60	71.68
		0.3	6.82	3.59	27.77
		0.5	4.27	2.77	6.59
		0.7	3.70	1.77	10.52
		0.9	1.53	1.11	2.21
1.3	3	0.1	87.13	5.09	243.78
		0.3	23.30	3.98	153.78
		0.5	6.70	2.83	63.86
		0.7	2.52	1.78	3.44
		0.9	1.37	1.10	1.63
	5	0.1	28.76	6.81	81.07
		0.3	7.78	4.50	16.00
		0.5	5.29	3.56	7.72
		0.7	3.49	2.08	6.01
		0.9	1.65	1.17	2.57
1.5	3	0.1	93.95	5.95	253.36
		0.3	25.32	4.51	161.71
		0.5	7.38	3.29	66.64
		0.7	2.73	1.90	3.75
		0.9	1.42	1.16	1.69
	5	0.1	23.19	8.07	63.99
		0.3	8.76	5.56	14.41
		0.5	5.95	3.61	8.38
		0.7	3.68	2.14	5.72
		0.9	1.60	1.14	2.43
2.0	3	0.1	98.59	7.21	264.45
		0.3	27.80	5.48	168.35
		0.5	8.31	3.82	68.82
		0.7	3.10	2.20	4.34
		0.9	1.49	1.21	1.86
	5	0.1	17.33	9.77	34.14
		0.3	10.61	7.12	14.71
		0.5	7.26	4.37	10.37
		0.7	4.12	2.56	5.78
		0.9	1.56	1.21	2.11

Appendix C

Appendix for Chapter 5

C.1 Alternative sensitivity analysis method

Lin, Psaty, and Kronmal (1998) proposed a sensitivity analysis method for assessing the sensitivity of the point estimates and their confidence intervals of exposure effect to the unmeasured confounding after adjusting for measured confounders. They assumed that the true exposure effect is represented by a regression model that includes the exposure as well as both the measured and unmeasured confounders. Among the five methods being considered described in Section 3.2, Andersen, Hansen, and Klein (2004)'s pseudo-observation method and Tian, Zhao, and Wei (2014)'s ANCOVA-type model which directly relate the RMST to confounders can be applied to the sensitivity analysis method proposed by Lin, Psaty, and Kronmal (1998). Here, we apply the sensitivity analysis method of Lin, Psaty, and Kronmal (1998) only to ANCOVA-type regression model to consider unmeasured confounding for evaluating the estimate of the difference in adjusted RMST. Pseudo-observation method can be also applied to sensitivity analysis method in Lin, Psaty, and Kronmal (1998) in a similar way.

We first assume that if we adjusted for a single unmeasured confounder (which we denoted by U) along with measured confounders $\boldsymbol{L}$, then all confounding is removed. Also, assume that the expected value of the restricted survival time T_τ is related to A, $\boldsymbol{L}$, and U through the linear model (hereafter referred to this linear model as the full linear model)

$$E(T_\tau \mid A, \boldsymbol{L}, U) = \beta_0^* + \beta_1^* A + \boldsymbol{\beta_l^{*\intercal}} \boldsymbol{L} + \gamma_A U \tag{C.1.1}$$

where γ_0 and γ_1 pertain to the effect of U for the exposed and unexposed groups, respectively. Note that the parameterization $\beta_1^* A + \gamma_A U$ is equal to $\beta_1^* A + \gamma_0 U + (\gamma_1 - \gamma_0) AU$. If $\gamma_0 \neq \gamma_1$, then β_1^* is the main effect of A under the model in which there is an interaction term between A and U, so that β_1^* cannot be independently interpreted. For most practical purposes, it is suffices to set $\gamma_0 = \gamma_1$.

Since U is unmeasured, one can be forced to fit the reduced linear model

$$E(T_\tau \mid A, \boldsymbol{L}) = \beta_0 + \beta_1 A + \boldsymbol{\beta_l^\intercal} \boldsymbol{L} \tag{C.1.2}$$

where β_0, β_1, and $\boldsymbol{\beta_l^\intercal}$ are potentially different from β_0^*, β_1^*, and $\boldsymbol{\beta_l^{*\intercal}}$ in equation (C.1.1). Lin, Psaty, and Kronmal (1998) referred to β_1^* and β_1 as the true and apparent differences in RMST (i.e., true and apparent exposure effects), respectively. Since β_1^* cannot estimated from the observed data but β_1 can be directly estimated, it is of interest to identify the relation between β_1^* and β_1.

