Estimating Treatment Effects and Identifying Optimal Treatment Regimes to Prolong Patient Survival

by

Jincheng Shen

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biostatistics)
in The University of Michigan
2014

Doctoral Committee:

Professor Jeremy M.G. Taylor, Co-Chair
Assistant Professor Lu Wang, Co-Chair
Professor John D. Kalbfleisch
Associate Professor Chuanwu Xi

To Yiqin, my parents and my dogs

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere appreciation to my co-advisors Jeremy and Lu. Thank you very much for being so patient and caring. It is my fortune and honor to be your student. I learned so much from your knowledge, expertise and insightful ideas. This thesis could not be completed without your guidance and support. Thank you for bringing me to the area of cancer research and causal inference, which I have enjoyed very much and learned a lot from.

My deep gratitude also goes to Dr. John D. Kalbfleisch and Dr. Chuanwu Xi for providing valuable comments on my thesis and being on the committee. I always feel inspired and encouraged after discussing the research ideas with them. Especially, thank Dr. John D. Kalbfleisch for introducing me to the amazing field of survival analysis from his course BIOSTAT 675, I have enjoyed very much working in this field. Special thanks are also due to Dr. Matthew J. Schipper who has provided me a great opportunity to learn to work with physicians in the field of radiation oncology and it has been a very pleasant experience working as a research assistant under his supervision.

Special thank you also to my officemates, past and present, for putting up with me: Yong-Seok Park, Jared Foster, Edward Kennedy, Phil Boonstra, Oliver Lee, Laura Fernandes, Anna Conlon, John Rice, Daniel Muenz and Wenting Chen. I would also like to thank all faculty and staff members in the Department of Biostatistics at the University of Michigan. I feel extremely lucky and honored to be a graduate student here.

This thesis is dedicated to my inspiring family. Their boundless love and consistent encouragement grants me the power to move forward with my life. Thanks to Mr. Bobe for always being supportive and keep me company for all these years. Finally, to my girlfriend, Yiqin, thank you for always being by my side through the roughness, you have inspired me to become a better man. I definitely could not have gone this far without you.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

ABSTRACT

Estimating Treatment Effects and Identifying Optimal Treatment Regimes to Prolong Patient Survival

by

Jincheng Shen

Co-Chairs: Professor Jeremy M. G. Taylor and Assistant Professor Lu Wang

Motivated by an observational prostate cancer recurrence study, we investigate the effect of treatment on survival outcome. For studies such as these, it is important to properly handle the confounding effects, especially from longitudinal covariates. In addition, baseline covariates may also reflect the heterogeneity of the population in responding to the treatment. It is possible to recognize these differences and customize the treatment strategy accordingly.

In the first project, we formulate a generalized accelerated failure time (AFT) model to describe the treatment effect and the model includes a longitudinal covariate as a functional predictor, whose coefficient is a time-varying nonparametric function. We propose a spline-based sieve estimation for the time-varying coefficient of the functional predictor, and maximize the likelihood in the sieve space where we approximate the functional predictor and nonparametric coefficient using a B-spline basis. Under certain regularity conditions, the proposed estimator is consistent and semi-parametrically efficient. Simulation studies are conducted to demonstrate the

potential use of the proposed method to estimate the effect of a treatment assigned at baseline under various time-dependent confounding mechanisms.

We further consider the interaction between treatment and other covariates, and explore the heterogeneity of the treatment effect and approaches to personalize the treatment assignment to optimize the survival outcome. In the second project, using the causal inference framework, we consider the counterfactual outcome as if every patient follows a given treatment regimen and develop a method to identify the optimal dynamic treatment regime from observational longitudinal data. We propose to use Random Forest to model the regime adherence of each subject, and use inverse probability weights to adjust for non-adherence to obtain the regime specific survival distribution. The proposed estimator is consistent, and its finite sample performance is assessed through simulation studies.

In the third project, we consider the optimal treatment regime based on available baseline covariates in a target population. The available data to estimate the optimal regime is an observational study that includes these and some additional baseline covariates. Instead of optimizing over a pre-defined class of regimes as in Project 2, we consider a more general class of candidate regimes through flexible models of the outcomes. We propose to use Random Survival Forest plus an inverse probability weighted bootstrap to estimate the causal outcomes while marginalizing over the covariates that may not be of primary interest. By comparing the restricted mean survival times, the optimal regime can be estimated for the target population. We evaluate the performance of the proposed method through simulation studies, and demonstrate its advantage compared to some traditional approaches.

CHAPTER I

Introduction

Prostate cancer is the most commonly diagnosed cancers among American men (*Siegel et al.*, 2013). Often, after initial treatment, patients diagnosed with clinically localized prostate cancer are actively monitored for elevated and/or rising levels of prostate-specific antigen (PSA). The typical pattern of PSA after initial radiation therapy is well known and associated with some of the pre-treatment variables (*Proust-Lima et al.*, 2008). It decreases in everyone for about a year and then may never show a subsequent increase; if it does rise, it increases approximately exponentially with time. Figure 1.1 shows a typical trajectory of log(PSA+0.1). Rising values of PSA are indicative of an increased risk for the clinical recurrence of prostate cancer (*Zagars and von Eschenbach*, 1993). In these cases, patients sometimes receive additional new treatment (called salvage therapy) in order to prevent or delay recurrence. Figure 1.2 shows the four types of possible sequences of salvage therapy and recurrence. One such salvage therapy treatment is androgen deprivation therapy (ADT). Although salvage ADT is generally thought to be beneficial in delaying the recurrence of prostate cancer, the magnitude of this benefit is not well quantified.

Complicated confounding is involved between the salvage treatment which is given by indication, and the time to recurrence, which is a survival type of outcome. Elevated

and/or rising PSA levels are a risk factor for recurrence of prostate cancer but are also a predictor of treatment by salvage ADT (SADT), thus PSA and slope of PSA are (time-dependent) confounders in the relation between salvage ADT and prostate cancer recurrence. In general, this type of a relation between a time-dependent confounder and a time-varying treatment is typically present whenever there is "treatment by indication" (*Robins*, 1989a). Correctly estimating the effect of the time-dependent salvage treatment and the impact from the longitudinal covariates is of great importance for clinical practice. Along this line, *Kennedy et al.* (2010) focused on a multi-center prostate cancer recurrence study with prostate cancer patients who were initially treated with external beam radiation therapy (EBRT). Patients came from four cohorts: University of Michigan (Michigan, U.S.A.), Radiation Therapy Oncology Group, Peter MacCallum Cancer Centre (Melbourne, Australia), and William Beaumont Hospital (Michigan, U.S.A.). Patients were closely monitored for possible recurrence. Their PSA values were recorded at regular visits along with other clinical characteristics. Salvage ADT treatment was given to some of the patients to reduce the risk of cancer recurrence. More details of the dataset can be found in Chapter III or *Proust-Lima et al.* (2008); *Kennedy et al.* (2010). *Kennedy et al.* (2010) proposed two methods to estimate the treatment effect conditional on the other covariates, especially the time-dependent PSA. Two-stage method fits the longitudinal model for PSA in the first step and then estimates the treatment effect from a Cox model. The other method, so-called Sequential Stratification, constructs time-dependent stratum for each patient undergoing SADT to mimic a sequence of conditionally randomized SADT assignments. A stratified Cox model is fitted to obtain the estimation of the treatment effect. Taylor et al. (2013) further evaluated the performance of these methods and compared them with marginal structural models (MSMs) under inverse probability weighting (IPW) through series of simulation studies. The relationship between the subject specific treatment effect and the marginal treatment effect was

also investigated.

In Chapter II, we investigate the potential impact of PSA history on the current risk of recurrence. We consider a situation with a time-dependent covariate (such as PSA) and a time-independent covariate of interest, which could be treatment group indicator, while the PSA is considered as a potential predictor. The Cox proportional hazard model is the most frequently used and well-recognized approach in modeling time to event outcomes. The Cox model can incorporate both time-independent and time-dependent covariates. However, its proportional hazard assumption may limit its application in some situations. The accelerated failure time (AFT) model is an appealing alternative that can naturally incorporate the continuous impact of time dependent predictors. Furthermore, the coefficient for the time-independent covariate in an AFT model is quantified in a straightforward way. For example, hormone therapy is often thought of as delaying the prostate cancer recurrence by stretching the time scale with a multiplying factor. If we consider the baseline covariate to be an indicator of assigning patients to either hormone therapy or control group, then the AFT model will provide the desired summary measure for the treatment effect, and it can also handle the entire PSA history in a natural way. Thus we consider an AFT model where the coefficients for the time-dependent covariates may be time-varying. To obtain an unbiased estimator for the coefficients for the baseline covariates (e.g. the treatment effect for hormone therapy), we propose a spline-based efficient estimation procedure, and maximize the likelihood in the sieve space where we approximate the functional predictor and nonparametric coefficient using a B-spline basis. Asymptotic properties of the proposed estimator are developed. We derive the convergence rate for the sieve estimator of the nonparametric coefficient of the functional predictor, and establish $\sqrt{n}$-consistency of the regression coefficients of the parametric part, such as the treatment effect.

However, in practice the therapy may not be given at baseline, and it may be of more interest to know what would be the best time to initiate hormone therapy based on the patient's specific clinical history, so that therapy can be tailored for each patient. For example, guidelines such as "give SADT when PSA first goes above 4 ng/ml if age is less than 75" would be very useful to physicians and health care providers in the management of prostate cancer recurrence. The major challenges to provide such detailed guidelines are (i) it is very expensive to run a randomized trial to cover all possible treatment plans of interest; (ii) in observational studies, it is a problem on how to use the data efficiently as the "optimal treatment regime" would only have been followed by a very small subgroup of the whole dataset; (iii) in order to compare the treatment effect between groups with different treatment regimes in observational data, we need to correctly account for the artificial treatment effect, which arise, for example, because groups of patients with lower PSA levels tends to have delayed recurrences even without treatment.

Thus, in Chapter III, we investigate this question of identifying the optimal dynamic treatment regime from observational data employing techniques from causal inference. We extend the marginal structural model method to estimate the survival distribution for different counterfactual treatment regimes. The challenges come from the weight estimation which are derived from the model for the observed treatment assignment mechanism. These weights are needed at multiple time points for each patient in the marginal structural model methodology. Logistic models which are linear in the covariates $\boldsymbol{X}$ are frequently employed for this purpose (*Hernán et al.*, 2006; *Zhang et al.*, 2013). This may not always provide a satisfactory approximation as the underlying treatment mechanism varies from case to case, and a more sophisticated dependency may exist between the actual treatment assignment and the covariates. We propose to use Random Forests to model the treatment assignment mechanism

and estimate the time-dependent weights. The restricted mean survival time (RMST) is also employed to enable the comparison among various regimes and serves as the objective function in the optimization process.

Another issue in personalizing the treatment regime relates to the modeling of the outcomes. Identifying the optimal regime always involves some maximization/minimization of some summary statistics of the outcome. The optimization thus depends on the model assumption for the outcome, which is equivalent to performing the optimization within a pre-specified class of regimes. It is important that the outcome model matches the truth, or there could be bias for the final regime that is identified. For example, consider a randomized trial with a scalar covariate $X \sim \text{Uniform}(-1, 1)$, and A is the treatment indicator which takes value 1 if the patient is assigned to treatment group and 0 for being assigned to the control group, with $P(A = 1) = 0.5$. Consider a proportional hazard model for the outcome of the following form:

$$\begin{aligned}\lambda(t) &= \lambda_0 \exp\left(ax^2 - ax + \frac{1}{4}a\right) \\ &= \lambda_0 \exp\left\{a\left(x - \frac{1}{2}\right)^2\right\}\end{aligned}$$

with baseline hazard $\lambda_0 = 0.2$. If we do not consider censoring, then the true optimal regime will be $g_0^{\text{opt}}(X) \equiv 0$, i.e. always not giving treatment. However, if we fit the survival time with a Cox model using the following structure:

$$\begin{aligned}\lambda(t) &= \lambda_0(t) \exp\left(\beta_1 x + \beta_2 a + \beta_3 ax\right) \\ &= \lambda_0(t) \exp\left\{\beta_1 x + \left(\beta_2 + \beta_3 x\right) a\right\},\end{aligned}$$

and consider treatment regimes in the form of $\mathcal{G} = \{g(X) = I(\beta_2 + \beta_3 X < 0),\ \beta_2, \beta_3 \in \mathbb{R}\}$, for $-1 \leq X \leq 1$. It is easy to see that $g^{\text{opt}} \in \mathcal{G}$, however, as we generate data

for a cohort of $n = 1000$ patients with 1000 replicates, we get the average estimates from Cox model as $\hat{\beta}_2 = 0.53$ and $\hat{\beta}_3 = -1.00$, the estimated optimal regime would be $\hat{g}^{\text{opt}}(X) = I(0.53 - 1.00X < 0) = I(X > 0.53)$, which is clearly different than g^{opt}. Thus, even if the true optimal regime is in the considered regime space, the algorithm may pick a different one when the regime space under consideration mismatches the structure of the true outcome model. One possible solution here is to assume a model with enough flexibility, and then apply machine learning techniques to provide promising alternatives for this purpose. In Chapter IV, we propose to employ the Random Survival Forest (RSF) method to tackle this issue. As the treatment model is not the major concern in this case, we consider the situation where all the covariates $\boldsymbol{W} = (\boldsymbol{X}^T, \boldsymbol{Z}^T)$ are measured at baseline (pretreatment), and the treatment A is also assigned at baseline. Again, we consider the scenario of observational data, where $\boldsymbol{W}$ is associated with both the treatment and the survival outcome. Although $\boldsymbol{X}$ and $\boldsymbol{Z}$ may be highly correlated, the availability of $\boldsymbol{Z}$ is not guaranteed in the general population due to various reasons, so the optimal regime of interest need to be a series of generalizable treatment rules and need to be based on $\boldsymbol{X}$ only. We propose to use inverse probability weighting technique to account for the possible selection bias. By conducting a weighted bootstrap with Random Survival Forest, we are able to correctly model the counterfactual survival outcomes using observational data and identify the optimal treatment regime as the one which maximizes the restricted mean counterfactual survival time. We conclude the dissertation in Chapter V with a discussion of future work and related open problems.

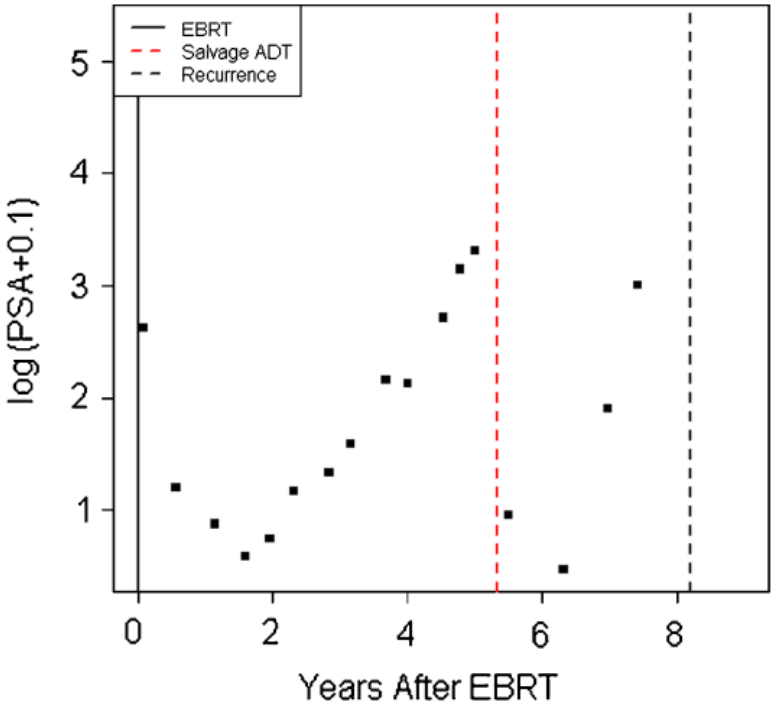


Figure 1.1: Typical log(PSA) Patterns

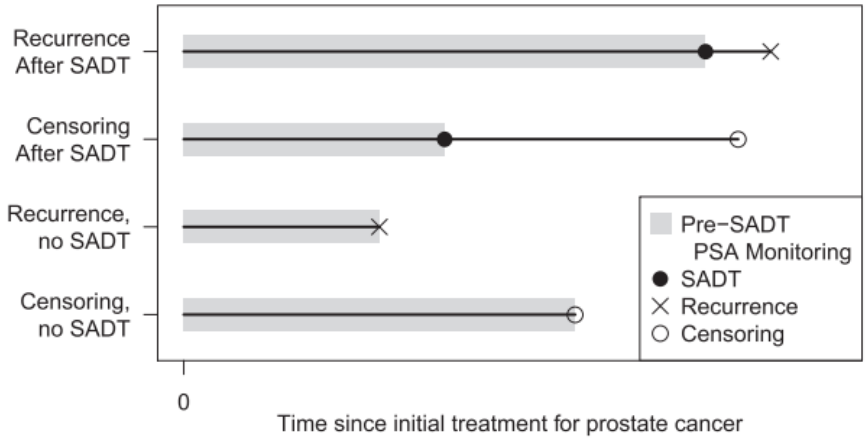


Figure 1.2: Structure of Longitudinal, Treatment, and Recurrence Data.

CHAPTER II

Efficient Estimation of the Treatment Effect Under the Impact of the Entire PSA History

The accelerated failure time (AFT) model is an approach that can be used to assess the continuous impact of time-dependent predictors in time-to-event analyses, where the effect of a baseline covariate can be quantified in a straightforward way. For example, the AFT model with an indicator of hormone therapy provides the desired summary measure for the treatment effect, which is often thought of as delaying prostate cancer recurrence by stretching the time scale with a multiplying factor. In this chapter, we consider an AFT model with both time-dependent covariates $\boldsymbol{X}(t)$ and time-independent covariates $\boldsymbol{Z}$. This model can handle the entire history of a longitudinal variable PSA as a functional predictor, and the coefficients for $\boldsymbol{X}(t)$ in our model are allowed to vary with time as well. We propose a spline-based sieve estimation of the time-varying coefficient of the functional predictor, and maximize the likelihood in the sieve space where we approximate the functional predictor and nonparametric coefficient using B-spline basis. The proposed estimator is proved to be asymptotically consistent, and the regression coefficients of the parametric part, such as treatment effect, is shown to be $\sqrt{n}$-consistent.

2.1 Introduction

In cancer research, it is often of primary interest to understand the association between a risk factor or a certain treatment with the patients' survival outcome. In many situations one big challenge is how to correctly adjust for the time-varying confounding variables. For example, in prostate cancer, PSA is a very important time-varying biomarker for disease progression, and is believed to be highly associated with the hazard of cancer recurrence. One would expect such a biomarker to have a continuous impact over the course of disease development, because as the disease progress to a different point, PSA may indicate a different level of cancer recurrence risk. In fact, the whole history of PSA plays an important prognostic role besides the point-wise health status of a prostate cancer patient. For instance, it is likely that two patients with different PSA trajectories respond very differently to the same therapy, even though their current PSA values are the same. Therefore, how to adjust for the longitudinal confounding effect of the whole PSA trajectory or PSA changes while evaluating the association of other baseline covariates, such as the treatment effect of hormone therapy on patient recurrence risk becomes complicated. In this chapter, we are proposing a semiparametric modeling approach to allow for more flexibility.

Statistical tools have been developed to study the cumulative effects of longitudinal factors, especially for studies with continuous or categorical outcomes (*Xia and Tong*, 2006; *Kong*, 2010). When the outcome is a time to event, the proportional hazards model (*Cox*, 1972, 1975) is the most frequently used approach in practice, due to the availability of efficient estimation procedures that are implemented in all statistical software packages. However, in order to obtain valid statistical inferences, these models require the proportional hazard assumption, which is quite strong for many practical applications and can be easily violated when there are continuous im-

pacts from time-varying predictors. Although the proportional hazard model can be extended by incorporating time-varying coefficients (*Murphy and Sen*, 1991), the statistical inference involves histogram-sieve estimation on the partial likelihood. *Chen and Zhou* (2007) proposed a local partial likelihood approach to directly estimate the relative risk function. The time-varying proportional hazard ratio assumption may not have a straightforward interpretation. On the other hand, the accelerated failure time (AFT) model (*Kalbfleisch and Prentice*, 2011), which directly relates the monotonically transformed survival time to the covariates of interests, is "in many ways more appealing because of its quite direct physical interpretation" (*Reid*, 1994) and thus frequently serves as an attractive alternative to the Cox proportional hazards model. Suppose the covariate vector is $\boldsymbol{W}$, then a common AFT model for the survival outcome T takes the log-linear form

$$\log T = -\boldsymbol{W}^T\boldsymbol{\gamma} + \epsilon \tag{2.1}$$

where ϵ is measurement error, which is assumed to have a particular distribution independent of $\boldsymbol{W}$. As in the cancer example, where T is the recurrence free survival time and one dimensional W is the treatment, then γ, the treatment effect estimated from model (2.1), would have a more straightforward interpretation as a multiplicative factor to accelerate or decelerate the disease course to cancer recurrence.

There has been a plethora of literature on both theoretical development and applications of AFT models. Log-normal model is one of the most popular parametric AFT models and has been used in a wide range of applications (*Royston*, 2001; *Longford*, 2009; *Köhler and Kowalski*, 2012; *Chapman et al.*, 2013). To relax the fully parametric restriction, semi-parametric AFT models has also been intensively studied. *Prentice* (1978) proposed the rank estimator for the baseline hazard function

based on the well-known weighted log-rank statistics, and soon after that, *Buckley and James* (1979) introduced a modified version of least-squares estimator to account for the censoring. Following this initial work, extensive theoretical development and extensions have been done (*Ritov*, 1990; *Tsiatis*, 1990; *Lai and Ying*, 1991; *Ying*, 1993; *Jin et al.*, 2003, 2006). *Ritov and Wellner* (1988) derived the semiparametric efficient score functions for the slope parameters in the linear regression model, which involves the derivative of the baseline hazard function (i.e., density of ϵ). *Zhang et al.* (2009) proposed to use monotone B-spline when there is interval censoring, and *Ding and Nan* (2011) employed a spline-based method when the parameters are bundled together. *Zeng and Lin* (2007) considered a more general setting, where $\boldsymbol{W}$ could be time-dependent, and proposed a kernel smoothing based efficient estimation procedure, but they restrict the effect of $\boldsymbol{W}$ to be constant, and the kernel smoothing procedure may not be straightforward to implement in biomedical studies. In practice, it may be more realistic to assume that the impact of a time-dependent covariate also varies with time. In this chapter, we consider a more general setting where $\boldsymbol{W}$ is a vector which contains both time-dependent covariates $\boldsymbol{X}(t)$ and time-independent covariates $\boldsymbol{Z}$, and the coefficients for $\boldsymbol{X}(t)$ could also vary with time. Specifically, we consider the following model

$$e^{\epsilon} = \int_0^T \exp\{\boldsymbol{X}(u)\boldsymbol{\beta}(u) + \boldsymbol{Z}\boldsymbol{\gamma}\}du \tag{2.2}$$

where ϵ is the error term similar as in model (2.1). Unlike *Zeng and Lin* (2007), our model specification in (2.2) allows the impact of PSA to vary over time, which represent a more realistic situation. For example, it is possible that the patient's PSA history within the last year would have more impact on his current health status than his PSA levels three years ago. In model (2.2), we also consider a baseline covariate, for example, the hormone therapy which is assigned to the patient at time

zero. We assume a parametric baseline hazard using a lognormal distribution. We propose a new approach by directly maximizing the log likelihood function in a sieve space, where the time-varying coefficients are approximated by B-splines. A resampling procedure is developed for the variance estimation. In addition, we investigate the asymptotic consistency and efficiency of the proposed estimators using empirical process theory.

The rest of the chapter is organized as follows. In Section 2.2, we introduce the notation and specify the structure for our AFT model. In Section 2.3 we propose the sieve estimator for both the time-varying and fixed coefficients. We investigate the asymptotic properties of our new estimators in Section 2.4, followed by comprehensive simulation studies presented in Section 2.5. Finally, we conclude with a brief discussion and propose some potential extensions in Section 2.6. Technical details are included in Section 2.7.

2.2 Notation and Model Specification

We consider a cohort of n patients. For patient i, let T_i denote the failure time and C_i denote the censoring time for subject i, where $i = 1, \cdots, n$. Define $Y_i = \min(T_i, C_i)$ and $\Delta_i = I(T_i \leq C_i)$. We assume that C_i is independent of T_i, conditional on covariates $\boldsymbol{W}_i = \{\boldsymbol{X}_i(\cdot), \boldsymbol{Z}_i\}$, To make the presentation easier, without loss of generality, we only consider the case when both $\boldsymbol{X}_i(\cdot)$ and $\boldsymbol{Z}_i$ are one-dimensional, throughout denoted as $X_i(\cdot)$ and Z_i respectively. Then the likelihood for a single observation of $(Y, \Delta, \boldsymbol{W})$ of model (2.2) is

$$L_{Y,\Delta,\boldsymbol{W}}(y, \delta, \boldsymbol{w}) = \lambda_{T|\boldsymbol{W}}(y))^{\delta} \exp\{-\Lambda_{T|\boldsymbol{W}}(y)\} H(y, \delta, \boldsymbol{w}),$$

where $\lambda_{T|\boldsymbol{W}}(t)$ and $\Lambda_{T|\boldsymbol{W}}(t)$ denote the conditional hazard and cumulative hazard functions of T given and $\boldsymbol{W}$. $H(y, \delta, \boldsymbol{w})$ includes both the conditional distribution of C given $\boldsymbol{W} = (X(\cdot), Z)$ and the marginal distribution of X and Z. We assume that H is free of λ, β and γ, so we ignore H from the likelihood function. We further assume a lognormal distribution with $\epsilon_i \sim N(0, \sigma^2)$. The log-likelihood for single observation $(Y, \Delta, \boldsymbol{W})$ becomes:

$$\begin{aligned} l(\beta, \gamma, \sigma) =& \Delta\{\beta(Y)X(Y)\} - \Delta R(Y;\beta) - \Delta \log \sigma \\ &+ \Delta \log \phi\left(\frac{R(Y;\beta) + \gamma Z}{\sigma}\right) + (1-\Delta)\log \Phi\left(-\frac{R(Y;\beta) + \gamma Z}{\sigma}\right), \end{aligned}$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the cdf and pdf of the standard normal distribution, and $R(Y;\beta) = \log \int_0^Y e^{\beta(u)X(u)} du$. Notice that the way we set up model (2.2) makes the current hazard depends on the current value of $X(\cdot)$, but also the history of $X(\cdot)$ up to the current time, i.e.

$$\begin{aligned} \lambda_{T|\boldsymbol{W}}(t|\boldsymbol{w}) &= \lambda_{T|\overline{X}(t),Z}(t|\overline{x}(t), z) \\ &= \phi\left(\frac{r(t;\beta) + \gamma z}{\sigma}\right) \Big/ \left\{\sigma e^{r(t;\beta)+\gamma z} \cdot \Phi\left(\frac{r(t;\beta) + \gamma z}{\sigma}\right)\right\}, \end{aligned}$$

for $t \in (0, \infty)$, where $\overline{x}(t) = \{x(u) : 0 < u \leq t\}$ is the observed history for $X(\cdot)$ up to time t, $r(t;\beta) = \log \int_0^t e^{\beta(u)x(u)} du$ is the realization of $R(t;\beta)$ which depends on $\overline{x}(t)$.

2.3 The Estimating Procedure

2.3.1 Spline Based Estimation

In order to perform estimation, we propose to use the spline-base sieve maximum likelihood estimation for Model (2.2). Spline techniques have been extensively used as an effective tool for dimension reduction in nonparametric estimation. *Stone* (1985, 1986) has proved in theory that a smooth unknown function can be well approximated

using a spline. Some further convergence results of spline-based sieve estimates have been developed by *Shen and Wong* (1994). Under the regularity conditions (C1)-(C3) stated in Section 2.4, let $(0, a]$ be an interval of interest, where $0 < a < \infty$. Let $0 < t_1 < \cdots < t_{K_n} < t_{K_n+1} = b$ be a partition of $(0, a]$ with K_n subintervals $I_0 = (0, t_1)$ and $I_j = [t_j, t_{j+1})$, $j = 1, \cdots, K_n - 1$ and $I_{K_n} = [t_{K_n}, t_{K_n+1}]$, with $K_n = O(n^\nu)$ and $\max_{1 \le j \le K_n+1} |t_j - t_{j-1}| = O(n^{-\nu})$ for some positive number $\nu \in (0, 1/2)$. Denote the set of partition points by $T_{K_n} = \{t_1, \cdots, t_{K_n}\}$. Following *Schumaker* (1981), Definition 4.1, let $\mathcal{S}_n(T_{K_n}, K_n, p)$ be the space of polynomial splines of order $p \ge 1$ consisting of function s satisfying (i) the restriction of s to I_j is a polynomial of order p (or equivalently, of degree $p - 1$) for $p \le K_n$; (ii) for $p \ge 2$ and $0 \le p' \le p - 2$, s is p' times continuously differentiable on $[a, b]$. According to Schumaker (*Schumaker* (1981), Corollary 4.10), there exists a set of B-spline basis functions $\{B_j, 1 \le j \le q_n\}$ with $q_n = K_n + p$ such that for any $s \in \mathcal{S}_n(T_{K_n}, K_n, p)$, we can write

$$s(t) = \sum_{j=1}^{q_n} \omega_j B_j(t), \tag{2.3}$$

where following *Shen and Wong* (1994) we require $\max_{j=1,\cdots,q_n} |\omega_j| \le c_n$ and c_n is allowed to grow with n slowly enough. Let $\boldsymbol{\omega} = \{\omega_1, \cdots, \omega_{q_n}\}$ be the collection of all the coefficients in the representation (2.3). Under suitable smoothness assumptions, the true parameter $\beta_0(\cdot)$ can be well approximated by some function in $\mathcal{S}_n(T_{K_n}, K_n, p)$.

2.3.2 Likelihood Approximation

Our goal is to seek a member of $\mathcal{S}_n(T_{K_n}, K_n, p)$ together with a value of $\gamma \in \Gamma$ and a value of $\sigma \in \Sigma$ that maximizes the log likelihood function. Specifically, let $(\hat{\gamma}_n, \hat{\sigma}_n)$ be the value that maximizes

$$l_n(\boldsymbol{\omega}, \gamma, \sigma) = n^{-1} \sum_{i=1}^{n} \left[\Delta_i \left\{ \sum_{j=1}^{q_n} \omega_j B_j(Y_i) X_i(Y_i) + \gamma Z_i \right\} + \Delta_i \log \lambda(e^{Q_i^*(\boldsymbol{\omega}, \gamma)}) - \Lambda(e^{Q_i^*(\boldsymbol{\omega}, \gamma)}) \right],$$

where

$$Q_i^*(\boldsymbol{\omega}, \gamma) = \log \int_0^{Y_i} \exp \left\{ \sum_{j=1}^{q_n} \omega_j B_j(Y_i) X_i(Y_i) + \gamma Z_i \right\} du,$$

and σ is a parameter in $\lambda(\cdot)$ and $\Lambda(\cdot)$. This then becomes a sieve maximum likelihood estimation problem where the unknown function in the log likelihood is approximated by a linear span of some known basis functions to form a sieve log likelihood. Following *Geman and Hwang* (1982), we can maximize the sieve log likelihood with respect to the unknown coefficients in the linear span to obtain a sieve maximum likelihood estimator. This will significantly reduces the dimensionality of the maximization since the number of basis functions needed to approximate the unknown function grows at a much slower rate as the sample size increases. Here the maximization can be solved by standard methods for the accelerated failure time model with parametric baseline hazards. We use R function *aftreg*() from *eha* package to fit for $(\hat{\gamma}_n, \hat{\sigma}_n)$ and $\hat{\boldsymbol{\omega}}$. Then the time-varying coefficient can be estimated as $\hat{\beta}(t) = \sum_{j=1}^{q_n} \hat{\omega}_j B_j(t)$.

2.4 Asymptotic Theory

In this section, we investigate the asymptotic properties of the proposed estimator. To obtain the consistency of $\hat{\gamma}_n$ and $\hat{\beta}(t)$, and the asymptotic normality result for $\hat{\gamma}_n$, we need to impose the following regularity conditions:

(C1) $X_i(t)$ is bounded for $t \in R^+$, i.e. $\forall t$, $\exists C$ s.t. $\mathbb{P}(\|X_i(t)\| \leq C) = 1$;

(C2) $\sigma \in \Sigma$, where Σ is a compact subset of R^+. We assume there exists a small $\kappa > 0$, s.t. $\sigma > \kappa$;

(C3) $\gamma \in \Gamma$, where Γ is a compact subset of R^-. We assume there exist a large $M > 0$, s.t. $-M < \gamma \leq 0$;

(C4) $\forall t$, $\exists C$ s.t. $\mathbb{P}(0 \leq \|\beta(t) X_i(t)\| \leq C) = 1$;

(C5) Let $\mathcal{B}$ denote the collection of bounded functions β on $(0, a]$ with bounded derivatives $\beta^{(j)}$, $j = 1, \cdots, k$, and we assume that the kth derivative $\beta^{(k)}$ satisfies the

following Lipschitz continuity condition:

$$\left|\beta^{(k)}(s) - \beta^{(k)}(t)\right| \leq L\,|s-t|^m \text{ for } s,t \in (0,a]$$

where k is a positive integer and $m \in (0,1]$ such that $p = k + m \geq 3$, and $L < \infty$ is an unknown constant. The true weight $\beta_0 \in \mathcal{B}^p$.

Condition (C1)-(C4) are common regularity assumptions that have been imposed in the literature for similar problems (*Zhang et al.*, 2009), while conditions (C5) is assumed here to provide desirable controls of the spline approximation error rates of the first and second derivatives of β_0. We can see in the following theorems that these conditions are sufficient to guarantee our estimator $\hat{\boldsymbol{\theta}}_n = (\hat{\beta}_n, \hat{\gamma}_n, \hat{\sigma}_n)$ to be asymptotic consistent.

