Statistical Methods for Evaluating Effects of Health Service Utilization on Survival Times with Linked Administrative Data, with Application to Opioid Use Disorder Treatment Management

by

Trevor James Thomson

M.Sc., Simon Fraser University, 2017
B.Sc., University of Winnipeg, 2015

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
Department of Statistics and Actuarial Science
Faculty of Science

SIMON FRASER UNIVERSITY
Summer 2023

Declaration of Committee

Name: Trevor James Thomson

Degree: **Doctor of Philosophy**

Thesis title: **Statistical Methods for Evaluating Effects of Health Service Utilization on Survival Times with Linked Administrative Data, with Application to Opioid Use Disorder Treatment Management**

Committee:

Chair: Liangliang Wang
Associate Professor, Statistics and Actuarial Science

X. Joan Hu
Supervisor
Professor, Statistics and Actuarial Science

Thomas Loughin
Committee Member
Professor, Statistics and Actuarial Science

Bohdan Nosyk
Committee Member
Professor, Health Sciences

Rhonda J. Rosychuk
Committee Member
Adjunct Professor, Statistics and Actuarial Science

Himchan Jeong
Examiner
Assistant Professor, Statistics and Actuarial Science

Paul S. Albert
External Examiner
Senior Investigator, Biostatistics Branch, National Cancer Institute - Division of Cancer Epidemiology & Genetics
National Institutes of Health

Abstract

Due to their collection of rich information over time, administrative databases have become a popular data source to conduct population-based health research. Motivated by estimating the long-term effect of an opioid agonist treatment (OAT) on mortality risk, this dissertation develops methodology for estimating effects of health service utilization on survival times. We frame the health service utilization of interest as an internal covariate under extended Cox regression models. An internal covariate obstructs the conventional relationship between the hazard and survivor functions. This not only invalidates likelihood methods to infer model parameters, but also makes the associated survival probabilities redundant. We review two general approaches to overcome challenges brought on by internal covariates: avoid all inference procedures that rely on the survivor function, or somehow rectify the relationship between the hazard and survivor functions.

We conduct a preliminary analysis with the motivating data by jointly modelling time-to-death and the health service utilization records. This is done by summarizing the health records with either one-jump processes, functional principal scores, or a random effect. We demonstrate how this approach greatly reduces the computational complexity that currently plagues traditional joint models, and is capable of providing survival predictions. Our results not only reveal the OAT dispensation effect to be time-varying, but younger individuals receive the greatest protective effect of OAT against mortality, despite having the smallest OAT dispensation rates.

To utilize the entire observed internal covariate history in a survival model, we propose a generalized Cox regression model and conduct time-dependent stratification, where the strata are defined by the history of health records. An estimating equation based procedure is adopted to estimate model parameters, and a testing procedure is proposed to update the the stratification variable. The proposed approach is examined both asymptotically through modern empirical process theory, and numerically via simulation. Our analysis shows the effect on mortality risk decreases in successive OAT attempts, in which two risk classes based on an individual's treatment episode number are established: (i) 1-3 OAT episodes, and (ii) 4+ OAT episodes.

We revisit our modelling from the preliminary analysis to address the apparent bias upon directly replacing an internal covariate process with a model based summary. Specifically, we extend the conditional score approach of Tsiatis and Davidian (2001) by allowing successive observations in the health service utilization process to be correlated. Moreover, since the conventional relationship between the hazard and survivor functions is held intact, the proposed method is also capable of producing model based survival predictions. Through a simulation study, our proposed method is demonstrated to produce consistent estimates, whereas naively ignoring the autocorrelation within the data underestimates the true effect. We additionally address confounding by age within our data application by weighting birth generation specific estimates by their relative group size. This procedure up-weights contributions made by younger individuals, which reveals an overall protective treatment effect against mortality.

We then extend our developed methodology to allow the health service utilization process to be multivariate. Specifically, we estimate the effect of a multivariate internal covariate process on the mortality hazard, and extend the conditional score approach in hopes of obtaining survival predictions. In the context of our data application, we extract the OAT dispensation indicator, OAT type, and dosage level from the OAT dispensation process,

Although the proposed research is motivated within the context of opioid use disorder management through health service utilization records, we anticipate the methodology to have broad applications, and the proposed methods are intuitive and simple to implement.

Keywords: Administrative service utilization records; Cox regression; Internal covariate; Joint modelling; Opioid agonist treatment; Time-dependent stratification

Acknowledgements

I would first like to express my heartfelt appreciation to Professor Joan Hu, for her guidance and support throughout my time as a graduate student. Her mentorship has been instrumental in making this research possible, and shaped me into the researcher that I am today. I am fortunate to have been under her tutelage since entering SFU as an MSc. student. I want to thank Professor Bohdan Nosyk and his HERU team at St. Paul's Hospital for the opportunity to use their study program to motivate my thesis research, and their assistance to address questions whenever they arose. I would also like to thank my thesis committee members for their insightful comments and suggestions to further improve the final product.

I would like to express my appreciation to the faculty and staff within the Department of Statistics and Actuarial Science for fostering an inclusive and engaging environment to work. Special thanks goes to my fellow graduate students for their support and friendship. In particular, I want to thank Anqi (Angela) Chen, JinCheol Choi, Renny Doig, Tianyu Guan, Lulu Guo, Khalif Halani, Grace Hsu, Megan Kurz, Dongdong Li, Yan (Lillian) Lin, Shaun McDonald, Payman Nickchi, Kangyi (Ken) Peng, William Ruth, Chenlu Shi, Nikola Surjanovic, Qi (Emma) Wen, Yi Xiong, Yuping Yang, and Haoxuan (Charlie) Zhou.

I would also like to thank my parents, Brenda and Lorne; and my brothers, Jonathon and Shawn, for their everlasting support and encouragement throughout this journey. I would additionally like to thank members of the Simon Fraser University Concert Orchestra (SFUCO) for their continued friendship and invaluable solace during tough times. I would also like to thank additional members of my "second" family: Leon, Lindley, Nathan, and Milo Chow for their care and hospitality, especially during the COVID-19 pandemic.

Last but not least, I would like to thank my beautiful soulmate, Caitlin Chow. It is impossible to summarize in a few sentences how much her love and support has helped me throughout this journey. Beyond her unyielding belief in me, Caitlin was the heart and soul of my emotional support system throughout this journey. I eagerly await the adventures that lie ahead, knowing that any challenge can be conquered with you by my side.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Administrative databases have become an increasingly popular data source to conduct population-based health research (Hinds *et al.* 2016) due in part to its rich collection of service utilization records (Iron and Manuel 2007). Although these databases were not initially intended for research purposes, the wealth of information provides researchers an opportunity to explore how a clinically meaningful event is associated with service utilization records; see, for example, Bharat *et al.* (2021), Barbieri *et al.* (2022), and Clair *et al.* (2022). When risk factors are time-independent, the conventional approach to do this is to specify a Cox regression model (Cox 1972) that relates the risk factors to the instantaneous rate of the event, that is, the conditional hazard function of the event time. The analysis results would typically be summarized by the fitted regression model, and survival predictions are readily available. However, as many risk factors from administrative databases are derived from an individual's service utilization, these risk factors are not only time-varying, but their mere existence implies an individual's survivorship. This obstructs the standard relationship between the hazard and survivor functions, and traditional methods from survival analysis are no longer applicable. This thesis explores methods to overcome this hurdle, and develops procedures to estimate effects of time-varying risk factors on an event time of interest. This provides a tool to assess effects of different medical decisions, and the analysis outcome can guide clinicians to optimally select an action that best addresses their patient's specific needs.

1.1 Background and Motivation

1.1.1 Motivating Example

Due to a rise of opioid-related deaths following the introduction of fentanyl and its analogues into the illicit drug supply, a public health emergency was declared in April of 2016, in British Columbia (BC), Canada (Province of British Columbia 2018). One of the major public health responses to this emergency was the increased delivery of opioid agonist

treatments (OATs) for individuals with an opioid use disorder. An OAT is the first-line treatment for moderate to severe opioid use disorders, and works to eliminate withdrawal symptoms and block euphoric effects of other opioids. *Methadone*, a synthetic opioid agonist with high μ-opioid receptor binding affinity (Tetrault and Fiellin 2012), has been the most commonly prescribed OAT since its introduction in Canada in 1959 (Paulus and Halliday 1967). The primary alternative to methadone is *buprenorphine*, a partial agonist with high affinity at the μ-opioid receptor and an antagonist at the κ-opioid receptor (Lewis 1985), and was first prescribed in Canada in 2007 (Sud *et al.* 2022). Compared to methadone, prior studies have shown buprenorphine to have a shorter induction period, milder side effects and withdrawal symptoms, fewer drug interactions, a decreased risk of overdose due to a partial agonist "ceiling effect", and reduced risks of respiratory depression (British Columbia Centre on Substance Use 2017). Other OAT types that have recently been introduced in British Columbia include *slow-release oral morphine* (Government of British Columbia 2018, Laing *et al.* 2018) in 2014, and injectable forms of OAT, such as *diacetylmorphine* (pharmaceutical-grade heroin) and *hydromorphone* in 2016 and 2019, respectively (Health Canada 2016, Health Canada 2019).

Randomized controlled trials and observational studies have consistently demonstrated protective effects of OAT usage against mortality; see Sordo *et al.* (2017) and Santo Jr *et al.* (2021) for systematic reviews and meta-analyses, and Larney *et al.* (2023) for a recent example. Nonetheless, a consensus study report from National Academies of Sciences, Engineering, Medicine (2019) highlighted the need to determine the most appropriate medication for key population subgroups, and the comparative effectiveness between medications over long term exposure. In an ideal setting, these questions can be addressed by designing an experiment where individuals are randomly assigned to treatment arms, and survival outcomes are then contrasted in order to estimate treatment effects. This approach however encounters many challenges arising from ethical and economic constraints; however, administrative service utilization records provides an alternative data source that bypasses such hurdles. The health economic research unit at St. Paul's Hospital led by Bohdan Nosyk extracted all the administrative service utilization records between 01/01/1996 to 10/01/2018 in BC, Canada. The data include drug dispensations (British Columbia Ministry of Health 2018*c*), hospital and emergency department admissions (British Columbia Ministry of Health 2018*a*), physician billing records (British Columbia Ministry of Health 2018*b*), incarceration records (Ministry of Public Safety and Solicitor General 2018), and deaths (British Columbia Vital Statistics Agency 2018), for individuals identified with an opioid use disorder. These health records provide valuable information on the overall quality of care for this population, and in particular, to help address clinical questions pertaining to OAT usage within a real-world setting (Piske *et al.* 2020). Since prescribed treatments were based on treatment availability, current clinical guidelines, and individual clinician expertise, we cannot deduce any form of treatment randomization taking place. Instead, the data can be viewed as if it were collected

from a hypothetical observational study, wherein the type of OAT and dosage can vary over time based on a clinician's discretion. Furthermore, even an individual's OAT dispensation indicator is time-varying, as it is dependent on them interacting with the healthcare system. Given an individual's OAT dispensation indicator implies their survivorship, estimating the association between the OAT dispensation process and mortality risk is challenging.

1.1.2 Survival Analysis: Cox Regression with Time-Independent Covariates

Let T be a nonnegative random variable representing the time to an event of interest (e.g. death). For our purpose, assume that T is a continuous random variable. Suppose we are interested in $q \geq 1$ time-independent risk factors, denoted by $\boldsymbol{X} = (X_1, \cdots, X_q)'$. The objective is to quantify the association of T with $\boldsymbol{X}$. The standard approach to accomplish this task is to (i) obtain a random sample of size n from the target population, $\{(T_i, \boldsymbol{X}_i) : i = 1, \cdots, n\}$, and then (ii) proceed to estimate the conditional distribution of T given $\boldsymbol{X}$. In survival analysis however, we do not typically observe T_i for all i, but instead, observe the pair (T_i^*, δ_i), where $T_i^* = T_i \wedge C_i$ is the minimum between T_i and a non-informative censoring time C_i, and $\delta_i = I(T_i \leq C_i)$ indicates if the event of interest is observed. In other words, the available data takes the form $\{(T_i^*, \delta_i, \boldsymbol{X}_i) : i = 1, \cdots, n\}$. Furthermore, it is conventional to estimate the conditional hazard function of T:

$$\lambda(t; \boldsymbol{X}) = \lim_{\Delta t \to 0^+} \frac{1}{\Delta t} P(t \leq T < t + \Delta t | T \geq t, \boldsymbol{X}), \tag{1.1}$$

which is interpreted as the expected instantaneous rate of the event occurring, given that the event has not yet occurred at time t and $\boldsymbol{X}$. For fixed $t > 0$, we can show that $\lambda(t; \boldsymbol{X})$ is connected to $P(T > t | \boldsymbol{X})$ by partitioning the interval $[0, t]$ as $0 = u_0 < u_1 < \cdots < u_R = t$ with $\Delta u_{r-1} = u_r - u_{r-1}$ such that $\max\{\Delta u_{r-1}\} \to 0$ as $R \to \infty$. From (1.1), we have

$$\begin{aligned} P(T \in [t, t + \Delta t) | T \geq t, \boldsymbol{X}) &= \lambda(t; \boldsymbol{X}) \Delta t + o(\Delta t) \\ P(T \notin [t, t + \Delta t) | T \geq t, \boldsymbol{X}) &= 1 - \lambda(t; \boldsymbol{X}) \Delta t + o(\Delta t), \end{aligned} \tag{1.2}$$

where we use "little-o" notation to represent $o(a)/a \to 0$ as $a \to 0$. It follows that

$$P(T > t | \boldsymbol{X}) = \lim_{R \to \infty} \prod_{r=1}^{R} P(T \notin [u_{r-1}, u_r) | T \geq u_{r-1}, \boldsymbol{X}) \tag{1.3}$$

$$= \lim_{R \to \infty} \prod_{r=1}^{R} (1 - \lambda(u_{r-1}; \boldsymbol{X}) \Delta u_{r-1} + o(\Delta u_{r-1})) \tag{1.4}$$

$$= \exp\{- \int_0^t \lambda(u; \boldsymbol{X}) du\}, \tag{1.5}$$

where (1.3) is obtained from the joint probability of the event not occurring in any $[u_{r-1}, u_r)$ interval ($r = 1, \cdots, R$), (1.4) follows from (1.2), and (1.5) is the so-called product limit of (1.4). The relationship between the hazard and survivor functions in (1.5) shows us that we can naturally estimate survival probabilities upon estimating the conditional hazard function in (1.1). This is often done with a Cox proportional hazards model (Cox 1972):

$$\lambda(t; \boldsymbol{X}) = \lambda_0(t) \exp\{\boldsymbol{\theta}' \boldsymbol{X}\}, \tag{1.6}$$

where $\lambda_0(t) = \lambda(t; \mathbf{0})$ is an unspecified function referred to as the baseline hazard function, and $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_q)'$ is an unknown parameter that quantifies the association of T with $\boldsymbol{X}$. The model in (1.6) is referred to as a *proportional hazards model* since the hazard ratio between individuals i and j

$$\frac{\lambda(t; \boldsymbol{X}_i)}{\lambda(t; \boldsymbol{X}_j)} = \exp\{\boldsymbol{\theta}'(\boldsymbol{X}_i - \boldsymbol{X}_j)\}, \tag{1.7}$$

which we see is constant over time. In fact, if $\boldsymbol{X}_j = \mathbf{0}$ so that (1.7) becomes $\exp\{\boldsymbol{\theta}' \boldsymbol{X}_i\}$, θ_k being positive or negative implies a heightened or reduced risk of the event occurring if the kth variable in $\boldsymbol{X}$ is increased from zero, respectively.

With the available data $\{(T_i^*, \delta_i, \boldsymbol{X}_i) : i = 1, \cdots, n\}$, we can estimate $\lambda_0(\cdot)$ and $\boldsymbol{\theta}$ by maximizing the observed data likelihood function

$$L(\lambda_0(\cdot), \boldsymbol{\theta}) = \prod_{i=1}^{n} \lambda(T_i^*; \boldsymbol{X}_i)^{\delta_i} \exp\{-\int_0^{T_i^*} \lambda(u; \boldsymbol{X}_i) du\}, \tag{1.8}$$

where we used the result from (1.5) to obtain (1.8). Since $\lambda_0(\cdot)$ is an infinite-dimensional "parameter", one may specify it as a piece-wise constant function between uncensored failure times observed in the data, so that there are only a finite number of "parameters" attributed to this function. Letting $N_i(t) = I(T_i \leq t)$ and $Y_i(t) = I(T_i^* \geq t)$, and $\Lambda_0(t) = \int_0^t \lambda_0(u) du$, we can differentiate the logarithm of (1.8) with respect to the unknown "parameters" in $\Lambda_0(\cdot)$. This process yields the following estimator for $d\Lambda_0(t)$:

$$d\hat{\Lambda}_0(t; \boldsymbol{\theta}) = \sum_{i=1}^{n} \frac{Y_i(t) dN_i(t)}{E_0(t, \boldsymbol{\theta})}, \tag{1.9}$$

where

$$E_r(t, \boldsymbol{\theta}) = \sum_{j=1}^{n} E_{rj}(t, \boldsymbol{\theta}) = \sum_{j=1}^{n} Y_j(t) \exp\{\boldsymbol{\theta}' \boldsymbol{X}_j\} \boldsymbol{X}_j^{\otimes r},$$

with $\boldsymbol{a}^{\otimes 0} = 1$, and $\boldsymbol{a}^{\otimes 1} = \boldsymbol{a}$. By differentiating the logarithm of (1.8) with respect to $\boldsymbol{\theta}$ and substituting (1.9) for the unknown term $d\Lambda_0(t)$, we obtain the following score function for

$\boldsymbol{\theta}$:

$$U(\boldsymbol{\theta}) = \sum_{i=1}^{n} \int_0^{\infty} Y_i(t) \left(\boldsymbol{X}_i - \frac{E_1(t, \boldsymbol{\theta})}{E_0(t, \boldsymbol{\theta})} \right) dN_i(t). \tag{1.10}$$

We can then estimate $\boldsymbol{\theta}$ with $\hat{\boldsymbol{\theta}}$, which is the solution to $U(\boldsymbol{\theta}) = \mathbf{0}$, and estimate $d\Lambda_0(t)$ with $d\hat{\Lambda}_0(t; \hat{\boldsymbol{\theta}})$. Note that we could have obtained $U(\boldsymbol{\theta})$ as the derivative of the so-called partial likelihood function of $\boldsymbol{\theta}$, in which large sample properties of the estimators $\hat{\boldsymbol{\theta}}$ and $d\hat{\Lambda}_0(t; \hat{\boldsymbol{\theta}})$ have been established from this aspect (Tsiatis 1981, Andersen and Gill 1982).

1.1.3 Survival Analysis: Cox Regression with Time-Varying Covariates

In addition to $\boldsymbol{X}$, suppose we observe a risk factor changing over time, denoted by $Z(\cdot)$, and the history of $Z(\cdot)$ up to time t is denoted by $\mathcal{Z}(t) = \{Z(u) : 0 \leq u \leq t\}$. The available data can be represented as $\{(T_i^*, \delta_i, \boldsymbol{X}_i, \mathcal{Z}_i(T_i^*)) : i = 1, \cdots, n\}$. Consider the following conditional hazard function of T:

$$\lambda(t; \mathcal{Z}(t), \boldsymbol{X}) = \lim_{\Delta t \to 0^+} \frac{1}{\Delta t} P(t \leq T < t + \Delta t | T \geq t, \mathcal{Z}(t), \boldsymbol{X}). \tag{1.11}$$

Note that (1.11) conditions on the event not occurring before time t, the covariates $\boldsymbol{X}$, and the history of the time-varying covariate up to t. Before proceeding as in Section 1.1.2 and specifying a model for (1.11), it is important to distinguish two types of time-varying covariates: *external* and *internal*. A time-varying covariate $Z(\cdot)$ with history $\mathcal{Z}(\cdot)$ that satisfies the following equation (Kalbfleisch and Prentice 2002) is an *external* (time-varying) covariate:

$$P(T \in [u, u + \Delta u) | T \geq u, \mathcal{Z}(t)) = P(T \in [u, u + \Delta u) | T \geq u, \mathcal{Z}(u)) \tag{1.12}$$

for all u and t with $0 < u < t$. Since the left- and right-hand side of (1.12) conditions on the covariate history up to time t and u, respectively, (1.12) allows us to deduce $Z(\cdot)$ to be external if the future path of $Z(\cdot)$ is not affected by the occurrence of the event at time u. Generally, external covariates belong to one of the following three categories.