Binary confounder

Let $F(u \mid A, \boldsymbol{L})$ be the distribution function of U given A and $\boldsymbol{L}$. Then, by the law of conditional expectation,

$$E(T_\tau \mid A, \boldsymbol{L}) = \int_{-\infty}^{\infty} E(T_\tau \mid A, \boldsymbol{L}, u) \, \mathrm{d}F(u \mid A, \boldsymbol{L}). \tag{C.1.3}$$

Under the full linear model (C.1.1), the conditional expectation (C.1.3) becomes

$$E(T_\tau \mid A, \boldsymbol{L}) = \beta_0^* + \beta_1^* A + \boldsymbol{\beta}_{\boldsymbol{l}}^{*\intercal} \boldsymbol{L} + \int_{-\infty}^{\infty} \gamma_A u \, \mathrm{d}F(u \mid A, \boldsymbol{L}). \tag{C.1.4}$$

Assume that the unmeasured confounder U is binary such that $F(u \mid A, \boldsymbol{L})$ is a Bernoulli distribution with success probability $P_{A,\boldsymbol{L}} = P(U = 1 \mid A, \boldsymbol{L})$. Then, equation (C.1.4) becomes

$$\begin{aligned} E(T_\tau \mid A, \boldsymbol{L}) &= \beta_0^* + \beta_1^* A + \boldsymbol{\beta}_{\boldsymbol{l}}^{*\intercal} \boldsymbol{L} + \gamma_A P_{A,\boldsymbol{L}} \\ &= \beta_0^* + \boldsymbol{\beta}_{\boldsymbol{l}}^{*\intercal} \boldsymbol{L} + (\gamma_0 P_{0,\boldsymbol{L}} + \{\beta_1^* + \gamma_1 P_{1,\boldsymbol{L}} - \gamma_0 P_{0,\boldsymbol{L}}\} A). \end{aligned} \tag{C.1.5}$$

If we assume that U is independent of $\boldsymbol{L}$ conditional on A (i.e., $U \perp\!\!\!\perp \boldsymbol{L} \mid A$, so $P_{A,\boldsymbol{L}} = P(U = 1 \mid A)$) and let $P_A = P(U = 1 \mid A)$, then (C.1.5) is

$$E(T_\tau \mid A, \boldsymbol{L}) = \beta_0^* + \boldsymbol{\beta}_{\boldsymbol{l}}^{*\intercal} \boldsymbol{L} + (\gamma_0 P_0 + \{\beta_1^* + \gamma_1 P_1 - \gamma_0 P_0\} A). \tag{C.1.6}$$

Considering both (C.1.2) and (C.1.6), the true difference in RMST is

$$\beta_1^* = \beta_1 - (\gamma_1 P_1 - \gamma_0 P_0). \tag{C.1.7}$$

When $\gamma_0 = \gamma_1 = \gamma$, (C.1.7) can be reduced by

$$\beta_1^* = \beta_1 - \gamma(P_1 - P_0).$$

Normal confounder

Suppose that conditional on A and $\boldsymbol{L}$, the unmeasured confounder U is normally distributed with mean $\mu_{A,\boldsymbol{L}}$ and variance one. In this case, the true difference in RMST can be calculated similarly to the case of binary confounder. If U is independent of $\boldsymbol{L}$ given A, then $E(U \mid A, \boldsymbol{L}) = \mu_{A,\boldsymbol{L}} = \mu_A$. Thus, the true

difference in RMST is

$$\beta_1^* = \beta_1 - (\gamma_1\mu_1 - \gamma_0\mu_0). \tag{C.1.8}$$

If $\gamma_0 = \gamma_1 = \gamma$, then (C.1.8) can be reduced by

$$\beta_1^* = \beta_1 - \gamma(\mu_1 - \mu_0). \tag{C.1.9}$$

Note that (C.1.9) does not require the conditional independence of U and $\boldsymbol{L}$ given A, but the effects of A and $\boldsymbol{L}$ on the mean of U have to be additive. For example, $\mu_{A,\boldsymbol{L}} = \mu_A + q(\boldsymbol{L})$ for normal unmeasured confounder U, where q is any arbitrary function of $\boldsymbol{L}$.