Theorem 2.1. Let $K_n = O(n^\nu)$, where ν satisfies the restriction $\frac{1}{2(1+p)} < \nu < \frac{1}{2p}$ with p being the smoothness parameter defined in Condition (C5). Suppose Condition (C1)-(C5) hold and the failure time T follows Model (2.2), then

$$d(\hat{\boldsymbol{\theta}}_n, \boldsymbol{\theta}_0) = O_p\left[n^{-\min\{p\nu,(1-\nu)/2\}}\right],$$

where $\boldsymbol{\theta} = \{\beta(\cdot), \gamma, \sigma\}$ is the vector of parameters, $d(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \{||\beta_2 - \beta_1||_Y^2 + |\gamma_2 - \gamma_1|^2 + |\sigma_2 - \sigma_1|^2\}^{1/2}$ and we define $\|\beta_2 - \beta_1\|_Y = \int_0^Y |\beta_2(u) - \beta_1(u)|\,du$.

Theorem 2.1 establishes the consistency of our proposed estimators. But $\hat{\beta}_n(\cdot)$ converges to the true $\beta_0(\cdot)$ in a slower rate compared to $\hat{\gamma}_n$ and $\hat{\sigma}_n$. With a non-parametric estimator $\hat{\beta}_n(\cdot)$, one would be interested to see whether that will affect the parametric convergence rate of $\hat{\gamma}_n$. In Theorem 2.2, we show the asymptotic normality for our

proposed estimator $\hat{\gamma}_n$, and also develop the estimation efficiency bound.

Theorem 2.2. Under Model (2.2), the efficient score function of estimating γ is:

$$\begin{aligned}\dot{l}^*_\gamma(\beta,\gamma,\sigma) =& \frac{\int_0^Y \{Z - h^*(u)X(u)\}e^{\beta(u)X(u)}du}{\int_0^Y e^{\beta(u)X(u)}du} \\ & \times \frac{1}{\sigma}\left\{\Delta\frac{\phi^{(1)}(\{R(Y;\beta_0)+\gamma Z\}/\sigma)}{\phi(\{R(Y;\beta_0)+\gamma Z\}/\sigma)} - (1-\Delta)\frac{\phi(\{R(Y;\beta_0)+\gamma Z\}/\sigma)}{1-\Phi(\{R(Y;\beta_0)+\gamma Z\}/\sigma)}\right\}, \\ \text{with } h^*(u) =& \frac{E_{\widetilde{Y},\widetilde{\Delta},\widetilde{\boldsymbol{W}}}\{\widetilde{Z}\widetilde{X}(\widetilde{Y})|\widetilde{R} = \log\int_0^t e^{\beta(u)X_(u)}du, \widetilde{\Delta}=1\}}{E_{\widetilde{Y},\widetilde{\Delta},\widetilde{\boldsymbol{W}}}\{\widetilde{X}^2(\widetilde{Y})|\widetilde{R} = \log\int_0^t e^{\beta(u)X_(u)}du, \widetilde{\Delta}=1\}}\end{aligned}$$

Notice here $\phi^{(1)}(\cdot)$ is the first derivative of $\phi(\cdot)$. Suppose that the conditions in Theorem 2.1 hold and $I(\gamma_0) = E\left\{l^*_{\gamma_0}(Y,\Delta,Z)^{\otimes 2}\right\}$ is nonsingular, we have

$$\sqrt{n}(\hat{\gamma}_n - \gamma_0) \xrightarrow{d} N(0,\ I^{-1}(\gamma_0))$$

Theorem 2.2 shows that as $n \to \infty$, if we properly control the increasing rate of K_n, our proposed estimator for γ_0, $\hat{\gamma}_n$, can achieve the $\sqrt{n}$-consistency. In fact, the proposed estimating procedure can provide efficient estimation of γ_0, which means the asymptotic variance of $\hat{\gamma}_n$ achieves the efficiency bound under Model (2.2), $I^{-1}(\gamma_0) = \left[E\left\{l^*_{\gamma_0}(Y,\Delta,\boldsymbol{W})^{\otimes 2}\right\}\right]^{-1}$. More detials of the proof of these theorems are given in the Appendix, as well as the development of the efficient score function to estimate γ.

2.5 Simulation Study

Simulation studies are carried out to evaluate the finite sample performance of the proposed method. Data are generated as follows: for each subject, we independently

generate $(Y_i, \Delta_i, X_i(\cdot), Z_i)$, where the treatment-free longitudinal covariate

$$X_i(t) = (\alpha_0 + a_{i0}) + (\alpha_1 + a_{i1})t, \tag{2.4}$$

(α_0, α_1) are fixed effect parameters, and (a_{i0}, a_{i1}) are subject-specific random effects. We assume the longitudinal covariate (e.g.,PSA in the case of prostate cancer) is linear over time, and the random effects $(a_{i0}, a_{i1}) \sim \text{MVN}(0, \Sigma)$.

The time-independent covariate Z_i is generated from $Z_i \sim \text{Bernoulli}(p)$. Given the trajectory of $X_i(t)$ and Z_i, we generate the survival time according to a log-normal distribution

$$e^{\epsilon_i} = \int_0^{T_i} \exp\left[\beta(u)X_i(u) + \gamma Z_i\right] du, \tag{2.5}$$

where $\epsilon_i \sim N(0,\, \sigma^2)$, γ is the treatment effect, the true values of these parameters are as following:

$$\alpha_0 = -0.5, \alpha_1 = -0.3, \Sigma = \begin{pmatrix} 1.0 & 0.5 \\ 0.5 & 0.25 \end{pmatrix}, \gamma = -2.0.$$

$\beta(t)$ is the weight for $X_i(t)$, for $t \in [0, K]$, where K is the longest follow-up time under investigation and we consider three different scenarios with different $\beta(t)$ functions:
(I) $\beta(t) = 0.6 \cdot g(t/K)$ where $g(x)$ is the density function of Beta(3,3) distribution, i.e. $g \sim \text{Beta}(3, 3)$. This reflects the situation where the weight $\beta(t)$ is symmetric over $t = K/2$ and the $X(t)$'s at the middle time range (round $t = K/2$) has the biggest impact. Here, we choose $K = 35, 15$ to generate the case of 15% and 25% censoring, respectively.
(II) $\beta(t) = 0.6 \cdot g(t/K)$ where $g(x)$ is the density function of Beta(2,4) distribution, i.e. $g \sim \text{Beta}(2, 4)$. This is a situation where the $X(t)$'s at earlier time have bigger impact

to the survival outcome then the ones at later time. Again we choose $K = 35, 15$ respectively to generate 15% and 25% censoring.

(III) $\beta(t) = 0.4 \cdot g(t/K)$ where we choose $g(x) = \sin\{(2x+1)\pi\}$. This is a situation where the impact of $X(t)$'s are negative (reducing the hazard) at earlier time and positive (increasing the hazard) at later time. Here we choose $K = 8, 7$ respectively to generate 15% and 25% censoring.

We only consider administrative censoring at $C = K$, then the observed survival outcome can be generated as $Y_i = T_i \wedge C$ and $\Delta_i = I(T_i > C)$. In each scenario, K is chosen such that about 15% censoring will be generated.

To set up the spline, we need to choose a set of knots. These can be viewed as parameters that have to be estimated according to a goodness-of-fit criterion, e.g. the Akaike information criterion (AIC) (*McCullagh and Nelder*, 1989). Several methods have been described in the literature and are in general referred to as adaptive knot selection (*Hastie and Tibshirani*, 1990; *Friedman*, 1991). Here we use a simple approach to choose knots at the corresponding quantiles of the covariate. Three different numbers of interior knots for the B-splines are tried, which are 2, 3 and 4. The results are quite similar and we present the results for the case with 2 interior knots.

For all scenarios, we simulate 150 datasets each with 500 subjects, and we use bootstrap with $B = 100$ for each dataset to estimate the variances. We choose to use cubic B-splines to approximate the log hazard function. From Table 2.1, which summarizes the results on the estimation of the treatment effect, we can see that the estimates for γ from all scenarios are close to the truth ($\gamma_0 = -2.0$), while the situations with lower censoring rates tend to have smaller bias. And the coverage rates are close to 95%. As the coverage rates are calculated from bootstrap resampling, it is possible

that larger B for bootstrap could further improve the coverage rates (to be more close to 95%). From Figure 2.1 and Figure 2.2, we can also see that the estimated $\hat{\beta}(t)$'s (solid red lines) are close to the true curves (dash line in blue), and the true curves are within the 95% confidence intervals calculated from the bootstrap sampling (dotted lines in green).

Table 2.1: The estimation for the time-independent coefficient γ: the estimation $\hat{\gamma}_n$ under different setting of true $\beta_0(t)$ are listed here (true $\gamma_0 = -2.0$). For Scenario I, the true $\beta_0(t)$ has a symmetric shape ($\sim Beta(3,3)$) over the time range considered, in Scenario II $\beta_0(t)$ has the shape of $Beta(2,4)$, and in Scenario III, $\beta_0(t)$ has the shape of a sin function ($\sim \sin(2\pi t/K)$). The mean estimation is average over the 150 replicates, while the empirical standard error is the estimate of the variance from the bootstrap procedure ($B = 100$), averaged over the 150 replicates. The mean bias is the averaged bias for the point estimation, and the coverage rate is also calculated from the bootstrap results in each replicate and average over the 150 rounds.

censoring	Shape of $\beta(t)$	Mean Est.	bsSE	Mean Bias	Coverage %
15%	Scenario I	-1.939	0.164	0.061	92.0
	Scenario II	-1.885	0.188	0.115	93.3
	Scenario III	-1.927	0.182	0.073	96.0
25%	Scenario I	-1.926	0.186	0.074	91.0
	Scenario II	-1.867	0.200	0.133	92.5
	Scenario III	-1.925	0.188	0.075	93.6

Note: bsSE is the mean of the bootstrapped standard errors, Coverage is the coverage probability of the 95% confidence interval.

2.6 Discussion

In this Chapter, we propose an AFT model to account for possible cumulative effect of the time-dependent covariates $\boldsymbol{X}(t)$. We assume a rather general form of the model to allow $\boldsymbol{\beta(t)}$, the coefficients of $\boldsymbol{X}(t)$, to be time-varying. too. The estimating procedure is statistically efficient and computationally feasible. In this chapter, we assume a parametric distribution for the error term in model (2.2). Instead of a lognormal distribution, there are several other parametric distribution models that

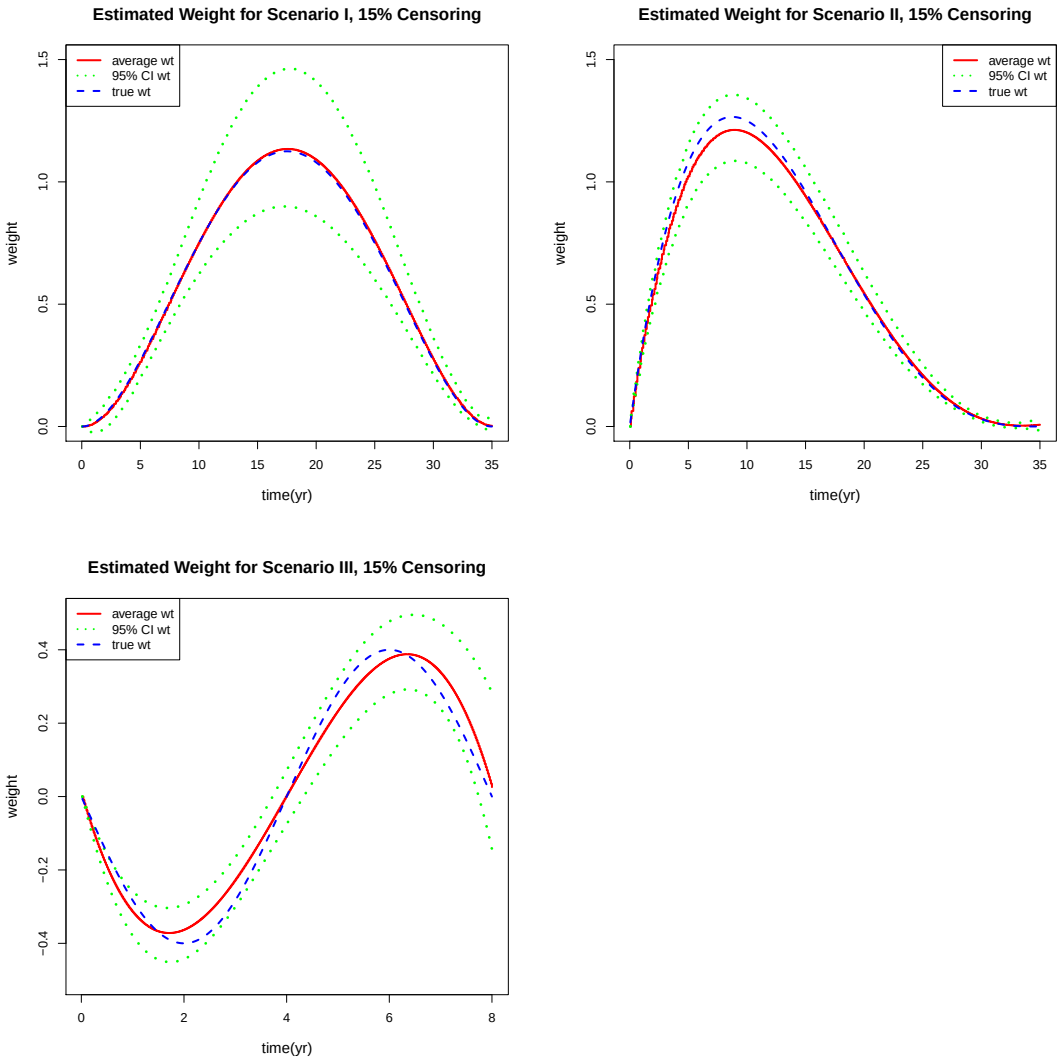


Figure 2.1: Estimated Time-Varying Coefficient for $X(t)$ for 15% censoring cases. Three scenarios are plotted here, where the upper left panel is for Scenario I, where the true $\beta_0(t)$ has a symmetric shape ($\sim Beta(3,3)$) over the time range considered, the upper right panel is for Scenario II, where the earlier $X(t)$ values have more impacts ($\sim Beta(2,4)$), and the lower left panel is for Scenario III, where $X(t)$ values have a positive effect (increase hazard) at beginning and negative effect (reduce hazard) at later time ($\sim \sin(2\pi t/K)$). In all cases, the dashed line in blue is the truth $\beta_0(t)$. the solid line in red is the estimation averaged over 150 datasets, and the dotted lines in green show the averaged the 95% confidence interval which come from the bootstrap procedure.

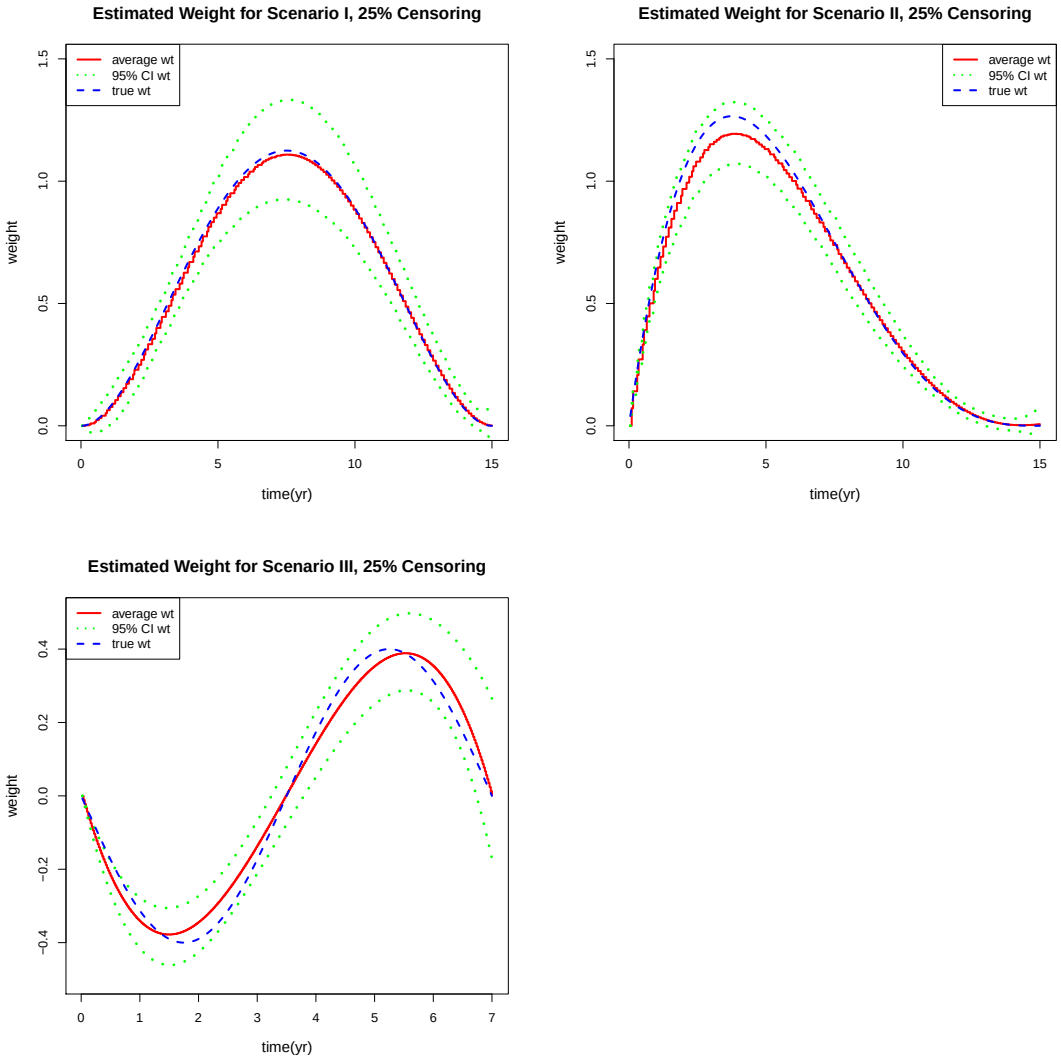


Figure 2.2: Estimated Time-Varying Coefficient for $X(t)$ for 25% censoring cases. Similar as in Figure 2.1 Scenario I - III are plotted. In all cases, the dashed line in blue is the truth $\beta_0(t)$. the solid line in red is the estimation averaged over 150 datasets, and the dotted lines in green show the averaged the 95% confidence interval which come from the bootstrap procedure.

are frequently used in survival analysis. For example, under certain circumstances, a Weibull distribution or log-logistic model are preferred (*Kalbfleisch and Prentice*, 2011). It is straightforward to extend the proposed estimation procedure to models with other parametric distribution assumptions. Furthermore, the baseline cumulative hazard function $\Lambda_0(t)$ can also be estimated non-parametrically. However, it would be challenging both theoretically and numerically to accommodate two non-parametric terms in the same model.

In model (2.2), we have both time-dependent and fixed effects, in practice, it allows the more complicated association between the survival time and time-dependent variable. However, one need to be cautious in interpreting γ and $\beta(t)$. For example, if $X(t)$ is the PSA measured over time, and Z is some treatment decision made at baseline, then γ would have the interpretation of a treatment effect, given the same trajectory of $X(t)$, the treatment will prolong the survival time multiplicatively by a factor of $e^{-\gamma}$. The interpretation for $\beta(t)$ would be even harder, especially for practitioners who are used to deal with the hazard ratios in Cox proportional hazard model. AFT model do provide a more attractive alternative in a lot of cases, for example, in some scenarios, when the outcome is data with censoring, like measurement with lower detect boundary, the outcomes will be censored on the left as they are under certain threshold that the equipment can detect, then the hazard function may not have very good practical meaning, while the interpretation of coefficients in AFT model may be more preferred.

In the model setting, we assume that $\beta(t)$ is only indexed by t. Suppose we are now looking at time τ, it is more realistic to assume that the effect of PSA ($X(t)$, $t \in [0, \tau]$) on the current hazard $\lambda(\tau)$ not only depends on t, but also depends on τ, i.e. $\beta(t, \tau)$. An explanation is that the magnitude of the effect of $X(t)$ on $\lambda(\tau)$

depends on t, and also how close this t is to the point τ. For example, we would expect the PSA value at 2 years to have more impact on the hazard at 3 year than on the hazard at 10 year. This would require splines in two-dimensions and becomes harder to study its properties.

In this chapter, we consider the case where the indicator for treatment is time-independent. In many biomedical studies, it is common to that the treatment initiation time for subject i, denoted as S_i is also random. So another extension of possible interest is that we allow the treatment indicator to be a time-dependent covariate, so we would have two time-dependent covariates, one with time varying coefficient and one with a time-independent coefficient. For this we would consider the following model

$$e^{\epsilon_i} = \int_0^{T_i} \exp\{\beta(u)X_i(u) + \gamma Z_i(u)\}du$$

where $Z_i(u)$ is a step function that jumps from 0 to 1 at $u = S_i$.

2.7 Appendix: Proof of the Technical Results

2.7.1 The Derivatives

In order to work on the asymptotic properties, first, we derive the form of the first and second derivatives of the log likelihood function. As mentioned above, the log likelihood for a single observation is:

$$\begin{aligned} l(\beta,\gamma,\sigma) =& \Delta\{\beta(Y)X(Y)\} - \Delta R(Y;\beta) - \Delta\log\sigma \\ &+ \Delta\log\phi\left(\frac{R(Y;\beta)+\gamma Z}{\sigma}\right) + (1-\Delta)\log\Phi\left(-\frac{R(Y;\beta)+\gamma Z}{\sigma}\right) \end{aligned}$$

For any fixed $\beta(\cdot) \in \Omega$, let $\{\beta_{(\eta)}(\cdot) : \eta$ in a neighborhood of $0 \in R\}$ be a smooth curve in Ω running through $\beta(\cdot)$ at $\eta = 0$, i.e., $\beta_{(\eta)}(\cdot)|_{\eta=0} = \beta(\cdot)$. and $\frac{\partial\beta_{(\eta)}}{\partial\eta}|_{\eta=0} = h$.

Furthermore, we write

$$R = R(Y, \beta) = \log \int_0^Y e^{\beta(u)X(u)} du$$

$$R_1 = \log \int_0^Y h(u)X(u)e^{\beta(u)X(u)} du$$

$$R_2 = \log \int_0^Y h_1(u)h_2(u)X^2(u)e^{\beta(u)X(u)} du$$

and

$$1 - \Phi = \Phi(-(R + \gamma Z)/\sigma)$$

$$\phi = \phi((R + \gamma Z)/\sigma) = \phi(-(R + \gamma Z)/\sigma)$$

$$\phi^{(1)} = \phi^{(1)}((R + \gamma Z)/\sigma)$$

$$\phi^{(2)} = \phi^{(2)}((R + \gamma Z)/\sigma)$$

then we can write out the derivatives as

$$\dot{l}_1(\beta, \gamma, \sigma)[h] = \Delta h(Y)X(Y) - \Delta \frac{e^{R_1}}{e^R} + \frac{\Delta}{\sigma} \frac{\phi^{(1)}}{\phi} \frac{e^{R_1}}{e^R} - \frac{(1 - \Delta)}{\sigma} \frac{\phi}{(1 - \Phi)} \frac{e^{R_1}}{e^R}$$

$$\dot{l}_2(\beta, \gamma, \sigma) = \frac{\Delta Z}{\sigma} \frac{\phi^{(1)}}{\phi} - \frac{(1 - \Delta)Z}{\sigma} \frac{\phi}{(1 - \Phi)}$$

$$\dot{l}_3(\beta, \gamma, \sigma) = -\frac{\Delta}{\sigma} - \frac{\Delta(R + \gamma Z)}{\sigma^2} \frac{\phi^{(1)}}{\phi} - \frac{(1 - \Delta)(R + \gamma Z)}{\sigma^2} \frac{\phi}{1 - \Phi}$$

The second derivatives are:

$$\begin{aligned}\ddot{l}_{11}(\beta,\gamma,\sigma)[h_1,h_2] =& -\Delta\frac{e^{R_2}}{e^R}+\Delta\frac{e^{2R_1}}{e^{2R}}+\frac{\Delta}{\sigma^2}\frac{e^{2R_1}}{e^{2R}}\frac{\phi^{(2)}}{\phi}+\frac{\Delta}{\sigma}\frac{e^{R_2}}{e^R}\frac{\phi^{(1)}}{\phi}-\frac{\Delta}{\sigma^2}\frac{e^{2R_1}}{e^{2R}}\left(\frac{\phi^{(1)}}{\phi}\right)^2\\ &-\frac{\Delta}{\sigma}\frac{e^{2R_1}}{e^{2R}}\frac{\phi^{(1)}}{\phi}-\frac{(1-\Delta)}{\sigma^2}\frac{e^{2R_1}}{e^{2R}}\frac{\phi^{(1)}}{(1-\Phi)}-\frac{(1-\Delta)}{\sigma}\frac{e^{R_2}}{e^R}\frac{\phi}{(1-\Phi)}\\ &-\frac{(1-\Delta)}{\sigma^2}\frac{e^{2R_1}}{e^{2R}}\left(\frac{\phi}{(1-\Phi)}\right)^2+\frac{(1-\Delta)}{\sigma}\frac{e^{2R_1}}{e^{2R}}\frac{\phi}{(1-\Phi)}\end{aligned}$$

$$\ddot{l}_{22}(\beta,\gamma,\sigma) =\frac{\Delta Z^2}{\sigma^2}\frac{\phi^{(2)}}{\phi}-\frac{\Delta Z^2}{\sigma^2}\left(\frac{\phi^{(1)}}{\phi}\right)^2-\frac{(1-\Delta)Z^2}{\sigma^2}\frac{\phi^{(1)}}{1-\Phi}-\frac{(1-\Delta)Z^2}{\sigma^2}\left(\frac{\phi}{1-\Phi}\right)^2$$

$$\begin{aligned}\ddot{l}_{33}(\beta,\gamma,\sigma) =&\frac{\Delta}{\sigma^2}+\frac{\Delta(R+\gamma Z)^2}{\sigma^4}\frac{\phi^{(2)}}{\phi}+\frac{2\Delta(R+\gamma Z)}{\sigma^3}\frac{\phi^{(1)}}{\phi}-\frac{\Delta(R+\gamma Z)^2}{\sigma^4}\left(\frac{\phi^{(1)}}{\phi}\right)^2\\ &+\frac{(1-\Delta)(R+\gamma Z)^2}{\sigma^4}\frac{\phi^{(1)}}{1-\Phi}+\frac{2(1-\Delta)(R+\gamma Z)}{\sigma^3}\frac{\phi}{1-\Phi}\\ &+\frac{(1-\Delta)(R+\gamma Z)^2}{\sigma^4}\left(\frac{\phi}{1-\Phi}\right)^2\end{aligned}$$

$$\begin{aligned}\ddot{l}_{12}(\beta,\gamma,\sigma)[h] =& \frac{\Delta Z}{\sigma^2}\frac{e^{R_1}}{e^R}\frac{\phi^{(2)}}{\phi} - \frac{\Delta Z}{\sigma^2}\frac{e^{R_1}}{e^R}\left(\frac{\phi^{(1)}}{\phi}\right)^2 - \frac{(1-\Delta)Z}{\sigma^2}\frac{e^{R_1}}{e^R}\frac{\phi^{(1)}}{1-\Phi} \\ & - \frac{(1-\Delta)Z}{\sigma^2}\frac{e^{R_1}}{e^R}\left(\frac{\phi}{1-\Phi}\right)^2\end{aligned}$$

$$\begin{aligned}\ddot{l}_{13}(\beta,\gamma,\sigma)[h] =& -\frac{\Delta(R+\gamma Z)}{\sigma^3}\frac{e^{R_1}}{e^R}\frac{\phi^{(2)}}{\phi} - \frac{\Delta}{\sigma^2}\frac{e^{R_1}}{e^R}\frac{\phi^{(1)}}{\phi} + \frac{\Delta(R+\gamma Z)}{\sigma^3}\frac{e^{R_1}}{e^R}\left(\frac{\phi^{(1)}}{\phi}\right)^2 \\ & + \frac{(1-\Delta)(R+\gamma Z)}{\sigma^3}\frac{e^{R_1}}{e^R}\frac{\phi^{(1)}}{1-\Phi} + \frac{(1-\Delta)}{\sigma^2}\frac{e^{R_1}}{e^R}\frac{\phi}{1-\Phi} \\ & + \frac{(1-\Delta)(R+\gamma Z)}{\sigma^3}\frac{e^{R_1}}{e^R}\left(\frac{\phi}{1-\Phi}\right)^2\end{aligned}$$

$$\begin{aligned}\ddot{l}_{23}(\beta,\gamma,\sigma) =& -\frac{\Delta Z(R+\gamma Z)}{\sigma^3}\frac{\phi^{(2)}}{\phi} - \frac{\Delta Z}{\sigma^2}\frac{\phi^{(1)}}{\phi} + \frac{\Delta Z(R+\gamma Z)}{\sigma^3}\left(\frac{\phi^{(1)}}{\phi}\right)^2 \\ & + \frac{(1-\Delta)Z(R+\gamma Z)}{\sigma^3}\frac{\phi^{(1)}}{1-\Phi} + \frac{(1-\Delta)Z}{\sigma^2}\frac{\phi}{1-\Phi} \\ & + \frac{(1-\Delta)Z(R+\gamma Z)}{\sigma^3}\left(\frac{\phi}{1-\Phi}\right)^2\end{aligned}$$

2.7.2 Outline of the Proofs

In the section, we provide some details on the theorems. The following conditions are sufficient to guarantee the forthcoming properties.

(C1) $X_i(t)$ is bounded for $t \in R^+$, i.e. $\forall t$, $\exists C$ s.t. $\mathbb{P}(||X_i(t)|| \leq C) = 1$

(C2) $\gamma \in \Gamma$, $\sigma \in \Sigma$, where Γ is a compact subset of R, and Σ is a compact subset of R^+.

(C3) $\beta(\cdot) \geq 0$, $\forall t \in R^+$.

(C4) $\forall t$, $\exists C$ s.t. $\mathbb{P}(0 \leq ||\beta(t)X_i(t)|| \leq C) = 1$

(C5) Let $\mathcal{B}$ denote the collection of bounded functions β on $(0, a]$ with bounded derivatives $\beta^{(j)}$, $j = 1, \cdots, k$, and the kth derivative $\beta^{(k)}$ satisfies the following Lips-

chitz continuity condition:

$$\left|\beta^{(k)}(s) - \beta^{(k)}(t)\right| \leq L\,|s-t|^m \text{ for } s,t \in (0,a]$$

We have $\boldsymbol{\theta} = (\beta, \gamma, \sigma)^T$, $\beta \in \mathcal{B}$, $\gamma \in \Gamma$ and $\sigma \in \Sigma$, Let $\boldsymbol{\Theta} = \Omega \times \Gamma \times \Sigma$ be the parameter space of $\boldsymbol{\theta}$. We define a distance between $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \boldsymbol{\Theta}$ by

$$d(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \{||\beta_2 - \beta_1||^2 + |\gamma_2 - \gamma_1|^2 + |\sigma_2 - \sigma_2|^2\}^{1/2}$$

where $|\cdot|$ is the Euclidean distance and $\|\cdot\|$ is some norm. Let $\boldsymbol{\Theta}_n$ be the sieve parameter space, a sequence of increasing subsets of the parameter space growing dense in $\boldsymbol{\Theta}$ as $n \to \infty$.