Time-Independent: Although redundant to introduce an index over time, a time-independent covariate can be viewed as a time-dependent covariate, in which $Z(t) \equiv Z(0) \equiv Z$ for all $t > 0$. Note that $\mathcal{Z}(t) = Z$, (1.11) is then $\lambda(t; Z, \boldsymbol{X})$, and it is trivial that (1.12) holds.

Defined: For any time $t > 0$, suppose we can construct $\mathcal{Z}(t)$ solely based on time t, and information at time $t = 0$. For example, let $Z(t)$ denote an individual's age at time t. We can write $Z(t) = a_0 + t$, where a_0 is the age of the individual at time $t = 0$, and in principle, $Z(t)$ can be computed for any $t > 0$, provided that a_0 is known.

Since an individual's future age at time t would not affect their current mortality risk at time u, $Z(\cdot)$ satisfies (1.12).

Ancillary: Suppose the time-varying covariate is a realization of a stochastic process that is external to individuals. For example, suppose $Z(t)$ is the temperature at time t. Suppose $\mathcal{Z}(t)$ is available, and we are interested in assessing an individual's mortality risk at time $u < t$. We would conclude that (1.12) holds since only the temperature by time u could have an effect on the event occurring at time u.

A time-varying covariate $Z(\cdot)$ with history $\mathcal{Z}(\cdot)$ is *internal* if (1.12) does not hold. The interpretation of an internal covariate of an event is that the future path of $Z(\cdot)$ is influenced by the occurrence of the event of interest at time u, say if $Z(\cdot)$ can only be measured whenever the event has not yet occurred. In our application, this becomes pertinent when considering T as an individual's time to death since their initial OAT dispensation record in the administrative database, and $Z(t)$ as the individual's OAT dispensation indicator at time t. We can deduce that $Z(\cdot)$ is an internal covariate since having knowledge of an individual's OAT dispensation indicator at time t implies their survivorship up to that point. Consequently, the left-hand side of (1.12) must be zero, whereas the right-hand side of (1.12) lies in the interval $(0, 1)$. As we will discuss further, this fact presents challenges when applying conventional survival analysis approaches in the application.

To elaborate the distinction between external and internal covariates, suppose we are interested in measuring the association between $\mathcal{Z}(\cdot)$ and T. The natural strategy following Section 1.1.2 would be to specify an extended Cox regression model, such as

$$\lambda(t; \mathcal{Z}(t), \boldsymbol{X}) = \lambda_0(t) \exp\{\gamma Z(t) + \boldsymbol{\theta}' \boldsymbol{X}\}, \tag{1.13}$$

where $\lambda_0(\cdot)$, γ, and $\boldsymbol{\theta}$ are unknown. We note that (1.13) explicitly assumes that $Z(t)$ is the only quantity from $\mathcal{Z}(t)$ that effects the conditional hazard. Given the observed data $\{(T_i^*, \delta_i, \boldsymbol{X}_i, \mathcal{Z}_i(T_i^*)) : i = 1, \cdots, n\}$, one might naively follow our reasoning from Section 1.1.2, and deduce that the likelihood function of $\lambda_0(\cdot)$, γ, $\boldsymbol{\theta}$ is

$$L(\lambda_0(\cdot), \gamma, \boldsymbol{\theta}) = \prod_{i=1}^{n} \lambda(T_i^*; \mathcal{Z}(T_i^*), \boldsymbol{X}_i)^{\delta_i} \exp\{-\int_0^{T_i^*} \lambda(u; \mathcal{Z}(u), \boldsymbol{X}_i) du\}. \tag{1.14}$$

For some fixed $t > 0$ however, consider the following partition of $[0, t]$: let $0 = u_0 < u_1 < \cdots < u_R = t$ with $\Delta u_{r-1} = u_r - u_{r-1}$ such that $\max\{\Delta u_{r-1}\} \to 0$ as $R \to \infty$. From (1.11), we have

$$\begin{aligned} P(T \in [t, t + \Delta t) | T \geq t, \mathcal{Z}(t), \boldsymbol{X}) &= \lambda(t; \mathcal{Z}(t), \boldsymbol{X}) \Delta t + o(\Delta t), \text{ and} \\ P(T \notin [t, t + \Delta t) | T \geq t, \mathcal{Z}(t), \boldsymbol{X}) &= 1 - \lambda(t; \mathcal{Z}(t), \boldsymbol{X}) \Delta t + o(\Delta t). \end{aligned}$$

Letting $R \to \infty$, it follows that

$$P(T > t|\mathcal{Z}(t), \boldsymbol{X}) = \lim_{R\to\infty} \prod_{r=1}^{R} P(T \notin [u_{r-1}, u_r)|T \geq u_{r-1}, \mathcal{Z}(t), \boldsymbol{X})$$

$$= \begin{cases} \lim_{R\to\infty} \prod_{r=1}^{R} P(T \notin [u_{r-1}, u_r)|T \geq u_{r-1}, \mathcal{Z}(u_{r-1}), \boldsymbol{X}) & \text{if } Z(\cdot) \text{ is external} \\ \lim_{R\to\infty} \prod_{r=1}^{R} P(T \notin [u_{r-1}, u_r)|T \geq u_{r-1}, \mathcal{Z}(t), \boldsymbol{X}) & \text{if } Z(\cdot) \text{ is internal} \end{cases} \tag{1.15}$$

$$= \begin{cases} \lim_{R\to\infty} \prod_{r=1}^{R} 1 - \lambda(u_{r-1}; \mathcal{Z}(u_{r-1}), \boldsymbol{X})\Delta u_{r-1} + o(\Delta u_{r-1}) & \text{if } Z(\cdot) \text{ is external} \\ 1 & \text{if } Z(\cdot) \text{ is internal} \end{cases}$$

$$= \begin{cases} \exp\left\{-\int_0^t \lambda(u; \mathcal{Z}(u), \boldsymbol{X})du\right\} & \text{if } Z(\cdot) \text{ is external} \\ 1 & \text{if } Z(\cdot) \text{ is internal} \end{cases}. \tag{1.16}$$

In the case where $Z(\cdot)$ is external, the result in (1.16) essentially follows the same derivation to obtain (1.5). In fact, (1.16) shows that $L(\lambda_0(\cdot), \gamma, \boldsymbol{\theta})$ takes the form presented in (1.14) if $Z(\cdot)$ is an external covariate; if $Z(\cdot)$ is internal however, the conventional relationship between the hazard and survivor functions no longer holds. In particular, (1.15) conditions on both $\mathcal{Z}(t)$ and the event $\{T \geq u_{r-1}\}$, but since observing $\mathcal{Z}(t)$ implies $\{T \geq u_R\}$, the probability shown in (1.15) must be one. Consequently, the likelihood function is not as presented in (1.14), but rather

$$L(\lambda_0(\cdot), \gamma, \boldsymbol{\theta}) = \prod_{i=1}^{n} \lambda(T_i^*; \mathcal{Z}(T_i^*), \boldsymbol{X}_i)^{\delta_i}, \tag{1.17}$$

where only non-surviving individuals contribute to the likelihood function. As it is well known that non-survivors form a biased representation of a population, directly maximizing (1.17) to estimate the model parameters will result in biased estimates for $\lambda_0(\cdot)$, γ, and $\boldsymbol{\theta}$. Therefore, alternative strategies and inference procedures are required whenever $Z(\cdot)$ is internal.

1.2 Literature Review

The discussion from Section 1.1.3 reveals that we cannot adopt the standard likelihood based procedure to estimate parameters in a regression model for the hazard when internal time-varying covariates are included in the model. The core issue is that internal covariates obstruct the conventional relationship between the hazard and survivor functions, and this thesis is to develop an inference procedure that overcomes this issue. There are two general approaches we can adopt to infer parameters in (1.11): (i) adopt a procedure that does not

rely the survivor function; or (ii) "modify" either the response process or $\mathcal{Z}(t)$ in some way to ensure the conventional relationship between the hazard and survivor function holds, in order to apply a likelihood inference procedure. We review relevant literature that adopts these two general approaches.

1.2.1 Non-Likelihood Based Estimation Procedures in Survival Analysis

Following Godambe (1991) suppose W has a probability model $f(w;\boldsymbol{\theta})$, where $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_q)'$ is a parameter vector, and $U(\boldsymbol{\theta}; w)$ is a $q \times 1$ estimating function of $\boldsymbol{\theta}$ for a given w. Define

$$\boldsymbol{\Psi}(\boldsymbol{\theta}) = \mathbb{E}\left\{\frac{\partial U(\boldsymbol{\theta}; W)}{\partial \boldsymbol{\theta}}\right\}, \quad \boldsymbol{\Phi}(\boldsymbol{\theta}) = \mathbb{E}\{U(\boldsymbol{\theta}; W)U(\boldsymbol{\theta}; W)'\},$$

where $\mathbb{E}\{A\}$ is the expected value of A, taken with respect to $f(w;\boldsymbol{\theta})$. Let $\boldsymbol{\theta}_0$ denote the true value of $\boldsymbol{\theta}$, and assume that $\mathbb{E}\{U(\boldsymbol{\theta}; W)\} = \mathbf{0}$. Given a random sample $\{W_i : i = 1, \cdots, n\}$ from the population, suppose the following equation has a unique solution, denoted by $\hat{\boldsymbol{\theta}}$:

$$U(\boldsymbol{\theta}) = \sum_{i=1}^{n} U(\boldsymbol{\theta}; W_i) = \mathbf{0}. \tag{1.18}$$

Then under regularity conditions, $\hat{\boldsymbol{\theta}} \xrightarrow[n\to\infty]{p} \boldsymbol{\theta}_0$, and $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow[n\to\infty]{d} \mathcal{N}(\mathbf{0}, \boldsymbol{\Psi}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{\Phi}(\boldsymbol{\theta}_0)\boldsymbol{\Psi}^{-1}(\boldsymbol{\theta}_0))$. The function $U(\boldsymbol{\theta})$ is referred to as an *estimating function* for $\boldsymbol{\theta}$, and (1.18) is an *estimating equation*. The estimator $\hat{\boldsymbol{\theta}}$ is referred to as a *Z-estimator* for $\boldsymbol{\theta}$, as $U(\hat{\boldsymbol{\theta}}) = 0$.

When specifying a Cox regression model with time-independent covariates, as in (1.6), we again solve the equation $U(\boldsymbol{\theta}) = \mathbf{0}$ to obtain the Z-estimator $\hat{\boldsymbol{\theta}}$, where $U(\boldsymbol{\theta})$ is shown in (1.10), the partial score function of $\boldsymbol{\theta}$ associated with the likelihood function. However, the conventional justification in using $U(\boldsymbol{\theta}) = \mathbf{0}$ is problematic when we consider internal time-varying covariates.

Robins and Tsiatis (1992) considered an accelerated failure time model with time-dependent covariates, in which the parameter $\boldsymbol{\theta}$ is estimated by solving $U(\boldsymbol{\theta}) = \mathbf{0}$, where $U(\boldsymbol{\theta})$ takes the form of a weighted log rank test that conducts the test $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. The large sample properties of the resulting estimators were then established by Lin and Ying (1995). Since then, this approach seems to have fallen out of favour within the literature, but time-varying covariates in accelerated failure time models has remained relevant particularly within the causal inference literature. Hernan *et al.* (2005) consider a popular model termed a *structural nested accelerated failure time model*, and estimate model parameters with the solution to $U(\boldsymbol{\theta}) = \mathbf{0}$ where $U(\boldsymbol{\theta})$ is derived by semiparametric efficiency theory (Tsiatis 2006, Yang *et al.* 2020). Its theoretical complexity and general lack of implementation in available software programs has hindered its adoption among statisticians (Vansteelandt and Joffe 2014, Simoneau *et al.* 2020).

Instead, one may consider the model in (1.13), and consider the following estimating function

$$U(\gamma, \boldsymbol{\theta}) = \sum_{i=1}^{n} \int_0^{\infty} Y_i(t) \left((Z_i(t), \boldsymbol{X}_i')' - \frac{E_1(t, \gamma, \boldsymbol{\theta})}{E_0(t, \gamma, \boldsymbol{\theta})} \right) dN_i(t),$$

where

$$E_r(t, \gamma, \boldsymbol{\theta}) = \sum_{j=1}^{n} E_{rj}(t, \gamma, \boldsymbol{\theta}) = \sum_{j=1}^{n} Y_j(t) \exp\{\gamma Z_j(t) + \boldsymbol{\theta}' \boldsymbol{X}_j\} (Z_j(t), \boldsymbol{X}_j')'^{\otimes r}.$$

Under (1.13), $U(\gamma, \boldsymbol{\theta}) = \mathbf{0}$ is an unbiased estimating equation, regardless if $Z(\cdot)$ is external or internal, and it is reasonable to estimate $(\gamma, \boldsymbol{\theta}')'$ with the solution to $U(\gamma, \boldsymbol{\theta}) = \mathbf{0}$. In fact, we could show that $U(\gamma, \boldsymbol{\theta})$ corresponds to the partial score function of $(\gamma, \boldsymbol{\theta}')'$ if $Z(\cdot)$ is external. When $Z(\cdot)$ is internal, we view $U(\gamma, \boldsymbol{\theta})$ simply as an estimating function of $(\gamma, \boldsymbol{\theta}')'$. Unfortunately, $U(\gamma, \boldsymbol{\theta})$ is commonly (naively) regarded as the partial score function of $(\gamma, \boldsymbol{\theta}')'$, as the asymptotic results of Tsiatis (1981) and Andersen and Gill (1982) are based on the partial likelihood function. By viewing $U(\gamma, \boldsymbol{\theta})$ as an estimating function, we can leverage modern empirical process theory to establish its large sample properties (Lin *et al.* 2000); this concept will be further explored in Chapter 3.

1.2.2 Rectifying the Relationship Between the Hazard and Survivor Functions

We review existing methods on how researchers summarize $\mathcal{Z}(\cdot)$ or the event of interest in some fashion in order to make likelihood based approaches applicable. Compared to adopting a non-likelihood inference procedure, this has been the more popular strategy. We review the relevant strategies within the literature.

1.2.2.1 Modelling the Joint Distribution of the Event Time and Covariate Process

In principle, if we can avoid specifying how the event time is directly associated with the entire internal covariate process up to the current time as in (1.11), we are able to bypass the associated challenges. The most natural way is to model the joint distribution of T and $\mathcal{Z}(\cdot)$ given $\boldsymbol{X}$, which we denote as $[T, \mathcal{Z}(\cdot)|\boldsymbol{X}]$. If $Z(\cdot)$ is a categorical variable, researchers have proposed to model $[T, \mathcal{Z}(\cdot)|\boldsymbol{X}]$ within a multistate framework, where transient states of the process jointly correspond to different levels of $Z(\cdot)$ and the event not occurring, and an absorbing state corresponds to the event occurring (Beyersmann and Schumacher 2008, Cortese and Andersen 2009, Brookmeyer and Abdalla 2019, Cook *et al.* 2022). Models for the intensity functions between states are then specified, and a likelihood based inference can be implemented due to $\boldsymbol{X}$ being an external covariate.

In the situation where $Z(\cdot)$ is not a categorical variable, note that the joint distribution of T and $\mathcal{Z}(\cdot)$ can be expressed as $[T, \mathcal{Z}(\cdot)|\boldsymbol{X}] = [T|\mathcal{Z}(\cdot), \boldsymbol{X}] \times [\mathcal{Z}(\cdot)|\boldsymbol{X}]$. Here, a sub-model for $[T|\mathcal{Z}(\cdot), \boldsymbol{X}]$ could then be specified, and another sub-model is specified for $[\mathcal{Z}(\cdot)|\boldsymbol{X}]$. By linking the two sub-models together, this approach has been referred to as *joint-modelling*, and was first used as a tool to infer parameters in (1.13) whenever $Z(\cdot)$ is measured with error. That is, rather than observing the process $\mathcal{Z}(\cdot)$, the process $\mathcal{Z}^*(t) = \{Z^*(u) : 0 \leq u \leq t\}$ is observed instead, where

$$Z^*(t) = Z(t) + e(t), \tag{1.19}$$

and $e(t)$ is the measurement error at time t with mean zero. By specifying the hazard model (1.13), Tsiatis *et al.* (1995) showed that

$$\lambda(t; \mathcal{Z}^*(t), \boldsymbol{X}) = \lambda_0(t) \exp\{\boldsymbol{\theta}'\boldsymbol{X}\}\mathbb{E}\{\exp\{\gamma Z(t)\}|\mathcal{Z}^*(t)\}. \tag{1.20}$$

They proposed the following two-stage inference approach to conduct their statistical inference:

Step 1: Specify the evolution of $Z(t)$'s realization associated with subject i over time, for example, as

$$Z_i(t) = \nu_{i0} + \nu_{i1}t. \tag{1.21}$$

Suppose the available observations on $Z_i^*(\cdot)$ are $\{Z_i^*(t_{ij}) : j = 1, \cdots, m_i\}$, in which (1.19) is a linear mixed effect model (Laird and Ware 1982). Assume that for any given i

$$(\nu_{i0}, \nu_{i1})' \overset{iid}{\sim} \mathcal{N}\left((\nu_0, \nu_1)', \begin{bmatrix} \sigma_{\nu_0\nu_0} & \sigma_{\nu_0\nu_1} \\ \sigma_{\nu_1\nu_0} & \sigma_{\nu_1\nu_1} \end{bmatrix}\right),$$

$e_i(t) \sim N(0, \sigma^2)$, $e_i(s) \perp\!\!\!\perp e_i(t)$ for $t \neq s$, and $\{\nu_{i0}, \nu_{i1}\} \perp\!\!\!\perp e_i(t)$.

Step 2: Calculate $\mathbb{E}\{\cdot\}$ in (1.20), estimate γ and $\boldsymbol{\theta}$ by constructing a partial likelihood function, and estimate $\Lambda_0(t) = \int_0^t \lambda_0(u)du$ with a Breslow-like estimator.

Further developments in joint modelling have been made to improve either the modelling of $Z_i(t)$ in **Step 1**, or the inference procedure in **Step 2**. Wulfsohn and Tsiatis (1997) modified **Step 2** by viewing $(\nu_{i0}, \nu_{i1})'$ as "missing" and estimated parameters in (1.13) with the expectation-maximization (EM) algorithm. Wang and Taylor (1997) specified $\{e_i(t_{ij}) : j = 1, \cdots, m_i\}$ to follow an integrated Ornstein-Uhlenbeck process in **Step 1**, and took a Bayesian approach and developed a Markov chain Monte Carlo (MCMC) algorithm for **Step 2**. Henderson *et al.* (2000) and Xu and Zeger (2001) both specified $\{e_i(t_{ij}) : j = 1, \cdots, m_i\}$ to be a stationary Gaussian process in **Step 1**, and in **Step 2**, Henderson *et al.* (2000) used

the EM algorithm, whereas Xu and Zeger (2001) proposed an MCMC algorithm. Tsiatis and Davidian (2001) avoided to specify the distribution of $(\nu_{i0}, \nu_{i1})'$ in **Step 1** by estimating them with $(\hat{\nu}_{i0}, \hat{\nu}_{i1})'$ under the model (1.21) based on subject i's available information. By viewing the estimates as "noisy measurements" of $(\nu_{i0}, \nu_{i1})'$, they extended the conditional score approach for regression under generalized linear models of Stefanski and Carroll (1987) to the Cox regression model (1.13).