C.2 Limitation of alternative method

Lin, Psaty, and Kronmal (1998)'s sensitivity analysis method has some limitations. First, their method deals only with a single unmeasured confounder. When there are multiple unmeasured confounders, one should substitute multiple unmeasured confounders with a composite of them. Also, the unmeasured confounder should follow either a Bernoulli distribution or a normal distribution. Additionally, to apply their method to real applications, the true exposure effect should be represented by a regression model.

However, our sensitivity analysis method requires only one sensitivity parameter Λ. Also, our method can be used regardless of the distribution of unmeasured confounder and regardless of model for the exposure (i.e., regardless of PS estimation method).

국 문 초 록

생존 분석에서 제한된 평균 생존 시간(restricted mean survival time; RMST)의 차이는 위험 비율(hazard ratio)에 대한 대안 척도로 점점 더 많이 사용되고 있다. 위험 비율과 같은 상대적 효과 측도(relative effect measure)와 달리, RMST 차이는 직관적으로 해석 가능한 절대 위험(absolute risk)에 대한 정보를 제공하며 비례 위험 가정에 관계없이 로버스트한 것으로 알려져 있다.

무작위대조시험에서는 Kaplan-Meier 곡선 아래의 면적을 특정 시점까지 적분하여 RMST를 계산하고, 두 그룹 간의 RMST 차이를 노출(exposure)의 인과효과로 사용한다. 이에 반해, 관찰 연구에서는 비무작위 노출 할당으로 인한 교란 편향 때문에 표준적인 Kaplan-Meier 추정량을 RMST 계산에 직접 사용할 수 없다. 이러한 교란 편향을 보정한 RMST의 차이를 계산하는 방법으로는 직접 RMST 회귀, 역 확률 가중치 (inverse probability weighting), G-computation 등이 있다. 모든 모델이 올바르게 지정된 복수의 시뮬레이션을 통해 우리는 고려한 모든 방법이 비편향추정값(unbiased estimate)을 제공하고 백분위수 붓스트랩(percentile bootstrap) 신뢰구간이 명목표함확률(nominal coverage probability)을 달성함을 확인했다.

관찰 연구에서 교란에 대해 보정된 RMST의 차이를 평가하기 위한 몇 가지 방법이 개발되었지만, 측정되지 않은 교란의 민감도 분석(sensitivity analysis)에 대한 연구는 아직까지 없다. 따라서, 우리는 보정된 RMST의 차이를 평가하기 위해 측정되지 않은 교란을 고려하는 새로운 민감도 분석 방법을 제안한다. 사용자 지정 민감도 매개변수가 주어지면, 편향 조정된 RMST 차이(bias-adjusted difference in RMST)의 추정치에 대한 민감도 범위(sensitivity range)와 신뢰구간을 얻을 수 있다. 민감도 범위와 신뢰구간을 얻기 위해서는 복잡한 최적화 문제를 풀어야 하지만, 특별한 경우를 제외하고는 분석적 해가 존재하지 않는다. 최적화 문제의 해를 L-BFGS-B와 같은 최적화 알고리즘을 사용하여 구할 수 있지만 (직접 최적화 방법), 이 경우 해를 구하기 위해 상당한 계산 시간이 소요된다.

따라서, 우리는 편향과 계산시간 모두에서 직접 최적화 방법보다 열등하지 않은 근사 최적화 방법을 제안했고, 집약적인 Monte Carlo 시뮬레이션을 통해 제안한 근사 최적화 방법이 실용적인 대안책이 될 수 있음을 보였다. 민감도 분석을 실제 문제에 적용할 때, 우리는 중도절단률(censoring rate)이 0.7 미만인 경우 근사 최적화 방법을 사용하고, 중도절단률이 0.7 이상인 경우 직접 최적화 방법을 사용하는 것을 권고한다.

주요어: 제한된 평균 생존 시간, 인과 추론, 생존 분석, 관찰 연구, 민감도 분석, 측정되지 않은 교란

학 번: 2020-30064