2.7.2.1 Proof of Theorem 2.1

For Theorem 2.1, in order to give the convergence rate for $\hat{\boldsymbol{\theta}}_n$, we first prove the asymptotic consistency, i.e. : $d(\hat{\boldsymbol{\theta}}_n, \boldsymbol{\theta}_0) \longrightarrow 0$. Let $M(\boldsymbol{\theta}) = Pl(\boldsymbol{\theta}; Y, \Delta, \boldsymbol{W})$ and $\mathbb{M}_n(\boldsymbol{\theta}) = \mathbb{P}_n l(\boldsymbol{\theta}; Y, \Delta, \boldsymbol{W})$. Hence $\mathbb{M}_n(\boldsymbol{\theta}) - M(\boldsymbol{\theta}) = (\mathbb{P}_n - P) l(\boldsymbol{\theta}; Y, \Delta, \boldsymbol{W})$. According to Theorem 5.7 in van der Vaart (1998), this would be equivalent to check the following the conditions:

(A1) Let $\mathcal{L}_1 = \{l(\boldsymbol{\theta}; Y, \Delta, \boldsymbol{W}) : \boldsymbol{\theta} \in \boldsymbol{\Theta}_n\}$ then the c-bracketing number for $\mathcal{L}_1$ with $L_1(P)$-norm is bounded by some $C(1/\epsilon)^{n^a}$ thus $\mathcal{L}_1$ is Glivenko-Cantelli and $\sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_n} |\mathbb{M}_n(\boldsymbol{\theta}) \ - M(\boldsymbol{\theta})| \to_p 0$

(A2) $\mathbb{M}(\boldsymbol{\theta}_0) - M(\boldsymbol{\theta}) \geq 0$

(A3) For $\beta_0 \in \mathcal{B}$, there exists a $\beta_{0,n} \in \mathbf{B}_n$ of order $m \geq p+2$ such that $\|\beta_{0,n} -$

$\beta_0\|_\infty \leq Cq_n^{-p} = O(n^{-p\nu})$ then let $\boldsymbol{\theta}_{0,n} = (\beta_{0,n}, \gamma_0, \sigma_0)$ then we can have

$$\begin{aligned}
\mathbb{M}_n(\hat{\boldsymbol{\theta}}_n) - M(\boldsymbol{\theta_0}) &= \mathbb{M}_n(\hat{\boldsymbol{\theta}}_n) - \mathbb{M}_n(\boldsymbol{\theta}_{0,n}) + \mathbb{M}_n(\boldsymbol{\theta}_{0,n}) - M(\boldsymbol{\theta}_0) \\
&\geq \mathbb{P}_n\{l(\boldsymbol{\theta}_{0,n}; Y, \Delta, \boldsymbol{W}) - l(\boldsymbol{\theta}_0; Y, \Delta, \boldsymbol{W})\} \\
&\geq -o_p(1)
\end{aligned}$$

thus $Cd^2(\boldsymbol{\theta}_0, \boldsymbol{\theta}_{0,n}) \to 0$

(A4) Furthermore we want

$$\begin{aligned}
\mathbb{M}_n(\hat{\boldsymbol{\theta}}_n) - M(\boldsymbol{\theta}_0) \geq & (\mathbb{P}_n - P)\{l(\beta_{0,n}, \gamma_0, \sigma_0; Y, \Delta, \boldsymbol{W}) - l(\beta_0, \gamma_0, \sigma_0; Y, \Delta, \boldsymbol{W})\} \\
& + P\{l(\beta_{0,n}, \gamma_0, \sigma_0; Y, \Delta, \boldsymbol{W}) - l(\beta_0, \gamma_0, \sigma_0; Y, \Delta, \boldsymbol{W})\} \\
= & - O_p(n^{-2p\nu}) \\
= & - O_p\left(n^{-min(p\nu, (1-\nu)/2)}\right)
\end{aligned}$$

Proof of (A1)

Let $\mathcal{L}_1 = \{l(\boldsymbol{\theta}; Y, \Delta, \boldsymbol{W}) : \boldsymbol{\theta} \in \boldsymbol{\Theta}_n\}$. by the calculation of *Shen and Wong* (1994), page 597, $\forall \epsilon > 0$, there exists a set of brackets $\left\{\left[\beta_j^L, \beta_j^U\right] : j = 1, 2, \cdots, [(1/\epsilon)^{C_1 q_n}]\right\}$ such that for any $\beta \in \mathbf{B}_n$, one has $\beta_j^L(u) \leq \beta(u) \leq \beta_j^U(u)$ for some $1 \leq j \leq [(1/\epsilon)^{C_1 q_n}]$ and all $u \in [0, b]$, and $P|\beta_j^U(Y, \Delta, \boldsymbol{W}) - \beta_j^U(Y, \Delta, \boldsymbol{W})| \leq \epsilon$. Since $\Gamma \in R$ is compact, Γ can be covered by $[C_2/\epsilon]$ balls with radius ϵ; that is, for any $\gamma \in \Gamma$, there exists an $1 \leq s_1 \leq [C_2/\epsilon]$ such that $|\gamma - \gamma_{s_1}| \leq \epsilon$. Similarly, for any $\sigma \in \Sigma$, there exists an $1 \leq s_2 \leq [C_3/\epsilon]$ such that $|\sigma - \sigma_{s_2}| \leq \epsilon$. Then $\forall l(\beta, \gamma, \sigma; Y, \Delta, \boldsymbol{W}) \in \mathcal{L}_1$, $\exists s_1, s_2, j$,

s.t. $l \in [l^L_{s_1,s_2,j}(Y,\Delta,\boldsymbol{W}),\ l^U_{s_1,s_2,j}(Y,\Delta,\boldsymbol{W})]$, for any sample point $(Y,\Delta,\boldsymbol{W})$, where

$$\begin{aligned} l^L_{s_1,s_2,j}(Y,\Delta,\boldsymbol{W}) =& \Delta_i \left\{\beta^L_j(Y_i)X_i(Y_i) + (\gamma_{s_1}-\epsilon)Z_i\right\} \\ & - \Delta_i \log \int_0^{Y_i} e^{\beta^U_j(u)X_i(u)}du - (\gamma_{s_1}+\epsilon)Z_i - \Delta_i \log(\sigma^2_{s_2}+\epsilon) \\ & + \Delta_i \log\phi\left(\frac{\log\int_0^{Y_i} e^{\beta^L_j(u)X_i(u)}du + (\gamma_{s_1}-\epsilon)Z_i}{\sigma^2_{s_2}+\epsilon}\right) \\ & + (1-\Delta_i)\log\Phi\left(-\frac{\log\int_0^{Y_i} e^{\beta^U_j(u)X_i(u)}du + (\gamma_{s_1}+\epsilon)Z_i}{\sigma^2_{s_2}-\epsilon}\right) \end{aligned}$$

$$\begin{aligned} l^U_{s_1,s_2,j}(Y,\Delta,\boldsymbol{W}) =& \Delta_i \{\beta^U_j(Y_i)X_i(Y_i) + (\gamma_{s_1}+\epsilon)Z_i\} \\ & - \Delta_i \log \int_0^{Y_i} e^{\omega^L_j(u)X_i(u)}du - (\gamma_{s_1}-\epsilon)Z_i - \Delta_i \log(\sigma^2_{s_2}-\epsilon) \\ & + \Delta_i \log\phi\left(\frac{\log\int_0^{Y_i} e^{\beta^U_j(u)X_i(u)}du + (\gamma_{s_1}+\epsilon)Z_i}{\sigma^2_{s_2}-\epsilon}\right) \\ & + (1-\Delta_i)\log\Phi\left(-\frac{\log\int_0^{Y_i} e^{\beta^L_j(u)X_i(u)}du + (\gamma_{s_1}-\epsilon)Z_i}{\sigma^2_{s_2}+\epsilon}\right) \end{aligned}$$

By Taylor expansion and some calculation, we have for all s_1, s_2, s_3 and i

$$P(|(l^U_{s_1,s_2,s_3,i} - l^L_{s_1,s_2,s_3,i})(Y,\Delta,\boldsymbol{W})|) \le C_1'\epsilon$$

thus the c-bracketing number for $\mathcal{L}_1 = \{l(\boldsymbol{\theta};Y,\Delta,\boldsymbol{W}) : \boldsymbol{\theta} \in \boldsymbol{\Theta}_n\}$ with $L_1(P)$-norm is bounded by $C(1/\epsilon)^{C_1 q_n+3}$, so $\mathcal{L}_1$ is Glivenko-Cantelli, $\sup_{\theta\in\Theta_n}|\mathbb{M}_n(\boldsymbol{\theta}) - M(\boldsymbol{\theta})| \to_p 0$

Proof of (A2)

Since $Pl(\beta,\gamma,\sigma)$ is maximized at $(\beta_0,\gamma_0,\sigma_0)$, so its derivatives at $(\beta_0,\gamma_0,\sigma_0)$ are equal

to 0. so by Taylor expansion,

$$\begin{aligned}
M(\boldsymbol{\theta}_0) - M(\boldsymbol{\theta}) =& Pl(\boldsymbol{\theta}_0; Y, \Delta, \boldsymbol{W}) - Pl(\boldsymbol{\theta}; Y, \Delta, \boldsymbol{W}) \\
=& \frac{1}{2} P \Big\{ \ddot{l}_{11}(\boldsymbol{\theta}_0)[\beta - \beta_0, \beta - \beta_0] + \ddot{l}_{22}(\boldsymbol{\theta}_0)(\gamma - \gamma_0)^2 + \ddot{l}_{33}(\boldsymbol{\theta}_0)(\sigma - \sigma_0)^2 \\
& + 2\ddot{l}_{12}(\boldsymbol{\theta}_0)(\gamma - \gamma_0)[\beta - \beta_0] + 2\ddot{l}_{13}(\boldsymbol{\theta}_0)(\sigma - \sigma_0)[\beta - \beta_0] \\
& + 2\ddot{l}_{23}(\boldsymbol{\theta}_0)(\gamma - \gamma_0)(\sigma - \sigma_0) \Big\} + o\left(d^2(\boldsymbol{\theta}, \boldsymbol{\theta}_0)\right) \\
=& A + o\left(d^2(\boldsymbol{\theta}, \boldsymbol{\theta}_0)\right)
\end{aligned}$$

Before we proceed to find the upper bound for A, first, by Cauchy-Schwarz Inequality, we have

$$\begin{aligned}
& \int_0^Y \left[\{\beta(u) - \beta_0(u)\} X(u)\right]^2 \cdot \exp\left\{\beta_0(u) X(u)\right\} du \\
\leq & \int_0^Y \left[\{\beta(u) - \beta_0(u)\} X(u)\right]^2 du \cdot \int_0^Y \exp\left\{\beta_0(u) X(u)\right\} du \\
\lesssim & e^R \int_0^Y \left[\{\beta(u) - \beta_0(u)\} X(u)\right]^2 du \\
\lesssim & e^R \|\beta - \beta_0\|_Y^2
\end{aligned}$$

where $\lesssim$ denotes that the left-hand side is bounded above by a constant times the right-hand side, similarly,

$$\begin{aligned}
& \int_0^Y \{\beta(u) - \beta_0(u)\} X(u) \cdot \exp\left\{\beta_0(u) X(u)\right\} du \\
\leq & \int_0^Y \{\beta(u) - \beta_0(u)\} X(u) du \cdot \int_0^Y \exp\left\{\beta_0(u) X(u)\right\} du \\
\lesssim & e^R \|\beta - \beta_0\|_Y
\end{aligned}$$

then we have $0 \leq e^{R_2} \lesssim e^R \cdot \|\beta - \beta_0\|_Y^2$ and $0 \leq (e^{R_1})^2 \lesssim (e^R)^2 \cdot \|\beta - \beta_0\|_Y^2$ thus

$$\begin{aligned}
\ddot{l}_{11}(\beta_0, \gamma_0, \sigma_0)[\beta - \beta_0, \beta - \beta_0] = & -\Delta \frac{e^{R_2}}{e^R} + \Delta \frac{e^{2R_1}}{e^{2R}} + \frac{\Delta}{\sigma_0^2} \frac{e^{2R_1}}{e^{2R}} \frac{\phi^{(2)}}{\phi} + \frac{\Delta}{\sigma_0} \frac{e^{R_2}}{e^R} \frac{\phi^{(1)}}{\phi} \\
& - \frac{\Delta}{\sigma_0^2} \frac{e^{2R_1}}{e^{2R}} \left(\frac{\phi^{(1)}}{\phi} \right)^2 - \frac{\Delta}{\sigma_0} \frac{e^{2R_1}}{e^{2R}} \frac{\phi^{(1)}}{\phi} - \frac{(1-\Delta)}{\sigma_0^2} \frac{e^{2R_1}}{e^{2R}} \frac{\phi^{(1)}}{(1-\Phi)} \\
& - \frac{(1-\Delta)}{\sigma_0} \frac{e^{R_2}}{e^R} \frac{\phi}{(1-\Phi)} - \frac{(1-\Delta)}{\sigma_0^2} \frac{e^{2R_1}}{e^{2R}} \left(\frac{\phi}{1-\Phi} \right)^2 \\
& + \frac{(1-\Delta)}{\sigma_0} \frac{e^{2R_1}}{e^{2R}} \frac{\phi}{(1-\Phi)} \\
\lesssim & \left\{ \frac{\Delta}{\sigma_0^2} \cdot \frac{\left|\phi^{(2)}\right|}{\phi} + \frac{\Delta}{\sigma_0} \cdot \frac{2\left|\phi^{(1)}\right|}{\phi} + \frac{(1-\Delta)}{\sigma_0^2} \frac{\left|\phi^{(1)}\right|}{1-\Phi} \right. \\
& \left. + \frac{(1-\Delta)}{\sigma_0} \frac{|\phi|}{1-\Phi} \right\} \cdot \|\beta - \beta_0\|_Y^2
\end{aligned}$$

similarly, we can have

$$\begin{aligned}
& \ddot{l}_{12}(\beta_0, \gamma_0, \sigma_0^2)[\beta - \beta_0] = \ddot{l}_{21}(\beta_0, \gamma_0, \sigma_0^2)[\beta - \beta_0] \\
= & \frac{\Delta Z}{\sigma_0^2} \frac{e^{R_1}}{e^R} \frac{\phi^{(2)}}{\phi} - \frac{\Delta Z}{\sigma_0^2} \frac{e^{R_1}}{e^R} \left(\frac{\phi^{(1)}}{\phi} \right)^2 - \frac{(1-\Delta)Z}{\sigma_0^2} \frac{e^{R_1}}{e^R} \frac{\phi^{(1)}}{1-\Phi} - \frac{(1-\Delta)Z}{\sigma_0^2} \frac{e^{R_1}}{e^R} \left(\frac{\phi}{1-\Phi} \right)^2 \\
\lesssim & \left\{ \frac{\Delta}{\sigma_0^3} \frac{\left|\phi^{(2)}\right|}{\phi} + \frac{\Delta}{\sigma_0^2} \frac{(\phi^{(1)})^2}{\phi^2} + \frac{(1-\Delta)}{\sigma_0^2} \frac{\left|\phi^{(1)}\right|}{1-\Phi} \right\} |Z| \cdot \|\beta - \beta_0\|_Y
\end{aligned}$$

$$
\begin{aligned}
&\ddot{l}_{13}(\beta_0,\gamma_0,\sigma_0^2)[\beta-\beta_0] = \ddot{l}_{31}(\beta_0,\gamma_0,\sigma_0^2)[\beta-\beta_0] \\
=& -\frac{\Delta(R+\gamma_0 Z)}{\sigma_0^3}\frac{e^{R_1}}{e^R}\frac{\phi^{(2)}}{\phi} - \frac{\Delta}{\sigma_0^2}\frac{e^{R_1}}{e^R}\frac{\phi^{(1)}}{\phi} + \frac{\Delta(R+\gamma_0 Z)}{\sigma_0^3}\frac{e^{R_1}}{e^R}\left(\frac{\phi^{(1)}}{\phi}\right)^2 \\
&+ \frac{(1-\Delta)(R+\gamma_0 Z)}{\sigma_0^3}\frac{e^{R_1}}{e^R}\frac{\phi^{(1)}}{1-\Phi} + \frac{(1-\Delta)}{\sigma_0^2}\frac{e^{R_1}}{e^R}\frac{\phi}{1-\Phi} + \frac{(1-\Delta)(R+\gamma_0 Z)}{\sigma_0^3}\frac{e^{R_1}}{e^R}\left(\frac{\phi}{1-\Phi}\right)^2 \\
\lesssim& \left\{ \frac{\Delta}{\sigma_0^3}\frac{\left|\phi^{(2)}\right|}{\phi}|R| + \frac{\Delta}{\sigma_0^3}\frac{\left|\phi^{(2)}\right|}{\phi}|\gamma_0||Z| + \frac{\Delta}{\sigma_0^2}\frac{\left|\phi^{(1)}\right|}{\phi} + \frac{\Delta}{\sigma_0^3}\frac{(\phi^{(1)})^2}{\phi^2}|R| + \frac{\Delta}{\sigma_0^3}\frac{(\phi^{(1)})^2}{\phi^2}|\gamma_0||Z| \right. \\
&+ \frac{(1-\Delta)\left|\phi^{(1)}\right|}{\sigma_0^3(1-\Phi)}|R| + \frac{(1-\Delta)\left|\phi^{(1)}\right|}{\sigma_0^3(1-\Phi)}|\gamma_0||Z| + \frac{(1-\Delta)}{\sigma_0^2}\frac{\phi}{1-\Phi} + \frac{(1-\Delta)}{\sigma_0^3}\frac{\phi^2}{(1-\Phi)^2}|R| \\
&\left. + \frac{(1-\Delta)}{\sigma_0^3}\frac{\phi^2}{(1-\Phi)^2}|\gamma_0||Z| \right\} \cdot \|\beta-\beta_0\|_Y
\end{aligned}
$$

Thus, we can calculate the following quantity

$$
\begin{aligned}
D_1 =& \ddot{l}_{11}(\boldsymbol{\theta}_0)[\beta-\beta_0,\beta-\beta_0]+\ddot{l}_{22}(\boldsymbol{\theta}_0)(\gamma-\gamma_0)^2+\ddot{l}_{33}(\boldsymbol{\theta}_0)(\sigma-\sigma_0)^2 \\
&+2\ddot{l}_{12}(\boldsymbol{\theta}_0)[\beta-\beta_0](\gamma-\gamma_0)+2\ddot{l}_{13}(\boldsymbol{\theta}_0)[\beta-\beta_0](\sigma-\sigma_0)+2\ddot{l}_{23}(\boldsymbol{\theta}_0)(\gamma-\gamma_0)(\sigma-\sigma_0) \\
\lesssim& \left\{\frac{\Delta}{\sigma_0^2}\cdot\frac{\left|\phi^{(2)}\right|}{\phi}+\frac{\Delta}{\sigma_0}\cdot\frac{2\left|\phi^{(1)}\right|}{\phi}+\Delta+\frac{(1-\Delta)}{\sigma_0^2}\frac{\left|\phi^{(1)}\right|}{1-\Phi}+\frac{(1-\Delta)}{\sigma_0}\frac{|\phi|}{1-\Phi}\right\} \\
&\cdot\|\beta-\beta_0\|_Y^2+\left\{\frac{\Delta Z^2}{\sigma_0^2}\frac{\phi^{(2)}}{\phi}-\frac{\Delta Z^2}{\sigma_0^2}\left(\frac{\phi^{(1)}}{\phi}\right)^2-\frac{(1-\Delta)Z^2}{\sigma_0^2}\frac{\phi^{(1)}}{1-\Phi}\right. \\
&\left.-\frac{(1-\Delta)Z^2}{\sigma_0^2}\left(\frac{\phi}{1-\Phi}\right)^2\right\}\cdot(\gamma-\gamma_0)^2+\left\{\frac{\Delta}{\sigma_0^2}+\frac{\Delta(R+\gamma_0Z)^2}{\sigma_0^4}\frac{\phi^{(2)}}{\phi}+\frac{2\Delta(R+\gamma_0Z)}{\sigma_0^3}\frac{\phi^{(1)}}{\phi}\right. \\
&-\frac{\Delta(R+\gamma_0Z)^2}{\sigma_0^4}\left(\frac{\phi^{(1)}}{\phi}\right)^2+\frac{(1-\Delta)(R+\gamma_0Z)^2}{\sigma_0^4}\frac{\phi^{(1)}}{1-\Phi}+\frac{2(1-\Delta)(R+\gamma_0Z)}{\sigma_0^3}\frac{\phi}{1-\Phi} \\
&\left.+\frac{(1-\Delta)(R+\gamma_0Z)^2}{\sigma_0^4}\left(\frac{\phi}{1-\Phi}\right)^2\right\}\cdot(\sigma-\sigma_0)^2+2\left\{\frac{\Delta}{\sigma_0^3}\frac{\left|\phi^{(2)}\right|}{\phi}+\frac{\Delta}{\sigma_0^2}\frac{(\phi^{(1)})^2}{\phi^2}\right. \\
&\left.+\frac{(1-\Delta)}{\sigma_0^2}\frac{\left|\phi^{(1)}\right|}{1-\Phi}\right\}\cdot\|\beta-\beta_0\|_Y(\gamma-\gamma_0)+2\left\{\frac{\Delta}{\sigma_0^3}\frac{\left|\phi^{(2)}\right|}{\phi}|R|+\frac{\Delta}{\sigma_0^3}\frac{\left|\phi^{(2)}\right|}{\phi}|\gamma_0||Z|\right. \\
&+\frac{\Delta}{\sigma_0^2}\frac{\left|\phi^{(1)}\right|}{\phi}+\frac{\Delta}{\sigma_0^3}\frac{(\phi^{(1)})^2}{\phi^2}|R|+\frac{\Delta}{\sigma_0^3}\frac{(\phi^{(1)})^2}{\phi^2}|\gamma_0||Z|+\frac{(1-\Delta)\left|\phi^{(1)}\right|}{\sigma_0^3(1-\Phi)}|R|+\frac{(1-\Delta)\left|\phi^{(1)}\right|}{\sigma_0^3(1-\Phi)}|\gamma_0||Z| \\
&\left.+\frac{(1-\Delta)}{\sigma_0^2}\frac{\phi}{1-\Phi}+\frac{(1-\Delta)}{\sigma_0^3}\frac{\phi^2}{(1-\Phi)^2}|R|+\frac{(1-\Delta)}{\sigma_0^3}\frac{\phi^2}{(1-\Phi)^2}|\gamma_0||Z|\right\}\cdot\|\beta-\beta_0\|_Y(\sigma-\sigma_0) \\
&+2\left\{-\frac{\Delta Z(R+\gamma_0Z)}{\sigma_0^3}\frac{\phi^{(2)}}{\phi}-\frac{\Delta Z}{\sigma_0^2}\frac{\phi^{(1)}}{\phi}+\frac{\Delta Z(R+\gamma_0Z)}{\sigma_0^3}\left(\frac{\phi^{(1)}}{\phi}\right)^2\right. \\
&\left.+\frac{(1-\Delta)Z(R+\gamma_0Z)}{\sigma_0^3}\frac{\phi^{(1)}}{1-\Phi}+\frac{(1-\Delta)Z}{\sigma_0^2}\frac{\phi}{1-\Phi}+\frac{(1-\Delta)Z(R+\gamma_0Z)}{\sigma_0^3}\left(\frac{\phi}{1-\Phi}\right)^2\right\} \\
&\cdot(\gamma-\gamma_0)(\sigma-\sigma_0) \\
=&\Delta N_1+(1-\Delta)N_2
\end{aligned}
$$

where we denote

$$
\begin{aligned}
N_1 =& \left\{ \frac{1}{\sigma_0^2} \cdot \frac{\left|\phi^{(2)}\right|}{\phi} + \frac{1}{\sigma_0} \cdot \frac{2\left|\phi^{(1)}\right|}{\phi} + 1 \right\} \cdot \|\beta - \beta_0\|_Y^2 + \left\{ \frac{Z^2}{\sigma_0^2} \frac{\phi^{(2)}}{\phi} - \frac{Z^2}{\sigma_0^2} \left(\frac{\phi^{(1)}}{\phi} \right)^2 \right\} \cdot (\gamma - \gamma_0)^2 \\
&+ \left\{ \frac{1}{\sigma_0^2} + \frac{(R + \gamma_0 Z)^2}{\sigma_0^4} \frac{\phi^{(2)}}{\phi} + \frac{2(R + \gamma_0 Z)}{\sigma_0^3} \frac{\phi^{(1)}}{\phi} - \frac{(R + \gamma_0 Z)^2}{\sigma_0^4} \left(\frac{\phi^{(1)}}{\phi} \right)^2 \right\} \cdot (\sigma - \sigma_0)^2 \\
&+ 2 \left\{ \frac{1}{\sigma_0^3} \frac{\left|\phi^{(2)}\right|}{\phi} + \frac{1}{\sigma_0^2} \frac{(\phi^{(1)})^2}{\phi^2} \right\} \cdot \|\beta - \beta_0\|_Y (\gamma - \gamma_0) + 2 \left\{ \frac{1}{\sigma_0^3} \frac{\left|\phi^{(2)}\right|}{\phi} |R| + \frac{1}{\sigma_0^3} \frac{\left|\phi^{(2)}\right|}{\phi} |\gamma_0| \, |Z| \right. \\
&\left. + \frac{1}{\sigma_0^2} \frac{\left|\phi^{(1)}\right|}{\phi} + \frac{1}{\sigma_0^3} \frac{(\phi^{(1)})^2}{\phi^2} |R| + \frac{1}{\sigma_0^3} \frac{(\phi^{(1)})^2}{\phi^2} |\gamma_0| \, |Z| \right\} \cdot \|\beta - \beta_0\|_Y (\sigma - \sigma_0) \\
&+ 2 \left\{ -\frac{Z(R + \gamma_0 Z)}{\sigma_0^3} \frac{\phi^{(2)}}{\phi} - \frac{Z}{\sigma_0^2} \frac{\phi^{(1)}}{\phi} + \frac{Z(R + \gamma_0 Z)}{\sigma_0^3} \left(\frac{\phi^{(1)}}{\phi} \right)^2 \right\} (\gamma - \gamma_0)(\sigma - \sigma_0)
\end{aligned}
$$

After some calculation and by CauchySchwarz inequality, we have

$$
\begin{aligned}
N_1 \lesssim & C_1 \|\beta - \beta_0\|_Y^2 + C_2 (\gamma - \gamma_0)^2 + C_3 (\sigma - \sigma_0)^2 \\
\lesssim & \|\beta - \beta_0\|_Y^2 + (\gamma - \gamma_0)^2 + (\sigma - \sigma_0)^2
\end{aligned}
$$

where C_1, C_2 and C_3 are constant with respect to $\boldsymbol{\theta} = (\beta(\cdot), \gamma, \sigma)$. Similarly, we can have

$$
\begin{aligned}
N_2 =& \left\{ \frac{1}{\sigma_0^2} \frac{\left|\phi^{(1)}\right|}{1 - \Phi} + \frac{1}{\sigma_0} \frac{|\phi|}{1 - \Phi} \right\} \cdot \|\beta - \beta_0\|_Y^2 + \left\{ -\frac{Z^2}{\sigma_0^2} \frac{\phi^{(1)}}{1 - \Phi} - \frac{Z^2}{\sigma_0^2} \left(\frac{\phi}{1 - \Phi} \right)^2 \right\} \cdot (\gamma - \gamma_0)^2 \\
&+ \left\{ \frac{(R + \gamma_0 Z)^2}{\sigma_0^4} \frac{\phi^{(1)}}{1 - \Phi} + \frac{2(R + \gamma_0 Z)}{\sigma_0^3} \frac{\phi}{1 - \Phi} + \frac{(R + \gamma_0 Z)^2}{\sigma_0^4} \left(\frac{\phi}{1 - \Phi} \right)^2 \right\} \cdot (\sigma - \sigma_0)^2 \\
&+ 2 \frac{1}{\sigma_0^2} \frac{\left|\phi^{(1)}\right|}{1 - \Phi} \cdot \|\beta - \beta_0\|_Y (\gamma - \gamma_0) + 2 \left\{ \frac{\left|\phi^{(1)}\right|}{\sigma_0^3 (1 - \Phi)} |R| + \frac{\left|\phi^{(1)}\right|}{\sigma_0^3 (1 - \Phi)} |\gamma_0| \, |Z| \right. \\
&\left. + \frac{1}{\sigma_0^2} \frac{\phi}{1 - \Phi} + \frac{1}{\sigma_0^3} \frac{\phi^2}{(1 - \Phi)^2} |R| + \frac{1}{\sigma_0^3} \frac{\phi^2}{(1 - \Phi)^2} |\gamma_0| \, |Z| \right\} \cdot \|\beta - \beta_0\|_Y (\sigma - \sigma_0) \\
&+ 2 \left\{ \frac{Z(R + \gamma_0 Z)}{\sigma_0^3} \frac{\phi^{(1)}}{1 - \Phi} + \frac{Z}{\sigma_0^2} \frac{\phi}{1 - \Phi} + \frac{Z(R + \gamma_0 Z)}{\sigma_0^3} \left(\frac{\phi}{1 - \Phi} \right)^2 \right\} \cdot (\gamma - \gamma_0)(\sigma - \sigma_0) \\
\lesssim & \|\beta - \beta_0\|_Y^2 + (\gamma - \gamma_0)^2 + (\sigma - \sigma_0)^2
\end{aligned}
$$

Thus

$$\begin{aligned} D_1 \lesssim & \Delta N_1 + (1-\Delta)N_2 \\ \lesssim & \|\beta - \beta_0\|_Y^2 + (\gamma - \gamma_0)^2 + (\sigma - \sigma_0)^2 \\ = & d^2(\boldsymbol{\theta}, \boldsymbol{\theta}_0), \end{aligned}$$

we can then have $A = \frac{1}{2}P\{D_1\} \leq Cd^2(\boldsymbol{\theta}, \boldsymbol{\theta}_0)$ with some C for which is bounded.

Proof of (A3)

$$\begin{aligned} \mathbb{M}_n(\hat{\boldsymbol{\theta}}_n) - \mathbb{M}_n(\boldsymbol{\theta}_0) &= \mathbb{M}_n(\hat{\boldsymbol{\theta}}_n) - \mathbb{M}_n(\boldsymbol{\theta}_{0.n}) + \mathbb{M}_n(\boldsymbol{\theta}_{0,n}) - \mathbb{M}_n(\boldsymbol{\theta}_0) \\ &\geq \mathbb{P}_n l(\boldsymbol{\theta}_{0,n}; Y, \Delta, \boldsymbol{W}) - \mathbb{P}_n l(\boldsymbol{\theta}_0; Y, \Delta, \boldsymbol{W}) \\ &= (\mathbb{P}_n - P)\left\{l(\boldsymbol{\theta}_{0,n}; Y, \Delta, \boldsymbol{W}) - l(\boldsymbol{\theta}_0; Y, \Delta, \boldsymbol{W})\right\} + \mathbb{M}(\boldsymbol{\theta}_{0,n}) - \mathbb{M}(\boldsymbol{\theta}_0) \end{aligned}$$

Define $\mathcal{L}_2 = \{l(\beta, \gamma_0, \sigma_0) - l(\beta_0, \gamma_0, \sigma_0) : \beta \in \mathbf{B}_n, \text{ and } \|\beta - \beta_0\| \leq Cn^{-p\nu}\}$ with the ϵ-bracketing number associated $L_2(P)$-norm bounded by $(1/\epsilon)^{Cq_n}$. Thus $\mathcal{L}_2$ is P-Donsker. So

$$P\left\{l(\beta, \gamma_0, \sigma_0; Y, \Delta, \boldsymbol{W}) - l(\beta_0, \gamma_0, \sigma_0; Y, \Delta, \boldsymbol{W})\right\}^2 \to 0 \text{ as } n \to \infty$$

Hence $(\mathbb{P}_n - P)\left\{l(\boldsymbol{\theta}_{0,n}; Y, \Delta, \boldsymbol{W}) - l(\boldsymbol{\theta}_0; Y, \Delta, \boldsymbol{W})\right\} = o_p\left(n^{-1/2}\right)$, by the relationship between P-Donsker and asymptotic equicontinuity (*van der Vaart and Wellner* (1996) Corollary 2.3.12). By the Dominated Convergence Theorem, $M(\boldsymbol{\theta}_{0,n}) - M(\boldsymbol{\theta}_0) > -o(1)$ as $n \to \infty$. Therefore,

$$\mathbb{M}_n(\hat{\boldsymbol{\theta}}_n) - M(\boldsymbol{\theta}_0) \geq o_p(n^{-1/2}) - o(1) = -o_p(1)$$

Combine (A1) - (A3), we have the consistency result, i.e. $d(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \to 0$ in probability. And Condition (A4) further gives the convergence rate.

Proof of (A4)

Here, we need to verify the conditions of Theorem 3.2.5 of *van der Vaart and Wellner* (1996) in order to derive the convergence rate. We have

$$\mathbb{M}_n(\hat{\boldsymbol{\theta}}_n) - \mathbb{M}_n(\boldsymbol{\theta}_0) \leq I_{1,n} + I_{2,n}$$

where $I_{1,n} = (\mathbb{P}_n - P)\{l(\beta_{0,n}, \gamma_0, \sigma_0; Y, \Delta, \boldsymbol{W}) - l(\beta_0, \gamma_0, \sigma_0; Y, \Delta, \boldsymbol{W})\}$ and $I_{2,n} = P\{l(\beta_{0,n}, \gamma_0, \sigma_0; Y, \Delta, \boldsymbol{W}) - l(\beta_0, \gamma_0, \sigma_0; Y, \Delta, \boldsymbol{W})\}$. By Taylor expansion, we have

$$\begin{aligned} I_{1,n} &= (\mathbb{P}_n - P)\left\{\dot{l}_2(\tilde{\beta}, \gamma_0, \sigma_0; Y, \Delta, \boldsymbol{W})(\beta_{0,n} - \beta_0)\right\} \\ &= n^{-p\nu+\epsilon}(\mathbb{P}_n - P)\left\{\dot{l}_2(\tilde{\beta}, \gamma_0, \sigma_0; Y, \Delta, \boldsymbol{W})\frac{\beta_{0,n} - \beta_0}{n^{-p\nu+\epsilon}}\right\} \end{aligned}$$

for any $0 < \epsilon < 1/2 - p\nu$. Because $\|\beta_{0,n} - \beta_0\|_\infty = O(n^{-p\nu})$ and $\dot{l}_2(\tilde{\beta}, \gamma_0, \sigma_0; Y, \Delta, \boldsymbol{W})$ is uniformly bounded due to Conditions (C1) - (C5), we can easily obtain that

$$P\left\{\dot{l}_2(\tilde{\beta}, \gamma_0, \sigma_0; Y, \Delta, \boldsymbol{W})\frac{\beta_{0,n} - \beta_0}{n^{-p\nu+\epsilon}}\right\}^2 \to 0$$

Due to $\mathcal{L}_2$ being P-Donsker, using Corollary 2.3.12 of *van der Vaart and Wellner* (1996) again, we can conclude that $(\mathbb{P}_n - P)\left\{\dot{l}_2(\tilde{\beta}, \gamma_0, \sigma_0; Y, \Delta, \boldsymbol{W})\frac{\beta_{0,n} - \beta_0}{n^{-p\nu+\epsilon}}\right\} = o_p(n^{-1/2})$. Hence $I_{1,n} = o_p(n^{-p\nu+\epsilon-1/2}) = o_p(n^{-2p\nu})$.