Perhaps unknowingly at the time, considering a certain model for $Z(\cdot)$ may address the problem brought on by internal covariates. This is because a model such as in (1.21) essentially defines $Z_i(t)$ over time (see Section 1.1.3). This has motivated researchers to utilize joint modelling as a tool to conduct so-called "dynamic prediction", which is to construct $\mathcal{Z}_i(t)$ for all $t > 0$ with (1.21), and estimate

$$P(T_i \geq t + s | T_i > t, \mathcal{Z}_i(t + s), \boldsymbol{X}_i) = \exp\left\{ - \int_t^{t+s} \lambda(u; \mathcal{Z}_i(u), \boldsymbol{X}_i) du \right\}. \tag{1.22}$$

Rizopoulos (2011) considered the procedure developed by Wulfsohn and Tsiatis (1997) to estimate parameters under the joint model, and adopted a Bayesian framework to predict (1.22). More of the recent literature has gone towards improving the accuracy of prediction for (1.22). Rizopoulos *et al.* (2014) specified multiple longitudinal sub-models in the form (1.21) with different covariance structures for $(\nu_{i0}, \nu_{i1})'$, and then conduct Bayesian model averaging. Other improvements have been made by specifying a more flexible model than (1.21), such as using spline functions (Barrett and Su 2017, Andrinopoulou *et al.* 2018).

1.2.2.2 Partly Conditional Modelling

Rather than conditioning on $\mathcal{Z}(t)$ in (1.11), consider "freezing" the process at some time point $s < t$, so that the mortality hazard function becomes

$$\lambda(t; \mathcal{Z}(s), \boldsymbol{X}) = \lim_{\Delta t \to 0^+} \frac{1}{\Delta t} P(t \leq T < t + \Delta t | T \geq t, \mathcal{Z}(s), \boldsymbol{X}). \tag{1.23}$$

Since the covariate process no longer varies over time (after time s), then clearly (1.12) must hold for any $t > s$. The idea is to specify a model for (1.23), select K *landmark* time points $\{s_1, \cdots, s_K\}$ to obtain K datasets $\{(T_i^* - s_k, \delta_i, \mathcal{Z}_i(s_k), \boldsymbol{X}) : i = 1, \cdots, n\}$, and conduct statistical inference by pooling the K datasets together. Note that because (1.12) holds, estimating functions can then be derived from the likelihood function. Specifying a model for (1.23) is referred to as a *partly conditional model*, for only part of the covariate history is being conditioned on. Zheng and Heagerty (2005) specified landmarks as observation times of the time-dependent covariate, whereas Schaubel *et al.* (2009) specified landmarks based on the observed time points a categorical time-dependent covariate changed its value. Gong and Schaubel (2013) specified landmarks to be equally spaced out in time, and proposed an inverse probability weighting method to address dependent censoring in T. Since the con-

ventional relationship between the hazard and survivor function holds, survival probabilities can be computed as

$$P(T > t|T > s, \mathcal{Z}(s), \boldsymbol{X}) = \frac{P(T > t|\mathcal{Z}(s), \boldsymbol{X})}{P(T > s|\mathcal{Z}(s), \boldsymbol{X})} = \exp\left\{-\int_s^t \lambda(u; \mathcal{Z}(s), \boldsymbol{X})du\right\} \quad (1.24)$$

(van Houwelingen 2007). Noting the similarities between (1.24) and (1.22) has led researchers to compare the predictive performance between partly conditional modelling and joint modelling. Generally, researchers have declared joint modelling as the preferred approach (Rizopoulos *et al.* 2017, Suresh *et al.* 2017). Nevertheless, there has been a recent attempt to improve partly conditional modelling (Putter and van Houwelingen 2022). It is intuitive to understand the inferiority of partly conditional modelling relative to joint modelling: (1.23) is inherently different from (1.11). The reason why partly conditional modelling gained popularity in the first place is mainly to its simplicity, compared to high computational costs that joint modelling suffers from. This issue is further discussed in Chapter 2. Our proposed methodology overcomes the challenge.

1.3 Notation and Framework

This section presents notation to be used throughout this dissertation. Additional notation will be introduced at the beginning of each chapter when needed.

Keeping consistent with Section 1.1, let T denote an individual's event time, where the time scale is generic. Specific time scales we consider will be discussed in later chapters. We do not necessarily observe T, but rather the pair (T^*, δ), where C is a right-censoring time, $T^* = T \wedge C$ is the minimum of T and C, and $\delta = I(T \leq C)$.

We let $\boldsymbol{Z}(t)$ denote internal covariates at time t, and $\boldsymbol{\mathcal{Z}}(t) = \{\boldsymbol{Z}(u) : 0 \leq u \leq t\}$ denote its history up to (and including) time t. With regards to the OAT dispensation records that motivated this research, $\boldsymbol{Z}(t)$ denotes all relevant information pertaining to the OAT dispensation at time t, including (i) dispensation indicator, (ii) OAT type, (iii) OAT dosage, (iv) pharmaceutical formulation of the dispensed OAT, (v) relevant prescriber factors, etc. When having one element pertaining to $\boldsymbol{Z}(\cdot)$, we instead replace $\boldsymbol{Z}(t)$ with $Z(t)$, and let $\mathcal{Z}(t) = \{Z(u) : 0 \leq u \leq t\}$ denote its history up to time t. This will be relevant in the forthcoming chapters when we focus on individuals' OAT dispensation indicators over time t. We let $\boldsymbol{X}(t) = (X_1(t), \cdots, X_q(t))'$ denote external time-varying covariates and time-independent characteristics of an individual, and $\boldsymbol{\mathcal{X}}(t) = \{\boldsymbol{X}(u) : 0 \leq u \leq t\}$. For fixed $k = 1, \cdots, q$, the covariate $X_k(t)$ is time-independent if $X_k(0) \equiv X_k(t)$, for all $t > 0$. Covariates included in our real data analyses will be specified in Chapter 2, when we further explore the relevant information captured from the administrative databases pertaining to individuals identified with an opioid use disorder.

The general objective of this thesis is to develop methods to estimate the effect of $\boldsymbol{\mathcal{Z}}(\cdot)$ in the presence of $\boldsymbol{\mathcal{X}}(\cdot)$ on T. This is typically done by modelling the hazard function

$$\lambda(t; \boldsymbol{\mathcal{Z}}(t), \boldsymbol{\mathcal{X}}(t)) = \lim_{\Delta t \to 0^+} \frac{1}{\Delta t} P(t \leq T < t + \Delta t | T \geq t, \boldsymbol{\mathcal{Z}}(t), \boldsymbol{\mathcal{X}}(t)).$$

The specific model we choose will depend on the specific objective we aim to achieve, and the inference procedure will be developed to overcome challenges brought on by internal covariates. We aim to develop approaches that are easy to implement in practice, and in the same time, adequately capturing complex relationships between $\boldsymbol{\mathcal{Z}}(\cdot)$ and T.

1.4 Thesis Outline

When the objective is to investigate the association of a response with a set of predictor variables, the traditional strategy is to specify a model for their association, and then obtain a random sample from the target population to estimate model parameters. The approaches can be applied when the response is a time-to-event, but challenges arise when predictors are internal time-varying covariates. Methods referenced in Section 1.2 have been developed to achieve this objective, but their implementation has been lacking due to restrictive assumptions or heavy computation costs. Motivated by addressing a real-world problem, this thesis aims to overcome challenges in estimating effects of internal covariate(s) on time-to-death.

We specify our objective into the following three aims:

Aim 1: To estimate the effect of an internal covariate process on the mortality hazard.

Aim 2: To develop a predictive survival model based on information of an internal covariate.

Aim 3: To extend **Aim 1** and **Aim 2** to accommodate situations with multivariate internal covariate(s).

Our methodology is motivated and illustrated with administrative service utilization records pertaining to opioid use disorder management. The proposed methods can in principle be implemented to any setting that shares the same aims specified above.

The remainder of this thesis is organized as follows. In Chapter 2, we provide a further description of the administrative service utilization records, and a preliminary analysis. The analysis jointly models time-to-death and the OAT dispensation indicator process. Chapter 3 proposes an estimating equation based procedure for estimating the OAT dispensation indicator process effect on the mortality hazard. In order to utilize the entire observed covariate history, we propose a generalized Cox regression model and conduct time-dependent stratification, where the strata are defined by the overall treatment history. Chapter 4 revisits the joint model from Chapter 2, and adapts the conditional score approach of Tsiatis and

Davidian (2001) to allow for correlated successive observations in the OAT dispensation indicator process. Chapter 5 extends the methods developed in Chapters 3 and 4 to situations with multivariate internal covariate(s), where we include OAT type and dose as additional features of the OAT dispensation process. A summary of this dissertation's contribution and discussions about future investigation are provided in Chapter 6. All relevant figures and tables are provided at the end of each respective chapter. All datasets were constructed by SAS (SAS Institute Inc. 2013). Analyses were conducted by R (R Core Team 2022), where estimating functions were coded in C++ and exported into R with the `Rcpp` R package to improve the computation speed (Eddelbuettel and François 2011, Eddelbuettel 2013, Eddelbuettel and Balamuta 2018).

Chapter 2

Preliminary Analysis of Administrative Service Utilization Records Pertaining to the Management of Opioid Use Disorder

2.1 Introduction

This chapter provides further details of the administrative service utilization records that motivated this research, and describes how the available information will be used in the forthcoming analyses. Descriptive statistics of follow-up times and potential risk factors are presented, as well as a preliminary data analysis based on a joint model of mortality risk and the OAT dispensation process. The required assumptions and analysis results will serve as a motivation for the methodological development in the forthcoming chapters.

2.2 Description of Administrative Service Utilization Records for Individuals Identified with Opioid Use Disorders

To address the 2016 public health emergency declared in BC, the Ministry of Mental Health and Addictions has requested the Health Economic Research Unit (PI: B. Nosyk) at St. Paul's Hospital to assess the quality of care for people identified with opioid use disorders. Administrative service utilization records, as discussed in Section 1.1.1, were provided to the research group to assess the service utilization for this cohort. Specifically, during a data extraction window, health service records for any individual with an OAT dispensation in BC-PharmaNet (see Table 2.1 for drug identification numbers) or a health record attributed to opioid use (see Table 2.2 for ICD-CA-9/10 codes) were provided.

The beginning of the data extraction window pertains to the start date of data collection (01/01/1996) but the end date varies, as the linked administrative records are updated biannually; the closing date of the data extraction window used in the forthcoming analyses is 10/01/2018. In total, 91,876 individuals were identified to have an opioid use disorder if they had at least (i) one OAT dispensation, (ii) one OAT-related hospitalization, or (iii) three opioid use disorder related physician claims. By removing the individuals without a recorded OAT dispensation, this reduces the cohort to 55,347 individuals. Furthermore, the adult ($\geq$ 18 years old) cohort with at least one OAT dispensation in BC from 01/01/1996 to 10/01/2018 includes $n = 54{,}739$ individuals.

To illustrate the information present in the database, Figure 2.1 illustrates the date of a hypothetical adult's ($\geq$ 18 years old) first recorded OAT dispensation in BC-PharmaNet, between the data extraction window. In terms of an individual's death date, there are two scenarios that can occur: **(a)** it is recorded in BC Vital Statistics and taken as evidence that they died on a particular date, or **(b)** it is not recorded and the death time is treated as right-censored, where the individual's drop-out date is taken as the date following their final record in any of the administrative databases. The OAT type and dosage dispensed on a particular date an individual received is derived from the drug identification numbers pertaining to OATs are presented in Table 2.1. We emphasize that the dispensation records have limited information on individuals adhering to their dispensed treatment. In the situation where the drug identification number in Table 2.1 implies an "interaction" with a medical personnel or pharmacist, or if the number of days the recorded dispensation is low (e.g. 1 or 2 days), we can confidently assume that the individual adhered to the dispensation record. This confidence diminishes as the number of days a recorded dispensation covers increases, especially if the number of days is large (e.g. $\geq$ 7 days). Although the majority of dispensation records are daily, interpreting the *OAT dispensation indicator* as the *on OAT indicator* explicitly assumes individuals perfectly adhere to their prescription. Wherever possible, we refer the *OAT usage process* as the *OAT dispensation process*.

Inpatient hospitalization is a setting where an individual likely received OAT, especially if their past or future dispensation records indicate potential OAT use, but the dispensation record is unavailable. We follow Pearce *et al.* (2020), and assume that an individual was dispensed OAT during their (entire) hospitalization if they either have an OAT dispensation record within five days (i) before their admission date, or (ii) after their discharge date. This issue also brings up another issue: the initial recorded OAT dispensation date may not necessarily be the individual's actual first OAT dispensation date. For example, the available dispensation records (i) do not include records outside the data extraction window, (ii) may not record dispensations outside of British Columbia, and as referenced before, (iii) do not record dispensations within acute care settings. This highlights the difference between two time scales regarding the definition of time zero. The first recorded OAT dispensation date is dependent on the collected data, whereas the first *true* OAT dispensation date is

a clinically relevant event. To circumvent issues related to the definition of time zero, we consider *age* as the time scale, in which time zero is defined as the time of birth. In the forthcoming analyses, we present our results under both *age* and the *time since first recorded OAT dispensation* time scales. Unless otherwise stated, the default time scale within our methodological development is the *time since first recorded OAT dispensation*, and the default time unit is in *days*.

As mentioned in Chapter 1, administrative databases have the advantage of recording a broad range of clinical characteristics. This allows us to include a variety of risk factors in our analysis that might otherwise be difficult or expensive to obtain, such as an individual's characteristics or factors related to their interaction with the BC healthcare system. Measured characteristics on an individual's first recorded OAT dispensation date can be considered as baseline covariates in our regression models. These characteristics include *sex* (male vs. female), *birth generation* (indicators of birth year intervals: 1901-1945 vs. 1946-1964 vs. 1965-1980 vs. 1981+), *health authority* (indicators of residence region: Fraser Health vs. Interior vs. Vancouver Coastal vs. Vancouver Island vs. Northern), and *year category* (category corresponding to first recorded OAT dispensation date: 1996-2000 vs. 2001-2006 vs. 2007-2012 vs. 2013-2018). Although the variable *health authority* can in principle change over time, this variable is rather stable, and therefore treated as a time-independent categorical variable. Some individuals had changes in their reported *sex* over time, but we found that most of the observed changes were due to data entry errors. Therefore, we kept *sex* as a time-independent variable in our analyses. In terms of other risk factors, Table 2.3 displays codes used to identify if certain conditions are satisfied. We identify an individual satisfying a particular condition if (i) their purpose of hospitalization matches the condition, (ii) they have at least three physician claims that matches the condition, or (iii) they have a pharmaceutical dispensation pertaining to the condition. Since it is difficult to gauge the severity of conditions strictly from these health records, we restrict our attention to the presence of these conditions. That is, each condition listed in Table 2.3 is specified as time-varying one-jump binary processes, where the change of such a process corresponds to the first observed date when the condition is satisfied. Individuals that do not have a health record pertaining to a condition are classified to not satisfying the condition in question. We incorporate information pertaining to an individual's socioeconomic status by assessing if they ever received financial assistance, through BC PharmaCare Plans C or G to cover medication costs, as well as incarceration records through their incarceration status.

In terms of the notation provided in Section 1.3, the available information we have can be expressed as $\{(T_i^*, \delta_i, \boldsymbol{\mathcal{Z}}_i(T_i^*), \boldsymbol{\mathcal{X}}_i(T_i^*)) : i = 1, \cdots, n\}$. Our aim is to estimate the parameters in a regression model using the available data to assess the association of T with $\boldsymbol{\mathcal{Z}}(\cdot)$ while controlling for additional factors in $\boldsymbol{\mathcal{X}}(\cdot)$.

2.3 Preliminary Analysis with Joint Modelling the OAT Dispensation and Mortality Risk Processes

To simplify our task, we restrict the OAT dispensation process to be the OAT dispensation indicator. Our objective is to quantify the association of T with $\mathcal{Z}(\cdot)$ while controlling for additional factors in $\boldsymbol{\mathcal{X}}(\cdot)$. The conventional approach to achieve this is to model the conditional hazard function of T (at time t), given the processes $\mathcal{Z}(t)$ and $\boldsymbol{\mathcal{X}}(t)$:

$$\lambda(t; \mathcal{Z}(t), \mathcal{X}(t)) = \lim_{\Delta t \to 0^+} \frac{1}{\Delta t} P(t \leq T < t + \Delta t | T \geq t, \mathcal{Z}(t), \boldsymbol{\mathcal{X}}(t)). \tag{2.1}$$

Before specifying an appropriate model, we need to address two questions: (a) "what information from $\mathcal{Z}(t)$ do we want to include in the hazard model at time t?" and (b) "how do we infer model parameters when conditioning on an internal covariate?" Regarding (a), the two extreme quantities are (i) only $Z(t)$ from $\mathcal{Z}(t)$, the current OAT dispensation indicator for an individual at time t is used, and (ii) $R(t) = \int_0^t Z(u)du/t$, the average proportion of time an individual is dispensed OAT from their first recorded OAT dispensation up to time t, which summarizes the information from $\mathcal{Z}(t)$. Given the conviction that a more comprehensive summary of the data is preferable, we proceed to work with $R(t)$. Following the discussion in Chapter 1, our options to address (b) are to (i) adopt a non-likelihood inference procedure, or to (ii) summarize $\mathcal{Z}(t)$ in some way so that the conventional relationship between the hazard and survivor functions is preserved. We proceed with the latter option and summarize the internal covariate with some quantity that does not imply an individual's survival status. That is the popular approach in the literature.

For example, consider the following model:

$$R_{ij} = \nu_i + \varepsilon_{ij} \quad \text{for } i = 1, \cdots, n \text{ and } j = 1, \cdots, m_i, \tag{2.2}$$

where $\{t_{ij} : j = 1, \cdots, m_i\}$ are observation times of the OAT dispensation process with $T_i^* = t_{i,m_i}$, $R_{ij} \equiv R_i(t_{ij})$, ν_i is a subject-specific (unknown) quantity, and ε_{ij} is a mean-zero error term. If we are willing to make the assumption that $T_i \perp\!\!\!\perp \mathcal{Z}_i(\infty) \mid \nu_i$, we can then replace $\mathcal{Z}_i(\cdot)$ with ν_i in the conditional hazard function (2.1). The following specification of (2.1) could then be used to achieve our objective:

$$\lambda(t; \mathcal{Z}_i(t), \boldsymbol{\mathcal{X}}_i(t)) = \lambda_0(t) \exp\{\gamma \nu_i + \boldsymbol{\theta}' \boldsymbol{X}_i(t)\} \quad \text{for } i = 1, \cdots, n, \tag{2.3}$$

where $\lambda_0(t)$ is an unspecified baseline hazard function, and $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_q)'$ and γ are unknown regression parameters with γ being the parameter of primary interest. Since ν_i is time-independent, the conventional relationship between the hazard and survivor functions is thus preserved. We can therefore estimate $\lambda_0(\cdot)$, $\boldsymbol{\theta}$, and γ with a likelihood based inference procedure. The main challenge we face is that ν_i is unobservable. Since ν_i is a summary

of $\mathcal{Z}(\cdot)$, it is natural to estimate ν_i under (2.2) with $\hat{\nu}_i = \sum_{j=1}^{m_i} R_{ij}/m_i$, the sample mean of dispensation proportions. We refer ν_i and $\hat{\nu}_i$ as the *OAT dispensation rate* and the *observed OAT dispensation rate* for individual i, respectively.

Table 2.4 provides descriptive statistics of T^*, age on first recorded OAT dispensation date, and potential time-independent risk factors to include in (2.3). Of the n = 54,739 individuals from our study, there are 7,008 (12.8%) deaths. By stratifying individuals into survivors and non-survivors, Table 2.4 showcases some characteristics that differ between the groups. In particular, non-survivors were generally older than survivors on their first recorded OAT dispensation date and had a higher *observed OAT dispensation rate* compared to survivors. This latter observation will be further discussed shortly. As the *incarceration status* over time is also an alternating-binary process, we used a strategy similar to summarizing the OAT dispensation process with (2.2) to summarize an individual's incarceration process. However, since the majority of individuals were never incarcerated during the study period, specifying a model as in (2.2) to summarize someone's incarceration process is challenging. This motivated us to summarize this process in an alternative way by pooling all of the incarceration processes together, and then let the data identify important features by conducting a dimension reduction procedure. By viewing each individual's incarceration process as a realization from a functional binary process, we applied sparse logistic functional principal component (FPC) analysis for binary data (Zhong *et al.* 2021, Zhong and Zhang 2022), in which 94.1% of the overall variability is attributed to the first principal component.