Similarly, by Taylor expansion, we can argue that $\mathbb{M}(\boldsymbol{\theta}_0) - \mathbb{M}(\boldsymbol{\theta}_{0,n}) \leq C\|\beta_{0,n} - \beta_0\|^2 = O(n^{-2p\nu})$, which implies that $I_{2,n} = \mathbb{M}(\boldsymbol{\theta}_{0,n}) - \mathbb{M}(\boldsymbol{\theta}_0) \geq -O(n^{-2p\nu})$. Thus we conclude that

$$\mathbb{M}_n(\hat{\boldsymbol{\theta}}_n) - \mathbb{M}_n(\boldsymbol{\theta}_0) \geq -O_p(n^{-2p\nu}) = O_p(n^{-2\min(p\nu,(1-\nu)/2)})$$

Let $\mathcal{L}_3(\eta) = \{l(\boldsymbol{\theta}) - l(\boldsymbol{\theta_0}) : \ \beta \in \mathbb{B}_n \text{ and } d(\boldsymbol{\theta}, \boldsymbol{\theta}_0) \leq \eta\}$. Using the same argument as that in the proof of consistency, we obtain that the logarithm of the ϵ-bracketing number of $\mathcal{L}_3(\eta)$, $\log N_{[]}\{\epsilon, \mathcal{L}_3(\eta), L_2(P)\}$ is bounded by $Cq_n \log(\eta/\epsilon)$. This leads to

$$J_{[]}\{\epsilon, \mathcal{L}_3(\eta), L_2(P)\} = \int_0^{\eta} \sqrt{1 + \log N_{[]}\{\epsilon, \mathcal{L}_3(\eta), L_2(P)\}} d\epsilon \leq C q_n^{1/2} \eta$$

Because Condition (C1) and (C3) guarantee the uniform boundedness of $l(\boldsymbol{\theta})$, using Theorem 3.4.1 of *van der Vaart and Wellner* (1996), the key function $\beta_n(\eta)$ in Theorem 3.2.5 of *van der Vaart and Wellner* (1996) is given by $\beta_n(\eta) = q_n^{1/2}\eta + q_n/n^{1/2}$. Note that

$$n^{2p\nu}\beta_n(1/n^{p\nu}) = n^{p\nu}n^{\nu/2} + n^{2p\nu}n^{\nu} + n^{2p\nu}n^{\nu}/n^{1/2} = n^{1/2}\left\{n^{p\nu-(1-\nu)/2} + n^{2p\nu-(1-\nu)}\right\}$$

Therefore, if $p\nu \leq (1-\nu)/2$, $n^{2p\nu}\beta_n(1/n^{p\nu}) \leq n^{1/2}$. This implies that if we choose $r_n = n^{\min(p\nu,(1-\nu)/2)}$, it follows that $r_n^2\beta_n(1/r_n) \leq n^{1/2}$ and $\mathbb{M}_n(\hat{\boldsymbol{\theta}}_n) - \mathbb{M}_n(\boldsymbol{\theta}_0) \geq -O_p(r_n^{-2})$. Hence $d(\hat{\boldsymbol{\theta}}_n, \boldsymbol{\theta}_0) = O_p(r_n^{-1}) = O_p\left[n^{-\min\{p\nu,(1-\nu)/2\}}\right]$.

2.7.2.2 Proof of Theorem 2.2

For Theorem 2.2, given the efficient score for γ in the censored linear model derived as follows:

$$\begin{aligned} \dot{l}^*_{\gamma}(\boldsymbol{\theta}) = &\left[Z - \frac{E_{\widetilde{Y},\widetilde{\Delta},\widetilde{\boldsymbol{W}}}\{\widetilde{Z}\widetilde{X}(\widetilde{Y})|\widetilde{R} = \log\int_0^t e^{\beta_0(u)X_{(}u)}du, \widetilde{\Delta} = 1\}}{E_{\widetilde{Y},\widetilde{\Delta},\widetilde{\boldsymbol{W}}}\{\widetilde{X}^2(\widetilde{Y})|\widetilde{R} = \log\int_0^t e^{\beta_0(u)X_{(}u)}du, \widetilde{\Delta} = 1\}}\right] \\ &\times \frac{1}{\sigma}\left\{\Delta\frac{\phi'(\{R(Y;\beta_0) + \gamma Z\}/\sigma)}{\phi(\{R(Y;\beta_0) + \gamma Z\}/\sigma)} - (1-\Delta)\frac{\phi(-\{R(Y;\beta_0) + \gamma Z\}/\sigma)}{\Phi(-\{R(Y;\beta_0) + \gamma Z\}/\sigma)}\right\} \end{aligned}$$

where $R(s;\beta_0) = \log\int_0^s \exp\{\beta_0(u)X(u)\}du$. Suppose that the conditions in Theorem 1 hold and $I(\boldsymbol{\theta}_0) = E_{Y,\Delta,\boldsymbol{W}}\left\{\dot{l}^*_{\gamma}(\boldsymbol{\theta}_0)^{\otimes 2}\right\}$ is nonsingular, then in order to derive the

asymptotic normality for $\hat{\gamma}$, we need to verify the conditions of the general theorem given in Appendix B of *Zhang et al.* (2009) (simplified version of the general theorem given in *Huang et al.* (1996)).

(B1) $\mathbb{P}_n \dot{l}_\gamma(\hat{\boldsymbol{\theta}}_n; Y, \Delta, \boldsymbol{W}) = o_p\left(n^{-1/2}\right)$ and $\mathbb{P}_n \dot{l}_\beta(\hat{\boldsymbol{\theta}}_n; Y, \Delta, \boldsymbol{W})(\xi_0) = o_p\left(n^{-1/2}\right)$

(B2) $(\mathbb{P}_n - P)\left\{l^*_\gamma(\hat{\boldsymbol{\theta}}_n; Y, \Delta, \boldsymbol{W}) - l^*_\gamma(\boldsymbol{\theta}_0; Y, \Delta, \boldsymbol{W})\right\} = o_p(n^{-1/2})$

(B3) $P\left\{l^*_\gamma(\hat{\boldsymbol{\theta}}_n; Y, \Delta, \boldsymbol{W}) - l^*_\gamma(\boldsymbol{\theta}_0; Y, \Delta, \boldsymbol{W})\right\} = -I(\boldsymbol{\theta}_0)(\hat{\gamma}_n - \gamma_0) + o_p(|\hat{\gamma}_n - \gamma_0|) + o_p\left(n^{-1/2}\right)$

For Condition (B1), we only need to verify that $\mathbb{P}_n \dot{l}_\beta(\hat{\boldsymbol{\theta}}_n)(\xi_0) = o_p\left(n^{-1/2}\right)$, since $\mathbb{P}_n \dot{l}_\gamma(\hat{\boldsymbol{\theta}}_n) \equiv 0$. Because ξ_0 has a bounded derivative, it is also a function with bounded variation. Then it can be easily shown using the argument in *Billingsley* (1986), that there exist a $\xi_{0,n} \in S_n(D_n, K_n, m)$ such that $\|\xi_{0,n} - \xi_0\| = O(q_n^{-1}) = O(n^{-\nu})$ and $\mathbb{P}_n \dot{l}_\beta(\hat{\boldsymbol{\theta}}_n)(\xi_{0,n}) = 0$. Therefore we can write $\mathbb{P}_n \dot{l}_\beta(\hat{\boldsymbol{\theta}}_n)(\xi_0) = I_{3,n} + I_{4,n}$, where

$$I_{3,n} = (\mathbb{P}_n - P)\, \dot{l}_\beta(\hat{\boldsymbol{\theta}}_n)(\xi_0 - \xi_{0,n})$$

$$I_{4,n} = P\left\{\dot{l}_\beta(\hat{\boldsymbol{\theta}}_n)(\xi_0 - \xi_{0,n}) - \dot{l}_\beta(\boldsymbol{\theta}_0)(\xi_0 - \xi_{0,n})\right\}$$

Let $\mathcal{L}_4 = \left\{\dot{l}_\beta(\boldsymbol{\theta})(\xi_0 - \xi) :\ \boldsymbol{\theta} \in \boldsymbol{\Theta}_n,\ \xi \in S_n(D_n, K_n, m)\ and\ \|\xi_0 - \xi\| \le n^{-\nu}\right\}$. It can be similarly argued that the ϵ-bracketing number associated with $L_2(P)$-norm is bounded by $C_3(1/\epsilon)^{d+C_4 q_n}$ which leads $\mathcal{L}_4$ being a P-Donsker due to Theorem 19.5 of *van der Vaart* (1998). Furthermore, for any $r(\boldsymbol{\theta}, \xi; x) \in \mathcal{L}_4$, $Pr^2 \to 0$ as $n \to \infty$. Hence $I_{3,n} = o_p(n^{-1/2})$ by Corollary 2.3.12 of *van der Vaart and Wellner* (1996). By Cauchy-Schwatz inequality and regularity conditions, it can be easily shown that

$$\begin{aligned} I_{4,n} \le Cd(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}_0)\|\xi_0 - \xi_{0,n}\| &= O_p\left(n^{-\min(nu,(1-\nu)/2)} n^{-\nu}\right) \\ &= O_p\left(n^{-\min(\nu(p+1),(1+\nu)/2)}\right) \\ &= o_p(n^{-1/2}) \end{aligned}$$

So (B1) holds.

Condition (B2) holds by similarly verifying that the class $\mathcal{L}_5(\eta) = \left\{ l^*_\beta(\boldsymbol{\theta}) - l^*_\beta(\boldsymbol{\theta}_0) : \boldsymbol{\theta} \in \boldsymbol{\Theta}_n \text{ and } d(\boldsymbol{\theta}, \boldsymbol{\theta}_0) \leq \eta \right\}$ is P-Donsker and for any $r(\boldsymbol{\theta}) \in \mathcal{L}_5(\eta)$, $Pr^2 \to 0$ as $\eta \to 0$.

Condition (B3) can be easily established using Taylor expansion and the convergence rate derived in Theorem 2.1.

2.7.3 The Efficient Score Functions

In this section, we study the semiparametric efficiency bound of the proposed estimator for finite dimensional parameter γ. The semiparametric efficient score for γ of a single observation can be written as

$$\dot{l}^*_\gamma(\boldsymbol{\theta}_0) = \dot{l}_\gamma(\boldsymbol{\theta}_0) - \dot{l}_\beta(\boldsymbol{\theta}_0)(h^*)$$

where

$$h^*(t) = \frac{E_{\widetilde{Y},\widetilde{\Delta},\widetilde{\boldsymbol{W}}}\left[\widetilde{Z}\widetilde{X}(\widetilde{Y})|\widetilde{R}_0 = \log\int_0^t e^{\beta_0(u)X(u)}du, \widetilde{\Delta}=1\right]}{E_{\widetilde{Y},\widetilde{\Delta},\widetilde{\boldsymbol{W}}}\left[\widetilde{X}(\widetilde{Y})\widetilde{X}(\widetilde{Y})|\widetilde{R}_0 = \log\int_0^t e^{\beta_0(u)X(u)}du, \widetilde{\Delta}=1\right]}$$

To see this more clearly, we only need to proof that for the submodel $\beta(t) + \eta h^*(t)$ with the direction $h^* \in \mathcal{H}$, the following equation will be held $\forall h \in \mathcal{H}$:

$$E_{Y,\Delta,\boldsymbol{W}}\left[\left\{\dot{l}_\gamma(\boldsymbol{\theta}_0) - \dot{l}_\beta(\boldsymbol{\theta}_0)(h^*)\right\}\dot{l}_\beta(\boldsymbol{\theta}_0)(h)\right] = 0 \tag{2.6}$$

From the form of score functions, we have

$$\dot{l}_\gamma(\boldsymbol{\theta}_0) - \dot{l}_\beta(\boldsymbol{\theta}_0)(h^*)$$

$$\begin{aligned}
=&\frac{\Delta Z}{\sigma_0}\frac{\phi_0^{(1)}}{\phi_0} - \frac{(1-\Delta)Z}{\sigma_0}\frac{\phi_0}{(1-\Phi_0)} + \Delta Z - \Delta\frac{\int_0^Y Ze^{\beta_0(u)X(u)}du}{\int_0^Y e^{\beta_0(u)X(u)}du} \\
&- \Delta h^*(Y)X(Y) + \Delta\frac{\int_0^Y h^*(u)X(u)e^{\beta_0(u)X(u)}du}{\int_0^Y e^{\beta_0(u)X(u)}du} \\
&- \frac{\Delta}{\sigma_0}\frac{\phi_0^{(1)}}{\phi_0}\frac{\int_0^Y h^*(u)X(u)e^{\beta_0(u)X(u)}du}{\int_0^Y e^{\beta_0(u)X(u)}du} + \frac{(1-\Delta)}{\sigma_0}\frac{\phi_0}{(1-\Phi_0)}\frac{\int_0^Y h^*(u)X(u)e^{\beta_0(u)X(u)}du}{\int_0^Y e^{\beta_0(u)X(u)}du}
\end{aligned}$$

$$\begin{aligned}
=&\Delta\{Z - h^*(Y)X(Y)\} \\
&+ \frac{\int_0^Y \{Z - h^*(u)X(u)\}e^{\beta_0(u)X(u)}du}{\int_0^Y e^{\beta_0(u)X(u)}du} \times \left\{-\Delta + \frac{\Delta}{\sigma_0}\frac{\phi_0^{(1)}}{\phi_0} - \frac{(1-\Delta)}{\sigma_0}\frac{\phi_0}{(1-\Phi_0)}\right\}
\end{aligned}$$

where $\Phi_0 = \Phi\left(\frac{R_0+\gamma_0 Z}{\sigma_0}\right)$, $\phi_0 = \phi\left(\frac{R_0+\gamma_0 Z}{\sigma_0}\right)$ and $\phi_0^{(1)} = \phi^{(1)}\left(\frac{R_0+\gamma_0 Z}{\sigma_0}\right)$, with $R_0 = \log\int_0^Y e^{\beta_0(u)X(u)}du$. Notice that

$$\begin{aligned}
&E_{Y,\Delta,\boldsymbol{W}}\{Z - h^*(t)X(t)\} \\
=&E\left(E\left\{Z - h^*(t)X(t)|R_0 = \log\int_0^t e^{\beta_0(u)X(u)}du, \Delta = 1\right\}\right) \\
=&E\left(E\left\{Z - \frac{E_{\widetilde{Y},\widetilde{\Delta},\widetilde{\boldsymbol{W}}}\left[\widetilde{Z}\widetilde{X}(t)|\widetilde{R}_0 = \log\int_0^t e^{\beta_0(u)X(u)}du, \widetilde{\Delta} = 1\right]}{E_{\widetilde{Y},\widetilde{\Delta},\widetilde{\boldsymbol{W}}}\left[\widetilde{X}(t)\widetilde{X}(t)|\widetilde{R}_0 = \log\int_0^t e^{\beta_0(u)X(u)}du, \widetilde{\Delta} = 1\right]}X(t)\right.\right. \\
&\left.\left.|R_0 = \log\int_0^t e^{\beta_0(u)X(u)}du, \Delta = 1\right\}\right) \\
=&0
\end{aligned}$$

Equation (2.6) then follows.

CHAPTER III

Estimation of the Optimal Dynamic Treatment Regime from Observational Data Using Flexible Weighting Models

For many chronic diseases, a patient usually has to undergo multiple stages of treatment. Therefore, the goal of identifying the optimal dynamic treatment regime is very appealing as it allows a patient to receive the most appropriate treatment and dose assignment based on his/her evolving history of disease status, treatment, and other time-dependent clinical covariates. The challenge is to find the best regime amongst a set of defined regimes from observational data, in which the actual regime being followed by each subject is not well characterized. Inverse probability weighting (IPW) based estimators are used in the estimation of causal parameters as an efficient way to utilize information from an observational study. In this chapter, we focus on the case where 1) the outcome is time-to-event, and 2) some of the covariates are time-varying, and possibly follow a complicated pattern. We consider a class of dynamic treatment regimes that are fully determined by the longitudinal covariates. A novel Random Forest based IPW scheme is proposed to adjust for the complexity in the mechanism of adherence to the regime of interest. The optimal regime is then identified as the one with the largest restricted mean survival time. The performance

of the proposed method is assessed through simulation studies, which are designed to mimic the study of salvage treatment in the management of recurrent prostate cancer. The proposed method can efficiently accommodate complex and possibly unknown treatment/adherence mechanisms, and it is robust to cases where the proportional hazard assumption is violated. We apply the method to a longitudinal prostate cancer study.

3.1 Introduction

Prostate cancer is the most commonly diagnosed cancer among American men. After initial treatment, patients with clinically localized prostate cancer are actively monitored for possible cancer recurrence. Levels of prostate-specific antigen (PSA) are repeatedly measured, and a dramatic rise of this prognostic marker is considered as an indicator for increased risk of clinical cancer recurrence (*Zagars and von Eschenbach*, 1993). Salvage androgen deprivation therapy (SADT) would be applied as an effective way to reduce the recurrence rate in these situations. Clinically, "when to start SADT" is usually determined by the physician's own experience and personal judgment. However, the current biological understanding of the effect of SADT is not sufficient to determine the optimal time to start because early initiation of SADT has both risks and benefits. If SADT is given too early when PSA values are still low, it is wasted during the time when the patient is at low risk while later on the beneficial effect wears off as the patient develops resistance. On the other hand if the patient waits to start SADT until PSA is very high, it becomes less effective because the cancer is already well established and may have already spread to other sites by that time.

In this chapter, we try to address the above question and make recommendations on when would be the optimal time to start SADT in terms of prolonging patient's cancer recurrence free survival using flexible weighting models based on the longi-

tudinally collected observational data. This can be framed as a dynamic treatment regime (DTR) (*Murphy*, 2003; *Robins*, 2004), in which dose or treatment is continuously modified according to a patient's current history and disease status. Identifying such optimal dynamic decision rules offers an effective vehicle for personalized management of chronic diseases, for which a patient typically has to be treated at multiple stages. DTRs help physicians to adapt the type, dosage, and timing of the treatment at each stage to the evolving disease status, treatment and clinical history, and thus provide better care by tailoring the treatment individually based on clinical evidence (*Wagner et al.*, 2001).

In recent years, a large literature has been developed on both designs and analytic tools for DTRs. Although data from sequential multiple assignment randomized trials (SMARTs) are desirable (*Murphy*, 2005), observational studies are the most common source of data for complex disease studies, and a great deal of effort in statistical research has concentrated on how to make best use of observational data to construct DTRs (*Wang et al.*, 2012). Careful thoughts and assumptions are required to make valid causal inference, especially on how the observational data may restrict the set of DTRs that can be assessed, which are called the feasible (*Robins*, 1994) or viable (*Wang et al.*, 2012) DTRs. *Murphy* (2003) and *Robins* (2004) generalized the G-estimation method of structural nested models (*Robins*, 1986, 1989b, 1997, 2002) for optimal treatment regime estimation and this approach is efficient when all models are correctly specified. Q- and A-learning (*Watkins and Dayan*, 1992; *Robins*, 2004; *Murphy*, 2005) provides a powerful solution especially when the decision rules depend on multiple covariates. However, the computational burden would increase as the number of decision-making stages increases. For survival outcomes, treatment decisions are essentially made on a continuous time base. In these situations, methods based on inverse probability weighting (IPW) are easier to conduct and provide cer-

tain robustness against model mis-specifications (*Hernán et al.*, 2006; *Robins et al.*, 2008; *Orellana et al.*, 2010a,b). But the validity of these approaches still rely on expert knowledge to comprehensively understand the treatment assignment mechanism. This is challenging in our prostate cancer example, because the treatment assignment in this observational study is not completely understood, and the typical proportional hazard assumption is likely to be violated as well when comparing different regimes. Therefore, we propose a method which incorporates flexible modeling to account for the unknown and potentially complicated treatment assignment mechanism. Our method does not require model assumptions for the recurrence hazard either and can estimate the optimal DTR even when the hazards are not proportional.

Specifically, our proposed method starts with a class of pre-specified viable dynamic treatment regimes, and evaluates them based on observational data with a time to event outcome, while imposing minimal assumptions on the structure of the models. We proceed by artificially censoring subjects when they become noncompliant with a defined regime under investigation. This censoring potentially induces a bias which we correct by using a modified version of Inverse Probability of Censoring Weighting (IPCW) (*Robins*, 1993). We focus on the survival distribution as the target quantity of interest and use a weighted version of Nelson Aalen estimator with a flexible data driven weighting scheme to (1) accurately estimate the survival distribution under a pre-defined dynamic treatment regime of interest; and (2) compare the survival distribution under different viable dynamic treatment regimes.

The rest of this chapter is organized as follows. In Section 3.2, we introduce the notation under the framework of causal inference, and define the dynamic treatment regimes of interest. In Section 3.3, we establish our method of the weighted Nelson Aalen estimator and show asymptotic properties of the estimator. In Section 3.4,

we demonstrate the validity of the proposed method through a simulation study, followed by an application illustration in Section 3.5 by applying the proposed method to a prostate cancer recurrence dataset with the goal of comparing several clinically meaningful SADT regimens. Finally, we conclude with a discussion in Section 3.6, some technical details are summarized in Section 3.7.

3.2 Notation and Dynamic Treatment Regimes

Suppose that there are N patients in a clinical trial or an observational study, each of whom was observed at baseline $t_0 = 0$, and longitudinally at regular intervals $t_1, t_2, \cdots, t_k, \cdots$ until the end of K time intervals t_K (study end) or until the event of interest (e.g. cancer recurrence or death in our motivating example) occurs, whichever is earlier. To simplify the problem, we assume that there is no censoring (e.g. possible loss to follow-up) other than the administrative censoring at time t_K. Before having the event or dropping out, patients will come to the clinic once during each of the K intervals and have their time dependent covariates (e.g., the PSA level) measured. Treatment decisions, i.e. whether to start SADT, were made soon after each clinic visit and at no other time. Assume that the subjects in the cohort are a random sample from a large population of interest. For patient i at time t_k, with $i = 1, \cdots, N$ and $k = 0, \cdots, K$, let $\boldsymbol{L}_{ik}$ denote the time dependent covariates observed at the kth clinic visit. When $k = 0$, we follow the convention in the literature and use $\boldsymbol{L}_{i0}$ to denote all the baseline covariates. Let R_{ik} denote a binary indicator for event occurrence. R_{ik} takes the value 1 if the patient has experienced the event of interest by time t_k and 0 otherwise. Let A_{ik} denote the kth-specific treatment prescription which we assume takes values in a finite set $\mathcal{A}_k = \{0, 1\}$. Since our main interest is when would be the optimal timing to start the treatment, we assume that the patient will stay on the treatment once it is initiated. That is, the treatment can only go from 0 to 1, but not 1 to 0.

For a given subject i at time t_k, $k < K$ (if no event observed before the kth visit), the observational data would be A_{ik} and $\boldsymbol{O}_{ik} = (\boldsymbol{L}_{ik}, R_{ik})$. For simplicity, we will suppress the subject index i in the future when no confusion exists, and use overbars to denote the history of the variable up to the indexed time. Furthermore, we use capital letters to refer to random variables or vectors, while lower-case letters are used to denote the observed values of the corresponding random variables. For example, the observational data up to time t_k is denoted as $(\overline{\boldsymbol{A}}_{k-1}, \overline{\boldsymbol{O}}_k) = (\boldsymbol{L}_0, A_0, \boldsymbol{L}_1, R_1, A_1, \boldsymbol{L}_2, R_2, \cdots, A_{k-1}, \boldsymbol{L}_k, R_k)$ and a possible observed treatment history up to time t_k is denoted as $\overline{\boldsymbol{a}}_k = (a_0, \cdots, a_k) \in \mathcal{A}_0 \times \cdots \times \mathcal{A}_k = \overline{\boldsymbol{\mathcal{A}}}_k$.

Here we consider a causal framework with treatment regime specific counterfactual outcomes (*Robins*, 1986). For each patient, $\boldsymbol{L}_0$ is measured before the first treatment decision, so it is always observed. Let $\boldsymbol{L}_k^C(\overline{\boldsymbol{a}}_{k-1})$ denote the counterfactual covariate information that would be observed at time t_k were the patient to receive treatment history $\overline{\boldsymbol{a}}_{k-1}$ regardless what treatment sequence he actually followed up to t_k, and similarly let $R_k^C(\overline{\boldsymbol{a}}_{k-1})$ denote the corresponding counterfactual event status at time t_k under treatment history $\overline{\boldsymbol{a}}_{k-1}$. Then the counterfactual observations are denoted as

$$\boldsymbol{W}^C = \left\{ \boldsymbol{L}_0, \boldsymbol{L}_1^C(a_0), R_1^C(a_0), \cdots, \boldsymbol{L}_K^C(\overline{\boldsymbol{a}}_{K-1}), R_K^C(\overline{\boldsymbol{a}}_{K-1}), \text{ for all } \overline{\boldsymbol{a}}_{K-1} \in \overline{\boldsymbol{\mathcal{A}}}_{K-1} \right\}$$

Notice $\boldsymbol{W}^C$ includes all the counterfactual observation up to time t_K. However, in practice, the counterfactual observation would only be meaningful up to the time of the counterfactual event, if the counterfactual event is before t_K. In fact, the pieces after that time will never be used in the methods that we describe here and we only include them in $\boldsymbol{W}^C$ for ease of notation.

A dynamic treatment regime $g = \{g_k : k = 0, \cdots, K-1\}$ is a sequential rule for determining the next treatment prescription A_k at each time t_k. The rule $g_k(\overline{\boldsymbol{o}}_k) \in \mathcal{A}_k$ for $k = 0, \ldots, K-1$ may depend on part or all of the recorded health information about the patient's health up to and including time t_k. The optimal regime would be a regime that maximizes the expected utility function if all patients in the population follow this rule. Note the expected utility can depend on $\overline{\boldsymbol{o}}_k$ as well as the regime g, thus it provides a personalized treatment decision. For time to event outcomes, it may be unrealistic to expect the proportional hazard assumption to hold across different regimes, thus here we propose to use the restricted mean survival time as the utility function. If we denote the survival time by T, then for some arbitrary time bound $T_{\max}$, the restricted survival time is defined as $\min(T, T_{\max})$, and the restricted mean lifetime can be represented as $E\{\min(T, T_{\max})\}$. It can be shown that the restricted mean survival time can be represented as the area under the survival curve up to $T_{\max}$, $\mu = \int_0^{T_{\max}} S(t)dt$. Here $T_{\max}$ can be chosen as the administrative censoring time, or longest follow-up time, in our case, t_K. For our specific example, $L_k = \text{PSA}_k$ for $k = 1, 2, \cdots, K$, while $\boldsymbol{L}_0 = (\text{PSA}_0, V_0)$ where V_0 is the baseline covariate (e.g. T-stage of the patient at time t_0). In reality, following the initial treatment of prostate cancer with radiation therapy, PSA will typically decrease, then either remain stable or follow an increasing trend. A simple treatment regime could be that the patient starts SADT the first time his PSA level is above a threshold b, after the decreasing phase has past (practically, we require that the current PSA value to be larger than the value at the previous visit). To formalize it, we only consider the class of regimes $\mathcal{G} \equiv \left\{g^b : b \in \mathcal{R}\right\} = \left\{(g_0^b, \cdots, g_{K-1}^b) : b \in \mathcal{R}\right\}$ where the treatment indicator at time t_k is

$$g_k^b(\overline{\boldsymbol{O}}_k^C) = \begin{cases} 0 & \text{if } A_{k-1}^C = 0\,, \mathrm{PSA}_k^C \leq \mathrm{PSA}_{k-1}^C\,, \text{or } \mathrm{PSA}_k^C \leq b. \\ 1 & \text{if } A_{k-1}^C = 0\,, \mathrm{PSA}_k^C > \mathrm{PSA}_{k-1}^C\,, \text{and } \mathrm{PSA}_k^C > b. \\ 1 & \text{if } A_{k-1}^C = 1 \end{cases} \tag{3.1}$$

In this setting, a treatment regimen g^b is fully defined by cut-off value b. The counterfactual data used in the definition of g_k^b in (3.1) is specific to the case where all patients follow g^b. If we denote the restricted mean lifetime under this regime g^b to be μ^b, the optimal regime is $g^{\text{opt}} = \arg\max_{\{g^b \in \mathcal{G}\}} \mu^b$.

Definition (3.1) is based on the assumption that we observe the counterfactual data under all regimes $g^b \in \mathcal{G}$. In practice, not all of them can be observed for each patient, because each patient is observed to experience one and only one treatment history. So instead of calculating μ^b from $\overline{\boldsymbol{O}}_K^C(g^b)$ as if everyone follows g^b, we need to estimate it from the observed data $\overline{\boldsymbol{O}}_K$. To make this possible, we follow *Orellana et al.* (2010a) and make the following standard assumptions. (i) The consistency assumption states that $\boldsymbol{O}_k = \boldsymbol{O}_k^C(\overline{\boldsymbol{A}}_{k-1}) = \sum_{\overline{\boldsymbol{a}}_{k-1} \in \overline{\boldsymbol{\mathcal{A}}}_{k-1}} \boldsymbol{O}_k^C(\overline{\boldsymbol{a}}_{k-1}) I(\overline{\boldsymbol{A}}_{k-1} = \overline{\boldsymbol{a}}_{k-1})$ for $k = 1, \cdots, K$; that is, a patient's observed covariates and outcomes are the same as the potential ones he would exhibit under the treatment history actually received. (ii) No unmeasured confounder assumption (NUCA) implies that $\boldsymbol{W}^C$ is independent of A_k conditional on $(\overline{\boldsymbol{O}}_k, \overline{\boldsymbol{A}}_{k-1})$ for $k = 1, \cdots, K$. (iii) The positivity assumption says that $\Pr(A_k(g) | \overline{\boldsymbol{O}}_k(g), \overline{\boldsymbol{A}}_{k-1}(g)) \geq \epsilon > 0$ for $k = 1, \cdots, K$ with probability 1 for an arbitrary small positive constant ϵ; it basically guarantees that in the counterfactual world where everyone follows regime g, if there were patients with history $\overline{\boldsymbol{o}}_k$ and $\overline{\boldsymbol{a}}_{k-1}$ that would be assigned to treatment a_k, then, in the observational world, there must be some patients with the same history ($\overline{\boldsymbol{o}}_k$ and $\overline{\boldsymbol{a}}_{k-1}$) who are observed to take a_k. Using these assumptions, we have $p_{\boldsymbol{L}_k^C(\overline{\boldsymbol{a}}_{k-1}) | \overline{\boldsymbol{L}}_{k-1}^C(\overline{\boldsymbol{a}}_{k-2}), \overline{\boldsymbol{R}}_{k-1}^C(\overline{\boldsymbol{a}}_{k-2})}(\boldsymbol{l}_k | \bar{\boldsymbol{l}}_{k-1}, \overline{\boldsymbol{r}}_{k-1}) = p_{\boldsymbol{L}_k | \overline{\boldsymbol{L}}_{k-1}, \overline{\boldsymbol{R}}_{k-1}, \overline{\boldsymbol{A}}_{k-1}}(\boldsymbol{l}_k | \bar{\boldsymbol{l}}_{k-1}, \overline{\boldsymbol{r}}_{k-1}, \overline{\boldsymbol{a}}_{k-1})$ and $p_{R_k^C(\overline{\boldsymbol{a}}_{k-1}) | \overline{\boldsymbol{L}}_k^C(\overline{\boldsymbol{a}}_{k-1}), \overline{\boldsymbol{R}}_{k-1}^C(\overline{\boldsymbol{a}}_{k-2})}(r_k | \bar{\boldsymbol{l}}_k, \overline{\boldsymbol{r}}_{k-1}) =$

$p_{R_k|\overline{\boldsymbol{L}}_k,\overline{\boldsymbol{R}}_{k-1},\overline{\boldsymbol{A}}_{k-1}}(r_k|\overline{\boldsymbol{l}}_k,\overline{\boldsymbol{r}}_{k-1},\overline{\boldsymbol{a}}_{k-1})$, where $p(\cdot)$ denotes the probability function. Therefore we are able to make inference about μ^b using only the observed data $(\overline{\boldsymbol{A}}_{K-1},\overline{\boldsymbol{O}}_K)$ (see Section 3.7 for details of the proof).

3.3 Method

For most subjects in the observational data, we do not know if they are adhering to a treatment regime, and if they are following a specific regime, we may not know what that regime is, other than what can be inferred from the observed data. However, we wish to estimate the specific survival experience that the whole cohort of subjects would have had if they had truly been adherent to g^b.

3.3.1 Inverse Probability of Adherence Weighting

We proceed by artificially censoring subjects at their first non-adherent visit. For a specific regime g^b, let $C_k^b = A_k g_k^b(\overline{\boldsymbol{O}}_k) + (1-A_k)\{1-g_k^b(\overline{\boldsymbol{O}}_k)\}$ be the indicator of adherence at time $k = 0, \cdots, K-1$, which is 1 if the patient's observed treatment status at time k is the same as if he is following regime g^b (adherent), and is 0 if the observed and regime g^b specific treatment statuses are different. The patient would follow the regime until time k if $\overline{\boldsymbol{C}}_k^b = \bar{\mathbf{1}}$, where we use overbars to represent the history of a covariate, for example, $\overline{\boldsymbol{C}}_t^b = \{C_u^b;\ 0 \le u < t\}$ is a subject's adherence history up to time t. $\bar{\mathbf{1}}$ is a vector of 1's the same length as $\overline{\boldsymbol{C}}_k^b$. The patient is censored at time $t = \min_{t_k}\{C_k^b = 0 \text{ for } k = 0, \cdots, K\}$ to generate the regime g^b adherence dataset, i.e. we only consider the time period when a patient's observed treatment is consistent with regime g^b. Notice here, for a patient who partially follows the regime of interest, we will include him in the regime g^b adherence dataset, but only up to the first time he stops to follow the regime.