In terms of potential time-varying covariates to include in (2.3), we first examined for potential multicollinearity within the one-jump processes. In Table 2.5, we present the proportion of individuals who meet one condition compared to another. This motivated us to merge processes that appear to be highly correlated, specifically, (i) *alcohol use disorder* and *other substance use disorders*, (ii) *poor mental health* and *chronic pain*, and (iii) *Hepatitis C Virus* and *HIV/AIDS*. Table 2.6 summarizes the one-jump processes and the number of observed incarcerations (to be used in Chapters 3 & 5) by the end of follow-up. We can see the proportion of individuals satisfying the listed conditions are higher in the *non-survivor* cohort, relative to *survivors*. We would hence anticipate the estimated effect for these variables in (2.3) to be positive.

With the available data $\{(T_i^*, \delta_i, \boldsymbol{\mathcal{X}}_i(T_i^*)) : i = 1, \cdots, n\}$, we estimate the parameters in (2.3) by maximizing the following likelihood function:

$$
\begin{aligned}
&L(\lambda_0(\cdot), \boldsymbol{\theta}, \gamma) \\
&= \prod_{i=1}^{n} \int (\lambda_0(T_i^*) \exp\{\gamma\nu_i + \boldsymbol{\theta}' \boldsymbol{X}_i(T_i^*)\})^{\delta_i} \exp\{-\int_0^{T_i^*} \lambda_0(t) \exp\{\gamma\nu_i + \boldsymbol{\theta}' \boldsymbol{X}_i(t)\} dt\} dF(\nu_i),
\end{aligned}
\tag{2.4}
$$

where we assume $\nu_i \sim F(\cdot)$. Often, ν_i is specified to follow a normal distribution, and a numerical integration technique, such as Monte Carlo or Gaussian quadrature, is implemented to evaluate (2.4) (Tsiatis and Davidian 2004, Rizopoulos 2011). We can see that the resulting estimator under (2.4) not only depends on the correct specification of $F(\cdot)$, but is also computationally expensive to obtain, especially if the dimension of ν_i is large. Prior research has shown the resulting estimator under (2.4) to be fairly robust against misspecifying ν_i to follow a normal distribution (Song *et al.* 2002, Tsiatis and Davidian 2004, Hsieh *et al.* 2006), while attempts to reduce the computational costs with joint modelling (e.g. Ding and Wang (2008)) cannot avoid the integration in (2.4). If we can avoid integrating out ν_i in (2.4) however, this would solve the computational challenge. If we knew ν_i for each $i = 1, \cdots, n$, the likelihood function is

$$\begin{aligned} &L(\lambda_0(\cdot), \boldsymbol{\theta}, \gamma) \\ &= \prod_{i=1}^{n} (\lambda_0(T_i^*) \exp\{\gamma\nu_i + \boldsymbol{\theta}'\boldsymbol{X}_i(T_i^*)\})^{\delta_i} \exp\{-\int_0^{T_i^*} \lambda_0(t) \exp\{\gamma\nu_i + \boldsymbol{\theta}'\boldsymbol{X}_i(t)\}dt\}. \end{aligned}$$

Evaluating its associated maximum likelihood estimate of the parameters clearly has a lower computational cost compared to (2.4). If we replace ν_i with its estimate $\hat{\nu}_i$, the likelihood function becomes

$$\begin{aligned} &L(\lambda_0(\cdot), \boldsymbol{\theta}, \gamma) \\ &= \prod_{i=1}^{n} (\lambda_0(T_i^*) \exp\{\gamma\hat{\nu}_i + \boldsymbol{\theta}'\boldsymbol{X}_i(T_i^*)\})^{\delta_i} \exp\{-\int_0^{T_i^*} \lambda_0(t) \exp\{\gamma\hat{\nu}_i + \boldsymbol{\theta}'\boldsymbol{X}_i(t)\}dt\}. \end{aligned} \tag{2.5}$$

Note that (2.5) is a special case of (2.4) with $F(\nu_i) = I(\nu_i \geq \hat{\nu}_i)$, which places all the mass on the estimate $\hat{\nu}_i$ for ν_i. The function in (2.5) is similar to that of (1.14), and with $\boldsymbol{X}(\cdot)$ being an external covariate vector, one may proceed to estimate the model parameters by maximizing (2.5).

Table 2.7 presents the resulting estimates of γ and $\boldsymbol{\theta}$ under the model (2.3). The estimates are presented under two time scales: (i) *time since first recorded OAT dispensation record*, and (ii) *age*. The purpose of including *age* as a time scale is to overcome the subjectivity of time zero under *time since first recorded OAT dispensation record*. Generally, the estimates under both time scales corroborate with the descriptive statistics presented in Tables 2.4 and 2.6. However, we anticipated the effects of *ill mental health or chronic pain* and *ever on PharmaCare plans C or G* to be positive in the hazard model. These discrepancies are likely due to additional multicollinearity within the one-jump processes, where we display two-way contingency tables for each combination of the processes in Table 2.5. Since the effect of primary interest is the *OAT dispensation rate*, we do not make any further attempts to address this issue. One approach to address the serial correlation within

the one-jump processes is to maximize a penalized likelihood function that would jointly estimate model parameters and conduct variable selection (Fan and Li 2002).

The estimated effect of *OAT dispensation rate* is positive (and significant). It indicates that individuals with a higher *observed OAT dispensation rate* have a higher risk of mortality relative to individuals with a lower dispensation rate. Although this agrees with the descriptive summary in Table 2.4, it surprised us nonetheless. As it is known that opioid tolerance varies with age (Zubieta *et al.* 1999, Zhao *et al.* 2012), Table 2.8 presents summary statistics for $\hat{\nu}$ across birth generations, in which we see the OAT dispensation pattern varies between survival groups. In particular, survivors have a lower *observed OAT dispensation rate* within the older generations, and the opposite is true for the younger generations. This finding discloses that the analysis results in Table 2.7 are likely confounded by age. We were then motivated to stratify individuals according to their respective birth generations, and obtain stratum-specific estimates under (2.2). The regression estimates are presented in Table 2.9 and the cumulative baseline hazard estimates are illustrated in Figure 2.2. We can see in Table 2.9 that the estimated effects of *OAT dispensation rate* decrease by birth generation, and align with the descriptive statistics in Table 2.8. We can also see in Figure 2.2 that the mortality risks, as expected, are higher for older birth generations.

We extended (2.3) to accommodate potential time-varying effect for the *OAT dispensation rate*:

$$\lambda(t; \mathcal{Z}_i(t), \boldsymbol{\mathcal{X}}_i(t)) = \lambda_0(t) \exp\{\gamma(t)\nu_i + \boldsymbol{\theta}' \boldsymbol{X}_i(t)\}, \tag{2.6}$$

where $\gamma(t)$ is an unknown function of time. One may approximate $\gamma(\cdot)$ with a linear combination of natural cubic spline basis functions with percentile-based knots (Green and Silverman 1994),

$$\gamma(t) \approx \sum_{k=0}^{s} \phi_k C_k(t).$$

Here, s determines the number of interior knots in the spline. Figure 2.3 shows the estimates of $\gamma(\cdot)$ in both *time since first recorded OAT dispensation* and *age* time scales, where the value of s was selected by the Bayesian information criterion (Rice and Wu 2001). Since the estimates of $\hat{\boldsymbol{\theta}}$ are similar to those presented in Table 2.7, we omit displaying the corresponding estimates under (2.6). In principle, the patterns in Figure 2.3 can provide important insights towards the *observed OAT dispensation rate* effect, we do not want to read too much into them due to not adjusting the estimates for the confounding variable, *age*. Figures 2.4 and 2.5 show the *observed OAT dispensation rate* effects and smoothed estimated baseline hazard functions, where we stratified individuals according to their birth generation. In the *age* time scale, the estimated effects are linear over time, where we note that only the effect corresponding to *Millennials & Generation Z* has a decreasing trend.

Using these estimates alongside the estimated baseline hazard functions, we can predict an individual's survival probability with $t > C_i$, analogous to (1.22) as

$$\hat{P}(T_i \geq C_i + t | T > C_i, \nu_i, \boldsymbol{\mathcal{X}}_i(C_i)) = \exp\left\{ - \int_{C_i}^{C_i+t} \exp\{\hat{\gamma}_b(u)\nu_i + \hat{\boldsymbol{\theta}}_b' \boldsymbol{X}_i(C_i)\} d\hat{\Lambda}_{0b}(u) \right\}$$
$$\text{if individual } i \text{ belongs to birth generation group } b \in \{1, 2, 3, 4\}. \tag{2.7}$$

The survival probabilities are under (2.6) for individuals stratified by birth generation first. We condition on the available information from $\mathcal{Z}_i(\cdot)$ and $\boldsymbol{\mathcal{X}}_i(\cdot)$. In particular, we need to summarize $\mathcal{Z}_i(\cdot)$ by the rate ν_i, and keep the one-jump binary processes static at their observed value at time C_i. We illustrate such predicted survival probabilities for eight randomly selected survivors (two from each of the four birth generation groups) in Figure 2.6. Overall, our analysis indicates that individuals belonging to the *Greatest & Silent Generations* are subject to the largest mortality risk among the four birth generation groups.

2.4 Outlook for Chapter 3

The approach taken within our preliminary data analysis jointly models the OAT dispensation process and mortality hazard risk. As discussed in Chapter 1, this has been the most popular approach within the literature towards handling challenges induced by internal time-varying covariates. This approach requires us to specify a longitudinal sub-model, the one specified in (2.2), or summarize the process with a dimension reduction procedure, such as obtaining functional principal component scores (Spreafico and Ieva 2021, Zhong *et al.* 2021). However, what if we want to directly estimate an internal covariate's effect on the mortality hazard? As shown in Section 1.1.3, the conventional relationship between the hazard and survivor functions is obstructed, and we hence need to consider an alternative procedure to conduct our statistical inference. This will be the focus and main contribution of Chapter 3.

Figure 2.1: A hypothetical study individual with their date of death is (**a**) recorded in the dataset, and (**b**) right-censored.

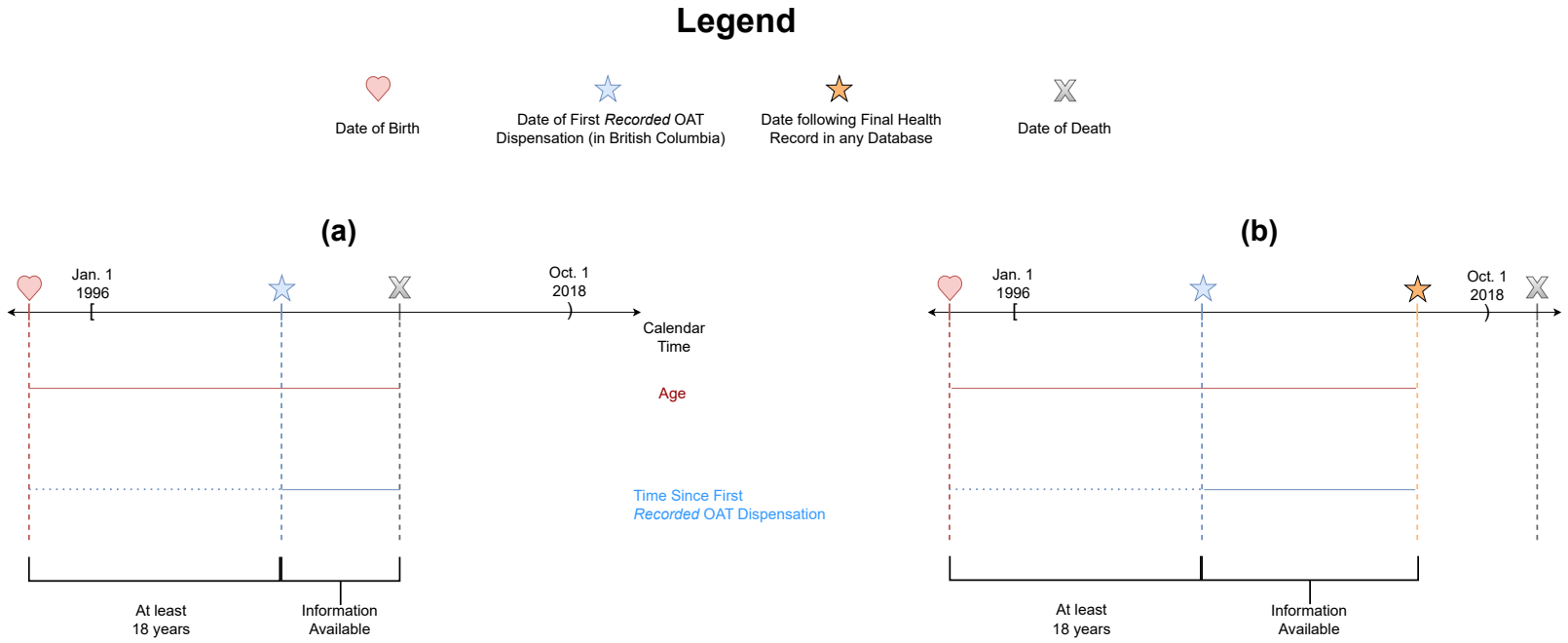
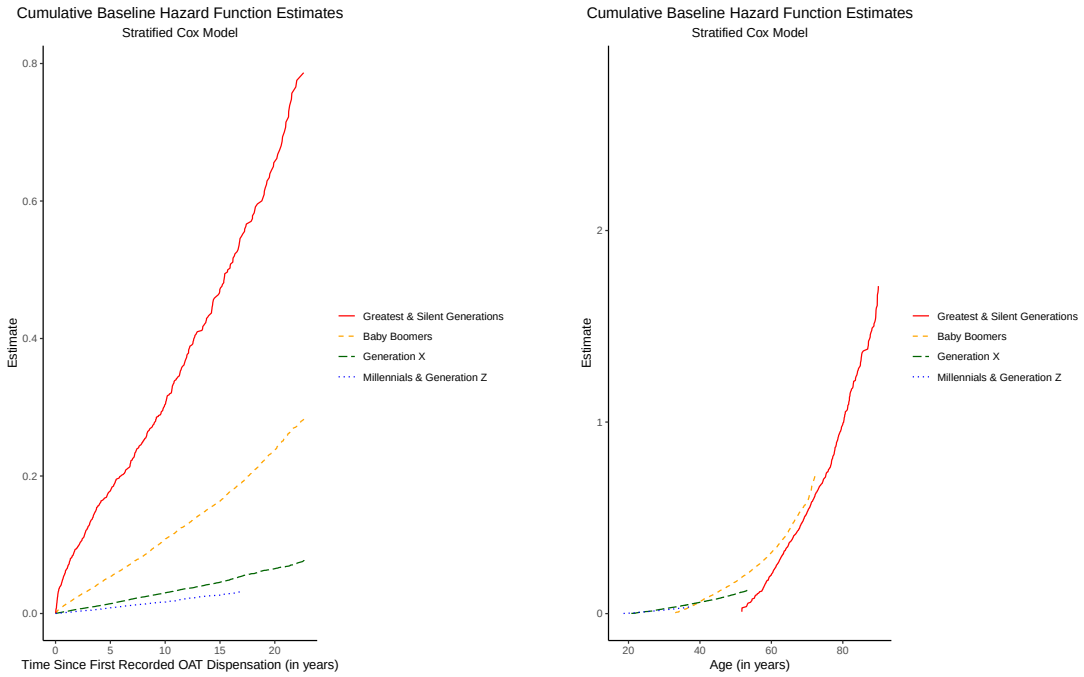
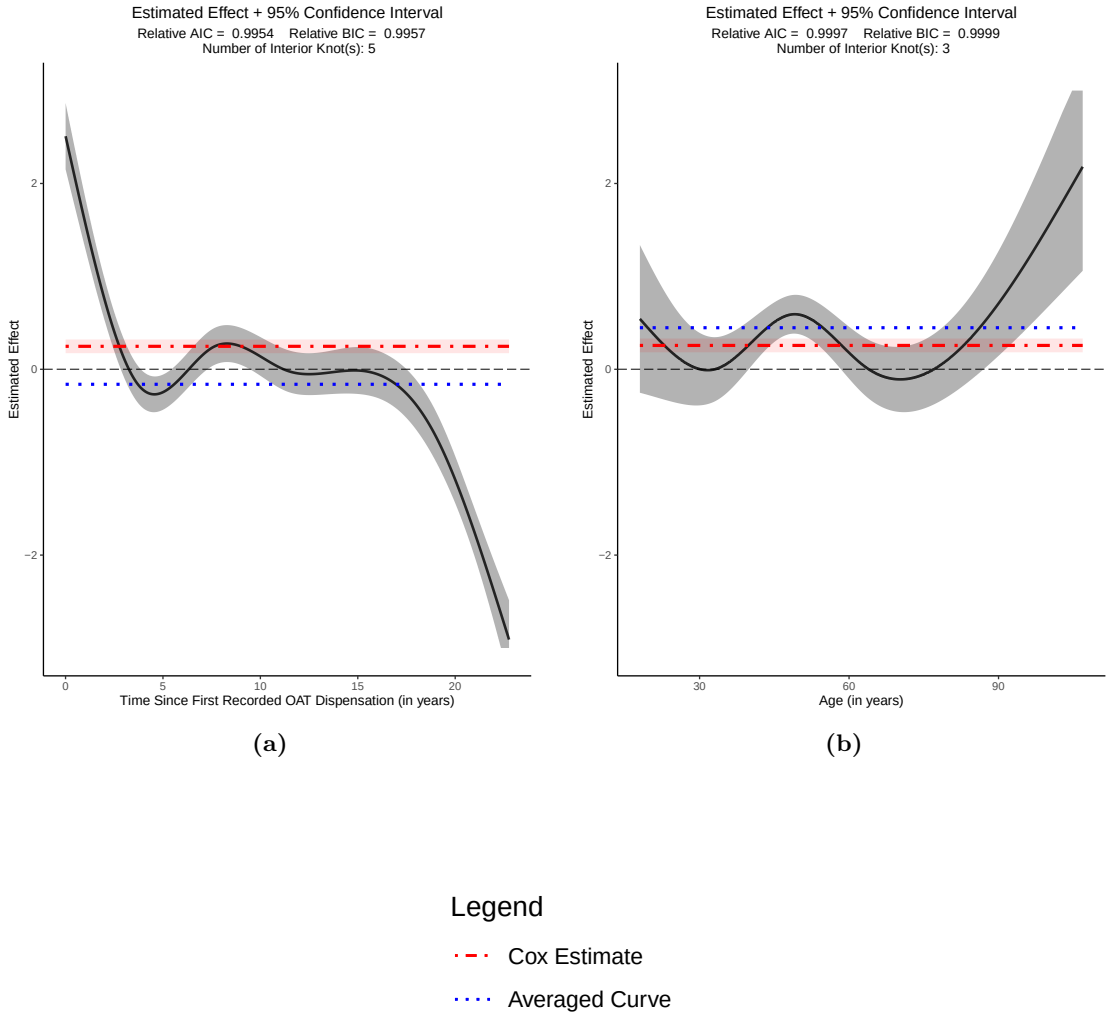


Figure 2.2: Estimates of the cumulative baseline hazard function under the Cox model (2.3), upon stratifying by birth generation levels, where the time scales are (**a**) *time since first recorded OAT dispensation* or (**b**) *age*.



(a) **(b)**

Figure 2.3: Estimates and 95% confidence intervals for the *OAT dispensation rate* estimate under the model (2.6), where the time scales are (**a**) *time since first recorded OAT dispensation* or (**b**) *age*. The average of the estimated curve (in blue) and corresponding estimate and 95% confidence interval under the Cox regression model from Table 2.7 (in red) are included as references. The AIC and BIC values are relative to the corresponding AIC and BIC values under model (2.3).



(a) **(b)**

Figure 2.4: Estimates and 95% confidence intervals for the *OAT dispensation rate* under the model (2.6), upon stratifying by birth generations, where the time scales are (**a**) *time since first recorded OAT dispensation* or (**b**) *age*. The average of the estimated curve (in blue) and corresponding estimate and 95% confidence interval under the (stratified) Cox regression model from Table 2.9 (in red) are included as references. The AIC and BIC values are relative to the corresponding AIC and BIC values under model (2.3).

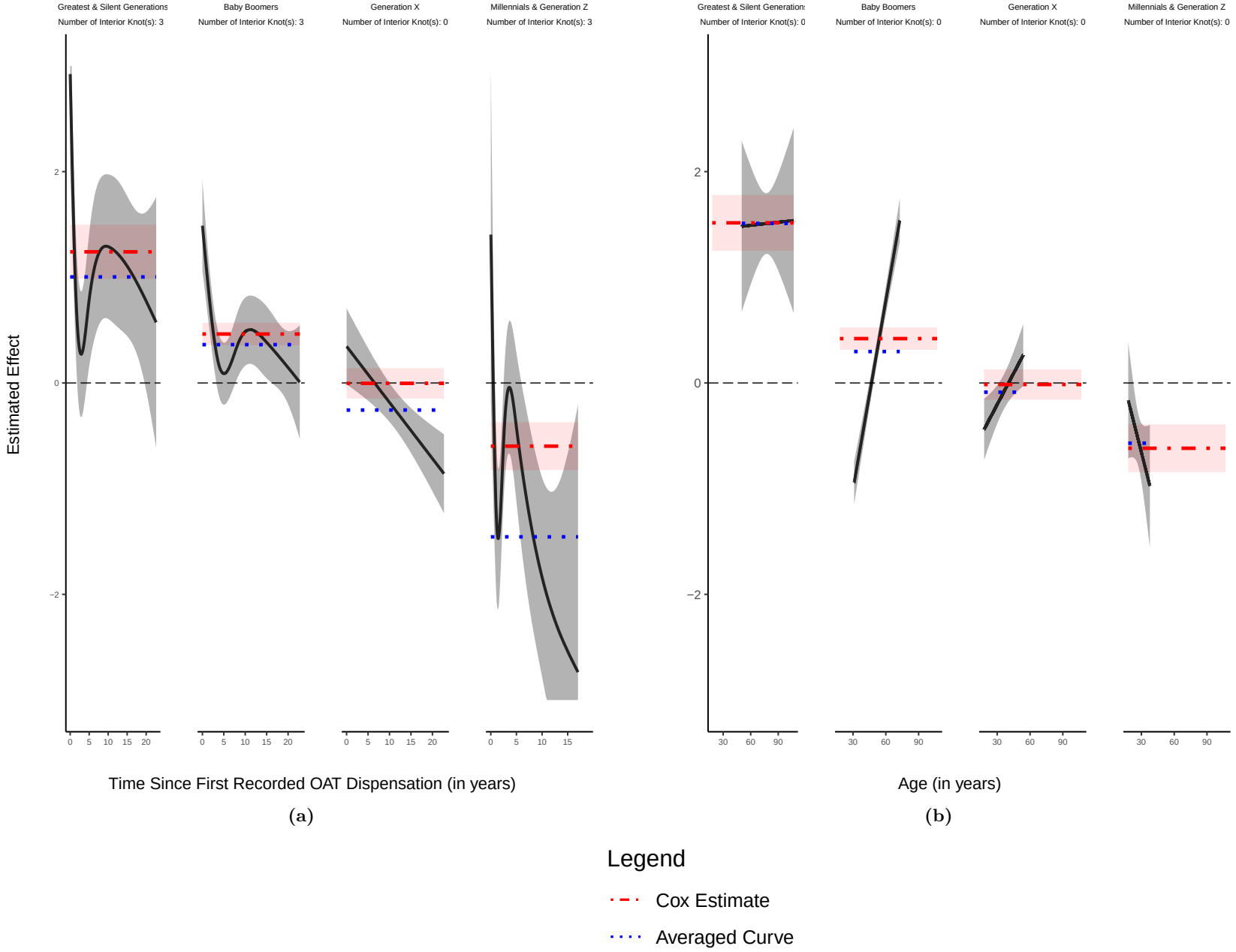
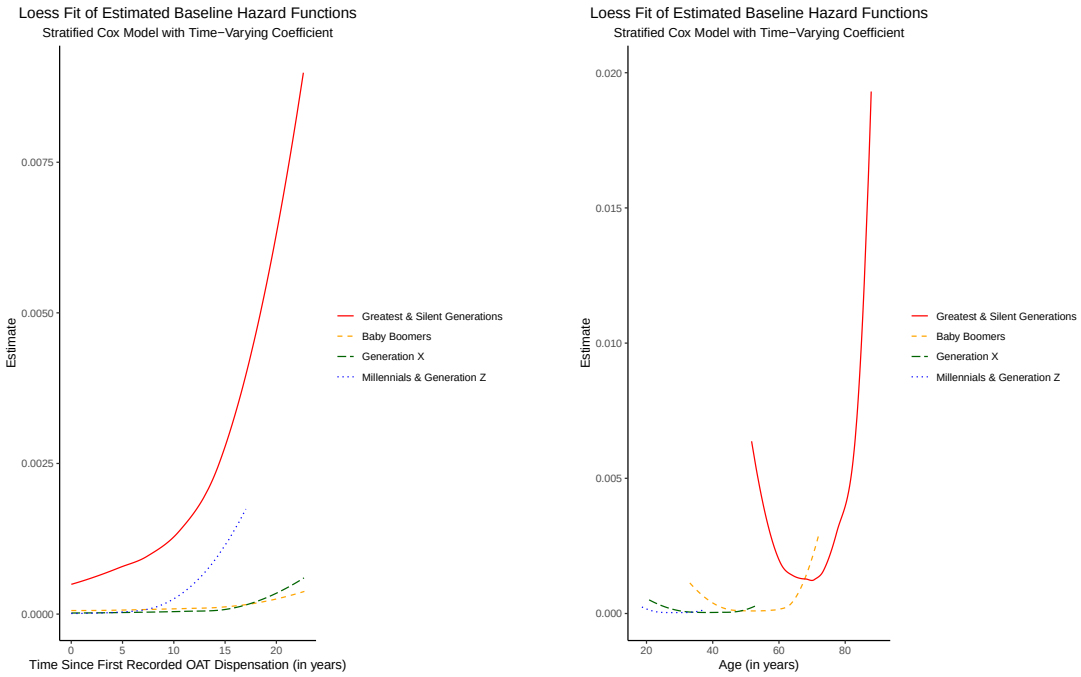


Figure 2.5: LOESS smoothed estimates of the baseline hazard functions under the Cox model (2.6), stratified by birth generations, where the time scales are (**a**) *time since first recorded OAT dispensation* or (**b**) *age*.



(a) **(b)**

Figure 2.6: Predicted survival probabilities for eight randomly selected survivors (two from each of the four birth generation groups) from (2.7), where the time scale is specified as (**a**) *time since first recorded OAT dispensation* or (**b**) *age*.

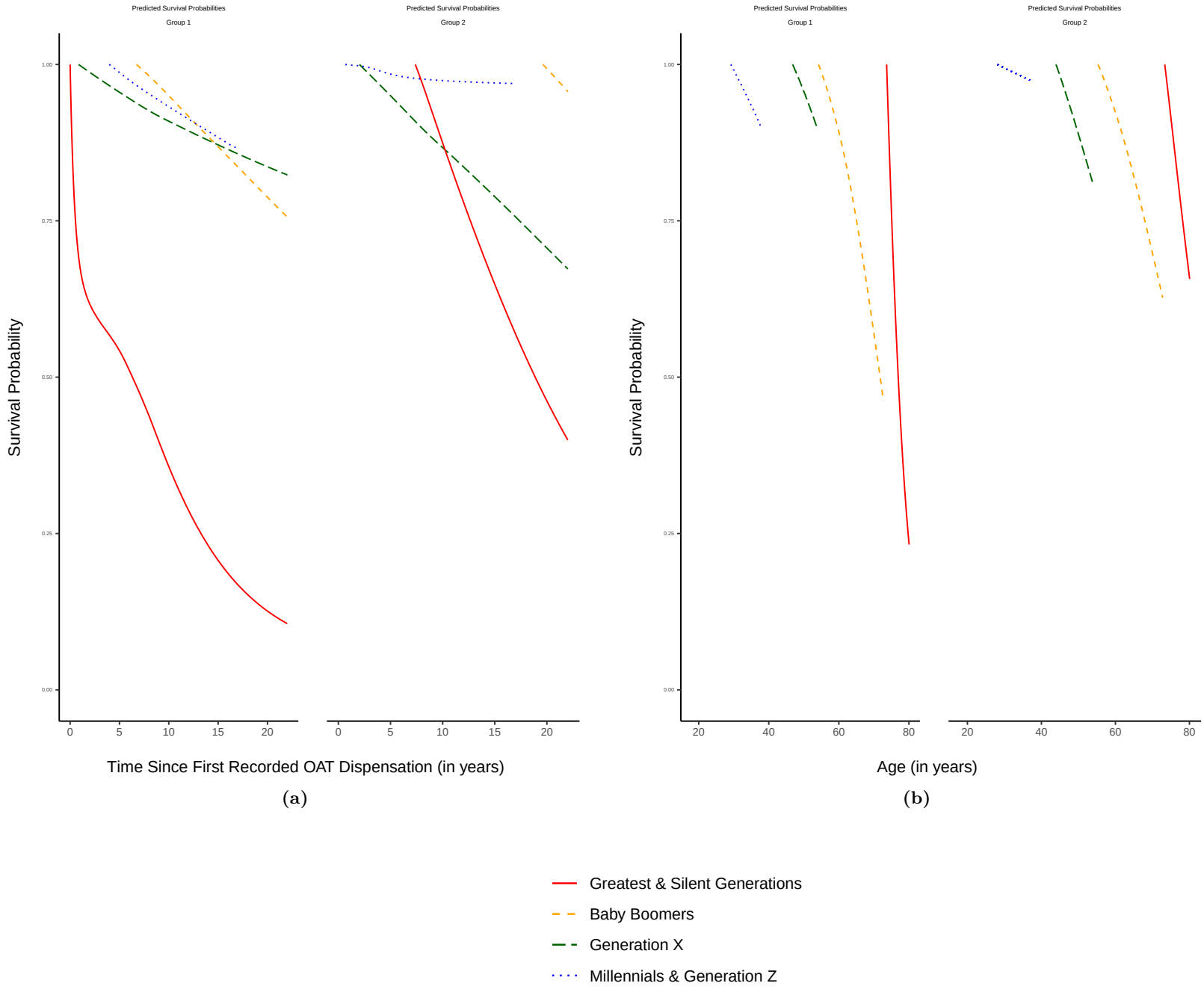


Table 2.1: Drug identification numbers (DINs) in BC-PharmaNet, and descriptions of the various OATs. The colours correspond to specific OATs: Methadone, Buprenorphine & Naloxone, Buprenorphine (only), Slow-Release Oral Morphine, and Injectable Hydromorphone.

Drug Identification Number	Description
999792	METHADONE (MAINTENANCE): 1 MG / ML
999793	METHADONE (MAINTENANCE): 2 MG / ML
66999990	METHADONE (MAINTENANCE): 1 MG / ML
66999991	METHADONE (MAINTENANCE): 1 MG / ML
66999992	METHADONE (MAINTENANCE): 2 MG / ML
66999993	METHADONE (MAINTENANCE): 2 MG / ML
66999997	METHADONE (METHADOSE): 10 MG / ML - DIRECT INTERACTION
66999998	METHADONE (METHADOSE): 10 MG / ML - DIRECT INTERACTION
66999999	METHADONE (METHADOSE): 10 MG / ML - DIRECT INTERACTION & DELIVERY
67000000	METHADONE (METHADOSE): 10 MG / ML - NO DIRECT INTERACTION & DELIVERY
67000001	METHADONE (SUGAR-FREE METHADOSE): 10 MG / ML - DIRECT INTERACTION
67000002	METHADONE (SUGAR-FREE METHADOSE): 10 MG / ML - NO DIRECT INTERACTION
67000003	METHADONE (SUGAR-FREE METHADOSE): 10 MG / ML - DIRECT INTERACT & DELIVERY
67000004	METHADONE (SUGAR-FREE METHADOSE): 10 MG / ML - NO DIRECT INTERACT & DELIVERY
67000005	METADOL-D: 10 MG / ML LIQUID METHADONE - DIRECT INTERACTION
67000006	METADOL-D: 10 MG / ML LIQUID METHADONE - NO DIRECT INTERACTION
67000007	METADOL-D: 10 MG / ML LIQUID METHADONE - DIRECT INTERACTION & DELIVERY
67000008	METADOL-D: 10 MG / ML LIQUID METHADONE - NO DIRECT INTERACTION & DELIVERY
2295695	SUBOXONE: 2 MG / 0.5 MG
2295709	SUBOXONE: 8 MG / 2 MG
2408090	MYLAN - BUPRENORPHINE / NALOXONE: 2 MG / 0.5 MG
2408104	MYLAN - BUPRENORPHINE / NALOXONE: 8 MG / 2 MG
2424851	PMS - BUPRENORPHINE / NALOXONE: 2 MG / 0.5 MG
2424878	PMS - BUPRENORPHINE / NALOXONE: 8 MG / 2 MG
2453908	ACT - BUPRENORPHINE / NALOXONE: 2 MG / 0.5 MG
2453916	ACT - BUPRENORPHINE / NALOXONE: 8 MG / 2 MG
2468085	SUBOXONE: 12 MG / 3 MG
2468093	SUBOXONE: 16 MG / 4 MG
2242962	SUBUTEX: 0.4 MG - TABLET
2242963	SUBUTEX: 2 MG - TABLET
2242964	SUBUTEX: 8 MG - TABLET
66999994	BUPRENORPHINE: 0.4 MG - TABLET
66999995	BUPRENORPHINE: 2 MG - TABLET
66999996	BUPRENORPHINE: 8 MG - TABLET
22123346	KADIAN: 20 MG - CAPSULE
22123347	KADIAN: 50 MG - CAPSULE
22123348	KADIAN: 100 MG - CAPSULE
22123349	KADIAN: 10 MG - CAPSULE
22123340	HYDROMORPHONE: 50 MG / ML - COMPOUNDED INJECTION (SALOME CLINICAL TRIAL)

Table 2.2: ICD-CA-9/10 codes affiliated with an opioid use disorder.

ICD-CA-9 Code	Description	ICD-CA-10 Code	Description
304.0	Opioid type dependence	F11	Mental and behavioural disorders due to use of opioids
304.7	Combinations of opioid type drug with any other drug dependence	T40.0	Poisoning by opium
305.5	Nondependent opioid abuse	T40.1	Poisoning by heroin
965.0	Poisoning by opiates and related narcotics	T40.2	Poisoning by other opioids
E850.0	Accidental poisoning by heroin	T40.3	Poisoning by methadone
E850.1	Accidental poisoning by methadone	T40.4	Poisoning by other synthetic narcotics
E850.2	Accidental poisoning by other opiates and related narcotics	T40.6	Poisoning by other and unspecified narcotics
		X42	Accidental poisoning by and exposure to narcotics and psychodysleptics (hallucinogens)
		X62	Intentional self-poisoning by and exposure to narcotics and psychodysleptics (hallucinogens)
		Y12	Poisoning by and exposure to narcotics and psychodysleptics (hallucinogens)

Table 2.3: ICD-CA-9/10 codes, drug identification numbers, and other codes used for classification.
Abbreviations: MSP: Medical Services Plan; AHFS: American Hospital Formulary Service.

Condition \ Diagnostic Codes	ICD-CA-9	ICD-CA-10	Drug Identification Number	Other
Alcohol Use Disorder	291, 303, 305.0, 357.5, 425.5, 535.3, 571.0-571.3, 655.4, 760.71,V65.42	F10, G31.2, G62.1, G72.1, I42.6, K29.2, K70, K86.0, O35.4, P04.3, Q86.0, Z50.2, Z71.4, Z72.1	2534, 2542, 2041375, 2041391, 2158655, 2213826, 2293269, 2444275, 2451883, 66124085, 66124087, 66124089	NA
Other Substance Use Disorders	292, 304.1-304.6, 304.8, 304.9, 305.2-305.4, 305.6-305.9, 648.3, 760.73, 760.75, 779.5, 969.4, 969.6, 969.7, 970.81, E853.2, E854.1, E854.2,	F12-F16, F19, P04.4, P96.1, T40.5, T40.7-T40.9, T42.4, T43.6, X42, X62, Y12, Z50.3, Z71.5, Z72.2	NA	NA
Mental Ill Health	295-298, 300, 301, 308, 309, 311, 314	F20-F25, F28-F34, F38-F43, F48.8, F48.9, F60-F61, F69, F90	NA	Additional MSP diagnostic code: 50B
Chronic Pain	307.80, 307.89, 338.0, 338.2, 338.4, 344.0, 344.1, 350, 352-357, 719.41, 719.45-719.47, 719.49, 720.0, 720.2, 720.9, 721.0-721.4, 721.6, 721.8, 721.9, 722, 723.0, 723.1, 723.3-723.9, 724.0-724.6, 724.70, 724.79, 724.8, 724.9, 729.0-729.2, 729.4, 729.5, 733.0, 733.7, 733.9, 781, 997.0	F45.4, G50, G52-G64, G82, G89.0, G89.2, G89.4, G97, M08.1, M25.50, M25.51, M25.55-M25.57, M43.2-M43.6, M45, M46.1, M46.3, M46.4, M46.9, M47, M48.0, M48.1, M48.8, M48.9, M50.8, M50.9, M51, M53.1-M53.3, M53.8, M53.9, M54, M60.8, M60.9, M63.3, M79.0-M79.2, M79.6, M79.7, M89, M96.1, R29	NA	NA
Hepatitis C Virus	70.41, 70.44, 70.51, 70.54, 70.7	B17.1, B18.2, B19.2	B20-B24, B97.35, F02.4, O98.7, Z21	NA
HIV/AIDS	042-044, 079.53, 795.8, V08	B20-B24, B97.35, F02.4, O98.7, Z21	NA	MSP fee item: 13015, 13105, 33645, 36370
Sedative Use	NA	NA	NA	AHFS category: 281204, 281208, 282404, 282408, 282492

Table 2.4: Descriptive Statistics of T^*, age on first recorded OAT dispensation date ($t = 0$), and potential time-independent covariates in (2.3).
Abbreviations:
S.D.: Sample standard deviation.

Follow-up Time (Time Since First Recorded OAT Dispensation Date) & Age						
Follow-up Time (in years)	Survivors	Non-Survivors	***Age at $t = 0$ (in years)***	Survivors	Non-Survivors	Total
Minimum	0.0027	0.0027	Minimum	18	18.0082	18
1st Quartile	1.4110	2.1856	1st Quartile	26.3671	32.6534	26.8384
Median	4.6055	5.9712	Median	32.8685	41.0932	33.7973
3rd Quartile	10.4055	11.4281	3rd Quartile	41.5644	48.9274	42.8411
Maximum	22.7644	22.7315	Maximum	99.9726	96.0712	99.9726
Mean	6.7950	7.3446	Mean	34.9387	41.7456	35.8102
S.D.	6.5416	5.9288	S.D.	11.0564	12.6521	11.5003
N (%)	47,731 (87.20)	7,008 (12.8)				
Risk Factors						
Observed OAT Dispensation Rate: $\hat{\nu}$	Survivors	Non-Survivors	Total	***Sex***	Survivors	Non-Survivors	Total
Minimum	0.0012	0.0014	0.0012		N (%)	N (%)	N (%)
1st Quartile	0.2699	0.2979	0.2731	Female	16,226 (33.99)	2,258 (32.22)	18,484 (33.77)
Median	0.6233	0.6700	0.6290	Male	31,505 (66.01)	4,750 (67.78)	36,255 (66.23)
3rd Quartile	0.9169	0.9439	0.9208	***Health Authority***	Survivors	Non-Survivors	Total
Maximum	1	1	1		N (%)	N (%)	N (%)
Mean	0.5804	0.6066	0.5837	Fraser Health	18,437 (33.11)	2,136 (30.48)	20,573 (37.58)
S.D.	0.3370	0.3376	0.3372	Interior	7,620 (15.96)	997 (14.23)	8,617 (15.74)
Incarcerated Rate	Survivors	Non-Survivors	Total	Vancouver Coastal	11,593 (24.29)	2,412 (34.42)	14,005 (25.59)
Median	0	0	0	Vancouver Island	8,085 (16.94)	1,217 (17.37)	9,302 (16.99)
75th Percentile	0.0010	0.0033	0.0013	Northern	1,996 (4.18)	246 (3.51)	2,242 (4.10)
85th Percentile	0.0201	0.0224	0.0205	***Year Category***	Survivors	Non-Survivors	Total
95th Percentile	0.2203	0.1531	0.2101		N (%)	N (%)	N (%)
Maximum	1	1	1	1996-2000	6,827 (14.30)	3,039 (43.36)	9,866 (18.02)
Mean	0.0358	0.0262	0.0345	2001-2006	6,337 (13.28)	1,656 (23.63)	7,993 (14.60)
S.D.	0.3370	0.3376	0.3372	2007-2013	12,150 (25.46)	1,373 (19.59)	13,523 (24.70)
				2013-2018	22,417 (46.97)	940 (13.41)	23,357 (42.67)
				Birth Generation	Survivors	Non-Survivors	Total
					N (%)	N (%)	N (%)
				Greatest & Silent: (1901-1945)	447 (0.94)	654 (9.33)	1,101 (2.01)
				Baby Boomers: (1946-1964)	9,904 (20.75)	3,575 (51.01)	13,479 (24.62)
				Generation X: (1965-1980)	18,861 (39.52)	1,998 (28.51)	20,859 (38.11)
				Millennials & Generation Z: (1981+)	18,519 (38.80)	781 (11.14)	19,300 (35.26)

Table 2.5: Proportions of individuals identified with the listed condition by the end of follow-up. **Bolded** cells specify processes that were merged together to fit the model in (2.3).
Abbreviations:
AUD: Alcohol use disorder;
SUD: Other substance use disorders;
MH: Ill mental health;
CP: Chronic pain;
HCV: Hepatitis C virus;
HIV: HIV/AIDS;
SED: Ever received a sedative;
Plan CG: Ever on PharmaCare plans C or G.