In order to correctly estimate the regime g^b specific counterfactual survival function

as well as μ^b, we propose to adjust for the bias induced from the artificial censoring by weighting each subject by their Inverse Probability of Adherence Weights,

$$w_k^b = \frac{I\left(\overline{\boldsymbol{C}}_k^b = \bar{\mathbf{1}}\right)}{\prod_{j=0}^{k} P\left(C_j^b = 1 | \overline{\boldsymbol{C}}_{j-1}^b = \bar{\mathbf{1}}, \overline{\boldsymbol{O}}_j = \overline{\boldsymbol{o}}_j\right)} \tag{3.2}$$

Briefly, at each time point, each adherent subject is weighted by the inverse of the probability that they remained adherent given their measured covariate history, and thus account for themselves as well as other similar subjects who were non-adherent and artificially censored. Because we are considering discretized visits, the probability of adherence for a patient at time k is then calculated as the multiplication of the conditional probabilities of adherence at each time point j ($j = 0, \cdots, k$) given that he remained adherent up to time $j-1$. For ease of notation, we define $C_{-1}^b = 1$ for all patients. Following *Robins* (1993), in practice, we use the stabilized version of the weights:

$$sw_k^b = \prod_{j=0}^{k} \frac{P\left(C_j^b = 1 | \overline{\boldsymbol{C}}_{j-1}^b = \bar{\mathbf{1}}, \boldsymbol{L}_0 = \boldsymbol{l}_0\right)}{P\left(C_j^b = 1 | \overline{\boldsymbol{C}}_{j-1}^b = \bar{\mathbf{1}}, \overline{\boldsymbol{O}}_j = \overline{\boldsymbol{o}}_j, \overline{\boldsymbol{A}}_{j-1} = \overline{\boldsymbol{a}}_{j-1}\right)} \tag{3.3}$$

The model in the numerator includes only baseline covariates and serves to stabilize (i.e. reduce the variability of) the weights.

3.3.2 Random Forest Regression

In order to make unbiased estimation based on these weights, it is important that the model for adherence is correctly specified. Traditionally, the probability models in the numerator and denominator of Equation (3.3) are estimated by fitting logistic regression models to the pooled data from all possible pieces of person-times (*Hernán et al.*, 2006). In our case, the regime rules are defined based on the PSA value, and since the PSA has a non-linear relationship with time, it would be hard to expect that a simple logistic regression model would capture the association between adher-

ence and covariates. We propose to use Random Forest regression to model these probablities.

Random Forests (*Breiman*, 2001) is a non-parametric classification and regression method. It employs a combination of resampling and ensembles single tree based models to give superior performance in both classification and regression. Compared to the parametric logistic regression, it provides more flexibility without imposing many structure assumptions. In detail, as the number of patients at risk decreases over time, it may not be efficient to fit separate conditional probability models at each time point. We proceed by pooling the data of all person-time pieces together to do the Random Forest regression, while putting in time as a covariate to account for the fact that these conditional probabilities vary over time. The numerator and denominator in Equation (3.3) will be modeled separately. For the denominator, we first fit model for observed treatment assignment mechanism, $P\left(A_k = 1|\overline{\boldsymbol{A}}_{k-1} = \overline{\mathbf{0}}, \overline{\boldsymbol{O}}_k = \overline{\boldsymbol{o}}_k\right)$ with all the observed data available up to the first time point when the patient is on treatment, i.e. person-time pieces up to t_m with $m = \max\{k : A_{k-1} = 0,\, k \leq K\}$ then the target probabilities can be calculated from

$$\begin{aligned}
&P\left(C_k^b = 1|\overline{\boldsymbol{C}}_{k-1}^b = \overline{\mathbf{1}}, \overline{\boldsymbol{O}}_k = \overline{\boldsymbol{o}}_k\right) \\
=&P\left(A_k = 1|\overline{\boldsymbol{A}}_{k-1} = \overline{\boldsymbol{a}}_{k-1}, \overline{\boldsymbol{O}}_k = \overline{\boldsymbol{o}}_k\right) I\{g_k^b(\overline{\boldsymbol{o}}_k) = 1\} \\
&+ P\left(A_k = 0|\overline{\boldsymbol{A}}_{k-1} = \overline{\boldsymbol{a}}_{k-1}, \overline{\boldsymbol{O}}_k = \overline{\boldsymbol{o}}_k\right) I\{g_k^b(\overline{\boldsymbol{o}}_k) = 0\},
\end{aligned}$$

Since the treatment mechanism model for the observational data is the same regardless of which regime is under investigation, this allows us to obtain the probability of adherence for various regimes while only fitting the pooled Random Forest model once. However, we may not be able to do the same thing for the numerator, as it requires that this model should not be related to the time-dependent treatment

status $\overline{\boldsymbol{A}}_{j-1}$. We proceed by directly fitting the model for regime specific adherence mechanism $P\left(C_j^b = 1 | \overline{\boldsymbol{C}}_{j-1}^b = \bar{\mathbf{1}}, \boldsymbol{L}_0 = \boldsymbol{l}_0\right)$ in the adherence cohort for each regime of interest. More discussion on this issue can be found in Section 3.6. Here, the Random Forest modeling is done using the R function *randomForest* with all default settings except for the number of trees per forest and number of variables included at each split, which we set at $ntree = 1000$ and $mtry = 1$. The adherence probabilities are given from the "out-of-bag" prediction.

3.3.3 Weighted Nelson Aalen Estimator

To reduce the sample selection bias due to adherence, we assign a time-dependent weight sw_{ik}^b for the ith subject at time t_k, while his data are compatible with regime g^b under consideration, $i = 1, \cdots, N$. Define the weighted number of events and the weighted number at risk at time t_k as

$$d_k^b = \sum_{i:T_i = t_k} sw_{ik}^b \delta_i^b \text{ and } \quad Y_k^b = \sum_{i:T_i \geq t_k} sw_{ik}^b$$

and then the formula that defines the weighted estimator for the regime g^b-specific survival function is $\hat{S}^b(t) = \exp\{-\hat{\Lambda}^b(t)\}$, where $\hat{\Lambda}^b(t) = \sum_{t_j \leq t} d_j^b / Y_j^b$ is the cumulative hazard function. The estimated restricted mean survival time for regime g can be calculated as $\hat{\mu}^b = \int_0^{t_k} \hat{S}^b(t) dt$. The optimal DTR is the one that maximize $\hat{\mu}^b$, that is $\hat{g}^{\text{opt}} = \arg\max_{g^b \in \mathcal{G}} \hat{\mu}^b$.

3.3.4 Property of the Estimator

To derive the large-sample properties of $\hat{\mu}^g(t)$, we need the following additional regularity conditions. Specifically, we require for $i = 1, \cdots, N$,

(a) The observed data $(\overline{\boldsymbol{A}}_{i,K-1}, \overline{\boldsymbol{O}}_{i,K})$ are independent and identically distributed.

(b) $\int_0^\tau \lambda_0^b(t) dt < \infty$ where $\lambda_0^b(t)$ is the true marginal hazard for any regime $g^b \in \mathcal{G}$.

(c) For the true marginal survival function $S_0^b(t)$, we assume there exist continuous first-order derivatives in t and bounded second partial derivatives (uniformly in $t \in (0, t_K]$).

Along with the assumptions from Section 3.2, we can show that the proposed estimator can give asymptotically consistent estimation to the counterfactual quantities of interest. The main results are summarized in the following two theorems.

THEOREM 3.1 For any patient $i = 1, \cdots, n$, and $k = 1, \cdots, K-1$, we have

$$P(C_k^b = 1 | \overline{\boldsymbol{C}}_{k-1}^b = \bar{\mathbf{1}}, \overline{\boldsymbol{O}}_k = \overline{\boldsymbol{o}}_k) = P(A_k = 1 | \overline{\boldsymbol{A}}_{k-1} = \overline{\boldsymbol{a}}_{k-1}, \overline{\boldsymbol{O}}_k = \overline{\boldsymbol{o}}_k) I\{g_k^b(\overline{\boldsymbol{o}}_k) = 1\}$$
$$+ P(A_k = 0 | \overline{\boldsymbol{A}}_{k-1} = \overline{\boldsymbol{a}}_{k-1}, \overline{\boldsymbol{O}}_k = \overline{\boldsymbol{o}}_k) I\{g_k^b(\overline{\boldsymbol{o}}_k) = 0\}$$

The proof of Theorem 3.1 is outlined in Section 3.7. It connects the treatment model and the model for the adherence to regime g^b. So the weight calculated in Section 3.3.2 can consistently estimate the weight of adherence. Instead of directly modeling the regime specific adherence mechanism for every regime, it only requires fitting one model for all the regimes of interest.

THEOREM 3.2. Under assumption (i) - (iii) in Section 3.2 and the regularity conditions (a) to (c), if the time dependent weights in Equation (3.3) are consistent estimators of the true weights, then $\hat{\mu}^b(t)$ converges almost surely and uniformly in $t \in (0, t_K]$ to $\mu_0^b(t)$.

Theorem 3.2 assures that as long as the weights are correctly estimated, the proposed weighted Nelson Aalen estimator will give a consistent estimate for the regime specific survival function in the counterfactual world, and the estimation of the utility

μ_0^b is also consistent. This establishes the consistency of the estimated optimal DTR by maximizing $\hat{\mu}^b(t)$ among $g^b \in \mathcal{G}$.

3.4 Simulation

In order to evaluate the performance of the proposed method, we conduct simulation studies where we have access to the fully adherent data for each defined regimes, and thus we can use these simulated counterfactual data to calculate the real causal outcome when everyone follows the defined regime as the "gold standard", and compare it to the proposed estimator. To illustrate the role of flexible modeling in the weight estimation, we compare the proposed method with an approach where the weights are estimated from pooled logistic regression (*Hernán et al.*, 2006). In detail, for the denominator of the weight in Equation (3.3), we will fit the treatment model as

$$\text{logit } P\left(A_k = 1 | \overline{\boldsymbol{A}}_{k-1} = \bar{\mathbf{0}}, \bar{\boldsymbol{O}}_k = \bar{\boldsymbol{o}}_k\right) = h_1(k) + \boldsymbol{\beta}_1 \boldsymbol{O}_k + \boldsymbol{\beta}_2^T \boldsymbol{V}_0$$

and for the numerator, we fit the regime specific adherence model as

$$\text{logit } P\left(C_j^b = 1 | \overline{\boldsymbol{C}}_{j-1}^b = \bar{\mathbf{1}}, \boldsymbol{V}_0 = \boldsymbol{v}_0\right) = h_2(k) + \boldsymbol{\beta}_3^T \boldsymbol{V}_0$$

where we consider time-dependent intercept for both models, i.e. $h_1(k)$ and $h_2(k)$. In the specific models below, $\overline{\boldsymbol{O}}_k$ is the valure of $\log \text{PSA}$ measured at time t_k, and $\boldsymbol{V}_0$ is the Tstage measured at baseline, in detail, we represent is as a 2-dimensional vector $\boldsymbol{V}_0 = (I(\text{Tstage} = 2), I(\text{Tstage} \geq 3))^T$. We use cubic spline with 2 internal knots to estimate the intercept terms.

We consider two scenarios: (1) a simple case where the true treatment is assigned

according to a logistic model, and (2) a more realistic model with a more complicated PSA model and mechanism for treatment assignment. For each scenario, we simulate 500 datasets each with 2000 subjects, respectively.

3.4.1 Scenario 1:

We first present a case with a PSA model where the logarithm of PSA is linear in time. The treatment would be generated such that it would be in favor of the pooled logistic model. From now on, other than the time-dependent PSA, we also consider the baseline covariate V_0 as the baseline T-stage.

3.4.1.1 Longitudinal PSA Values

Let $P_i(t)$ denote the observed PSA value, for subject i at $t \in (0, t_K]$ years after the start of follow-up (we choose $t_K = 15$ year). We measure PSA every year at $t = 0, 1, \cdots, 15$, and the observed PSA values are simulated from the following linear mixed model:

$$\text{logP}_i(t) = \text{logPSA}_i(t) + \epsilon_{it} = (\alpha_0 + a_{i0}) + (\alpha_1 + a_{i1})t + \epsilon_{it} \tag{3.4}$$

where $(\alpha_0, \alpha_1) = (-3.0, 0.3)$ are fixed effect parameters, (a_{i0}, a_{i1}) are subject-specific random effects. At a given time t, we assume the measurement error $\epsilon_{it} \sim N(0, \sigma^2)$ where $\sigma^2 = 0.1$, and we assume the random effects $(a_{i0}, a_{i1}) \sim \text{MVN}(0, \Sigma)$, where $\Sigma = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 0.25 \end{pmatrix}$. Here we also truncate a_{i1} at -0.1, such that the slope of $\log \text{PSA}_i(t)$ would not be negative, and treatment assignment according to the model below would be clinically reasonable.

Note that, given the random effects, and in the absence of any treatment after time $t = 0$, $\text{PSA}_i(t)$, which is referred to as the true PSA without treatment, is known and

non-random. In contrast, the observed PSA value, $P_i(t)$, are subject to measurement error, and not observed after the earliest of salvage time or recurrence.

3.4.1.2 Different Treatment Regimes and Observed Treatment Time

For the counterfactual outcomes, we consider a finite number of dynamic SADT treatment regimes as described in Definition (3.1). We consider 10 regimes with $\{b_1, b_2, \cdots, b_{10}\} = \{-0.5, 0, \cdots, 4.0\}$. Thus, according to subject i's PSA trajectory, we can calculate the regime specific time to initiate SADT as $U_i^{b_1}, U_i^{b_2}, \cdots, U_i^{b_{10}}$ for all 10 regimes. Here we assume that the observed treatment assignment for subject i follows one of the above 10 regimes with the threshold value B, where B follows a discrete uniform distribution for $\{b_1, b_2, \cdots, b_{10}\}$. The observed treatment time for subject i is then $U_i = U_i^B$.

3.4.1.3 Model for Recurrence and Fully Compliant Data

For subject i with treatment initiation time U_i, we simulate recurrence times according to a Cox model:

$$\lambda_i(t) = \lambda_0 \exp[\boldsymbol{\theta}_0^T \boldsymbol{V}_{0i} + \theta_1 \text{logPSA}_i(t) + \theta_2 \text{logPSA}_i'(t_k) + \gamma_i(t) I(t > U_i)] \tag{3.5}$$

where $\lambda_0 = 0.2$, $\boldsymbol{\theta}_0 = (0.2, 0.3)^T$, $\theta_1 = 0.3$ and $\theta_2 = 0$. $\boldsymbol{V}_{0i} = (I(\text{Tstage}_i = 2), I(\text{Tstage}_i \geq 3))^T$ with patient i's baseline T-stage sampled from $\{1, 2, 3, 4\}$ with probability $\boldsymbol{p} = (0.33, 0.59, 0.07, 0.01)$ (categorical distribution). The treatment effect

$$\gamma_i(t) = \begin{cases} \min\{[\gamma_{i0} + \beta_2(t - U_i)], 0\} & \text{if } \gamma_{i0} < 0 \text{ and } t > U_i \\ \max\{[\gamma_{i0} - \beta_2(t - U_i)], 0\} & \text{if } \gamma_{i0} > 0 \text{ and } t > U_i \\ 0 & \text{otherwise} \end{cases} \tag{3.6}$$

where the initial treatment effect $\gamma_{i0} = \beta_0 + \beta_1 \text{logPSA}_i(U_i)$, with $(\beta_0, \beta_1, \beta_2) = (-1.0, -0.4, 0.2)$. Thus, γ_{i0} linearly depends on the logPSA value at the time of treatment initiation, and then the magnitude of the treatment effect decays over time, until it shrinks to zero. The survival function is

$$S_i(t) = \exp\{-\int_0^t \lambda_i(s)ds\} \tag{3.7}$$

and the survival time for subject i is then generated as $T_i^* = S_i^{-1}(X)$, where $X \sim \text{Uniform}(0, 1)$, then T_i^* is rounded up to the closest visit time as T_i, or censored at 15 years as the study ends. Similarly, for each regime g^{b_j}, $j = 1, \cdots, 10$, we can calculate the survival time $T_i^{b_j}$ for subject i according to the counterfactual treatment initiation time $U_i^{b_j}$.

3.4.2 Scenario 2: More Complicated PSA Trajectory and Treatment Model

Here, we consider a more realistic PSA model, and the treatment regimes in this scenario are also set to be more complicated to mimic clinical practice.

3.4.2.1 PSA Models

In the absence of SADT, a typical trajectory of logPSA observed clinically would have three phases (0: post-therapy, 1: short-term evolution, 2: long-term evolution). Following *Proust-Lima et al.* (2008) and Taylor et al. (2013), for subject i at $t \in (0, t_K]$ years after the initial treatment, we simulate PSA values from the following mixed model:

$$\begin{aligned}\text{logP}_i(t) &= \text{logPSA}_i(t) + \epsilon_{it} \\ &= (\alpha_0 + a_{i0}) + (\alpha_{11} + \boldsymbol{\alpha}_{12}^T \boldsymbol{V}_{0i} + a_{i1}) f(t) + (\alpha_{21} + \boldsymbol{\alpha_{22}}^T \boldsymbol{V}_{0i} + a_{i2}) t + \epsilon_{it}\end{aligned}$$

where $f(t) = (1+t)^{-1.5} - 1$ is used to model the short-term decreasing trend of logPSA, and t is used to model the long-term increasing trend. $(\alpha_0, \alpha_{11}, \boldsymbol{\alpha}_{12}, \alpha_{21}, \boldsymbol{\alpha}_{22})$ are fixed effect parameters, we take $\alpha_0 = 1.0$, $\alpha_{11} = 1.5$, $\boldsymbol{\alpha}_{12} = (0.2, 0.2)^T$, $\alpha_{21} = 0.1$, and $\boldsymbol{\alpha}_{22} = (0.2, 0.5)^T$. (a_{i0}, a_{i1}, a_{i2}) are subject-specific random effects, and V_{0i} is the vector of baseline T-stage indicators as in scenario 1. At a given time t, we assume the measurement error $\epsilon_{it} \sim N(0, \sigma^2)$ with $\sigma^2 = 0.2$, and we assume the random effects $(a_{i0}, a_{i1}, a_{i2}) \sim \text{MVN}(0, \Sigma)$, where $\Sigma = \begin{pmatrix} 1.0 & 1.0 & 0.15 \\ 1.0 & 2.6 & 0.45 \\ 0.15 & 0.45 & 0.5 \end{pmatrix}$.

3.4.2.2 Different Treatment Regimes and Observed Treatment Time

We consider 10 different regimes with $\{b_1, b_2, \cdots, b_{10}\} = \{-1.5, -1.0, \cdots, 3.0\}$. Thus, according to Definition (3.1), we can calculate the regime specific treatment initiation times $U_i^{b_1}, U_i^{b_2}, \cdots, U_i^{b_{10}}$. The observed data are generated similarly as in Scenario 1, but with more possible regimes, the observed treatment assignment for subject i follows one of 100 regimes with the threshold value B, where B follows a discrete uniform distribution for 100 evenly gapped values $\{-1.95, -1.90, \cdots, 3.00\}$. The observed treatment time for subject i is then $U_i = U_i^B$.

3.4.2.3 Model for Recurrence and Fully Compliant Data

We will use the same hazard model to generate the treatment with slightly different parameter settings, i.e. for subject i with treatment initiation time U_i, the survival time T_i will be generated from equation (3.5) (3.6) and (3.7) where $\lambda_0 = 0.15$, $\boldsymbol{\theta}_0 = (0.8, 0.9)^T$, $\theta_1 = 0.1$, $\theta_2 = 0.1$, $\beta_0 = 10.0$, $\beta_1 = -10.0$ and $\beta_2 = 0.2$. Same models are also used to define the counterfactual survival time $T_i^{b_j}$ for each regime g^{b_j}, $j = 1, \cdots, 10$.

3.4.3 Simulation Results

Figure 3.1 shows the average Nelson Aalen survival curves in Scenario 1, where we present one of the regimes ($b_{10} = 2.0$). As we can see without weighting, the curve estimated from the adherent data without using weighting is biased from the counterfactual fully adherent curve, so the weighted estimator is needed to give good estimation for the counterfactual survival time. We can see both the pooled logistic model and proposed method can effectively reduce the bias, while the proposed method performs slightly better, we do see that the estimated curve in Figure 3.1a is closer to the fully adherence curve. Similar results are also observed for other regimes (data not shown).

In Scenario 2, the treatment is generated from a randomly selected regime and related to the PSA trajectory which has a complicated shape. So it becomes difficult for the pooled logistic method to correctly model the treatment assignment mechanism. As shown in Figure 3.2, for regimes $b_4 = -1.0$, the survival curve estimated by the proposed method is almost the same as the counterfactual fully adherent curve, while the curve estimated by the pooled logistic method shows notable bias to the truth.

For both scenarios, for any given b, we can calculate the true restricted mean survival time, $\mu_0^g(b)$ via Monte Carlo simulation with 10^6 replicates. Figure 3.3 gives $\mu_0^g(b)$ over different b for both scenarios. Figure 3.3a shows that for scenario 1, in the range considered, μ_0^g is maximized at around $b = -1.0$ with $\mu_0^g = 7.13$ year. Figure 3.3b shows similar uni-modal relationship of μ_0^g and b, thus the optimal DTR would have $b = 1.5$ which yields the maximum μ_0^g around 5.52 years.

Table 3.1 summarizes the estimated restricted mean survival time $\hat{\mu}$ for the regimes of interest. In both scenarios, the proposed estimator correctly identify the optimal

regime, and the $\hat{\mu}$ from the fully adherent cohort are close to μ_0^g in Figure 3.3. For Scenario 1, $\hat{\mu}$ from the unweighted dataset shows a monotone increasing trend as the threshold b becomes larger. Both the pooled logistic method and the proposed method recover the uni-modal trend and identify regime 7 with $b_7 = 2.5$ as the optimal one that yields the maximal $\hat{\mu}$. This is consistent with the results from full adherent data. For Scenario 2, although the unweighted estimator of μ also gives a uni-modal shape over b, it gives the optimal regime as regime 6 while the true optimal one is regime 7. As shown in the frequency of being identified as the optimal regime, over the 500 simulations, 99.4% of them shows regime 7 is the true optimal regime, while only 19.6% can correctly identify the optimal regime from the unweighted data. By employing the weights, the pooled logistic method can correctly identify regime 7 as the optimal, but the bias to the $\hat{\mu}$ for fully adherent data is also big. This bias becomes much smaller for the proposed method. And comparing to the pooled logistic method, the proposed method has much larger rate of identifying the right regime as the optimal one (71.0% vs. 50.4%). Thus we can see introducing weights can effectively reduce the bias when estimating the population survival outcome, and the proposed flexible model can further help to correctly estimate the regime specific outcomes in various scenarios.

3.5 Application to Real Prostate Cancer Data

We apply the proposed method to a multi-center prostate cancer study consisting of 2,781 patients with clinically localized prostate cancer, all of whom were initially treated with external beam radiation therapy (EBRT). In the dataset, PSA(ng/ml) and T-stage were recorded prior to initial EBRT, with PSA monitored at periodic visits throughout follow-up. A complete description of the data can be found in *Proust-Lima et al.* (2008). In this analysis, we restrict attention to patients who are being actively followed, and remove patients if the interval between adjacent PSA

measurements is more than 2 years. That gives a cohort of 2,427 patients. The longest follow-up time was then 15.6 years. We also use the last observation carried forward (LOCF) method to impute the missing PSA measures. There are 10.1% of the patients who received SADT. We consider DTRs g^b as defined earlier, where b is the cut-off for the logarithm of (PSA+0.1). We consider regimes with b ranging from 0.7 to 4.2, i.e $\{0.7, 0.8, \cdots, 4.2\}$. This is chosen following a data adaptive manner. For each regime under investigation, there are at least 3 patients who are fully adherent to the regime and receive SADT during the course of study. The weights are estimated using Random Forests for the probability of treatment/adherence with input covariates PSA, slope of PSA, baseline T-stage and time t. Figure 3.4a shows that the regime with $b = 0.9$ (2.36 ng/ml PSA) is identified as g^{opt} by the proposed method, i.e. it will be ideal to initiate SADT for patients with increasing PSA and first time has PSA level above 2.36 ng/ml. The corresponding estimated restricted mean survival time under this regime is $\hat{\mu}^g$ =14.80 year. Figure 3.4b gives the weighted Nelson-Aalen estimators for the optimal regime along with two other regimes.

3.6 Discussion

We show that the proposed method provides a powerful tool to identify the optimal DTR. Compared to other existing methods, our method possesses robustness in two ways: for the adherence mechanism, the Random Forest regression allows us to capture a large range of different treatment models, and for the survival outcome, the non-parametric estimation also allows us to put less structure on the estimator. Furthermore, the method is computationally feasible even for problems with relatively large number of time points at which it is possible to initiate treatment. So it is very useful in clinical studies and public health practice where the treatment or intervention is made dynamically and needs to be optimized.

In this chapter, we use the inverse probability weighting to correct for the selection bias from the observational data with arbitrary censoring. We pooled the person-time pieces from all time points to fit the model for adherence mechanism, since in most real applications, there may be few individuals that are observed to follow a given regime for a long time. Thus the adherence probability model at later time points may become very unstable due to the limited availability of data. Pooling the data across different time points can partially account for this problem, however, the traditional pooled logistic model may not always be satisfactory. Although allowing for a time-dependent intercept $h(t)$, it assumes the same linear dependence to covariates over different time points. Employing Random Forest method to the pooled sample may be preferable in this sense, as it automatically incorporates interactions between time and other covariates, which allows the association between adherence probability and covariates to vary over time.

In order to make the estimates more stable, we use the stabilized version of inverse probability weights in this chapter. However, we may need to be cautious here as the weights are now time-dependent, it is very important for us to make sure that the numerator for $sw_i^b(t)$'s are only models of baseline covariates. *Cain et al.* (2010) has shown that the stabilization procedures commonly used for static regimes are not valid for dynamic regimes. Here, we treat the models for numerator and denominator separately, the denominator model $P(C_t^b = 1|\overline{\boldsymbol{C}}_{t-1}^b = \bar{\mathbf{1}}, \bar{\boldsymbol{O}}_t = \bar{\boldsymbol{o}}_t, \overline{\boldsymbol{A}}_{t-1} = \overline{\boldsymbol{a}}_{t-1})$ could be vary according to the treatment history $\overline{\boldsymbol{A}}_{t-1}$, while the numerator model $P(C_t^b = 1|\overline{\boldsymbol{C}}_{t-1}^b = \bar{\mathbf{1}}, \boldsymbol{L}_0 = \boldsymbol{l}_0)$ and can calculated from the estimates of the treatment model. $P(A_t = 1|\overline{\boldsymbol{A}}_{t-1} = \bar{\mathbf{0}}, \overline{\boldsymbol{O}}_t = \overline{\boldsymbol{o}}_t)$. While for the numerator, we need to fit model for $P(C_t = 1|\overline{\boldsymbol{C}}_{t-1} = \bar{\mathbf{1}}, \boldsymbol{L}_0 = \boldsymbol{l}_0)$, which do not restrict to person-time pieces when $\overline{\boldsymbol{A}}_{t-1} = \bar{\mathbf{0}}$, so we choose to model the adherence C_t^b directly.

Pooled logistic regression has been a popular approach for estimating counterfactual outcome distributions for time to event data when using IPW methods. Previous literature has proved the equivalence of pooled logistic models and survival models (*D'Agostino et al.*, 1990). However, certain assumptions need to be satisfied to validate the use of the pooled logistic model. Recent development in Random Survival Forest has allowed the model to include time-varying covariates (*Bou-Hamad et al.*, 2011a) , and it would be interesting to employ this idea and directly fit a survival model in the weight estimation. Compared to the current approaches, this could further reduce the bias, while at the same time, provide more flexibility.

In the chapter, we try to make inference on the counterfactual outcomes from observational data. In practice, the estimation may only be consistent if the three basic assumptions for causal inference are satisfied. The consistency assumption may be the hardest one to verify. For the no unmeasured confounders assumption, it is very important that we include all the covariates that could possibly affect the treatment assignment and adherence, this would require thorough understanding of the problem and sufficient communication with the clinicians and practitioners. The positivity assumption would require that every patient should have a positive probability to follow all regimes under consideration. This is an assumption that may also be hard to verify in practice. One issue here is the number of treatment status or adherence status will grow exponentially as the number of stages or visits increases. Consider a study with visits at $t = 1, \cdots, K$, then if we want to make inference on every possible regime in $\overline{\mathcal{A}}_K = \{0, 1\}^K$, is would require that for any given $\overline{\boldsymbol{O}}_K = \overline{\boldsymbol{o}}_K$, the probability of being in any one of the 2^K status be positive. In practice, this would require a very large sample to make sure that every status would have a notable chance to be observed. For example, in Scenario 1 of our simulation, we consider $K = 15$, for a fixed $\overline{\boldsymbol{O}}_K = \overline{\boldsymbol{o}}_K$, if every status was observed at least once, this would require more

than 3×10^4 observations. This is also a problem in designing randomized trials to identify optimal DTR. One possible solution is to impose some dependency structure on the time series. For example, Markov assumption could be assumed such that the current treatment/adherence status of a patient only depends on his/her treatment/adherence status at previous time point, then the space of $\overline{\mathcal{A}}_K$ would reduce to be linear with the increase of K. Thus overall the methodology does requires large sample sizes and considerable heterogeneity in the patterns of treatment interaction in the observational data. Without these modeling assumptions the uncertainty in the results could be large. Meanwhile, the proposed method could also serve as a preliminary analysis, which can find out which regimes in $\overline{\mathcal{A}}_K$ are estimable (viable) in the given problem, then more accurate results could be obtained from randomized trials that are designed based on only the set of viable regimes.