		AUD		SUD		MH		CP		HCV		HIV		SED		Plan CG	
		No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
AUD	No	0.7096	0														
	Yes	0	0.2904														
SUD	No	**0.2346**	**0.0160**	0.2506	0												
	Yes	**0.4750**	**0.2744**	0	0.7494												
MH	No	0.2067	0.0183	0.1432	0.0819	0.2251	0										
	Yes	0.5028	0.2721	0.1074	0.6675	0	0.7749										
CP	No	0.2837	0.0510	0.1540	0.1807	**0.1611**	**0.1736**	0.3347	0								
	Yes	0.4259	0.2394	0.0966	0.5687	**0.0640**	**0.6013**	0	0.6653								
HCV	No	0.6471	0.2265	0.2428	0.6308	0.2129	0.6607	0.3155	0.5582	0.8737	0						
	Yes	0.0625	0.0639	0.0078	0.1186	0.0121	0.1142	0.0192	0.1071	0	0.1263						
HIV	No	0.6854	0.2679	0.2491	0.7041	0.2181	0.7351	0.3242	0.6291	**0.8514**	**0.1018**	0.9533	0				
	Yes	0.0242	0.0225	0.0015	0.0453	0.0069	0.0398	0.0105	0.0362	**0.0222**	**0.0245**	0	0.0467				
SED	No	0.6610	0.2250	0.2506	0.6354	0.2213	0.6647	0.3184	0.5676	0.7869	0.0991	0.8501	0.0359	0.8860	0		
	Yes	0.0486	0.0654	< 0.0001	0.1140	0.0038	0.1102	0.0163	0.0977	0.0868	0.0272	0.1032	0.0108	0	0.1140		
Plan CG	No	0.2576	0.0680	0.1524	0.1732	0.1258	0.1997	0.1467	0.1788	0.3062	0.0194	0.3205	0.0050	0.3056	0.0199	0.3256	0
	Yes	0.4520	0.2224	0.0982	0.5762	0.0992	0.5752	0.1880	0.4865	0.5675	0.1070	0.6328	0.0417	0.5804	0.0941	0	0.6744

Table 2.6: Summary statistics of demographic and clinical characteristics of study subjects by the end of follow-up.

		Survivors		Non-Survivors		Total	
		N	%	*N*	%	*N*	%
Demographic/Clinical Characteristics (by end of follow-up)		47,731	87.2	7,008	12.8	54,739	100.0
Other Substance Use or Alcohol Use Disorders		35,641	74.7	6,244	89.1	41,885	76.5
Ill Mental Health or Chronic Pain		39,595	83.0	6,326	90.3	45,921	83.9
Hepatitis C Virus or HIV/AIDS		5,760	12.1	2,418	34.5	8,178	14.9
Ever Received a Sedative		5,052	10.6	1,180	16.8	6,232	11.4
Ever on PharmaCare Plans C or G		31,753	66.5	5,165	73.7	36,918	67.4
Number of Incarcerations	0	35,445	74.3	4,799	68.5	40,244	73.5
	1	3,924	8.2	659	9.4	4,583	8.4
	2-3	3,392	7.1	619	8.8	4,011	7.3
	4-9	3,193	6.7	616	8.8	3,809	7.0
	10-20	1,337	2.8	238	3.4	1575	2.9
	21-30	292	0.6	53	0.8	345	0.6
	31+	143	0.3	24	0.3	172	0.3

Table 2.7: Parameter and standard error (S.E.) estimates under model (2.3) by maximizing the likelihood function in (2.5). **Bolded** estimates are statistically significant with the type 1 error rate set at 5%.

Time Scale	Time Since First Observed OAT Dispensation		Age	
Covariate Name	Estimate	S.E.	Estimate	S.E.
OAT Dispensation Rate	**0.2471**	0.0379	**0.2569**	0.0380
Incarceration FPC Score	**0.0283**	0.0072	**0.0318**	0.0072
Sex (vs. *Female*)	-	-	-	-
Male	**0.1842**	0.0260	**0.1686**	0.0262
Birth Generation (vs *Greatest & Silent Generations*)	-	-	-	-
Baby Boomers	**-1.3143**	0.0440	**-0.5687**	0.0653
Generation X	**-2.0660**	0.0473	**-0.8268**	0.0804
Millennials & Generation Z	**-2.1523**	0.0587	**-0.5277**	0.1057
Heath Authority (vs *Fraser Health*)	-	-	-	-
Interior	**0.2048**	0.0387	**0.1744**	0.0388
Vancouver Coastal	**0.1058**	0.0302	**0.0944**	0.0302
Vancouver Island	**0.0806**	0.0361	0.0626	0.0361
Northern	0.0499	0.0677	0.0532	0.0677
Year Category (vs. *1996-2000*)	-	-	-	-
2001-2006	**0.1629**	0.0331	0.0350	0.0315
2007-2012	**0.2432**	0.0391	0.0037	0.0359
2013-2018	**0.5914**	0.0491	**0.3576**	0.0433
Alcohol or Other Substance Use Disorders	**0.3588**	0.0423	**0.4049**	0.0431
Ill Mental Health or Chronic pain	**-0.1812**	0.0431	**-0.3047**	0.0428
Hepatitis C Virus or HIV/AIDS	**1.1533**	0.0273	**1.0923**	0.0271
Ever Received a Sedative	**0.4866**	0.0332	**0.4648**	0.0331
Ever on PharmaCare Plans C or G	**-0.2174**	0.0300	**-0.1599**	0.0306

Table 2.8: Summary statistics of the *observed OAT dispensation rate*, $\hat{\nu}_i$, across birth generations and survival status.
Abbreviations:
S.D.: Sample standard deviation.

Greatest & Silent Generations: 1901-1945				**Baby Boomers: 1946-1964**			
	Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total
Minimum	0.0012	0.0014	0.0012	Minimum	0.0012	0.0022	0.0012
1st Quartile	0.1009	0.2539	0.1677	1st Quartile	0.2857	0.4067	0.3123
Median	0.4244	0.7972	0.6648	Median	0.6991	0.8005	0.7244
3rd Quartile	0.9272	0.9855	0.9756	3rd Quartile	0.9550	0.9704	0.9610
Maximum	1	1	1	Maximum	1	1	1
Mean	0.4900	0.6360	0.5768	Mean	0.6121	0.6731	0.6282
S.D.	0.3836	0.3707	0.3826	S.D.	0.3483	0.3271	0.3439
N	447	654	1,101	N	9,904	3,575	13,479
Generation X: 1965-1980				**Millennials & Generation Z: 1981+**			
	Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total
Minimum	0.0013	0.0036	0.0013	Minimum	0.0014	0.0023	0.0014
1st Quartile	0.2750	0.2503	0.2727	1st Quartile	0.2631	0.1415	0.2574
Median	0.6146	0.5510	0.6076	Median	0.6001	0.4199	0.5929
3rd Quartile	0.9042	0.8322	0.8978	3rd Quartile	0.8998	0.7500	0.8951
Maximum	1	1	1	Maximum	1	1	1
Mean	0.5767	0.5375	0.5729	Mean	0.5693	0.4546	0.5647
S.D.	0.3320	0.3174	0.3308	S.D.	0.3335	0.3227	0.3338
N	18,861	1,998	20,859	N	18,519	781	19,300

Table 2.9: Parameter and standard error (S.E.) estimates under model (2.3) by maximizing the likelihood function in (2.5), upon stratifying individuals according to their birth generation. **Bolded** estimates are statistically significant with the type 1 error rate set at 5%.

	Greatest & Silent Generations: 1901-1945				**Baby Boomers: 1946-1964**			
Time Scale	Time Since First Observed OAT Dispensation		Age		Time Since First Observed OAT Dispensation		Age	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
On OAT Rate	**1.2411**	0.1309	**1.5154**	0.1347	**0.4634**	0.0545	**0.4200**	0.0543
Incarceration FPC Score	0.2709	0.1385	**0.3570**	0.1400	0.0225	0.0119	**0.0260**	0.0118
Sex (vs. *Female*)	-	-	-	-	-	-	-	-
Male	0.0269	0.0859	0.0787	0.0882	**0.1540**	0.0369	**0.1078**	0.0370
Heath Authority (vs *Fraser Health*)	-	-	-	-	-	-	-	-
Interior	**0.3651**	0.1216	**0.2658**	0.1237	**0.1505**	0.0569	**0.1332**	0.0569
Vancouver Coastal	0.1911	0.0999	0.1625	0.1007	0.0572	0.0426	0.0535	0.0425
Vancouver Island	0.0837	0.1278	0.1045	0.1282	0.0917	0.0498	0.0666	0.0498
Northern	-0.2900	0.2822	-0.1673	0.2833	0.1642	0.0937	**0.1858**	0.0937
Year Category (vs. *1996-2000*)	-	-	-	-	-	-	-	-
2001-2006	**0.5293**	0.1068	**0.4964**	0.1051	**0.1696**	0.0438	0.0616	0.0419
2007-2012	0.1327	0.1452	-0.0173	0.1444	**0.1897**	0.0568	-0.0836	0.0532
2013-2018	0.0067	0.1710	0.1698	0.1648	**0.4959**	0.0790	**0.1615**	0.0727
Alcohol or Other Substance Use Disorders	**-0.3996**	0.0959	**-0.4165**	0.0988	**0.2540**	0.0621	**0.2492**	0.0619
Ill Mental Health or Chronic pain	0.1898	0.1375	-0.1238	0.1339	**-0.1961**	0.0619	**-0.3272**	0.0614
Hepatitis C Virus or HIV/AIDS	**0.9876**	0.0997	**0.9033**	0.0978	**1.2545**	0.0359	**1.1909**	0.0356
Ever Received a Sedative	**0.3941**	0.1529	**0.4136**	0.1529	**0.3473**	0.0468	**0.3296**	0.0467
Ever on PharmaCare Plans C or G	**-0.5742**	0.0973	**-0.4854**	0.1019	**-0.2051**	0.0445	**-0.1177**	0.0450
	Generation X: 1964-1980				**Millennials & Generation Z: 1981+**			
Time Scale	Time Since First Observed OAT Dispensation		Age		Time Since First Observed OAT Dispensation		Age	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
On OAT Rate	-0.0034	0.0737	-0.0143	0.0735	**-0.5978**	0.1156	**-0.6176**	0.1156
Incarceration FPC Score	0.0251	0.0200	0.0276	0.0199	**0.0416**	0.0147	**0.0405**	0.0149
Sex (vs. *Female*)	-	-	-	-	-	-	-	-
Male	**0.1684**	0.0480	**0.1400**	0.0481	**0.5124**	0.0790	**0.4881**	0.0793
Heath Authority (vs *Fraser Health*)	-	-	-	-	-	-	-	-
Interior	**0.2387**	0.0726	**0.2274**	0.0726	**0.3091**	0.1034	**0.2942**	0.1034
Vancouver Coastal	**0.1495**	0.0557	**0.1385**	0.0557	**0.2062**	0.0941	**0.1972**	0.0941
Vancouver Island	0.0599	0.0687	0.0599	0.0687	0.1888	0.1082	0.1706	0.1082
Northern	-0.0767	0.1263	-0.0841	0.1262	0.0224	0.1908	-0.0058	0.1909
Year Category (vs. *1996-2000*)	-	-	-	-	-	-	-	-
2001-2006	0.0865	0.0616	0.0011	0.0586	0.7932	0.4694	0.7795	0.4622
2007-2012	**0.3967**	0.0682	**0.2256**	0.0613	**0.9263**	0.4698	0.8534	0.4530
2013-2018	**0.7952**	0.0895	**0.5286**	0.0786	**1.5084**	0.4728	**1.3604**	0.4522
Alcohol or Other Substance Use Disorders	**1.0797**	0.1146	**1.0791**	0.1146	**1.2189**	0.1514	**1.2356**	0.1512
Ill Mental Health or Chronic pain	**-0.3302**	0.0810	**-0.3795**	0.0809	**-0.4210**	0.1219	**-0.4297**	0.1217
Hepatitis C Virus or HIV/AIDS	**1.0655**	0.0515	**1.0141**	0.0519	**0.6221**	0.1414	**0.5876**	0.1413
Ever Received a Sedative	**0.6509**	0.0578	**0.6274**	0.0578	**0.7652**	0.0971	**0.7584**	0.0970
Ever on PharmaCare Plans C or G	-0.0871	0.0581	-0.0897	0.0581	**-0.3577**	0.0784	**-0.3543**	0.0784

Chapter 3

Estimating Effects of Time-Varying Exposures on Mortality Hazard Function

3.1 Introduction

This chapter focuses on estimating the effect of an internal covariate process on the mortality hazard. As likelihood inference procedures are no longer applicable, we consider estimating equation procedures. The statistical presentation is within the context of opioid use disorder management; we focus our attention to estimate the effect of a binary alternating process on mortality risk. Section 3.2 introduces notation used throughout this chapter. As individuals frequently changed their OAT dispensation indicator over time, we allow the association of mortality time with risk factors to depend on prior dispensation history. This motivates a generalized Cox regression model with time-dependent stratification (e.g. Hu *et al.* (2011)). We detail our estimation procedure in Section 3.3, and provide a straightforward testing procedure to dynamically update the time-dependent strata. We derive the asymptotic properties of our proposed estimator in Section 3.4, and apply the proposed approach to the provincial administrative service utilization records in Section 3.5. Section 3.6 summarizes the results from a simulation study based on the main findings from the data analysis. We conclude this chapter with further discussions in Section 3.8.

3.2 Notation and Modelling

3.2.1 Notation

Let T denote an individual's survival time since their first recorded OAT dispensation, and $Z(t) \in \{0, 1\}$ be an individual's OAT dispensation indicator at time $t \geq 0$. Let $\mathcal{Z}(t) = \{Z(u) : 0 \leq u \leq t\}$. Additional covariates are denoted by $\boldsymbol{X}(t)$ with $\boldsymbol{\mathcal{X}}(t) = \{\boldsymbol{X}(u) : 0 \leq u \leq t\}$ denoting its history up to time t. Let $\boldsymbol{W}(t) = (Z(t), \boldsymbol{X}(t)')'$ denote all covariates

at time t. Consider a study with observations on T subject to a right-censoring time C. That is, the available information of T is (T^*, δ), where $T^* = T \wedge C$ is the follow-up time of an individual, and $\delta = I(T \leq C)$ is the indicator for whether the survival time T is observed. Assume the study collects n independent and identically distributed realizations of $(T^*, \delta, \mathcal{Z}(T^*), \boldsymbol{\mathcal{X}}(T^*))$. Our objective is to estimate the association of T with $\mathcal{Z}(\cdot)$ upon adjusting for $\boldsymbol{\mathcal{X}}(\cdot)$. We assume that T and C are independent conditional on $\mathcal{Z}(\cdot)$ and $\boldsymbol{\mathcal{X}}(\cdot)$.

3.2.2 Modelling Mortality Hazard

Consider the following generalized Cox regression model for the hazard function of T: for $t > 0$,

$$\begin{aligned}\lambda(t; \mathcal{Z}(t), \boldsymbol{\mathcal{X}}(t)) &= \lim_{\Delta t \to 0^+} \frac{1}{\Delta t} P(t \leq T < t + \Delta t | T \geq t, \mathcal{Z}(t), \boldsymbol{\mathcal{X}}(t)) \\ &= \lambda_0(t; \mathcal{Z}(t)) \exp\{\boldsymbol{\theta}(\mathcal{Z}(t))' \boldsymbol{W}(t)\}, \end{aligned} \tag{3.1}$$

where $\lambda_0(t; \mathcal{Z}(t))$ is an arbitrary baseline hazard function, and $\boldsymbol{\theta}(\mathcal{Z}(t))$ is a known function up to finite dimensional parameters with the dimension of $\boldsymbol{\theta}(\mathcal{Z}(t))$ being the same as $\boldsymbol{W}(t)$. We explicitly permit both the regression parameter and baseline hazard function in (3.1) to vary in accordance with an individual's OAT dispensation history.

In an attempt to adequately quantify the association between the dispensation process and mortality risk, we stratify individuals into groups based on their dispensation history, which makes the stratification time-dependent. Let $g(\mathcal{Z}(t)) \in \{1, 2, \cdots, G\}$ denote a stratification variable that is fully determined by an individual's dispensation history up to time $t > 0$, where $G < \infty$ is known. This naturally leads us to consider the following three specifications:

> **Model A**: Suppose $\lambda_0(t; \mathcal{Z}(t)) = \lambda_{0g}(t)$ and $\boldsymbol{\theta}(\mathcal{Z}(t)) = \boldsymbol{\theta}_g = (\boldsymbol{\alpha}_g', \boldsymbol{\beta}')'$ when $g(\mathcal{Z}(t)) = g$, where $\boldsymbol{\theta}_g$ is a vector of unknown regression parameters, $\boldsymbol{\alpha}_g$ is a q_A-dimensional vector of stratum-specific effects, and $\boldsymbol{\beta}$ is a q_B-dimensional vector of shared effects across strata. Without loss of generality, we partition the covariates as $\boldsymbol{W}(t) = (\boldsymbol{W}^A(t)', \boldsymbol{W}^B(t)')'$, where $\boldsymbol{W}^A(t)$ and $\boldsymbol{W}^B(t)$ have the same dimensions as $\boldsymbol{\alpha}_g$ and $\boldsymbol{\beta}$, respectively. The model in (3.1) then becomes for $g(\mathcal{Z}(t)) = g$,
>
> $$\lambda(t; \mathcal{Z}(t), \boldsymbol{\mathcal{X}}(t)) = \lambda_{0g}(t) \exp\{\boldsymbol{\theta}_g' \boldsymbol{W}(t)\}. \tag{3.2}$$
>
> The model resembles an extended Cox regression model with a time-varying covariate $\boldsymbol{W}(t)$ and time-dependent strata (Hu *et al.* 2011).

Model B: Suppose $\lambda_0(t; \mathcal{Z}(t)) \equiv \lambda_0(t)$ and $\boldsymbol{\theta}(\mathcal{Z}(t)) = \boldsymbol{\theta}_g = (\boldsymbol{\alpha}_g', \boldsymbol{\beta}')'$ when $g(\mathcal{Z}(t)) = g$. The model in (3.1) then becomes for $g(\mathcal{Z}(t)) = g$,

$$\lambda(t; \mathcal{Z}(t), \boldsymbol{\mathcal{X}}(t)) = \lambda_0(t) \exp\{\boldsymbol{\theta}_g' \boldsymbol{W}(t)\}. \tag{3.3}$$

Clearly, (3.3) arises as a special case of (3.2) if $\lambda_{0g}(t) \equiv \lambda_0(t)$ for $g = 1, \cdots, G$. This can be done by fitting the model (3.2), plotting the estimates of $\lambda_{01}(t)$, $\cdots$, $\lambda_{0G}(t)$ over time, and visually assess if (3.3) is a better fit to the data. Alternatively, we could specify $\lambda_{0g}(t) = \lambda_{01}(t) \exp\{\gamma_g\}$ for $g = 2, \cdots, G$, and test whether $\gamma_g = 0$ for all $g = 2, \cdots, G$.