3.7 Appendix

3.7.1 The distribution of counterfactual and observational data

We demonstrate how to deduce the joint distribution $p_{\overline{\boldsymbol{O}}_k^C(\overline{\boldsymbol{a}}_{K-1})}(\overline{\boldsymbol{o}}_k)$ and conditional distributions

$p_{\boldsymbol{L}_k^C(\overline{\boldsymbol{a}}_{k-1})|\overline{\boldsymbol{L}}_{k-1}^C(\overline{\boldsymbol{a}}_{k-2}),\overline{\boldsymbol{R}}_{k-1}^C(\overline{\boldsymbol{a}}_{k-2})}(\boldsymbol{l}_k|\overline{\boldsymbol{l}}_{k-1},\overline{\boldsymbol{r}}_{k-1})$ and $p_{R_k^C(\overline{\boldsymbol{a}}_{k-1})|\overline{\boldsymbol{L}}_k^C(\overline{\boldsymbol{a}}_{k-1}),\overline{\boldsymbol{R}}_{k-1}^C(\overline{\boldsymbol{a}}_{k-2})}(r_k|\overline{\boldsymbol{l}}_k,\overline{\boldsymbol{r}}_{k-1})$ for a fixed $\overline{\boldsymbol{a}}_{k-1} \in \overline{\boldsymbol{A}}_{k-1}$, $k = 1, \cdots, K$ from the distribution of the observed data. Then under the consistency and no unmeasured confounders assumptions, the joint

density of $(\boldsymbol{W}^C, \overline{\boldsymbol{A}}_{K-1})$ is

$$\begin{aligned}
p_{\boldsymbol{W}^C,\overline{\boldsymbol{A}}_{K-1}}(\boldsymbol{w},\overline{\boldsymbol{a}}_{K-1}) &= p_{\boldsymbol{W}^C}(\boldsymbol{w}) p_{\overline{\boldsymbol{A}}_{K-1}|\boldsymbol{W}^C}(\overline{\boldsymbol{a}}_{K-1}|\boldsymbol{w}) \\
&= p_{\boldsymbol{W}^C}(\boldsymbol{w}) p_{A_0|\boldsymbol{W}^C}(a_0|\boldsymbol{w}) \prod_{j=1}^{K-1} p_{A_j|\overline{A}_{j-1},\boldsymbol{W}^C}(a_j|\overline{\boldsymbol{a}}_{j-1},\boldsymbol{w}) \\
&= p_{\boldsymbol{W}^C}(\boldsymbol{w}) p_{A_0|\boldsymbol{L}_0}(a_0|\boldsymbol{l}_0) \prod_{j=1}^{K-1} p_{A_j|\overline{\boldsymbol{A}}_{j-1},\overline{\boldsymbol{L}}_j,\overline{\boldsymbol{R}}_j}(a_j|\overline{\boldsymbol{a}}_{j-1},\overline{\boldsymbol{l}}_j,\overline{\boldsymbol{r}}_j)
\end{aligned}$$

Moreover,

$$\begin{aligned}
& p_{\boldsymbol{W}^C,\overline{\boldsymbol{A}}_{K-1}|\overline{\boldsymbol{A}}_{K-1},\overline{\boldsymbol{O}}_K}(\boldsymbol{w},\overline{\boldsymbol{a}}_{K-1}|\overline{\boldsymbol{a}}_{K-1},\overline{\boldsymbol{o}}_K) \\
=& \frac{p_{\boldsymbol{W}^C,\overline{\boldsymbol{A}}_{K-1}}(\boldsymbol{w},\overline{\boldsymbol{a}}_{K-1})}{\int_{\left\{u:\overline{\boldsymbol{L}}_K^C(\overline{\boldsymbol{a}}_{K-1})=\overline{\boldsymbol{l}}_K,\overline{\boldsymbol{R}}_K^C(\overline{\boldsymbol{a}}_{K-1})=\overline{\boldsymbol{r}}_K\right\}} p_{\boldsymbol{W}^C,\overline{\boldsymbol{A}}_{K-1}}(u,\overline{\boldsymbol{a}}_{K-1}) dv_{\boldsymbol{W}^C}(u)} \\
=& \frac{p_{\boldsymbol{W}^C}(\boldsymbol{w}) p_{A_0|\boldsymbol{L}_0}(a_0|\boldsymbol{l}_0) \prod_{j=1}^{K-1} p_{A_j|\overline{\boldsymbol{A}}_{j-1},\overline{\boldsymbol{L}}_j,\overline{\boldsymbol{R}}_j}(a_j|\overline{\boldsymbol{a}}_{j-1},\overline{\boldsymbol{l}}_j,\overline{\boldsymbol{r}}_j)}{\int_{\left\{u:\overline{\boldsymbol{L}}_K^C(\overline{\boldsymbol{a}}_{K-1})=\overline{\boldsymbol{l}}_K,\overline{\boldsymbol{R}}_K^C(\overline{\boldsymbol{a}}_{K-1})=\overline{\boldsymbol{r}}_K\right\}} p_{\boldsymbol{W}^C}(u) p_{A_0|\boldsymbol{L}_0}(a_0|\boldsymbol{l}_0) \prod_{j=1}^{K-1} p_{A_j|\overline{\boldsymbol{A}}_{j-1},\overline{\boldsymbol{L}}_j,\overline{\boldsymbol{R}}_j}(a_j|\overline{\boldsymbol{a}}_{j-1},\overline{\boldsymbol{l}}_j,\overline{\boldsymbol{r}}_j) dv_{\boldsymbol{W}^C}(u)} \\
=& \frac{p_{\boldsymbol{W}^C}(\boldsymbol{w})}{\int_{\left\{u:\overline{\boldsymbol{L}}_K^C(\overline{\boldsymbol{a}}_{K-1})=\overline{\boldsymbol{l}}_K,\overline{\boldsymbol{R}}_K^C(\overline{\boldsymbol{a}}_{K-1})=\overline{\boldsymbol{r}}_K\right\}} p_{\boldsymbol{W}^C}(u) dv_{\boldsymbol{W}^C}(u)} \\
=& p_{\boldsymbol{W}^C|\overline{\boldsymbol{O}}_K^C(\overline{\boldsymbol{a}}_{K-1})}(\boldsymbol{w}|\overline{\boldsymbol{o}}_K)
\end{aligned}$$

Thus,

$$\begin{aligned}
& p_{\overline{\boldsymbol{O}}_K^C(\overline{\boldsymbol{a}}_{K-1})}(\overline{\boldsymbol{o}}_K) = \frac{p_{\boldsymbol{W}^C}(\boldsymbol{w})}{p_{\boldsymbol{W}^C|\overline{\boldsymbol{O}}_K^C(\overline{\boldsymbol{a}}_{K-1})}(\boldsymbol{w}|\overline{\boldsymbol{o}}_K)} \\
=& \frac{p_{\boldsymbol{W}^C,\overline{\boldsymbol{A}}_{K-1}}(\boldsymbol{w},\overline{\boldsymbol{a}}_{K-1})}{p_{\boldsymbol{W}^C,\overline{\boldsymbol{A}}_{K-1}|\overline{\boldsymbol{A}}_{K-1},\overline{\boldsymbol{O}}_K^C(\overline{\boldsymbol{a}}_{K-1})}(\boldsymbol{w},\overline{\boldsymbol{a}}_{K-1}|\overline{\boldsymbol{a}}_{K-1},\overline{\boldsymbol{o}}_K)} \\
=& \frac{p_{\boldsymbol{W}^C,\overline{\boldsymbol{A}}_{K-1}}(\boldsymbol{w},\overline{\boldsymbol{a}}_{K-1})}{p_{A_0|\boldsymbol{L}_0}(a_0|\boldsymbol{l}_0) \prod_{j=1}^{K-1} p_{A_j|\overline{\boldsymbol{A}}_{j-1},\overline{\boldsymbol{L}}_j,\overline{\boldsymbol{R}}_j}(a_j|\overline{\boldsymbol{a}}_{j-1},\overline{\boldsymbol{l}}_j,\overline{\boldsymbol{r}}_j) p_{\boldsymbol{W}^C,\overline{\boldsymbol{A}}_{K-1}|\overline{\boldsymbol{A}}_{K-1},\overline{\boldsymbol{O}}_K^C(\overline{\boldsymbol{a}}_{K-1})}(\boldsymbol{w},\overline{\boldsymbol{a}}_{K-1}|\overline{\boldsymbol{a}}_{K-1},\overline{\boldsymbol{o}}_K)} \\
=& p_{\boldsymbol{O}_K|\overline{\boldsymbol{O}}_{K-1},\overline{\boldsymbol{A}}_{K-1}}(\boldsymbol{o}_K|\overline{\boldsymbol{o}}_{K-1},\overline{\boldsymbol{a}}_{K-1}) p_{\boldsymbol{L}_0}(\boldsymbol{l}_0) \prod_{j=1}^{K-1} p_{\boldsymbol{O}_j|\overline{\boldsymbol{O}}_{j-1},\overline{\boldsymbol{A}}_{j-1}}(\boldsymbol{o}_j|\overline{\boldsymbol{o}}_{j-1},\overline{\boldsymbol{a}}_{j-1})
\end{aligned}$$

Let $\boldsymbol{W}_k = \left\{\boldsymbol{L}_0, \boldsymbol{L}_1^C(a_0), R_1^C(a_0), \cdots, \boldsymbol{L}_k^C(\overline{\boldsymbol{a}}_{k-1}), R_k^C(\overline{\boldsymbol{a}}_{k-1}), \forall\, \overline{\boldsymbol{a}}_{k-1} \in \overline{\boldsymbol{\mathcal{A}}}_{k-1}\right\}, k = 1, \cdots, K$. Using the same argument, $p_{\boldsymbol{W}_k, \overline{\boldsymbol{A}}_{k-1}}(\boldsymbol{w}, \overline{\boldsymbol{a}}_{k-1}) = p_{\boldsymbol{W}_k}(\boldsymbol{w}_k) p_{A_0|\boldsymbol{L}_0}(a_0|\boldsymbol{l}_0) \prod_{j=1}^{k-1} p_{\overline{\boldsymbol{A}}_j|\overline{\boldsymbol{A}}_{j-1}, \overline{\boldsymbol{O}}_j}(a_j|\overline{\boldsymbol{a}}_{j-1}, \overline{\boldsymbol{o}}_j)$ and $p_{\overline{\boldsymbol{O}}_k^C(\overline{\boldsymbol{a}}_{k-1})}(\overline{\boldsymbol{o}}_k) = p_{\boldsymbol{L}_0}(\boldsymbol{l}_0) \prod_{j=1}^{k} p_{\boldsymbol{O}_j|\overline{\boldsymbol{O}}_{j-1}, \overline{\boldsymbol{A}}_{j-1}}(\boldsymbol{o}_j|\overline{\boldsymbol{o}}_{j-1}, \overline{\boldsymbol{a}}_{j-1})$. It follows that

$$\begin{aligned} p_{\boldsymbol{O}_K^C(\overline{\boldsymbol{a}}_{K-1})|\overline{\boldsymbol{O}}_{K-1}^C(\overline{\boldsymbol{a}}_{K-2})}(\boldsymbol{o}_K|\overline{\boldsymbol{o}}_{K-1}) &= \frac{p_{\overline{\boldsymbol{O}}_K^C(\overline{\boldsymbol{a}}_{K-1})}(\overline{\boldsymbol{o}}_K)}{p_{\overline{\boldsymbol{O}}_{K-1}^C(\overline{\boldsymbol{a}}_{K-2})}(\overline{\boldsymbol{o}}_{K-1})} \\ &= \frac{p_{\boldsymbol{L}_0}(\boldsymbol{l}_0) \prod_{j=1}^{k} p_{\boldsymbol{O}_j|\overline{\boldsymbol{O}}_{j-1}, \overline{\boldsymbol{A}}_{j-1}}(\boldsymbol{o}_j|\overline{\boldsymbol{o}}_{j-1}, \overline{\boldsymbol{a}}_{j-1})}{p_{\boldsymbol{L}_0}(\boldsymbol{l}_0) \prod_{j=1}^{k-1} p_{\boldsymbol{O}_j|\overline{\boldsymbol{O}}_{j-1}, \overline{\boldsymbol{A}}_{j-1}}(\boldsymbol{o}_j|\overline{\boldsymbol{o}}_{j-1}, \overline{\boldsymbol{a}}_{j-1})} \\ &= p_{\boldsymbol{O}_K|\overline{\boldsymbol{O}}_{K-1}, \overline{\boldsymbol{A}}_{K-1}}(\boldsymbol{o}_K|\overline{\boldsymbol{o}}_{K-1}, \overline{\boldsymbol{a}}_{K-1}) \end{aligned}$$

Similarly, for $k = 2, \cdots, K$,

$$p_{\boldsymbol{O}_k^C(\overline{\boldsymbol{a}}_{k-1})|\overline{\boldsymbol{O}}_{k-1}^C(\overline{\boldsymbol{a}}_{k-2})}(\boldsymbol{o}_k|\overline{\boldsymbol{o}}_{k-1}) = p_{\boldsymbol{O}_k|\overline{\boldsymbol{O}}_{k-1}, \overline{\boldsymbol{A}}_{k-1}}(\boldsymbol{o}_k|\overline{\boldsymbol{o}}_{k-1}, \overline{\boldsymbol{a}}_{k-1})$$

Furthermore, for $k = 2, \cdots, K$

$$p_{\boldsymbol{L}_k^C(\overline{\boldsymbol{a}}_{k-1})|\overline{\boldsymbol{O}}_{k-1}^C(\overline{\boldsymbol{a}}_{k-2})}(\boldsymbol{l}_k|\overline{\boldsymbol{o}}_{k-1}) = p_{\boldsymbol{L}_k|\overline{\boldsymbol{O}}_{k-1}, \overline{\boldsymbol{A}}_{k-1}}(\boldsymbol{L}_k|\overline{\boldsymbol{o}}_{k-1}, \overline{\boldsymbol{a}}_{k-1})$$

and

$$p_{R_k^C(\overline{\boldsymbol{a}}_{k-1})|\overline{\boldsymbol{O}}_{k-1}^C(\overline{\boldsymbol{a}}_{k-2}), \boldsymbol{L}_k(\overline{\boldsymbol{a}}_{k-1})}(r_k|\overline{\boldsymbol{o}}_{k-1}, l_k) = p_{R_k|\overline{\boldsymbol{O}}_{k-1}, \boldsymbol{L}_k, \overline{\boldsymbol{A}}_{k-1}}(r_k|\overline{\boldsymbol{o}}_{k-1}, \boldsymbol{l}_k, \overline{\boldsymbol{a}}_{k-1})$$

3.7.2 Proof of Theorem 3.1

For a given regime g, $C_k = A_k g_k(\overline{\boldsymbol{O}}_k) + (1 - A_k)\{1 - g_k(\overline{\boldsymbol{O}}_k)\}$ for $k = 0, \cdots, K-1$, thus,

(i) For the case that $a_{k-1} = 0$

$$\begin{aligned}
P(C_k = 1|\overline{\boldsymbol{C}}_{k-1} = \bar{\mathbf{1}}, \overline{\boldsymbol{O}}_k = \overline{\boldsymbol{o}}_k) =& E\left\{I(C_k = 1|\overline{\boldsymbol{C}}_{k-1} = \bar{\mathbf{1}}, \overline{\boldsymbol{O}}_k = \overline{\boldsymbol{o}}_k)\right\} \\
=& E\left[I\left\{A_k g_k(\overline{\boldsymbol{O}}_k) + (1 - A_k)\{1 - g_k(\overline{\boldsymbol{O}}_k) = 1|\overline{\boldsymbol{C}}_{k-1} = \bar{\mathbf{1}}, \overline{\boldsymbol{O}}_k = \overline{\boldsymbol{o}}_k\right\}\right] \\
=& E\left[I\left\{A_k = 1|\overline{\boldsymbol{C}}_{k-1} = \bar{\mathbf{1}}, \overline{\boldsymbol{O}}_k = \overline{\boldsymbol{o}}_k, g(\overline{\boldsymbol{o}}_k) = 1\right\}\right] I\left\{g(\overline{\boldsymbol{o}}_k) = 1\right\} \\
&+ E\left[I\left\{A_k = 0|\overline{\boldsymbol{C}}_{k-1} = \bar{\mathbf{1}}, \overline{\boldsymbol{O}}_k = \overline{\boldsymbol{o}}_k, g(\overline{\boldsymbol{o}}_k) = 0\right\}\right] I\left\{g(\overline{\boldsymbol{o}}_k) = 0\right\} \\
=& P\left\{A_k = 1|\overline{\boldsymbol{A}}_{k-1} = \overline{\boldsymbol{a}}_{k-1}, \overline{\boldsymbol{O}}_k = \overline{\boldsymbol{o}}_k\right\} I\left\{g(\overline{\boldsymbol{o}}_k) = 1\right\} \\
&+ P\left\{A_k = 0|\overline{\boldsymbol{A}}_{k-1} = \overline{\boldsymbol{a}}_{k-1}, \overline{\boldsymbol{O}}_k = \overline{\boldsymbol{o}}_k\right\} I\left\{g(\overline{\boldsymbol{o}}_k) = 0\right\}
\end{aligned}$$

(ii) For the case that $a_{k-1} = 1$, we have $P(C_k = 1|\overline{\boldsymbol{C}}_{k-1} = \bar{\mathbf{1}}, \overline{\boldsymbol{O}}_k = \overline{\boldsymbol{o}}_k) = 1$, meanwhile,

$$\begin{aligned}
&P\left\{A_k = 1|\overline{\boldsymbol{A}}_{k-1} = \overline{\boldsymbol{a}}_{k-1}, \overline{\boldsymbol{O}}_k = \overline{\boldsymbol{o}}_k\right\} I\left\{g(\overline{\boldsymbol{o}}_k) = 1\right\} \\
&+ P\left\{A_k = 0|\overline{\boldsymbol{A}}_{k-1} = \overline{\boldsymbol{a}}_{k-1}, \overline{\boldsymbol{O}}_k = \overline{\boldsymbol{o}}_k\right\} I\left\{g(\overline{\boldsymbol{o}}_k) = 0\right\} \\
=& P\left\{A_k = 1|\overline{\boldsymbol{A}}_{k-1} = \overline{\boldsymbol{a}}_{k-1}, \overline{\boldsymbol{O}}_k = \overline{\boldsymbol{o}}_k\right\} I\left\{g(\overline{\boldsymbol{o}}_k) = 1\right\} = 1
\end{aligned}$$

Theorem 3.1 then follows by combining (i) and (ii).

3.7.3 Proof of Theorem 3.2

The strong consistency of $\hat{\mu}^g(t_K)$ can be proved by first proving the consistency of $\hat{\Lambda}^g(t)$ for $t \in (0, t_K]$. For subject i, let T_i be the event time and D_i be the censoring time for subject i. Let $X_i = \min\{T_i, D_i\}$ and $\delta_i = I(T_i \geq D_i)$. The observed event counting process is defined as $N_i(t) = \delta_i I(X_i \leq t)$, and denote the at risk indicator by $Y_i(t) = I(X_i \geq t)$. Let $\hat{w}_i(t)$ and $w_{0,i}(t)$ be the estimated and true weight for subject

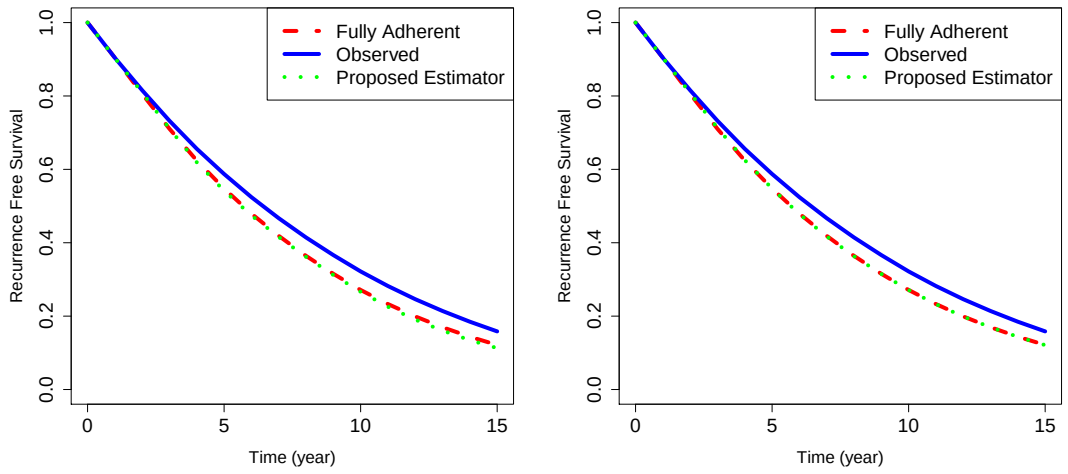
i at time t, respectively. Then the weighted Nelson Aalen estimator is

$$\hat{\Lambda}^g(t) = \frac{1}{n}\sum_{i=1}^{n}\int_0^t \frac{\hat{w}_i(s)}{n^{-1}\sum_{i=1}^{n} Y_i(s)\hat{w}_i(s)} dN_i(s)$$

Let $k = \max_j\{t_j \leq s\}$, using the fact that $\hat{w}_i(s) \rightarrow^{a.s.} w_{0,i}(s)$ as $n \rightarrow \infty$, and the Strong Law of Large Numbers (SLLN), one can obtain that $n^{-1}\sum_{i=1}^{n} Y_i(s)\hat{w}_i(s)$ converges almost surely to

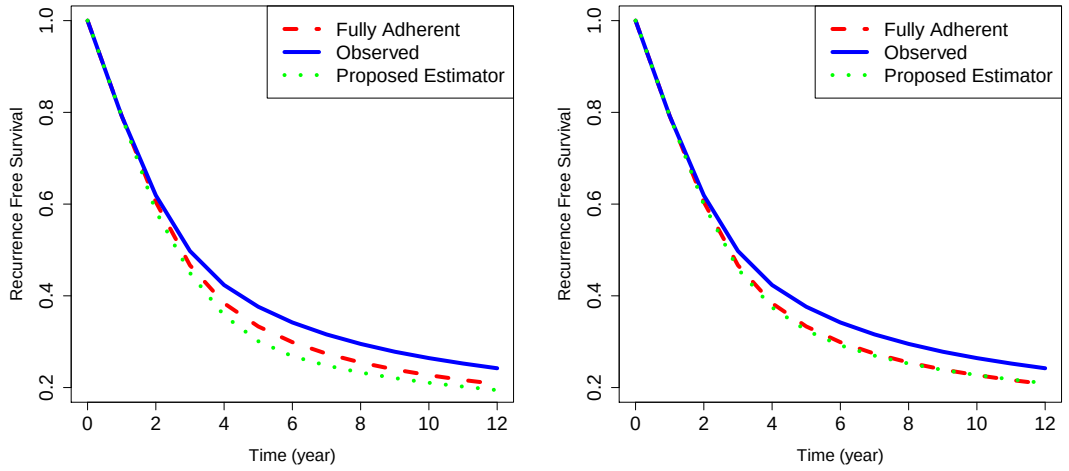
$$\begin{aligned}
E\left[Y_i(s)w_{0,i}(s)\right] &= E\left\{E\left[Y_i(s)w_{0,i}(s)|\overline{\boldsymbol{O}}_k, \overline{\boldsymbol{A}}_k\right]\right\} \\
&= E\left\{E\left[Y_i(s)\frac{I(\overline{\boldsymbol{C}}_{ik}=1|\overline{\boldsymbol{O}}_k, \overline{\boldsymbol{A}}_k)}{Pr(\overline{\boldsymbol{C}}_{ik}=1|\overline{\boldsymbol{O}}_k, \overline{\boldsymbol{A}}_k)}|\overline{\boldsymbol{O}}_k, \overline{\boldsymbol{A}}_k\right]\right\} \\
&= E\left\{\frac{Pr(\overline{\boldsymbol{C}}_{ik}=1|\overline{\boldsymbol{O}}_k=\overline{\boldsymbol{o}}_k, \overline{\boldsymbol{A}}_k=\overline{\boldsymbol{a}}_k)}{Pr(\overline{\boldsymbol{C}}_{ik}=1|\overline{\boldsymbol{O}}_k=\overline{\boldsymbol{o}}_k, \overline{\boldsymbol{A}}_k=\overline{\boldsymbol{a}}_k)}E\left[Y_i(s)|\overline{\boldsymbol{O}}_k=\overline{\boldsymbol{o}}_k, \overline{\boldsymbol{A}}_k=\overline{\boldsymbol{a}}_k\right]\right\} \\
&= E\left\{E\left[Y_i^C(s)|\overline{\boldsymbol{O}}_k^C(\overline{\boldsymbol{a}}_k)=\overline{\boldsymbol{o}}_k\right]\right\} \\
&= E\left[Y_i^C(s)\right] = Pr(D_i^C > s)S^g(s)
\end{aligned}$$

The second last equality holds by the results in Section 3.7.1. Using similar techniques, one can show that $n^{-1}\sum_{i=1}^{n}\hat{w}_i(s)dN_i(s)$ converges almost surely to $Pr(D_i^C > s)dF^g(s)$ as $n \rightarrow \infty$. The above listed results give $\hat{\Lambda}^g(t) \overset{a.s.}{\rightarrow} \Lambda_0^g(t)$. Therefore, $\hat{\mu}^g(t)$ converges to $\mu_0^g(t)$ almost surely as $n \rightarrow \infty$ using the continuous mapping theorem.



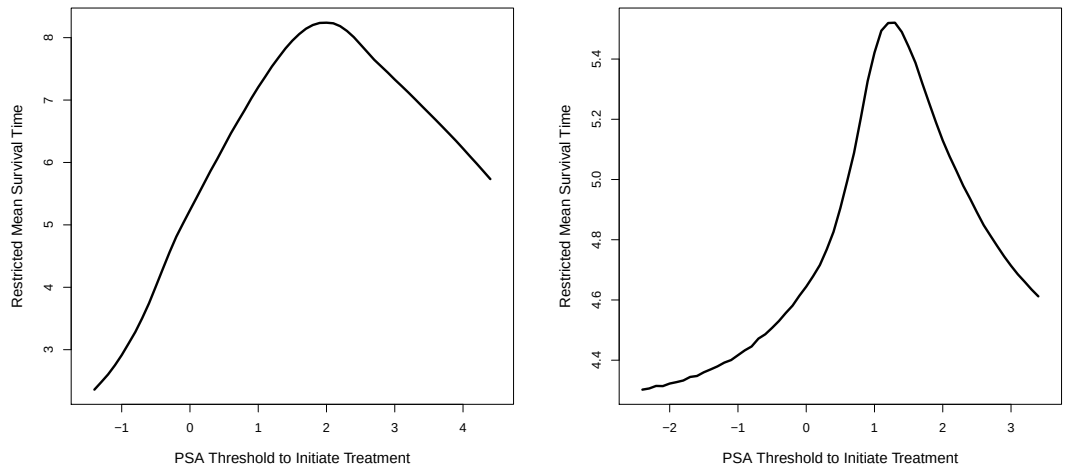
(a) Nelson-Aalen Estimator Using Logistic Model (b) Nelson-Aalen Estimator Using Proposed Method

Figure 3.1: The estimated survival curves for regime $b = 2.0$ in Scenario 1. In both plots, the Nelson-Aalen estimators of the recurrence free survival are plotted. The regime specific true curves (obtained from fully adherent counterfactual cohort) are shown in dashed lines. The solid lines are for the observational data (obtained by censoring subjects when they are no longer adherent with the regime), and the dotted lines are the weighted curves estimated by pooled logistic method (Figure 3.1(a)) or the proposed method (Figure 3.1(b)). All curves are obtained by averaging over 500 simulations.



(a) Nelson-Aalen Estimator Using Logistic Model (b) Nelson-Aalen Estimator Using Proposed Method

Figure 3.2: The Nelson-Aalen curves for regime $b = -1.0$ in Scenario 2. Similar as in Figure 3.1, the regime specific true curves are shown in dashed lines. The solid lines are for the observational data, and the dotted lines are the curve estimated by pooled logistic method (Figure 3.2(a)) or the proposed method (Figure 3.2(b)). All curves are obtained by averaging over 500 simulations.



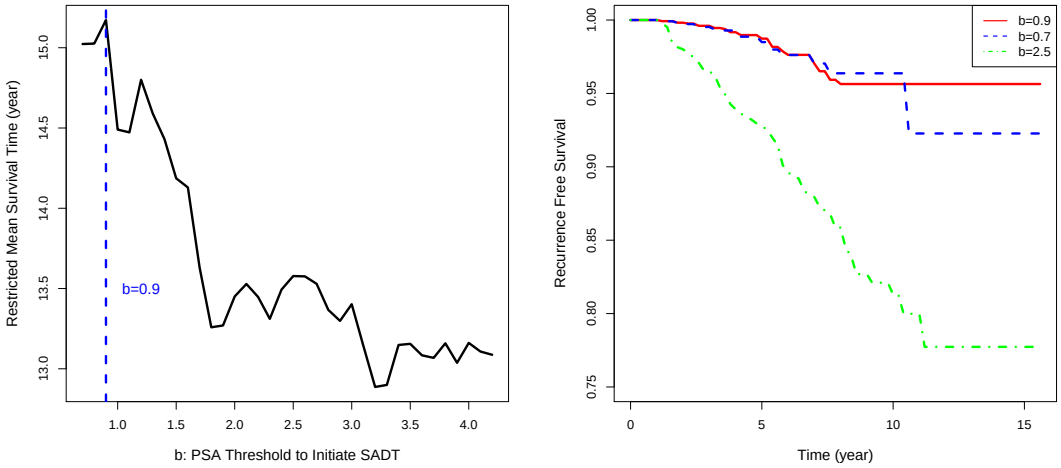
(a) (b)

Figure 3.3: True μ under the two simulation Scenario 1 (Figure 3.3(a)) and Scenario 2 (Figure 3.3(b)). Each point in the plots above are calculated from 10^6 Monte Carlo samples.

Table 3.1: Restricted Mean Survival Time for Regimes of Interest

	Regime	Full Adherent		Unweighted		Pooled Logistic		Proposed Method	
		$\hat{\mu}$	opt%	$\hat{\mu}$	opt%	$\hat{\mu}$	opt%	$\hat{\mu}$	opt%
Scenario 1	$b_1 = -0.5$	4.016 (0.067)	0%	4.953 (0.121)	0%	3.898 (0.071)	0%	4.047 (0.073)	0%
	$b_2 = -0.0$	4.876 (0.075)	0%	5.721 (0.115)	0%	4.676 (0.083)	0%	4.884 (0.080)	0%
	$b_3 = 0.5$	5.920 (0.091)	0%	6.578 (0.117)	0%	5.665 (0.102)	0%	5.918 (0.094)	0%
	$b_4 = 1.0$	6.787 (0.103)	0%	7.225 (0.124)	0%	6.575 (0.121)	0%	6.797 (0.107)	0%
	$b_5 = 1.5$	7.589 (0.113)	0%	7.785 (0.129)	0%	7.462 (0.133)	0%	7.600 (0.120)	0%
	$b_6 = 2.0$	8.340 (0.110)	0%	8.407 (0.128)	0%	8.296 (0.133)	0%	8.335 (0.131)	0%
	$b_7 = 2.5$	8.729 (0.114)	99.6%	8.919 (0.134)	3.8%	8.709 (0.148)	97.0%	8.710 (0.146)	93.6%
	$b_8 = 3.0$	8.503 (0.116)	0.4%	8.949 (0.143)	96.2%	8.415 (0.165)	3.0%	8.488 (0.158)	6.4%
	$b_9 = 3.5$	8.068 (0.114)	0%	8.668 (0.148)	0%	7.954 (0.177)	0%	8.045 (0.171)	0%
	$b_{10} = 4.0$	7.632 (0.111)	0%	8.324 (0.156)	0%	7.490 (0.220)	0%	7.472 (0.310)	0%
Scenario 2	$b_1 = -1.5$	4.872 (0.078)	0%	5.417 (0.096)	0%	4.153 (0.121)	0%	4.639 (0.103)	0%
	$b_2 = -1.0$	4.921 (0.079)	0%	5.402 (0.093)	0%	4.293 (0.109)	0%	4.687 (0.099)	0%
	$b_3 = -0.5$	4.991 (0.079)	0%	5.410 (0.092)	0%	4.544 (0.110)	0%	4.829 (0.098)	0%
	$b_4 = 0.0$	5.091 (0.078)	0%	5.456 (0.091)	0%	4.867 (0.106)	0%	5.047 (0.093)	0%
	$b_5 = 0.5$	5.294 (0.079)	0%	5.572 (0.091)	0%	5.298 (0.117)	0.2%	5.380 (0.095)	0%
	$b_6 = 1.0$	5.757 (0.082)	0.6%	5.766 (0.093)	80.4%	5.891 (0.153)	30.0%	5.862 (0.101)	29.0%
	$b_7 = 1.5$	5.848 (0.088)	99.4%	5.736 (0.096)	19.6%	5.927 (0.219)	50.4%	5.890 (0.108)	71.0%
	$b_8 = 2.0$	5.562 (0.084)	0%	5.488 (0.095)	0%	5.658 (0.336)	6.2%	5.563 (0.112)	0%
	$b_9 = 2.5$	5.315 (0.081)	0%	5.280 (0.092)	0%	5.499 (0.431)	4.8%	5.307 (0.120)	0%
	$b_{10} = 3.0$	5.138 (0.080)	0%	5.135 (0.092)	0%	5.372 (0.575)	8.4%	5.102 (0.134)	0%

Note: The values in parentheses are the empirical standard errors calculated from 500 MCMC replications,
%opt is the percentage for the given regime to be identified as the optimal regime among the 500 replicates



(a) Estimated $\hat{\mu}$ over different regimes g^b

(b) Estimated survival curves for selected regimes

Figure 3.4: The survival estimation for the prostate cancer dataset. Panel (a) shows the relationship between the restricted mean survival time estimated by the proposed method $\hat{\mu}$ and the logPSA threshold for SADT initiation b. Panel (b) shows the weighted Nelson Aalen curves estimated for three regimes, which includes the estimated optimal regime $b = 0.9$ (solid red line), along with two other regimes $b = 0.7$ (dashed blue line) and $b = 2.5$ (dotted green line).

CHAPTER IV

Identifying the Optimal Regime Using Random Survival Forest with Weighted Bootstrap

In many biomedical studies, investigating the performance of different interventions is a major goal, and in many cases, survival time is the primary outcome. Patients with different characteristics are likely to respond differently to an intervention. Thus, identifying the regime that defines who should receive which intervention in order to provide most benefit to the whole population is often of great interest. Recognizing treatment effect heterogeneity is essential in assigning the treatment to patients who can really benefit from it. In this chapter, we consider data from observational studies where some of the covariates affect both the survival outcome and the treatment assignment, while they may not all be available in the target populations. We propose to use Random Survival Forest (RSF) plus an inverse probability weighted bootstrap to estimate the causal outcome while marginalizing over the unavailable covariates. Furthermore, by comparing the restricted mean survival times, the optimal regime could be estimated for the target population based on the available covariates. The proposed method (1) provides a flexible model structure to account for the dependence of the survival outcome to the covariates, and (2) correctly estimates the counterfactual outcomes for each treatment group. Simulations illustrate that the proposed method performs reliably across a range of different scenarios.

4.1 Introduction

It has been shown that patients can exhibit significant heterogeneity in response to treatments in many different diseases (*Ishigooka et al.*, 2000; *Rothwell*, 2005). The emerging field of personalized medicine, which is focused on making treatment decisions for an individual patient based on his/her own clinical, genomic, and other information, has gained considerable interest, as it has the potential of maximizing the treatment benefit for each person and hence for the whole population (*Piquette-Miller and Grant*, 2007). Patients may be given a customized regime, with customized dose or customized treatment schedule according to his/her prognostic or genomic information. The therapy may only be assigned to the patients at the dose that can benefit that person the most, thus the optimal treatment effect could be achieved on the population level without spending the extra resources that would be used if everybody was treated with the same protocol.

Statistical approaches have been developed to find subgroups of patients who can really benefit from the treatment under investigation in randomized clinical trials (*Foster et al.*, 2011). This task becomes more challenging when the outcome of primary interest is survival time, which may be censored (*Kehl and Ulm*, 2006), as is common in phase III clinical trials. In practice, randomized trials are not always available, thus methods to find the optimal regime from the observational data are also needed(*Qian and Murphy*, 2011). In this chapter, we consider the case where the data came from an observational study, with a censored time-to-event outcome. There are two treatments, with treatment indicator A taking values either 0 or 1. Although all covariates are measured at baseline (pretreatment) $\boldsymbol{W} = (\boldsymbol{X}, \boldsymbol{Z})$, only a subset of the covariates $\boldsymbol{X}$ is eligible to be candidates for constructing an optimal regime which could be applied to other populations. The other covariates $\boldsymbol{Z}$ may also be related to both the treatment assignment and the survival outcome. These variables

may not be easy to obtain for the target population due to economic, logistical or other reasons. For example, the covariate $\boldsymbol{Z}$ might measure health insurance, which is not going to be relevant in the target population. The model of interest would be only conditional on the covariates that are generalizable to the target population. To enable the comparison between different treatment regimes, we adopt the causal inference framework *Rubin* (1974, 1990), and the question then is how to correctly estimate the mean for counterfactual outcomes conditional only on the covariates that will be available. Those covariates that are not applicable to the target population will need to be properly handled. A general approach in causal inference to adjust for the confounding and estimate the causal effects is inverse probability weighting (IPW), which has been widely used in survival analysis (*Van der Laan and Robins*, 2003; *Rotnitzky and Robins*, 2005).

The optimal regime is a treatment decision rule depending on a region of the covariate space. One common approach in identifying the optimal regime is to propose a linear combination of covariates, and if the linear combination is greater than a threshold then one treatment is preferred and if the linear combination is less than the threshold then the other treatment is preferred (*Zhang et al.*, 2012). However, it is very likely that the proposed space of optimal regimes does not include the truth, and will thus lead to biased estimates. In this chapter, we propose to use a Random Survival Forest (RSF) for the outcome model to provide the necessary flexibility in modeling, and use inverse probability weighting to account for the confounders $\boldsymbol{Z}$, where the weighting is implemented using a weighted bootstrap procedure. Since we are comparing survival outcomes under different counterfactual world, which have a different amount of follow-up time, we propose to define the optimal regime as the one that maximizes the restricted mean survival time. We introduce the notation and some details of the proposed method in the next section. Section 4.3 contains

results of simulation studies. We then conclude this chapter in Section 4.4 and some technical details are given in Section 4.5.

4.2 Notation and Method

Consider a cohort of n patients from an observational study, let $\boldsymbol{W}_i = (\boldsymbol{X}_i^T, \boldsymbol{Z}_i^T)^T$ denote a d-dimensional vector of baseline covariates for patient i, $i = 1, \cdots, n$, where $\boldsymbol{X}_i$ is a d_1-dimensional vector and $\boldsymbol{Z}_i$ is a d_2-dimensional (d_1+$d_2 = d$). $A_i = j$, $j = 0, 1$ is the indicator for observed treatment status, with 1 for patients who receive the new treatment, and 0 for patients who receive the standard treatment or no treatment. For the outcome, let T_i^0 denote the survival time if subject i did not receive the treatment, and T_i^1 denote the survival time if subject i receive the treatment. Let T_i denote the actual survival time and C_i the censoring time, the observed outcome is then (Y_i, Δ_i) with $Y_i = \min(T_i, C_i)$ and $\Delta_i = I(T_i \leq C_i)$. As in Chapter 3, we impose the same three basic assumptions in causal inference:

1. Consistency assumption: $T_i = A_i T_i^1 + (1 - A_i) T_i^0$;

2. Positivity assumption: $0 < P(A_i | \boldsymbol{W}_i) < 1$;

3. No unmeasured confounders assumption (NUCA): $T_i^a \amalg A_i | \boldsymbol{W}_i$, for $a = 0, 1$.

Note here that we are interested in the model for T conditional on X, so the above is a slightly stronger version of the assumptions required here. In addition, we also require the general assumption in survival analysis that $T_i^a \amalg C_i | \boldsymbol{W}_i$, for $a = 0, 1$, which will guarantee that $T_i \amalg C_i | \boldsymbol{W}_i$. Since X_i is the vector that is available for both the observed dataset and the target population, we are interested in deducing the optimal regime that is based only on $\boldsymbol{X}_i$, which is denoted by $g(\boldsymbol{X}_i)$ and takes the values of either 0 or 1 specifying which treatment should be taken.

4.2.1 Random Survival Forest

Machine learning algorithms have shown great potential in treatment heterogeneity and estimating optimal treatment regimes (*Foster et al.*, 2011; *Zhao et al.*, 2012). Although for survival variables with censoring, the Cox proportional hazard regression model and its extensions are frequently used (*Cox*, 1972), survival tree based methods have also been recognized as useful because of their flexible model structure (*LeBlanc and Crowley*, 1995). A tree is grown by dividing patients at each node into two groups, where the split is chosen to maximize certain criterion which measures the survival difference. It thus has the advantage of automatically identifying certain type of interaction without prespecifying the form of the interaction. The application of ensemble methods provides a simple yet ingenious solution to the instability of tree based methods (*Bou-Hamad et al.*, 2011b). The Random Survival Forest (RSF) is an ensemble tree method for analysis of right-censored survival data (*Ishwaran et al.*, 2008), which provides a flexible model structure that allows for complicated interactions especially between covariates and treatment.

4.2.2 Inverse Probability Weighting

We propose to fit a logistic model for the probability of a patient receiving treatment

$$\text{logit}P(A_i = 1|\boldsymbol{L}_i) = \eta_0 + \eta_1 \boldsymbol{L}_i,$$

where $\boldsymbol{L}_i$ is the vector of baseline covariates that will affect the treatment assignment. Here, $\boldsymbol{L}_i$ could have overlap with $\boldsymbol{W}_i$. In our case, we assume that $\boldsymbol{Z}_i$ is a subset of $\boldsymbol{L}_i$, and $\boldsymbol{L}_i$ is a subset of $\boldsymbol{W}_i$. Thus, the treatment decision for patient i in the observational data is made based on all or part of $\boldsymbol{W}_i$, which includes all $\boldsymbol{Z}_i$ and all or part of $\boldsymbol{X}_i$. Then the estimates $\hat{p}_i = \hat{P}(A_i = 1|\boldsymbol{L}_i)$ can be used to calculate the

estimated weights as

$$\hat{w}_i^1 = \frac{I(A_i = 1)}{\hat{p}_i} \text{ and } \hat{w}_i^0 = \frac{I(A_i = 0)}{1 - \hat{p}_i},$$

which reflects patient i's estimated weights in the so-called treatment sample ($\hat{w}_i^1$) and control sample ($\hat{w}_i^0$), separately. Essentially, for those who are treated in the observational data, they are weighted by $1/\hat{p}_i$ in the treatment sample, while weighted by 0 in the control sample. For those who are not treated, they are weighted by 0 in the treatment sample, while weighted by $1/(1-\hat{p}_i)$ in the control sample. With those weights, we are able to create pseudo-samples that mimic the real counterfactual world, it represent the whole population as if they were all treated in the same way, and the models for counterfactual outcomes can thus be built from these pseudo-samples.

4.2.3 Random Survival Forest in Weighted Bootstrap Samples

The goal is to build flexible models for the time to the event given $\boldsymbol{X}$ and A. Random Survival Forest (RSF) is chosen here as it does not require much knowledge about the underlying mechanism before model fitting (*Ishwaran et al.*, 2008). We will build separate models for the counterfactual outcomes of the treatment and of control samples. One major challenge here is that the original algorithm of Random Survival Forest does not consider weights for subjects, i.e. it assumes that every subject contributes the same to the final model. In order to incorporate the weights in the estimation procedure, we propose an additional layer of bootstrap sampling prior to applying the Random Survival Forest method. We propose to first draw a weighted bootstrap sample where the sampling probability for subject i is proportional to its estimated treatment weight, i.e. $\hat{w}_i^1/(\sum_{j=1}^n \hat{w}_j^1)$, and the Random Survival Forest models are then built on each sample as $(Y_i, \Delta_i) \sim \boldsymbol{X}_i + \boldsymbol{X}_i^2$, where

$\boldsymbol{X}_i^2 = \{X_{1i}^2, X_{2i}^2, \cdots\}$ contains all the squared terms of $\boldsymbol{X}_i$. Thus for the Random Survival Forest we allow both the covariates and their quadratic terms as inputs. The inclusion of $\boldsymbol{X}_i^2$ as covariates is not essential, but in numerical work we found that it commonly helps to improve the performance of the proposed method. The final prediction is then obtained by combining the predictions over all the weighted bootstrap samples. Several methods have been proposed to combine the results of survival trees from bootstrap samples (*Hothorn et al.*, 2004, 2006; *Ishwaran et al.*, 2008). Similar approaches can be employed to combine results from multiple survival forests, we propose to calculate the cumulative hazard function (CHF) by averaging all the CHFs from Random Survival Forest models from each weighted bootstrap samples. In detail, we use the R function $cforest()$ from the *party* package to build the Random Survival Forests, then the $treeresponse()$ function is used to obtain the survival probabilities $S^{1(m)}(t)$ (Kaplan-Meier estimator) for the RSF model from weighted bootstrap sample m, $m = 1, \cdots, B$. The number of bootstrap samples is set at $B = 100$. The survival estimate for the final model is obtained by averaging the cumulative hazard function (CHF) at the same time point over the bootstrap sample specific forests, i.e. the final survival function for the treatment group is estimated as

$$\hat{S}^1(t) = \exp\{-\frac{1}{B}\sum_{m=1}^{B} - \log \hat{S}^{1(m)}(t)\}.$$

The counterfactual mean model for the control group can be built in a similar way, where the weighted bootstrap sampling is done with sampling probability proportional to subject i's estimated weight of being not treated, i.e. $\hat{w}_i^0/(\sum_{j=1}^n \hat{w}_j^0)$. Then the final survival function for the control group $\hat{S}^0(t)$ is obtained using weighted bootstrap samples.

$$\hat{S}^0(t) = \exp\{-\frac{1}{B}\sum_{m=1}^{B} - \log \hat{S}^{0(m)}(t)\}.$$

A variation of this combining step would be to pool all the terminal nodes which contains subject i from all trees over all forests and all bootstrap samples, and calculate the subject-specific final survival function from this pooled sample.

4.2.4 The Optimal Treatment Regime

To identify the optimal regime, ideally, we would like to find the regime which gives the longest mean survival time $E\{T\}$ for the population. However, due to the censoring, we do not have that available, so we choose to compare the restricted mean survival time $\mu = E\{\min(T, \tau)\}$ for some $\tau > 0$, i.e. the optimal regime g^{opt} would be the one that gives the longest μ over the regime space $\mathcal{G}$, $g^{\text{opt}} = \arg\max_{g \in \mathcal{G}} \mu^g$, where the arbitrary time τ is chosen to be the longest follow-up time. μ^g is the restricted mean survival time corresponding to regime g, given by

$$\begin{aligned}
\mu^g &= E\{E\{\min(T^g, \tau)|\boldsymbol{X}\}\} = E\{E\{\min(g(\boldsymbol{X})T^1 + (1 - g(\boldsymbol{X}))T^0, \tau)|\boldsymbol{X}\}\} \\
&= E\{g(\boldsymbol{X})E\{\min(T^1, \tau)|\boldsymbol{X}\} + (1 - g(\boldsymbol{X}))E\{\min(T^0, \tau)|\boldsymbol{X}\}\} \\
&= E\{g(\boldsymbol{X})\mu^1(\boldsymbol{X}) + (1 - g(\boldsymbol{X}))\mu^0(\boldsymbol{X})\} \\
&= E\{\mu^0(\boldsymbol{X}) + g(\boldsymbol{X})(\mu^1(\boldsymbol{X}) - \mu^0(\boldsymbol{X}))\}
\end{aligned}$$

The optimal regime can then be written as $g^{\text{opt}}(\boldsymbol{X}) = I(\mu^1(\boldsymbol{X}) > \mu^0(\boldsymbol{X}))$, where $\mu^1(\boldsymbol{X})$ denotes the restricted mean survival time for the patients with $\boldsymbol{X}$ in the counterfactual world where everyone received the treatment, and $\mu^0(\boldsymbol{X})$ denotes the restricted mean counterfactual survival time where they were all assigned to the control group. Since we already have the estimated survival functions $\hat{S}^1$ and $\hat{S}^0$, the estimated conditional restricted survival time for the counterfactual outcome from the control and treatment groups can be calculated as

$$\hat{\mu}^0(\boldsymbol{X}) = \int_0^\tau \hat{S}^0(t)dt \text{ and } \hat{\mu}^1(\boldsymbol{X}) = \int_0^\tau \hat{S}^1(t)dt.$$

Only when $\hat{\mu}^1(\boldsymbol{X})$ is larger than $\hat{\mu}^0(\boldsymbol{X})$, should we assign the patient to be treated. Otherwise, we would leave the patient alone and not treat the patient, as we would expected mean survival to be higher in that situation. The optimal regime can then be estimated as $\hat{g}^{\text{opt}}(\boldsymbol{X}) = I(\hat{\mu}^1(\boldsymbol{X}) > \hat{\mu}^0(\boldsymbol{X}))$. Therefore, according to one patient's covariates $\boldsymbol{X}$, we can customize the decision on whether this individual should be treated or not.

4.3 Simulations

4.3.1 Other Methods

In order to assess the performance of our proposed method, we also consider three other methods to compare. As a standard and simple method, we consider the regular Cox model. We fit a Cox model $(Y_i, \Delta_i) \sim \boldsymbol{X}_i + A_i + \boldsymbol{X}_i \times A_i$, and use the final model to calculate $\hat{S}_i^1(t) = P(T_i > t | A_i = 1, \boldsymbol{X}_i)$ and $\hat{S}_i^0(t) = P(T_i > t | A_i = 0, \boldsymbol{X}_i)$. The conditional restricted mean survival time and optimal regime are then calculated in the same way as for the proposed method. Second, we consider a weighted version of the Cox model, where we fit separate Cox models $(Y_i, \Delta_i) \sim \boldsymbol{X}_i$ with weights $\hat{w}_i^1$ and $\hat{w}_i^0$. Furthermore, we also consider a standard Random Survival Forest procedure, where we fit an RSF (Y_i, Δ_i) with input covariates $\boldsymbol{X}_i$ and A_i, and then calculate $\hat{S}_i^1(t)$ and $\hat{S}_i^0(t)$ in a similar way as the unweighted Cox model method. If the actual treatment group for subject i is j, then $\hat{S}_i^j(t)$ is obtained from the out-of-bag estimate using the random forest, and $\hat{S}_i^{(1-j)}(t)$ is obtained by applying the random forest to that person's baseline covariates $\boldsymbol{X}_i$, with the opposite of the observed treatment group indicator $1-j$. Similar steps in estimating the optimal regime are then followed.

4.3.2 Simulation Schemes

We consider a moderate dimensional $\boldsymbol{X}$ with $d_1 = 20$ from independent $N(0, 1)$, and a scalar Z ($d_2 = 1$), which is correlated with $\boldsymbol{X}$, the treatment indicator is then generated from a logistic model for $\text{logit}(P(A = 1|X_1, X_2, Z))$. We generate the two counterfactual survival outcomes from lognormal distribution models of the general form

$$\log T^0 = \alpha + \boldsymbol{X}^T\beta + \gamma Z + \epsilon_0 \tag{4.1}$$

$$\log T^1 = \alpha + \boldsymbol{X}^T\beta + \gamma Z + h(\boldsymbol{X}, Z) + \epsilon_1, \tag{4.2}$$

where ϵ_0 and ϵ_1 are generated independently from $N(0, \sigma^2)$, and the observed survival time is $T = AT^1 + (1 - A)T^0$. For simplicity, we only consider uniformly distributed censoring time $C \sim \text{Uniform}(0, \tau)$. Thus we observe time $Y = \min(T, C)$ and the event indicator $\Delta = I(T \leq C)$. The true optimal regime depends on the form of $h(\boldsymbol{X}, Z)$.

The regime of interest is only based on $\boldsymbol{X}$. In order to identify the true optimal regime, we need to calculate the true model conditional on $\boldsymbol{X}$, which is obtained by integrating out Z from equations (4.1) and (4.2). Also notice that we compare the restricted mean survival time, so the choice of time limit τ may also influence the optimal regime. More details about this issue can be found in the Appendix.

4.3.3 Different Scenarios

We will set-up two sets of simulations with different optimal treatment regime settings. In the first set of simulations, we consider the model with $d = 21$ baseline covariates where $X_1, \cdots, X_{20}$ are i.i.d. normally distributed with mean 0 and variance 1, and Z is highly correlated to X_2, and generated from $Z|X_2 = x_2 \sim N(x_2, 1)$. For

treatment assignment, we consider the logistic model

$$\text{logit}P(A=1) = 1.0 + 0.5X_1 - 1.5X_2 + Z.$$

We generate the counterfactual survival outcome using a linear combination of $\boldsymbol{X}$ and Z:

$$\log T_i^0 = 0.3 + 0.5X_1 + 0.5X_2 + 1.1Z + \epsilon_{i0}$$

$$\log T_i^1 = -0.9 + 0.3X_1 + 2.5X_1^2 + 1.0X_2 + 0.3Z + \epsilon_{i1},$$

where ϵ_{i0} and ϵ_{i1} are generated from independent $N(0,\, 1)$. As shown in the Appendix, the conditional mean can be expressed as

$$E(\log T_i^0 | X_1, \cdots, X_{20}) = 0.3 + 0.5X_1 + 1.6X_2$$

$$E(\log T_i^1 | X_1, \cdots, X_{20}) = -0.9 + 0.3X_1 + 2.5X_1^2 + 1.3X_2.$$

The optimal regime would be to give the treatment when $E(T_i^1 | X_1, \cdots, X_{20}) > E(T_i^0 | X_1, \cdots, X_{20})$. Note that because the optimal is based on the expected value of T and not the expected value of $\log(T)$, the optimal regime is similar to but not exactly equals to $2.5X_1^2 - 0.2X_1 - 0.3X_2 - 1.2 > 0$. We consider censoring at a fixed time, i.e. $C = \tau$, with two different τ values chosen to give either 20% or 45% censoring.

In the second set of simulations, we consider the same distribution for $X_1, \cdots, X_{20}$

and Z. For the treatment assignment, we consider the logistic model

$$\text{logit}P(A = 1) = 0.1 - 0.5X_1 - 0.5X_2 + 2Z.$$

And we generate the counterfactual outcomes as follows:

$$\log T_i^0 = 0.5 - 0.2X_1 + 0.3X_2 + 0.5Z + \epsilon_{i0}$$

$$\log T_i^1 = 0.5 - 0.2X_1 + 0.3X_2 + 0.5Z + q(X_1, X_2) + \epsilon_{i1},$$

where ϵ_{i0} and ϵ_{i1} are generated from $N(0,\, 4)$. And we choose function $q(X_1, X_2) = -1+2.5\cdot I(X_1 > -0.5)\cdot I(X_2 < 0.5)$. It can be easily seen that, the true optimal regime (as $\tau \to \infty$) is then to give treatment to patients with $X_1 > -0.5$ and $X_2 < 0.5$. We generate censoring time $C \sim \text{uniform}(0, \tau)$, two τ values are chosen to create either 20% or 45% censoring.

4.3.4 Simulation Results

For each simulation setting, we generate observational data with $n = 1000$ patients, with 100 replicates. First, to evaluate the estimation of the optimal regime, we compare the estimated regimes with the true optimal regime in each dataset, and calculate the percentage of correctly predicted treatment assignment:

$$R_{\text{match}} = \frac{1}{n}\sum_{i=1}^{n} I(\hat{g}(\boldsymbol{X}_i) = g_0^{\text{opt}}(\boldsymbol{X}_i)),$$

where $\hat{g}(\cdot)$ is the optimal regime estimated by one of the methods under consideration, and $g_0^{\text{opt}}(\cdot)$ is the underlying true optimal regime. For each simulated dataset, R_{match} measures the fraction of patients where $\hat{g}$ and g_0^{opt} agree. Table 4.1 shows the mean and standard deviation of R_{match}. We can see that, in all the cases, RSF based

methods tend to give more patients the "right treatment" than the Cox model based methods. In both scenarios, the weighted methods tend to give better prediction than the unweighted method in the same category. This provides some evidence that the inverse probability weighting works in term of correcting the selection bias from observational data. It is interesting that in all cases, although the model structure is misspecified, the weighted Cox method shows slightly better prediction comparing to the standard Cox model.

The optimal treatment regime should be able to benefit the whole population, so we also study the performance of the estimated treatment regime from each method through the regime g specific restricted population mean survival time:

$$\mu^g = n^{-1} \sum_{i=1}^{n} [g(\boldsymbol{X}_i)\mu^0(\boldsymbol{X}_i) + \{1 - g(\boldsymbol{X}_i)\}\mu^1(\boldsymbol{X}_i)].$$

This can be interpreted as the true population average when everybody in the cohort follows regime g. A good regime will yield larger μ^g. From Table 4.2, we can see that in all scenarios, the true optimal regime yields the largest μ^g, the fact that it is bigger than the regime where everyone in the cohort is treated or everyone is not treated, suggests that the personalized regimes do have some advantage under certain circumstances. In both scenarios, the RSF based methods show larger μ^g than the Cox model based methods, and among the RSF based methods, the proposed method gives larger μ^g in all cases, suggesting that the weighted bootstrap can improve the results for such observational studies.

In both scenarios, the true regime is based only on X_1 and X_2, and not the other 18 covariates. Thus we can also project the estimated optimal regime onto the (X_1, X_2) plane, and look at how close the estimated optimal regime is to the truth on this

plane. Figure 4.1 and 4.2 show the plots for the scenario 1 with 20% and 45% censoring respectively, Figure 4.3 and 4.4 show the results for scenario 2 with 20% and 45% censoring respectively. In each figure, the combined results from pooling all 100 simulated datasets are plotted. We can see in both scenarios, both standard Cox model and weighted Cox model tend to give a linear partition of the treatment decision, while for both RSF based methods, the partition looks close to the true optimal, suggesting that the estimated treatment regimes for most of the patients are close to their true optimal treatment assignment. For both scenarios, comparing to the 45% censoring cases, the regimes estimated from 20% censoring cases are closer to the optimal plots, suggesting the proposed method will perform better with lower censoring rates.

Table 4.1: The Percentage of Subjects Who Would Be Correctly Treated With The Optimal Regime If They Followed The Estimated Optimal Regime. The mean and standard deviation (in parentheses) over the 100 replicates are recorded, separately. For each setting, 4 methods are compared, Cox: standard Cox model; wCox: Cox model with inverse probability weighting; RSF: standard RSF model; bsRSF: proposed method. For each method, the mean and empirical standard error of MSDs and recorded.

	Cox		wCox		RSF		bsRSF	
	mean %	(SD %)	mean %	(SD %)	mean %	(SD %)	mean %	(SD %)
Scenario 1								
20% censoring	49.08	(1.75)	49.87	(1.61)	70.56	(5.74)	77.38	(3.19)
45% censoring	44.73	(1.65)	46.03	(1.88)	58.74	(5.78)	63.23	(4.84)
Scenario 2								
20% censoring	62.68	(4.36)	63.53	(4.60)	68.51	(3.61)	72.58	(5.41)
45% censoring	61.83	(4.35)	61.07	(5.62)	62.20	(3.93)	67.52	(4.67)

4.4 Discussion

Identifying optimal treatment regime in personalized medicine is a very attractive idea as it maximizes the treatment effect at the population level while saving

Table 4.2: Regime Specific Restricted Population Mean Survival Time From Different Methods. The restricted population mean survival time μ^g for the optimal regime estimated for all four methods are listed, which includes standard Cox model (Cox); Cox model with inverse probability weighting (wCox); standard RSF model (RSF) and proposed method(bsRSF). Besides μ^g for the true optimal treatment regime (opt), and the regime where everybody will be in the control arm (all ctrl) and the regime where everybody will receive the treatment (all trt) are also calculated for each simulation setting.

	Cox		wCox		RSF		bsRSF		opt		all ctrl		all trt	
	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
Scenario 1														
20% censoring	8.34	(0.30)	8.23	(0.31)	8.77	(0.34)	8.99	(0.32)	9.89	(0.28)	5.32	(0.15)	8.28	(0.30)
45% censoring	1.80	(0.03)	1.79	(0.04)	1.90	(0.04)	1.94	(0.03)	2.04	(0.03)	1.58	(0.03)	1.79	(0.04)
Scenario 2														
20% censoring	12.06	(0.32)	11.63	(0.30)	12.19	(0.33)	12.47	(0.31)	14.53	(0.29)	9.17	(0.23)	11.02	(0.31)
45% censoring	4.26	(0.08)	4.19	(0.11)	4.26	(0.06)	4.33	(0.08)	4.79	(0.04)	3.63	(0.04)	3.99	(0.04)

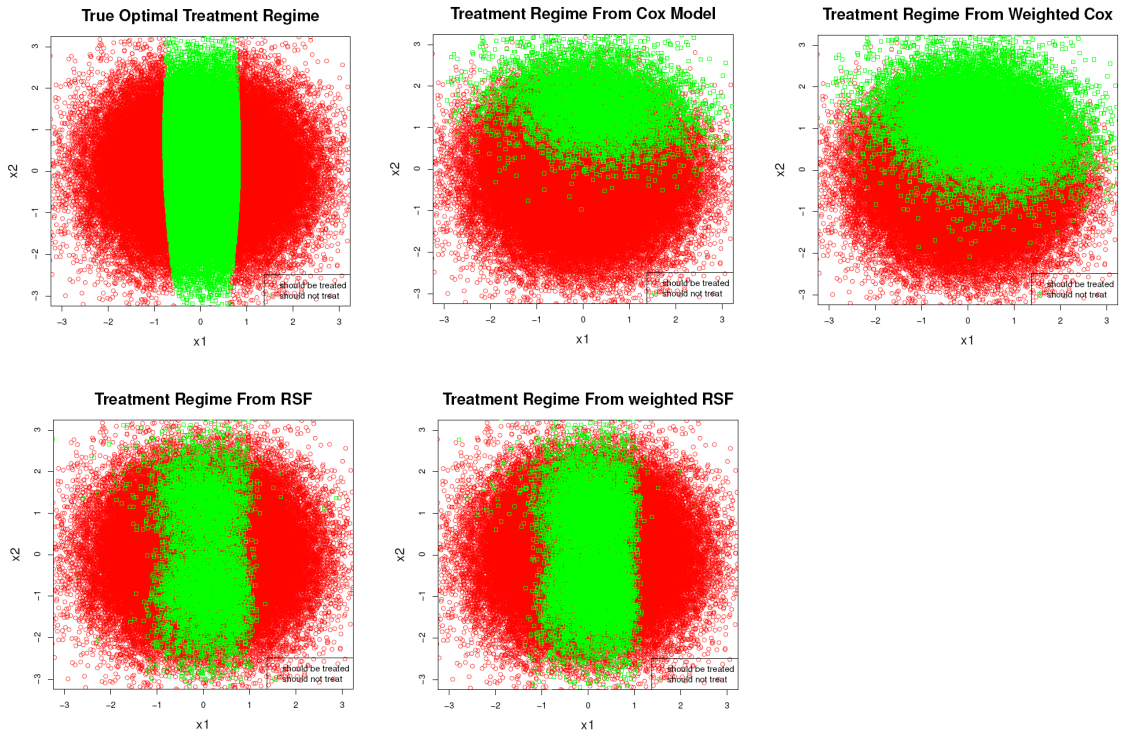


Figure 4.1: Cumulative Plot for Estimated Treatment Regime for Scenario 1 with 20% Censoring. In each plot, the estimated optimal treatment regime is plotted by showing the treatment assignment for all patients over the 100 replicates. The red dots are the ones should receive treatment according to the estimated regime, and green dots are the ones who should not be treated. The plot on top left is the true optimal regime, the middle panel on top is the optimal regime estimated from the standard Cox model, the top right panel is the optimal regime estimated from the weighted Cox model, the bottom left panel is the optimal regime estimated from the standard RSF model, the middle panel at the bottom is the optimal regime estimated from the proposed method.

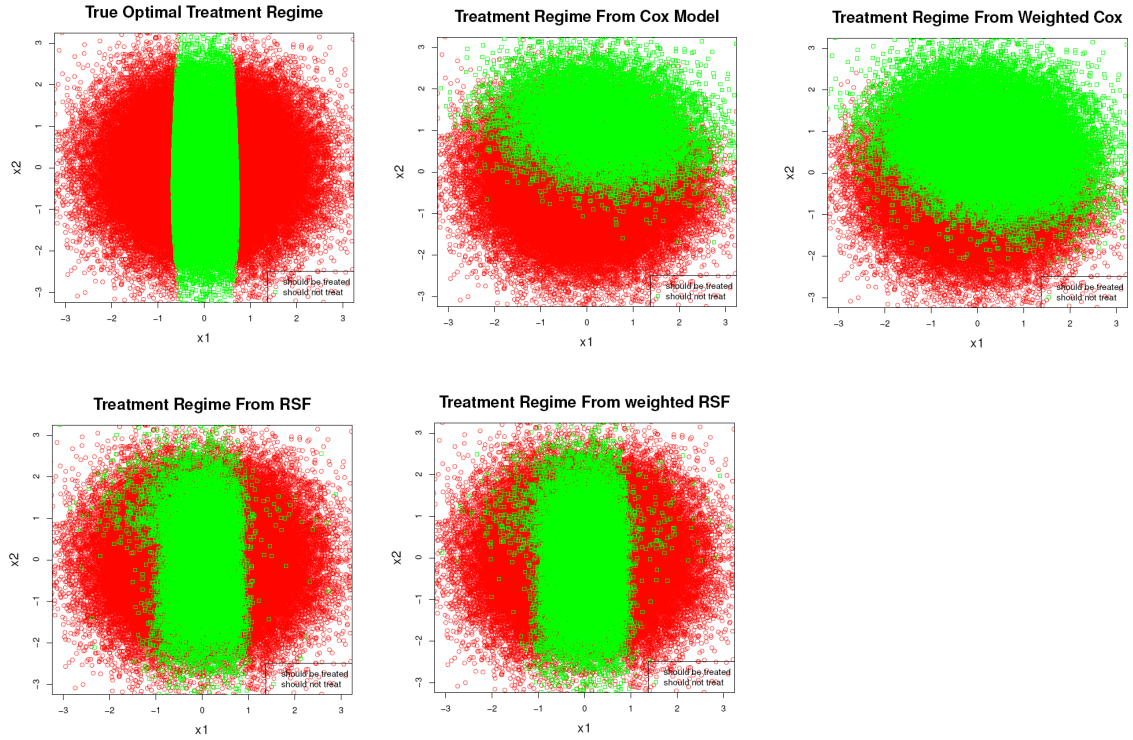


Figure 4.2: Cumulative Plot for Estimated Treatment Regime for Scenario 1 with 45% Censoring. Similarly, the estimated optimal treatment regime is plotted by showing the treatment assignment for all patients over the 100 replicates. The red dots are the ones should receive treatment according to the estimated regime, and green dots are the ones who should not be treated. The plot on top left is the true optimal regime, the middle panel on top is the optimal regime estimated from the standard Cox model, the top right panel is the optimal regime estimated from the weighted Cox model, the bottom left panel is the optimal regime estimated from the standard RSF model, the middle panel at the bottom is the optimal regime estimated from the proposed method.

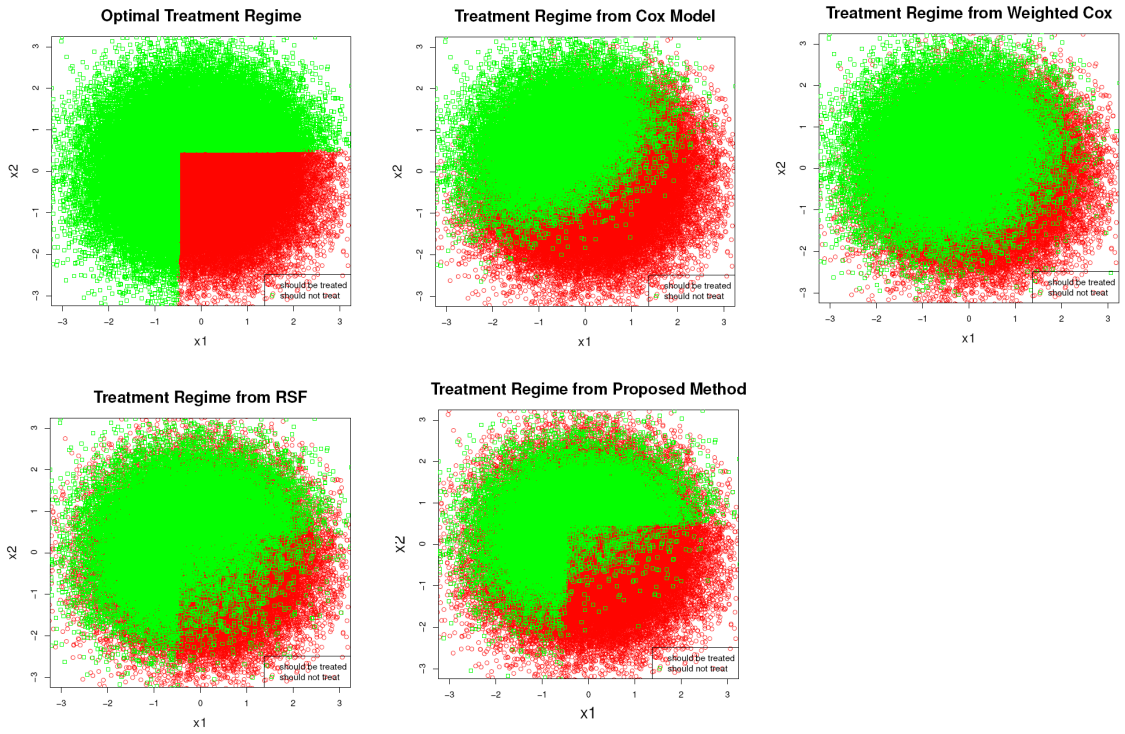


Figure 4.3: Cumulative Plot for Estimated Treatment Regime, Scenario 2 with 20% Censoring. Similarly, the estimated optimal treatment regime is plotted by showing the treatment assignment for all patients over the 100 replicates. The red dots are the ones should receive treatment according to the estimated regime, and green dots are the ones who should not be treated. The plot on top left is the true optimal regime, the middle panel on top is the optimal regime estimated from the standard Cox model, the top right panel is the optimal regime estimated from the weighted Cox model, the bottom left panel is the optimal regime estimated from the standard RSF model, the middle panel at the bottom is the optimal regime estimated from the proposed method.