Model C: Suppose $\lambda_0(t; \mathcal{Z}(t)) \equiv \lambda_0(t)$ and $\boldsymbol{\theta}(\mathcal{Z}(t)) \equiv \boldsymbol{\theta}$. The model in (3.1) reduces to the Cox regression model (Cox 1972) with time-varying covariates:

$$\lambda(t; \mathcal{Z}(t), \boldsymbol{\mathcal{X}}(t)) = \lambda_0(t) \exp\{\boldsymbol{\theta}' \boldsymbol{W}(t)\}. \tag{3.4}$$

Clearly, (3.4) is a special case of (3.3), when $\boldsymbol{\alpha}_1 = \cdots = \boldsymbol{\alpha}_G$.

We take (3.2) as the primary model, since it is the most general of the three models presented. The forthcoming estimation procedure is developed with model (3.2). It can estimate parameters under (3.4) by fixing $G \equiv 1$, and can estimate parameters under (3.3) by including time-dependent dummy variables pertaining to the original levels of $g(\cdot)$ in $\boldsymbol{W}(\cdot)$ and carrying out the inference procedure under (3.4).

3.3 Estimation Procedure

Let $N_i(t) = I(T_i \leq t)$, $Y_i(t) = I(T_i^* \geq t)$, and $\boldsymbol{\Theta} = (\boldsymbol{\alpha}_1', \cdots, \boldsymbol{\alpha}_G', \boldsymbol{\beta}')'$ be all the regression parameters in model (3.2), with $\boldsymbol{\Theta}_0$ denoting the true value of $\boldsymbol{\Theta}$.

3.3.1 Estimating Regression Parameters

Under (3.2), consider the estimating functions for $\boldsymbol{\Theta}$; $\boldsymbol{U}(\boldsymbol{\Theta}) = (U_1^A(\boldsymbol{\theta}_1)', \cdots, U_G^A(\boldsymbol{\theta}_G)', U^B(\boldsymbol{\Theta})')'$, where

$$
\begin{aligned}
U_g^A(\boldsymbol{\theta}_g) &= \int_0^\infty \sum_{i:g(\mathcal{Z}_i(t))=g} Y_i(t) \left[\boldsymbol{W}_i^A(t) - \frac{E_g^A(t, \boldsymbol{\theta}_g)}{E_g(t, \boldsymbol{\theta}_g)} \right] dN_i(t), \quad g = 1, \cdots, G, \\
U^B(\boldsymbol{\Theta}) &= \sum_{g=1}^G \int_0^\infty \sum_{i:g(\mathcal{Z}_i(t))=g} Y_i(t) \left[\boldsymbol{W}_i^B(t) - \frac{E_g^B(t, \boldsymbol{\theta}_g)}{E_g(t, \boldsymbol{\theta}_g)} \right] dN_i(t), \\
E_g(t, \boldsymbol{\theta}) &= \sum_{j:g(\mathcal{Z}_j(t))=g} Y_j(t) \exp\{\boldsymbol{\theta}' \boldsymbol{W}_j(t)\}, \text{ and} \\
E_g^A(t, \boldsymbol{\theta}) &= \sum_{j:g(\mathcal{Z}_j(t))=g} Y_j(t) \exp\{\boldsymbol{\theta}' \boldsymbol{W}_j(t)\} \boldsymbol{W}_j^A(t), \text{ and} \\
E_g^B(t, \boldsymbol{\theta}) &= \sum_{j:g(\mathcal{Z}_j(t))=g} Y_j(t) \exp\{\boldsymbol{\theta}' \boldsymbol{W}_j(t)\} \boldsymbol{W}_j^B(t).
\end{aligned}
$$

Note that $\boldsymbol{W}(t)$ may have internal or external covariate components, and $\boldsymbol{U}(\boldsymbol{\Theta})$ is well-defined regardless. In the case where all of the covariates are external, $\boldsymbol{U}(\boldsymbol{\Theta})$ corresponds to the partial score function of $\boldsymbol{\Theta}$ (Prentice *et al.* 1981). One can then adopt the inference procedure presented in Hu *et al.* (2011) to estimate model parameters. We show in Section 3.4.1 that $\boldsymbol{U}(\boldsymbol{\Theta}_0)$ is centred at zero asymptotically under (3.2). A reasonable estimator for $\boldsymbol{\Theta}$ is therefore the solution to the equation $\boldsymbol{U}(\boldsymbol{\Theta}) = \boldsymbol{0}$. In fact, we show in Section 3.4.2 that the estimator $\hat{\boldsymbol{\Theta}}$ converges almost surely to $\boldsymbol{\Theta}_0$ under some regularity conditions. We further establish the asymptotic distribution of $\sqrt{n}(\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0)$ in Section 3.4.3, in which the corresponding asymptotic variance can be consistently estimated with a Huber-like sandwich estimator $\widehat{\boldsymbol{AV}}(\hat{\boldsymbol{\Theta}}) = \hat{\boldsymbol{\Psi}}^{-1}(\hat{\boldsymbol{\Theta}}) \hat{\boldsymbol{\Phi}}(\hat{\boldsymbol{\Theta}}) \hat{\boldsymbol{\Psi}}^{-1}(\hat{\boldsymbol{\Theta}})$ with $\hat{\boldsymbol{\Psi}}(\boldsymbol{\Theta}) = -\frac{1}{n} \frac{\partial}{\partial \boldsymbol{\Theta}} \boldsymbol{U}(\boldsymbol{\Theta})$ and $\hat{\boldsymbol{\Phi}}(\boldsymbol{\Theta}) = \frac{1}{n} \sum_{i=1}^n \hat{\boldsymbol{\Omega}}_i(\boldsymbol{\Theta}) \hat{\boldsymbol{\Omega}}_i(\boldsymbol{\Theta})'$, where

$$
\begin{aligned}
\hat{\boldsymbol{\Omega}}_i(\boldsymbol{\Theta}) &= (\hat{\Omega}_{i1}^A(\boldsymbol{\theta}_1)', \cdots, \hat{\Omega}_{iG}^A(\boldsymbol{\theta}_G)', \hat{\Omega}_i^B(\boldsymbol{\Theta})')', \\
\hat{\Omega}_{ig}^A(\boldsymbol{\theta}) &= \int_0^\infty Y_i(t) I(g(\mathcal{Z}_i(t)) = g) \left[\boldsymbol{W}_i^A(t) - \frac{E_g^A(t, \boldsymbol{\theta})}{E_g(t, \boldsymbol{\theta})} \right] d\hat{M}_{ig}(t, \boldsymbol{\theta}), \\
\hat{\Omega}_i^B(\boldsymbol{\Theta}) &= \sum_{g=1}^G \int_0^\infty Y_i(t) I(g(\mathcal{Z}_i(t)) = g) \left[\boldsymbol{W}_i^B(t) - \frac{E_g^B(t, \boldsymbol{\theta}_g)}{E_g(t, \boldsymbol{\theta}_g)} \right] d\hat{M}_{ig}(t, \boldsymbol{\theta}_g), \text{ and} \\
\hat{M}_{ig}(t, \boldsymbol{\theta}) &= N_i(t) I(g(\mathcal{Z}_i(t)) = g) - \int_0^t Y_i(u) I(g(\mathcal{Z}_i(u)) = g) \exp\{\boldsymbol{\theta}' \boldsymbol{W}_i(u)\} d\hat{\Lambda}_{0g}(u)
\end{aligned}
$$

and $\hat{\Lambda}_{0g}(\cdot)$ is a consistent estimator for $\Lambda_{0g}(\cdot)$. If all of the covariates are indeed external, we recognize $\hat{\boldsymbol{\Psi}}(\hat{\boldsymbol{\Theta}})$ as the observed information matrix under (3.2), so the corresponding variance estimator would simplify to $\widehat{\boldsymbol{AV}}(\hat{\boldsymbol{\Theta}}) = \hat{\boldsymbol{\Psi}}^{-1}(\hat{\boldsymbol{\Theta}})$.

3.3.2 Estimating Baseline Hazard Function

For fixed $g \in \{1, \cdots, G\}$, we view $d\Lambda_{0g}(t) = \lambda_{0g}(t)dt$ as a finite-dimensional parameter upon treating $\lambda_{0g}(\cdot)$ as a piece-wise constant function with the jumps at the uncensored survival times. With $\boldsymbol{\theta}_g$ fixed, the following estimating equation is unbiased under (3.2):

$$\sum_{i:g(\mathcal{Z}_i(t))=g} Y_i(t)[dN_i(t) - \exp\{\boldsymbol{\theta}_g' \boldsymbol{W}_i(t)\} d\Lambda_{0g}(t)] = 0.$$

By solving the above equation for $d\Lambda_{0g}(t)$, this promotes the estimator

$$d\hat{\Lambda}_{0g}(t; \boldsymbol{\theta}_g) = \sum_{i:g(\mathcal{Z}_i(t))=g} \frac{Y_i(t) dN_i(t)}{\sum_{j:g(\mathcal{Z}_j(t))=g} Y_j(t) \exp\{\boldsymbol{\theta}_g' \boldsymbol{W}_j(t)\}}. \tag{3.5}$$

Here, we take the convention that $0/0 = 0$. By replacing the unknown $\boldsymbol{\theta}_g$ with its corresponding estimate, $\hat{\boldsymbol{\theta}}_g = (\hat{\boldsymbol{\alpha}}_g', \hat{\boldsymbol{\beta}}')'$, the baseline hazard function is estimated with a Breslow-type estimator $d\hat{\Lambda}_{0g}(t; \hat{\boldsymbol{\theta}}_g)$. Under the regularity conditions in Section 3.4, we show that $d\hat{\Lambda}_{0g}(t; \hat{\boldsymbol{\theta}}_g)$ converges almost surely to $d\Lambda_{0g}(t)$. Furthermore, we establish the weak convergence of $d\hat{\Lambda}_{0g}(t; \hat{\boldsymbol{\theta}}_g)$. Either $1 - \alpha^*$ pointwise confidence intervals or $1 - \alpha^*$ confidence bands for $d\Lambda_{0g}(t)$ can be constructed accordingly.

3.3.3 Dynamic Grouping Based on Wald-Type Testing

When the stratification variable is ordinal, one may question if the difference between successive groups are statistically significant. If the difference is not significant, we can simplify the model in (3.2) (or (3.3)) by merging groups $g - 1$ and g together and re-estimate $\boldsymbol{\Theta}$ with the updated groups; otherwise, we keep these two groups separate from each other. Proceeding in this manner would result in identifying $H \leq G$ data driven risk classes. This would provide a gain of efficiency by reducing the number of parameters to estimate.

To carry out the idea, we consider the following hypothesis test for a fixed $g \in \{2, \cdots, G\}$:

$$H_0 : \boldsymbol{\alpha}_g = \boldsymbol{\alpha}_{g-1} \quad \text{vs.} \quad H_a : \boldsymbol{\alpha}_g \neq \boldsymbol{\alpha}_{g-1}. \tag{3.6}$$

Here, we assume that $\lambda_{0g}(\cdot) \equiv \lambda_{0,g-1}(\cdot)$ under (3.2), so that group specific effects for groups $g - 1$ and g are captured by $\boldsymbol{\alpha}_{g-1}$ and $\boldsymbol{\alpha}_g$, respectively. Based on the asymptotic normality of $\hat{\boldsymbol{\Theta}}$, we construct a Wald test statistic

$$J_g = (\hat{\boldsymbol{\alpha}}_g - \hat{\boldsymbol{\alpha}}_{g-1})' \Big[\mathrm{Var}(\hat{\boldsymbol{\alpha}}_g - \hat{\boldsymbol{\alpha}}_{g-1}) \Big]^{-1} (\hat{\boldsymbol{\alpha}}_g - \hat{\boldsymbol{\alpha}}_{g-1}). \tag{3.7}$$

With the variance of $\hat{\boldsymbol{\Theta}}$ estimated with $\widehat{\boldsymbol{AV}}(\hat{\boldsymbol{\Theta}})$, we can therefore estimate $\mathrm{Var}(\hat{\boldsymbol{\alpha}}_g - \hat{\boldsymbol{\alpha}}_{g-1})$ by $\boldsymbol{C}_g \widehat{\boldsymbol{AV}}(\hat{\boldsymbol{\Theta}}) \boldsymbol{C}_g'$, where $\boldsymbol{C}_g$ is the constant vector such that $\boldsymbol{C}_g \boldsymbol{\Theta} = \boldsymbol{\alpha}_g - \boldsymbol{\alpha}_{g-1}$. Under H_0

in (3.6), $J_g \dot{\sim} F_{q_A}(\cdot)$, where $F_{q_A}(\cdot)$ denotes the χ^2-distribution function with q_A degrees of freedom. We reject H_0 if $J_g > F_{q_A}^{-1}(1-\alpha^*)$, where α^* is the predetermined type I error rate.

3.4 Asymptotic Properties

Assuming that the study collects n independent and identically distributed realizations of $(T^*, \delta, \mathcal{Z}(T^*), \boldsymbol{\mathcal{X}}(T^*))$, we proceed to establish the large sample properties claimed in Section 3.3. To do so, we impose the following regularity conditions:

(a) $P(C \geq \tau) > 0$, where τ is a predetermined constant. In our application, we can view τ as the largest follow-up time.

(b) $N_g/n \nrightarrow 0$ as $n \to \infty$, where $N_g = \max_{t>0} \#\{i : g(\mathcal{Z}_i(t)) = g\}$.

(c) There exists a constant K such that $|\boldsymbol{W}_{ij}(0)| + \int_0^\infty |Y_i(t) d\boldsymbol{W}_{ij}(t)| \leq K$, for all $i = 1, \cdots, n$ and $j = 1, \cdots, p+q$, where $\boldsymbol{W}_{ij}(\cdot)$ is the jth component of $\boldsymbol{W}_i(\cdot)$. That is, $\boldsymbol{W}_i(\cdot)$ has a bounded total variation.

(d) Let $\boldsymbol{\Theta}_0 = (\boldsymbol{\alpha}_{10}', \cdots, \boldsymbol{\alpha}_{G0}', \boldsymbol{\beta}_0')'$ denote the true value of $\boldsymbol{\Theta}$ under the model in (3.2). We assume that $\boldsymbol{\Psi}(\boldsymbol{\Theta}_0)$ is positive-definite, where

$$\boldsymbol{\Psi}(\boldsymbol{\Theta}) = \begin{bmatrix} \Psi_1(\boldsymbol{\theta}_1) & \mathbf{0} & \cdots & \mathbf{0} & \Psi_1^*(\boldsymbol{\theta}_1) \\ \mathbf{0} & \Psi_2(\boldsymbol{\theta}_2) & \cdots & \mathbf{0} & \Psi_2^*(\boldsymbol{\theta}_2) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \Psi_G(\boldsymbol{\theta}_G) & \Psi_G^*(\boldsymbol{\theta}_G) \\ \Psi_1^*(\boldsymbol{\theta}_1)' & \Psi_2^*(\boldsymbol{\theta}_2)' & \cdots & \Psi_G^*(\boldsymbol{\theta}_G)' & \Psi^*(\boldsymbol{\Theta}) \end{bmatrix},$$

$$\Psi_g(\boldsymbol{\theta}) = \int_0^\infty \left[\frac{\mathbb{E}\{E_g^{AA}(t,\boldsymbol{\theta})\}}{\mathbb{E}\{E_g(t,\boldsymbol{\theta})\}} - \frac{\mathbb{E}\{E_g^{A}(t,\boldsymbol{\theta})\}^{\otimes 2}}{\mathbb{E}\{E_g(t,\boldsymbol{\theta})\}^2} \right] \mathbb{E}\{Y(t)\lambda(t; \mathcal{Z}(t), \boldsymbol{\mathcal{X}}(t))\}\, dt,$$

$$\Psi^*(\boldsymbol{\Theta}) = \sum_{g=1}^{G} \int_0^\infty \left[\frac{\mathbb{E}\{E_g^{BB}(t,\boldsymbol{\theta}_g)\}}{\mathbb{E}\{E_g(t,\boldsymbol{\theta}_g)\}} - \frac{\mathbb{E}\{E_g^{B}(t,\boldsymbol{\theta}_g)\}^{\otimes 2}}{\mathbb{E}\{E_g(t,\boldsymbol{\theta}_g)\}^2} \right] \mathbb{E}\{Y(t)\lambda(t; \mathcal{Z}(t), \boldsymbol{\mathcal{X}}(t))\}\, dt,$$

$$\Psi_g^*(\boldsymbol{\theta}) = \int_0^\infty \left[\frac{\mathbb{E}\{E_g^{AB}(t,\boldsymbol{\theta})\}}{\mathbb{E}\{E_g(t,\boldsymbol{\theta})\}} - \frac{\mathbb{E}\{E_g^{A}(t,\boldsymbol{\theta})\}}{\mathbb{E}\{E_g(t,\boldsymbol{\theta})\}} \left\{ \frac{\mathbb{E}\{E_g^{B}(t,\boldsymbol{\theta})\}}{\mathbb{E}\{E_g(t,\boldsymbol{\theta})\}} \right\}' \right] \mathbb{E}\{Y(t)\lambda(t; \mathcal{Z}(t), \boldsymbol{\mathcal{X}}(t))\}\, dt, \text{ and}$$

$$E_g^{QR}(t,\boldsymbol{\theta}) = \sum_{j:g(\mathcal{Z}_j(t))=g} Y_j(t) \exp\{\boldsymbol{\theta}' \boldsymbol{W}_j(t)\} \boldsymbol{W}_j^Q(t) \boldsymbol{W}_j^R(t)' \quad Q, R \in \{A, B\}.$$

Here, $E_g(\cdot)$, $E_g^A(\cdot)$, and $E_g^B(\cdot)$ are as presented in Section 3.3, and $\mathbb{E}\{A\}$ denotes the expected value of A.

Essentially, condition (b) states that the size of group $g \in \{1, \cdots, G\}$ can be large enough if n is sufficiently large. Condition (d) holds if, at least for some interval of t, the distribution of $\boldsymbol{W}(t)$ conditional on $Y(t) = 1$ does not concentrate on a $(q_A G + q_B - 1)$-dimensional hyperplane.

3.4.1 Weak Convergence of Estimating Functions

Let $\boldsymbol{\theta}_{g0} = (\boldsymbol{\alpha}'_{g0}, \beta'_0)'$ denote the true value $\boldsymbol{\theta}_g$ under the model in (3.2). We proceed to establish the weak convergence of $U_g^A(t, \boldsymbol{\theta}_g)$, where

$$U_g^A(t, \boldsymbol{\theta}) = \int_0^t \sum_{i:g(\mathcal{Z}_i(u))=g} Y_i(u) \left[\boldsymbol{W}_i^A(u) - \frac{E_g^A(u, \boldsymbol{\theta})}{E_g(u, \boldsymbol{\theta})} \right] dN_i(u),$$

so that $U_g^A(\boldsymbol{\theta}) \equiv U_g^A(\infty, \boldsymbol{\theta})$. Letting

$$M_{ig}(t, \boldsymbol{\theta}_{g0}) = N_i(t) I(g(\mathcal{Z}_i(t)) = g) - \int_0^t Y_i(u) I(g(\mathcal{Z}_i(u)) = g) \exp\{\boldsymbol{\theta}'_{g0} \boldsymbol{W}_i(u)\} d\Lambda_{0g}(u),$$

we see after some algebra that

$$\begin{aligned} U_g^A(t, \boldsymbol{\theta}_{g0}) &= \int_0^t \sum_{i:g(\mathcal{Z}_i(u))=g} Y_i(u) \left[\boldsymbol{W}_i^A(u) - \frac{E_g^A(u, \boldsymbol{\theta}_{g0})}{E_g(u, \boldsymbol{\theta}_{g0})} \right] dM_{ig}(u, \boldsymbol{\theta}_{g0}) \\ &= \int_0^t \sum_{i:g(\mathcal{Z}_i(u))=g} Y_i(u) \boldsymbol{W}_i^A(u) dM_{ig}(u, \boldsymbol{\theta}_{g0}) - \int_0^t \sum_{i:g(\mathcal{Z}_i(u))=g} Y_i(u) \frac{E_g^A(u, \boldsymbol{\theta}_{g0})}{E_g(u, \boldsymbol{\theta}_{g0})} dM_{ig}(u, \boldsymbol{\theta}_{g0}) \\ &= \bar{M}_g^A(t, \boldsymbol{\theta}_{g0}) - \int_0^t \frac{E_g^A(u, \boldsymbol{\theta}_{g0})}{E_g(u, \boldsymbol{\theta}_{g0})} d\bar{M}_g(u, \boldsymbol{\theta}_{g0}), \end{aligned} \tag{3.8}$$

where $\bar{M}_g(t, \boldsymbol{\theta}) = \sum_{i:g(\mathcal{Z}_i(t))=g} Y_i(t) M_{ig}(t, \boldsymbol{\theta})$. Under the model in (3.2), both terms of (3.8) for fixed t are sums of independent and identically distributed zero-mean terms. Applying the multivariate central limit theorem, $(n^{-1/2}\bar{M}_g(\cdot), n^{-1/2}\bar{M}_g^A(\cdot))$ converges in finite dimensional distributions to a zero-mean Gaussian process, say $(\mathcal{G}_{M_g}(\cdot), \mathcal{G}_{M_g^A}(\cdot))$.