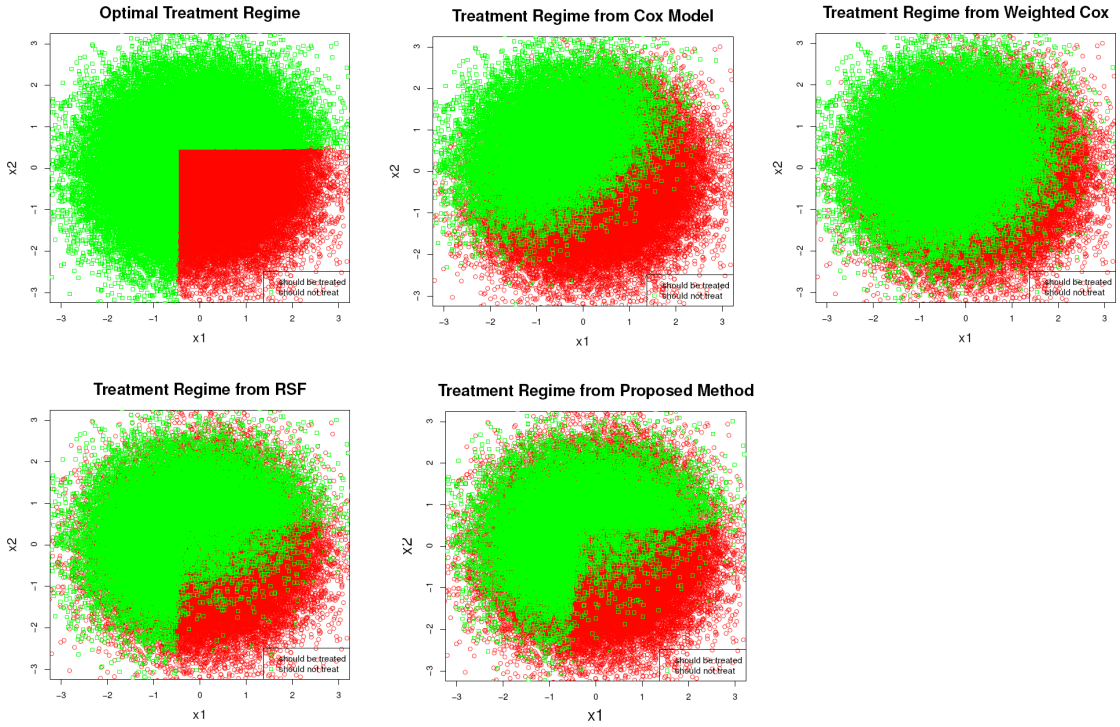


Figure 4.4: Cumulative Plot for Estimated Treatment Regime, Scenario 2 with 45% Censoring. Similarly, the estimated optimal treatment regime is plotted by showing the treatment assignment for all patients over the 100 replicates. The red dots are the ones should receive treatment according to the estimated regime, and green dots are the ones who should not be treated. The plot on top left is the true optimal regime, the middle panel on top is the optimal regime estimated from the standard Cox model, the top right panel is the optimal regime estimated from the weighted Cox model, the bottom left panel is the optimal regime estimated from the standard RSF model, the middle panel at the bottom is the optimal regime estimated from the proposed method.

the resources that will otherwise be allocated to patients who do not response well to the therapy. Even for normally distributed or binary outcome models, a lot of existing methods for estimating the optimal regime have a potential problem as the assumption about the regime space may not match the objective for optimization, this could lead to severe bias even if the proposed space contains the true optimal (*Qian and Murphy*, 2011; *Zhang et al.*, 2012). So it is useful to have a flexible model that does not restrict the space in which the optimal regime can lie. Tree based methods provide very good tools in this case as they allow a very rich model for the interaction between treatment and other covariates. Furthermore, RSF can also control the overfitting problem through its tree building processes. Similar to Random Forest, RSF can select the importance of the variables, which would be considered as a good feature as an optimal treatment regime is commonly more desirable if it requires the measurement of less baseline covariates.

In the simulation study, we consider standard and inverse probability weighted Cox model as alternatives to the RSF based methods, and show that the RSFs give better performance. A possible reason for the inferior performance of the Cox models is that they assume proportional hazards and linear combination of $\boldsymbol{X}$'s, neither of which is true. In practice, a user may consider other formulations within the Cox model framework, which may be a better approximation to the underlying true model, and thus may lead to better properties. However, the selection for the right Cox model would usually be a rather arbitrary process and require extensive experimentation and deep understanding of the underlying mechanism. On the other hand, the nature of RSF methods makes them adaptable to various underlying mechanisms automatically.

Here, we incorporate the inverse probability weighting into the survival modeling via a weighted bootstrap. It enables us to use existing RSF packages to facilitate the

implementation of the proposed method in R. An alternative would be to directly incorporate the inverse probability weights in the tree growing processes, and build each tree on weighted bootstrap samples to form the forest. The proposed method has two layers of bootstrap resampling, which introduces more randomness into the model, and thus would be expected to have better performance numerically. Meanwhile, it should be straightforward to adopt the weighted bootstrap approach to other machine learning methods, especially for the ones where weighted versions are not defined or developed.

One possible extension here would be to consider time-dependent covariates, and optimize not only who to give the treatment to but also when it should be given. The so-called dynamic treatment regime optimization has also drawn a lot of attention of statisticians (*Hernán et al.*, 2006; *Robins et al.*, 2008; *Orellana et al.*, 2010a,b; *Wang et al.*, 2012). Using a Random Survival Forest model with time-dependent covariates would be a promising direction for the development of effective tools to identify optimal dynamic treatment regimes from observational data.

4.5 Appendix: Some Technical Issues

4.5.1 Models Only Conditional on $\boldsymbol{X}$

To identify the optimal regime, ideally, we would like to find the regime which gives the longest conditional mean survival time $E\{T|\boldsymbol{X}\}$ for each subject. However, due to the censoring, we choose to compare the restricted mean survival time $\mu = E\{\min(T,\tau)|\boldsymbol{X}\}$ for some $\tau > 0$. Depend on the choice of τ, this may lead to a slightly different conclusion than direct comparison of the conditional mean survival, we will discuss this issue later. But the first step is to obtain the true marginal models, which are only conditional on $\boldsymbol{X}$. As the full model has both $\boldsymbol{X}$ and Z in it,

some calculation will be needed in order to marginalize over Z. In order to work out the formula for this, here, we will start with simple model where both X and Z are both one dimensional, consider a lognormal model

$$\log T = \beta_0 + \beta_1 X + \beta_2 Z + \epsilon$$

where $Z = \eta X + \epsilon_2$, $X \sim N(0, \sigma_1^2)$, $\epsilon_2 \sim N(0, \sigma_2^2)$, and $\epsilon \sim N(0, \sigma_0^2)$ are independently distributed. Now the quantity of interest is the conditional restricted mean survival $\mu = E\{\min(T, \tau)|X\}$. As the first step we can have the conditional mean for $\log T$ as

$$E\{\log T|X = x\} = \beta_0 + (\beta_1 + \eta\beta_2)x$$

The conditional mean survival time is then

$$\begin{aligned} E\{T|X = x\} &= \int\int e^{\beta_0+\beta_1 x+\beta_2 z+\epsilon} f_{z|x}(z) f_\epsilon(\epsilon) dz d\epsilon \\ &= \int\int e^{\beta_0+(\beta_1+\eta\beta_2)x+\beta_2\epsilon_2+\epsilon} f_{\epsilon_2}(\epsilon_2) f_\epsilon(\epsilon) d\epsilon_2 d\epsilon \\ &= e^{\beta_0+(\beta_1+\eta\beta_2)x} \int e^{\beta_2\epsilon_2}\phi(\epsilon_2)d\epsilon_2 \int e^{\epsilon}\phi(\epsilon)d\epsilon \end{aligned}$$

where

$$\int e^x \phi(x)dx = \int e^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{\sigma^2}} dx = e^{\frac{\sigma^2}{2}}$$

$$\int e^{\beta x} \phi(x)dx = \int e^{\beta x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{\sigma^2}} dx = \int e^{y} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{y^2}{\beta^2\sigma^2}} \frac{1}{\beta} dy = e^{\frac{\beta^2\sigma^2}{2}}$$

thus

$$E\{T|X = x\} = e^{\beta_0+(\beta_1+\eta\beta_2)x+\frac{1}{2}(\beta_2^2\sigma_2^2+\sigma_0^2)}$$

The restricted mean survival (conditional on $\boldsymbol{W} = (X, Z)^T$) is:

$$
\begin{aligned}
E\{\min(T, \tau)|X, Z\} &= \int_0^\infty \min(t, \tau) f(t) dt \\
&= \int_0^\tau t f(t) dt + \int_\tau^\infty \tau f(t) dt \\
&= \int_0^\tau t f(t) dt + \tau P(T > \tau | X, Z) \\
&= \left(e^{\beta_0 + \beta_1 x + \beta_2 z + \frac{1}{2}\sigma_0^2}\right) \Phi\left(\frac{\log \tau - (\beta_0 + \beta_1 x + \beta_2 z) - \sigma_0^2}{\sigma_0}\right) \\
&\quad + \tau \Phi\left(\frac{(\beta_0 + \beta_1 x + \beta_2 z) - \log \tau}{\sigma_0}\right)
\end{aligned}
$$

Next, we can calculate the conditional restricted mean survival (only conditional on X):

$$E\{\min(T,\tau)|X=x\} = \int_0^\tau t f(t)dt + \int_\tau^\infty \tau f(t)dt$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{u} e^{\beta_0+(\beta_1+\eta\beta_2)x+\beta_2\epsilon_2+\epsilon} f(\epsilon)d\epsilon f(\epsilon_2)d\epsilon_2 + \tau \int_{-\infty}^{+\infty}\int_{u}^{\infty} f(\epsilon)d\epsilon f(\epsilon_2)d\epsilon_2$$

$$= e^{\beta_0+(\beta_1+\eta\beta_2)x} \int_{-\infty}^{+\infty} e^{\beta_2\epsilon_2} \int_{-\infty}^{u} e^{\epsilon} f(\epsilon)d\epsilon f(\epsilon_2)d\epsilon_2 + \tau \int_{-\infty}^{+\infty}\int_{u}^{\infty} f(\epsilon)d\epsilon f(\epsilon_2)d\epsilon_2$$

$$= e^{\beta_0+(\beta_1+\eta\beta_2)x} \int_{-\infty}^{+\infty} e^{\beta_2\epsilon_2} e^{\frac{\sigma_0^2}{2}} \Phi(\frac{u-\sigma_0^2}{\sigma_0}) f(\epsilon_2)d\epsilon_2 + \tau \int_{-\infty}^{+\infty} \Phi(-\frac{u}{\sigma_0}) f(\epsilon_2)d\epsilon_2$$

$$= e^{\beta_0+(\beta_1+\eta\beta_2)x+\frac{\sigma_0^2}{2}} \int_{-\infty}^{+\infty} e^{\beta_2\epsilon_2} \Phi\left(\frac{\log\tau-\beta_0-(\beta_1+\eta\beta_2)x-\beta_2\epsilon_2-\sigma_0^2}{\sigma_0}\right) f_{\epsilon_2}(\epsilon_2)d\epsilon_2$$
$$+ \tau \int_{-\infty}^{+\infty} \Phi\left(-\frac{\log\tau-\beta_0-(\beta_1+\eta\beta_2)x-\beta_2\epsilon_2}{\sigma_0}\right) f_{\epsilon_2}(\epsilon_2)d\epsilon_2$$

$$= e^{\beta_0+(\beta_1+\eta\beta_2)x+\frac{\sigma_0^2}{2}} \int_{-\infty}^{+\infty} e^{\beta_2\sigma_2 z} \Phi\left(\frac{\log\tau-\beta_0-(\beta_1+\eta\beta_2)x-\beta_2\sigma_2 z-\sigma_0^2}{\sigma_0}\right) \phi(z)dz$$
$$+ \tau \int_{-\infty}^{+\infty} \Phi\left(-\frac{\log\tau-\beta_0-(\beta_1+\eta\beta_2)x-\beta_2\sigma_2 z}{\sigma_0}\right) \phi(z)dz$$

where $u = \log\tau - \beta_0 - (\beta_1+\eta\beta_2)x - \beta_2\epsilon_2$, and the integral here do not generally have a closed form, in practice, we use Gaussian quadrature based on nor-

mal distribution to numerically get the above quantity with the aid of function *gauss.quad.prob*() in the *statmod* package in R. The conditional restricted mean survival is then both a function of τ and X, so we will use notation $\mu^0(\tau, X)$ and $\mu^1(\tau; X)$ the survival outcome of interest for treated and untreated cases separately. To facilitate the comparison among regimes recommended by different methods, we will also calculate the empirical mean outcome given regime $g(X)$, that is $\mu^g(\tau) = \mathbb{P}_X\left[\mu^0(\tau; X)\{1 - g(X)\} + \mu^1(\tau; X)g(X)\right]$, which is the conditional restricted mean survival time for the targeted population when the whole population follows a given regime g.

4.5.2 The Choice of τ

As mentioned above, $\mu^0(\tau, X)$ and $\mu^1(\tau; X)$ will also depend on the choice of τ, to see this clearly, consider the case where the survival outcome for treated and untreated are

$$\log T^0 = \beta_0 + \beta_1 X + \beta_2 Z + \epsilon_0 \text{ and } \log T^1 = \beta_0 + \beta_1 X + h(X) + \beta_2 Z + \epsilon_1$$

where X and Z could be multi-dimensional. Then the optimal regime to give the longest conditional mean survival would be to give the treatment when $h(X) > 0$ (provided ϵ_0 and ϵ_1 have same distribution). However, the difference between the

restricted mean survivals is

$$
\begin{aligned}
&\mu^1(\tau;X)-\mu^0(\tau;X)\\
=&e^{\beta_0+x(\beta_1+\eta\beta_2)+h(x)+\frac{\sigma_0^2}{2}}\\
&\times\int_{-\infty}^{+\infty}e^{\beta_2\epsilon_{12}}\Phi\left(\frac{\log\tau-\beta_0-(\beta_1+\eta\beta_2)x-h(x)-\beta_2\epsilon_{12}}{\sigma_0}\right)\phi(\epsilon_{12})d\epsilon_{12}\\
&+\tau\int_{-\infty}^{+\infty}\Phi\left(-\frac{\log\tau-\beta_0-(\beta_1+\eta\beta_2)x-h(x)-\beta_2\epsilon_{12}}{\sigma_0}\right)\phi(\epsilon_{12})d\epsilon_{12}\\
&-e^{\beta_0+(\beta_1+\eta\beta_2)x+\frac{\sigma_0^2}{2}}\\
&\times\int_{-\infty}^{+\infty}e^{\beta_2\epsilon_{02}}\Phi\left(\frac{\log\tau-\beta_0-(\beta_1+\eta\beta_2)x-\beta_2\epsilon_{02}}{\sigma_0}\right)\phi(\epsilon_{02})d\epsilon_{02}\\
&-\tau\int_{-\infty}^{+\infty}\Phi\left(-\frac{\log\tau-\beta_0-(\beta_1+\eta\beta_2)x-\beta_2\epsilon_{02}}{\sigma_0}\right)\phi(\epsilon_{02})d\epsilon_{02}\\
=&A+B
\end{aligned}
$$

with

$$
\begin{aligned}
A=&e^{\beta_0+(\beta_1+\eta\beta_2)x+\frac{\sigma_0^2}{2}}\int_{-\infty}^{+\infty}\left\{e^{h(x)}\Phi\left(\frac{\log\tau-\beta_0-(\beta_1+\eta\beta_2)x-h(x)-\beta_2\epsilon_2}{\sigma_0}\right)\right.\\
&\left.-\Phi\left(\frac{\log\tau-\beta_0-(\beta_1+\eta\beta_2)x-\beta_2\epsilon_2}{\sigma_0}\right)\right\}\times e^{\beta_2\epsilon_2}\phi(\epsilon_2)d\epsilon_2
\end{aligned}
$$

$$
\begin{aligned}
B=&\tau\int_{-\infty}^{+\infty}\left\{\Phi\left(-\frac{\log\tau-\beta_0-(\beta_1+\eta\beta_2)x-h(x)-\beta_2\epsilon_2}{\sigma_0}\right)\right.\\
&\left.-\Phi\left(-\frac{\log\tau-\beta_0-(\beta_1+\eta\beta_2)x-\beta_2\epsilon_2}{\sigma_0}\right)\right\}\phi(\epsilon_2)d\epsilon_2
\end{aligned}
$$

we can see here, if $\tau\to\infty$, then $\mu^1(\tau;X)=\mu^0(\tau;X)$ if and only if $h(x)=0$. In practice, people usually choose time to study end or the longest follow-up time, i.e.

τ is always large, in such cases, approximately, the optimal regime is similar to give the treatment when $h(x) > 0$.

CHAPTER V

Discussion and Future Work

Understanding the effect of a treatment on survival time or time to a specific event for patients is one of the central tasks in clinical and health care research. The presence of time-varying covariates in such cases could add more complexity to the problem. There would be both theoretical and numerical challenges to extend existing methods into this area. Furthermore, there has been increasing interest in investigating how the treatment interacts with other pretreatment characteristics and thus to personalize the treatment according to these characteristics of each individual. When the covariates are time-dependent, for example like PSA in the prostate cancer recurrence study described in Chapter I, it is more difficult yet much more appealing to develop a treatment regime that will make treatment decision based on the real time monitored PSA values and optimize the outcome. In fact, the idea of tailoring the treatment dynamically according to the time-varying covariates has a wide spectrum of application. For example, in treating advanced non-small cell lung cancer (NSCLC), patients typically experience two or more lines of treatment, it would be very useful to find the optimal dynamic treatment schedule that can improve survival for patients (*Socinski and Stinchcombe*, 2007). In this dissertation, we focus on providing some insights to the current issues in this field.

In Chapter II, we proposed an estimation procedure to efficiently estimate the coefficient for baseline covariates under a general form of survival problem when there are mixed covariates (both time-dependent and time-independent). One direct application is that this approach can be used to estimate the treatment effect on recurrence free survival time for prostate cancer patients, when the hormone therapy is assigned at baseline and confounding effect of PSA is cumulative over time. It is reasonable to assume that the whole history of the time-dependent variable would impact the risk at current time. In optimal dynamic treatment regime (DTR) identification, the cumulative impact of the time-dependent variable may exist in both the time-dependent treatment assignment mechanism and counterfactual survival outcomes. Thus it may be more useful for a treatment regime to assign treatment based on the whole trajectory of the time-dependent covariates at the current time.

Machine learning methods do not heavily rely on pre-specified assumptions about the structure of the model, which is a desired feature when the underlying model structure is not very well understood. Tree-based methods are increasingly popular statistical tools since *Breiman et al.* (1984) introduced the classification and regression tree (CART) algorithm. The simplicity of the single tree based model gives the prediction rules very straightforward interpretations, however, it often yields poor prediction accuracy. The ideas of ensembles and randomization dramatically improve prediction accuracy, and make the tree-based methods become more and more popular in the research community. Among various ensemble tree-based algorithms, Random Forest is one of the widely used approaches after *Breiman* (2001) introduce its general framework. In Chapter III, we focus on employing Random Forest algorithm in modeling complex treatment mechanism and further developing personalized medicine from observational studies. We propose to model the time-dependent treatment assignment by considering series of conditional probability models at each visit.

The weights are then calculated as a cumulative product of probabilities. This is a common approach for handling time-varying weights. Although it gains efficiency by pooling conditional models at different time points together, it also forces these models to have similar structure which may not be realistic. As mentioned in Chapter III, we can directly model the time to treatment using tools for survival analysis, especially Random Survival Forest will be a very attractive alternative as it does not impose many assumptions. However, we may need to be cautious as we start to consider treatment over continuous time, stronger assumptions may be necessary to enable causal inference in such situations, for example, the positivity assumption. In practice, it may also require much larger populations in order to get reasonable estimation. On the other hand, it may be more straightforward to use RSF for discrete survival time in modeling the treatment mechanism when there are limited number of treatment stages.

In Chapter IV, we propose a modified version of Random Survival Forest method which allows different subjects to contribute differently to the model by giving them different weights. The weighted bootstrap procedure incorporates the inverse probability weights into the model to correct for the selection bias in the observational data. Moreover, it also provides additional randomness to the Random Survival Forest algorithm which is likely to improve the performance of the proposed method. One other possible improvement may be to include higher order terms and interaction terms of the covariates as input into the Random Forest procedure. Although, in theory, Random Survival Forest can automatically model the interactions, it has been shown in numerical work that the inclusion of extra terms like interaction or quadratics may improve the properties of the method in some circumstances (*Foster et al.*, 2011). Our goal in Chapter IV is to identify the optimal regime in the case when all covariates are measured pretreatment and at baseline, and the treatment

is also assigned at time zero. We focus on improving the performance of the modeling for the counterfactual outcomes, so we assume the true model for the treatment mechanism is known and can be correctly modeled. However, when the treatment mechanism is not clear, we may need to model the probability of receiving treatment non-parametrically. Meanwhile, variable selection would also be important as the pool of candidate predictors for the treatment probability could be very large. In such cases, machine learning methods similar to what we proposed in Chapter III would be useful to model the treatment. A more challenging problem would be to extend this approach to the optimal dynamic treatment regime (DTR) identification, where the weights are time-varying. Methods need to developed to allow each person-time piece to contribute differently when building the Random Survival Forest.

Random Forest, Random Survival Forest and their extensions provide very useful tools in detecting optimal DTR from observational data. Their flexible model structure is helpful in learning the treatment and survival outcome especially when the underlying mechanisms are not well understood. Meanwhile, the built-in variable importance measure gives us a tool for automatic variable selection. This is important since simple regimes with less covariates are preferable in practice. However, the original Random Survival Forest (*Ishwaran et al.*, 2008) and most of its variants only work for the case when the covariates are time independent. The major challenge lies in how to define splits on time-dependent covariates. One strategy that has been implemented replaced the time-varying covariate with a low-order polynomial approximation (*Segal*, 1992). In particular, linear summaries have been used where each time-varying covariate is first regressed against time within individuals. The intercept and slope for each individual are then used as covariates. This would highly depend on that the model assumption for the regression on the longitudinal covariate is correct, and as the regression is subject specific, it also requires that there

are enough observations per subject to make each regression accurate. Neither of the requirement would be easily satisfied in practice. Recently, *Bou-Hamad et al.* (2011a) proposed a new random survival forest method to accommodate simultaneously time-varying covariates and time-varying effects. Their basic idea is to handle time-varying covariates by decomposing each subject as subject-time pieces, or as they called them pseudo-subjects, according to the splitting rules. So one subject can be split apart across two children nodes. More work along this line can be done for covariates in continuous time, as well as its application in finding optimal DTRs.

BIBLIOGRAPHY

BIBLIOGRAPHY

Billingsley, P. (1986), Probability and measure, *Wiley Series in Probability and Mathematical Statistics*.

Bou-Hamad, I., D. Larocque, and H. Ben-Ameur (2011a), Discrete-time survival trees and forests with time-varying covariates application to bankruptcy data, *Statistical Modelling*, *11*(5), 429–446.

Bou-Hamad, I., D. Larocque, H. Ben-Ameur, et al. (2011b), A review of survival trees, *Statistics Surveys*, *5*, 44–71.

Breiman, L. (2001), Random forests, *Machine Learning*, *45*(1), 5–32.

Breiman, L., J. Friedman, R. Olshen, and C. Stone (1984), Classification and regression trees, *Wadsworth International Group*.

Buckley, J., and I. James (1979), Linear regression with censored data, *Biometrika*, *66*(3), 429–436.

Cain, L. E., J. M. Robins, E. Lanoy, R. Logan, D. Costagliola, and M. A. Hernán (2010), When to start treatment? a systematic approach to the comparison of dynamic regimes using observational data, *The International Journal of Biostatistics*, *6*(2).

Chapman, J., et al. (2013), Innovative estimation of survival using log-normal survival modelling on accent database, *British Journal of Cancer*, *108*(4), 784–790.

Chen, S., and L. Zhou (2007), Local partial likelihood estimation in proportional hazards regression, *The Annals of Statistics*, pp. 888–916.

Cox, D. R. (1972), Regression models and life tables (with discussion), *Journal of the Royal Statistical Society*, *34*, 187–220.

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

D'Agostino, R. B., M.-L. Lee, A. J. Belanger, L. A. Cupples, K. Anderson, and W. B. Kannel (1990), Relation of pooled logistic regression to time dependent cox regression analysis: the framingham heart study, *Statistics in Medicine*, *9*(12), 1501–1515.

Ding, Y., and B. Nan (2011), A sieve m-theorem for bundled parameters in semiparametric models, with application to the efficient estimation in a linear model for censored data, *The Annals of Statistics*, *39*(6), 3032–3061.

Foster, J. C., Taylor, J. M. G., and S. J. Ruberg (2011), Subgroup identification from randomized clinical trial data, *Statistics in Medicine*, *30*(24), 2867–2880.

Friedman, J. H. (1991), Multivariate adaptive regression splines, *The Annals of Statistics*, pp. 1–67.

Geman, S., and C.-R. Hwang (1982), Nonparametric maximum likelihood estimation by the method of sieves, *The Annals of Statistics*, pp. 401–414.

Hastie, T. J., and R. J. Tibshirani (1990), *Generalized additive models*, vol. 43, CRC Press.

Hernán, M. A., E. Lanoy, D. Costagliola, and J. M. Robins (2006), Comparison of dynamic treatment regimes via inverse probability weighting, *Basic Clinical Pharmacology & Toxicology*, *98*(3), 237–242.

Hothorn, T., B. Lausen, A. Benner, and M. Radespiel-Troger (2004), Bagging survival trees, *Statistics in Medicine*, *23*(1), 77–91.

Hothorn, T., P. Bühlmann, S. Dudoit, A. Molinaro, and M. J. Van Der Laan (2006), Survival ensembles, *Biostatistics*, *7*(3), 355–373.

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

Ishigooka, J., M. Murasaki, S. Miura, and The Olanzapine Late-Phase II Study Group (2000), Olanzapine optimal dose: Results of an open-label multicenter study in schizophrenic patients, *Psychiatry and clinical neurosciences*, *54*(4), 467–478.

Ishwaran, H., U. B. Kogalur, E. H. Blackstone, and M. S. Lauer (2008), Random survival forests, *The Annals of Applied Statistics*, pp. 841–860.

Jin, Z., D. Lin, L. Wei, and Z. Ying (2003), Rank-based inference for the accelerated failure time model, *Biometrika*, *90*(2), 341–353.

Jin, Z., D. Lin, and Z. Ying (2006), On least-squares regression with censored data, *Biometrika*, *93*(1), 147–161.

Kalbfleisch, J. D., and R. L. Prentice (2011), *The statistical analysis of failure time data*, vol. 360, John Wiley & Sons.

Kehl, V., and K. Ulm (2006), Responder identification in clinical trials with censored data, *Computational Statistics & Data Analysis*, *50*(5), 1338–1355.

Kennedy, E. H., Taylor, J. M. G., D. E. Schaubel, and S. Williams (2010), The effect of salvage therapy on survival in a longitudinal study with treatment by indication, *Statistics in Medicine*, *29*(25), 2569–2580.

Köhler, H. F., and L. P. Kowalski (2012), A critical appraisal of different survival techniques in oral cancer patients, *European Archives of Oto-Rhino-Laryngology*, *269*(1), 295–301.

Kong, E. (2010), Statistical modelling of nonlinear long-term cumulative effects, *Statistica Sinica*, *20*(3), 1097–1123.

Lai, T. L., and Z. Ying (1991), Large sample theory of a modified buckley-james estimator for regression analysis with censored data, *The Annals of Statistics*, pp. 1370–1402.

LeBlanc, M., and J. Crowley (1995), A review of tree-based prognostic models, in *Recent Advances in Clinical Trial Design and Analysis*, pp. 113–124, Springer.

Longford, N. T. (2009), Inference with the lognormal distribution, *Journal of Statistical Planning and Inference*, *139*(7), 2329–2340.

McCullagh, P., and J. A. Nelder (1989), *Generalized linear model*, vol. 37, Chapman & Hall/CRC.

Murphy, S. A. (2003), Optimal dynamic treatment regimes, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *65*(2), 331–355.

Murphy, S. A. (2005), An experimental design for the development of adaptive treatment strategies, *Statistics in Medicine*, *24*(10), 1455–1481.

Murphy, S. A., and P. K. Sen (1991), Time-dependent coefficients in a cox-type regression model, *Stochastic Processes and their Applications*, *39*(1), 153–180.

Orellana, L., A. Rotnitzky, and J. M. Robins (2010a), Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes, part i: main content, *The International Journal of Biostatistics*, *6*(2), 8.

Orellana, L., A. Rotnitzky, and J. M. Robins (2010b), Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes, part ii: Proofs of results, *The International Journal of Biostatistics*, *6*(2), 8.

Piquette-Miller, M., and D. Grant (2007), The art and science of personalized medicine, *Clinical Pharmacology and Therapeutics*, *81*(3), 311–315.

Prentice, R. L. (1978), Linear rank tests with right censored data, *Biometrika*, *65*(1), 167–179.

Proust-Lima, C., Taylor, J. M. G., S. G. Williams, D. P. Ankerst, N. Liu, L. L. Kestin, K. Bae, and H. M. Sandler (2008), Determinants of change in prostate-specific antigen over time and its association with recurrence after external beam radiation therapy for prostate cancer in five large cohorts, *International Journal of Radiation Oncology Biology Physics*, *72*(3), 782–791.

Qian, M., and S. A. Murphy (2011), Performance guarantees for individualized treatment rules, *Annals of Statistics*, *39*(2), 1180.

Reid, N. (1994), A conversation with sir david cox, *Statistical Science*, *9*(3), 439–455.

Ritov, Y. (1990), Estimation in a linear regression model with censored data, *The Annals of Statistics*, pp. 303–328.

Ritov, Y., and J. A. Wellner (1988), Censoring, martingales, and the cox model, *Contemp. Math*, *80*, 191–219.

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), 1393–1512.

Robins, J. (1989a), The control of confounding by intermediate variables, *Statistics in Medicine*, *8*(6), 679–701.

Robins, J., L. Orellana, and A. Rotnitzky (2008), Estimation and extrapolation of optimal treatment and testing strategies, *Statistics in Medicine*, *27*(23), 4678–4721.

Robins, J. M. (1989b), 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.

Robins, J. M. (1993), Information recovery and bias adjustment in proportional hazards regression analysis of randomized trials using surrogate markers, *Proceedings of the Biopharmaceutical Section, American Statistical Association*, *24*(3), 3.

Robins, J. M. (1994), Correcting for non-compliance in randomized trials using structural nested mean models, *Communications in Statistics-Theory and Methods*, *23*(8), 2379–2412.

Robins, J. M. (1997), Causal inference from complex longitudinal data, in *Latent Variable Modeling and Applications to Causality*, pp. 69–117, Springer.

Robins, J. M. (2002), Analytic methods for estimating hiv-treatment and cofactor effects, *Methodological Issues in AIDS Behavioral Research*, pp. 213–288.

Robins, J. M. (2004), Optimal structural nested models for optimal sequential decisions, in *Proceedings of the Second Seattle Symposium in Biostatistics*, pp. 189–326, Springer.

Rothwell, P. M. (2005), Subgroup analysis in randomised controlled trials: importance, indications, and interpretation, *The Lancet*, *365*(9454), 176–186.

Rotnitzky, A., and J. M. Robins (2005), Inverse probability weighting in survival analysis, *Encyclopedia of Biostatistics*.

Royston, P. (2001), The lognormal distribution as a model for survival time in cancer, with an emphasis on prognostic factors, *Statistica Neerlandica*, *55*(1), 89–104.

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

Rubin, D. B. (1990), Comment: Neyman (1923) and causal inference in experiments and observational studies, *Statistical Science*, *5*(4), 472–480.

Schumaker, L. L. (1981), *Spline functions: basic theory*, vol. 1981, Wiley New York.

Segal, M. R. (1992), Tree-structured methods for longitudinal data, *Journal of the American Statistical Association*, *87*(418), 407–418.

Shen, X., and W. H. Wong (1994), Convergence rate of sieve estimates, *The Annals of Statistics*, pp. 580–615.

Siegel, R., D. Naishadham, and A. Jemal (2013), Cancer statistics, 2013, *CA: a cancer journal for clinicians*, *63*(1), 11–30.

Socinski, M. A., and T. E. Stinchcombe (2007), Duration of first-line chemotherapy in advanced non–small-cell lung cancer: less is more in the era of effective subsequent therapies, *Journal of Clinical Oncology*, *25*(33), 5155–5157.

Stone, C. J. (1985), Additive regression and other nonparametric models, *The annals of Statistics*, pp. 689–705.

Stone, C. J. (1986), The dimensionality reduction principle for generalized additive models, *The Annals of Statistics*, pp. 590–606.

Taylor, J. M. G., J. Shen, E. H. Kennedy, L. Wang, and D. E. Schaubel (2013), Comparison of methods for estimating the effect of salvage therapy in prostate cancer when treatment is given by indication, *Statistics in Medicine*.

Tsiatis, A. A. (1990), Estimating regression parameters using linear rank tests for censored data, *The Annals of Statistics*, pp. 354–372.

Van der Laan, M. J., and J. M. Robins (2003), *Unified methods for censored longitudinal data and causality*, Springer.

van der Vaart, A., and J. Wellner (1996), *Weak Convergence and Empirical Processes: With Applications to Statistics*, Springer.

van der Vaart, A. W. (1998), *Asymptotic statistics*, Cambridge University Press (Cambridge, UK and New York, NY, USA).

Wagner, E. H., B. T. Austin, C. Davis, M. Hindmarsh, J. Schaefer, and A. Bonomi (2001), Improving chronic illness care: translating evidence into action, *Health Affairs*, *20*(6), 64–78.

Wang, L., A. Rotnitzky, X. Lin, R. E. Millikan, and P. F. Thall (2012), Evaluation of viable dynamic treatment regimes in a sequentially randomized trial of advanced prostate cancer, *Journal of the American Statistical Association*, *107*(498), 493–508.

Watkins, C. J., and P. Dayan (1992), Q-learning, *Machine Learning*, *8*(3-4), 279–292.

Xia, Y., and H. Tong (2006), Cumulative effects of air pollution on public health, *Statistics in Medicine*, *25*(20), 3548–3559.

Ying, Z. (1993), A large sample study of rank estimation for censored regression data, *The Annals of Statistics*, pp. 76–99.

Zagars, G. K., and A. C. von Eschenbach (1993), Prostate-specific antigen. an important marker for prostate cancer treated by external beam radiation therapy, *Cancer*, *72*(2), 538–548.

Zeng, D., and D. Lin (2007), Efficient estimation for the accelerated failure time model, *Journal of the American Statistical Association*, *102*(480), 1387–1396.

Zhang, B., A. A. Tsiatis, E. B. Laber, and M. Davidian (2012), A robust method for estimating optimal treatment regimes, *Biometrics*, *68*(4), 1010–1018.

Zhang, B., A. A. Tsiatis, E. B. Laber, and M. Davidian (2013), Robust estimation of optimal dynamic treatment regimes for sequential treatment decisions, *Biometrika*, *100*(2), 2.

Zhang, Y., J. Huang, and L. Hua (2009), An adaptive spline-based sieve semiparametric maximum likelihood estimation for the cox model with interval-censored data, *Tech. rep.*, the University of Iowa, Department of Biostatistics.

Zhao, Y., D. Zeng, A. J. Rush, and M. R. Kosorok (2012), Estimating individualized treatment rules using outcome weighted learning, *Journal of the American Statistical Association*, *107*(499), 1106–1118.