Note that individuals may change their OAT dispensation indicator throughout the study period. Consider a partition of the study period $[0, \tau]$: $0 = u_0 < u_1 < \cdots < u_R = \tau$, such that an individual can change their OAT dispensation indicator at most once in $(u_{r-1}, u_R]$, for $r = 1, \cdots, R$. For a fixed r, regularity condition (c) implies that for $t \in (u_{r-1}, u_r]$, $\boldsymbol{W}(t)$ is bounded, and $\boldsymbol{W}(t)$ has finite variation since $Z(t)$ can change at most once. Without loss of generality, we assume that each component of $\boldsymbol{W}^A(t)$ is non-negative, so that for $t \in (u_{r-1}, u_r]$, each of the q_A components of $\int_{u_{r-1}}^t \boldsymbol{W}^A(u) dM_g(u, \boldsymbol{\theta}_{g0})$ is the difference between two monotonic functions in t. Since monotone functions have a pseudo-dimension of one (Pollard (1990), Lemma A.2 of Bilias *et al.* (1997)), the processes $\{M_{ig}(t, \boldsymbol{\theta}_{g0}); i = 1, \cdots, n;\ t \in (u_{r-1}, u_r]\}$ and $\{\int_{u_{r-1}}^t \boldsymbol{W}_i^A(u) dM_{ig}(u, \boldsymbol{\theta}_{g0}); i = 1, \cdots, n\}$ are manageable (Pollard (1990), Lemma A.1 of Bilias *et al.* (1997)). It then follows from the functional central limit theorem (Pollard 1990) that $(n^{-1/2}\bar{M}_g(\cdot), n^{-1/2}\bar{M}_g^A(\cdot))$ is tight and converges weakly to $(\mathcal{G}_{M_g}(\cdot), \mathcal{G}_{M_g^A}(\cdot))$ for $t \in (u_{r-1}, u_r]$.

Furthermore, it can be shown that

$$\mathbb{E}\{\mathcal{G}_{M_g}(t) - \mathcal{G}_{M_g}(s)\}^4 \leq \tilde{K}_g(\Lambda_{0g}(u_{r-1}, t) - \Lambda_{0g}(u_{r-1}, s))^2, \quad s < t$$

for some constant $\tilde{K}_g > 0$, where $\Lambda_{0g}(u_{r-1}, t) = \int_{u_{r-1}}^{t} d\Lambda_{0g}(u)$. By the Kolmogorov-Centsov Theorem (Karatzas and Shreve 1988), $\mathcal{G}_{M_g}$ has continuous sample paths under Euclidean distance, and similarly, $\mathcal{G}_{M_g^A}$ also has continuous sample paths.

By the strong embedding theorem (Shorack and Wellner 1986), we see that almost surely, $(n^{-1/2}\bar{M}_g(\cdot), n^{-1/2}\bar{M}_g^A(\cdot), E_g(\cdot), E_g^A(\cdot))$ converges to $(\mathcal{G}_{M_g}(\cdot), \mathcal{G}_{M_g^A}(\cdot), \mathbb{E}\{E_g(\cdot)\}, \mathbb{E}\{E_g^A(\cdot)\})$ in a new probability space. By noting that $E_g(t, \boldsymbol{\theta}_{g0})$ and $E_g^A(t, \boldsymbol{\theta}_{g0})$ are monotonic functions in t, for $t \in (u_{r-1}, u_r]$, all of the conditions of Lemma A.1 of Lin *et al.* (2000) are satisfied. By applying the lemma, we have

$$n^{-1/2} \int_{u_{r-1}}^{t} \frac{E_g^A(u, \boldsymbol{\theta}_{g0})}{E_g(u, \boldsymbol{\theta}_{g0})} d\bar{M}_g(u) \to \int_{u_{r-1}}^{t} \frac{\mathbb{E}\{E_g^A(u, \boldsymbol{\theta}_{g0})\}}{\mathbb{E}\{E_g(u, \boldsymbol{\theta}_{g0})\}} d\mathcal{G}_{M_g}(u)$$

uniformly in t almost surely. Since this result holds for every $r = 1, \cdots, R$, we see that

$$n^{-1/2} \int_{0}^{t} \frac{E_g^A(u, \boldsymbol{\theta}_{g0})}{E_g(u, \boldsymbol{\theta}_{g0})} d\bar{M}_g(u) \to \int_{0}^{t} \frac{\mathbb{E}\{E_g^A(u, \boldsymbol{\theta}_{g0})\}}{\mathbb{E}\{E_g(u, \boldsymbol{\theta}_{g0})\}} d\mathcal{G}_{M_g}(u)$$

uniformly in t almost surely. Therefore, we have that $n^{-1/2}U_g^A(t, \boldsymbol{\theta}_{g0})$ uniformly converges to

$$\mathcal{G}_{M_g^A}(t) - \int_{0}^{t} \frac{\mathbb{E}\{E_g^A(u, \boldsymbol{\theta}_{g0})\}}{\mathbb{E}\{E_g(u, \boldsymbol{\theta}_{g0})\}} d\mathcal{G}_{M_g}(u)$$

almost surely in the new probability space, and thus weakly converges in the original probability space. Similar arguments apply to establish the weak convergence of

$$U^B(t, \boldsymbol{\Theta}_0) = \sum_{g=1}^{G} \int_0^t \sum_{i:g(\mathcal{Z}_i(u))=g} Y_i(u) \left[\boldsymbol{W}_i^B(u) - \frac{E_g^B(u, \boldsymbol{\theta}_{g0})}{E_g(u, \boldsymbol{\theta}_{g0})} \right] dM_i^g(u).$$

Therefore, we conclude that $n^{-1/2}\boldsymbol{U}(t, \boldsymbol{\Theta}_0)$ weakly converges to a Gaussian process with mean zero and covariance function $\boldsymbol{\Phi}(s, t, \boldsymbol{\Theta}_0) = \mathbb{E}\{\boldsymbol{\Omega}(s, \boldsymbol{\Theta}_0)\boldsymbol{\Omega}(t, \boldsymbol{\Theta}_0)'\}$, for $0 \leq s \leq t$, where

$$\begin{aligned}
\boldsymbol{\Omega}(t, \boldsymbol{\Theta}) &= (\Omega_1^A(t, \boldsymbol{\theta}_1)', \cdots, \Omega_G^A(t, \boldsymbol{\theta}_g)', \Omega^B(t, \boldsymbol{\Theta})')', \\
\Omega_g^A(t, \boldsymbol{\theta}) &= \int_0^t Y(u) I(g(\mathcal{Z}(u)) = g) \left[\boldsymbol{W}^A(u) - \frac{\mathbb{E}\{E_g^A(u, \boldsymbol{\theta})\}}{\mathbb{E}\{E_g(u, \boldsymbol{\theta})\}} \right] dM(u, \boldsymbol{\theta}), \text{ and} \\
\Omega^B(t, \boldsymbol{\Theta}) &= \sum_{g=1}^{G} \int_0^t Y(u) I(g(\mathcal{Z}(u)) = g) \left[\boldsymbol{W}^B(u) - \frac{\mathbb{E}\{E_g^B(u, \boldsymbol{\theta})\}}{\mathbb{E}\{E_g(u, \boldsymbol{\theta})\}} \right] dM(u, \boldsymbol{\theta}_g).
\end{aligned}$$

3.4.2 Consistency of Regression Parameter Estimator

To show the strong consistency of $\hat{\boldsymbol{\Theta}}$, we extend the presentation from Appendix A.1 of Lin *et al.* (2000). By again letting $\boldsymbol{\theta}_{g0} = (\boldsymbol{\alpha}_{g0}', \beta_0')'$ denote the true value $\boldsymbol{\theta}_g$ under (3.2), define $\pi_n(\boldsymbol{\Theta})$ as

$$\frac{1}{n}\sum_{g=1}^{G}\left(\int_0^\infty \sum_{i:g(\mathcal{Z}_i(t))=g} Y_i(t)(\boldsymbol{\theta}_g - \boldsymbol{\theta}_{g0})'\boldsymbol{W}_i(t)dN_i(t) - \int_0^\infty \sum_{i:g(\mathcal{Z}_i(t))=g} Y_i(t)\log\left(\frac{E_g(t,\boldsymbol{\theta}_g)}{E_g(t,\boldsymbol{\theta}_{g0})}\right)dN_i(t)\right).$$

Since $E_g(t,\boldsymbol{\theta}_g)$ has bounded variation by regularity condition (c), then by the strong law of large numbers,

$$\pi_n(\boldsymbol{\Theta}) \xrightarrow[n\to\infty]{a.s.} \Pi(\boldsymbol{\Theta}) \equiv \sum_{g=1}^{G}\mathbb{E}\left\{\int_0^\infty Y(t)(\boldsymbol{\theta}_g - \boldsymbol{\theta}_{g0})'\boldsymbol{W}(t)dN(t) - \int_0^\infty \log\left(\frac{\mathbb{E}\{E_g(t,\boldsymbol{\theta}_g)\}}{\mathbb{E}\{E_g(t,\boldsymbol{\theta}_{g0})\}}\right)dN(t)\right\}$$

for every $\boldsymbol{\Theta}$. By examining $\partial^2\pi_n(\boldsymbol{\Theta})/\partial\boldsymbol{\Theta}^2$ with components

$$\frac{\partial^2\pi_n(\boldsymbol{\Theta})}{\partial\boldsymbol{\alpha}_g^2} = -\frac{1}{n}\left(\int_0^\infty \sum_{i:g(\mathcal{Z}_i(t))=g} Y_i(t)\left[\frac{E_g^{AA}(t,\boldsymbol{\theta}_g)}{E_g(t,\boldsymbol{\theta}_g)} - \frac{E_g^{A}(t,\boldsymbol{\theta}_g)}{E_g(t,\boldsymbol{\theta}_g)}\left\{\frac{E_g^{A}(t,\boldsymbol{\theta}_g)}{E_g(t,\boldsymbol{\theta}_g)}\right\}'\right]dN_i(t)\right),$$

$$\frac{\partial^2\pi_n(\boldsymbol{\Theta})}{\partial\boldsymbol{\beta}^2} = -\sum_{g=1}^{G}\frac{1}{n}\left(\int_0^\infty \sum_{i:g(\mathcal{Z}_i(t))=g} Y_i(t)\left[\frac{E_g^{BB}(t,\boldsymbol{\theta}_g)}{E_g(t,\boldsymbol{\theta}_g)} - \frac{E_g^{B}(t,\boldsymbol{\theta}_g)}{E_g(t,\boldsymbol{\theta}_g)}\left\{\frac{E_g^{B}(t,\boldsymbol{\theta}_g)}{E_g(t,\boldsymbol{\theta}_g)}\right\}'\right]dN_i(t)\right),$$

$$\frac{\partial^2\pi_n(\boldsymbol{\Theta})}{\partial\boldsymbol{\alpha}_g\partial\boldsymbol{\alpha}_h} = \mathbf{0}, \quad \text{for } h \neq g,$$

$$\frac{\partial^2\pi_n(\boldsymbol{\Theta})}{\partial\boldsymbol{\alpha}_g\partial\boldsymbol{\beta}} = -\frac{1}{n}\left(\int_0^\infty \sum_{i:g(\mathcal{Z}_i(t))=g} Y_i(t)\left[\frac{E_g^{AB}(t,\boldsymbol{\theta}_g)}{E_g(t,\boldsymbol{\theta}_g)} - \frac{E_g^{A}(t,\boldsymbol{\theta}_g)}{E_g(t,\boldsymbol{\theta}_g)}\left\{\frac{E_g^{B}(t,\boldsymbol{\theta}_g)}{E_g(t,\boldsymbol{\theta}_g)}\right\}'\right]dN_i(t)\right),$$

$$\frac{\partial^2\pi_n(\boldsymbol{\Theta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\alpha}_g} = \left(\frac{\partial^2\pi_n(\boldsymbol{\Theta})}{\partial\boldsymbol{\alpha}_g\partial\boldsymbol{\beta}}\right)',$$

we can show that $\partial^2\pi_n(\boldsymbol{\Theta})/\partial\boldsymbol{\Theta}^2$ is negative semi-definite, and thus, $\pi_n(\boldsymbol{\Theta})$ is concave. This implies the uniform convergence of $\pi_n(\boldsymbol{\Theta})$ to $\Pi(\boldsymbol{\Theta})$ on any compact set of $\boldsymbol{\Theta}$ (Theorem 10.8 of Rockafeller (1970)). In particular, for $\mathcal{T}_\varepsilon = \{\boldsymbol{\Theta} : |\boldsymbol{\Theta} - \boldsymbol{\Theta}_0| \leq \varepsilon\}$ for $\varepsilon > 0$, we have

$$\sup_{\boldsymbol{\Theta}\in\mathcal{T}_\varepsilon} \|\pi_n(\boldsymbol{\Theta}) - \Pi(\boldsymbol{\Theta})\| \xrightarrow[n\to\infty]{a.s.} 0. \tag{3.9}$$

In total, we have under (3.2) that (i) $\Pi(\boldsymbol{\Theta})$ is a concave function, (ii) $\frac{\partial}{\partial\boldsymbol{\Theta}}\Pi(\boldsymbol{\Theta}_0) = \mathbf{0}$, and (iii) $\partial^2\Pi(\boldsymbol{\Theta}_0)/\partial\boldsymbol{\Theta}^2 = -\boldsymbol{\Psi}(\boldsymbol{\Theta}_0)$, where $\boldsymbol{\Psi}(\boldsymbol{\Theta})$ is as defined in regularity condition (d). Since $\boldsymbol{\Psi}(\boldsymbol{\Theta}_0)$ is assumed to be positive definite, $\boldsymbol{\Theta}_0$ must then be the unique maximum of $\Pi(\boldsymbol{\Theta})$. Using the result from (3.9), this implies that $\pi_n(\boldsymbol{\Theta}) \leq \pi_n(\boldsymbol{\Theta}_0)$, for all $\boldsymbol{\Theta} \in \mathcal{T}_\varepsilon$ and $n >> 1$. Therefore, there must be a maximizer of $\pi_n(\boldsymbol{\Theta})$, and a solution to $\partial\pi_n(\boldsymbol{\Theta})/\partial\boldsymbol{\Theta} = \mathbf{0}$. Letting $\hat{\boldsymbol{\Theta}}$ denote the solution to this equation, it must lie in the interior of $\mathcal{T}_\varepsilon$. Due to

regularity condition (d), we can apply the arguments of Jacobsen (1989) to show the (global) uniqueness of $\hat{\boldsymbol{\Theta}}$. Since ε can be chosen arbitrarily small, we conclude that $\hat{\boldsymbol{\Theta}}$ must converge to $\boldsymbol{\Theta}_0$ almost surely.

3.4.3 Asymptotic Normality of Regression Parameter Estimator

Taking the first-order Taylor expansion of $\boldsymbol{U}(\hat{\boldsymbol{\Theta}})$ around $\boldsymbol{\Theta}_0$ yields

$$\boldsymbol{U}(\hat{\boldsymbol{\Theta}}) = \boldsymbol{U}(\boldsymbol{\Theta}_0) + \frac{\partial}{\partial \boldsymbol{\Theta}} \boldsymbol{U}(\tilde{\boldsymbol{\Theta}}) \times (\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0) = \mathbf{0},$$

where $\tilde{\boldsymbol{\Theta}}$ is on the line segment between $\hat{\boldsymbol{\Theta}}$ and $\boldsymbol{\Theta}_0$. Thus $\sqrt{n}(\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0) = \hat{\boldsymbol{\Psi}}^{-1}(\tilde{\boldsymbol{\Theta}}) \frac{1}{\sqrt{n}} \boldsymbol{U}(\boldsymbol{\Theta}_0)$. Specifically, we see that $\hat{\boldsymbol{\Psi}}(\boldsymbol{\Theta}) = -\frac{1}{n} \frac{\partial}{\partial \boldsymbol{\Theta}} \boldsymbol{U}(\boldsymbol{\Theta})$ with components

$$\begin{aligned}
\frac{\partial}{\partial \boldsymbol{\alpha}_g} U_g^A(\boldsymbol{\theta}_g) &= \frac{\partial}{\partial \boldsymbol{\alpha}_g} \left(\int_0^\infty \sum_{i:g(\mathcal{Z}_i(t))=g} Y_i(t) \left[\boldsymbol{W}_i^A(t) - \frac{E_g^A(t, \boldsymbol{\theta}_g)}{E_g(t, \boldsymbol{\theta}_g)} \right] dN_i(t) \right) \\
&= -\int_0^\infty \sum_{i:g(\mathcal{Z}_i(t))=g} Y_i(t) \left[\frac{E_g^{AA}(t, \boldsymbol{\theta}_g)}{E_g(t, \boldsymbol{\theta}_g)} - \frac{E_g^A(t, \boldsymbol{\theta}_g)}{E_g(t, \boldsymbol{\theta}_g)} \left\{ \frac{E_g^A(t, \boldsymbol{\theta}_g)}{E_g(t, \boldsymbol{\theta}_g)} \right\}' \right] dN_i(t), \\
\frac{\partial}{\partial \boldsymbol{\beta}} U^B(\boldsymbol{\Theta}) &= \frac{\partial}{\partial \boldsymbol{\beta}} \left(\sum_{g=1}^G \int_0^\infty \sum_{i:g(\mathcal{Z}_i(t))=g} Y_i(t) \left[\boldsymbol{W}_i^B(t) - \frac{E_g^B(t, \boldsymbol{\theta}_g)}{E_g(t, \boldsymbol{\theta}_g)} \right] dN_i(t) \right) \\
&= -\sum_{g=1}^G \int_0^\infty \sum_{i:g(\mathcal{Z}_i(t))=g} Y_i(t) \left[\frac{E_g^{BB}(t, \boldsymbol{\theta}_g)}{E_g(t, \boldsymbol{\theta}_g)} - \frac{E_g^B(t, \boldsymbol{\theta}_g)}{E_g(t, \boldsymbol{\theta}_g)} \left\{ \frac{E_g^B(t, \boldsymbol{\theta}_g)}{E_g(t, \boldsymbol{\theta}_g)} \right\}' \right] dN_i(t), \\
\frac{\partial}{\partial \boldsymbol{\beta}} U_g^A(\boldsymbol{\theta}_g) &= \frac{\partial}{\partial \boldsymbol{\beta}} \left(\int_0^\infty \sum_{i:g(\mathcal{Z}_i(t))=g} Y_i(t) \left[\boldsymbol{W}_i^A(t) - \frac{E_g^A(t, \boldsymbol{\theta}_g)}{E_g(t, \boldsymbol{\theta}_g)} \right] dN_i(t) \right) \\
&= -\int_0^\infty \sum_{i:g(\mathcal{Z}_i(t))=g} Y_i(t) \left[\frac{E_g^{AB}(t, \boldsymbol{\theta}_g)}{E_g(t, \boldsymbol{\theta}_g)} - \frac{E_g^A(t, \boldsymbol{\theta}_g)}{E_g(t, \boldsymbol{\theta}_g)} \left\{ \frac{E_g^B(t, \boldsymbol{\theta}_g)}{E_g(t, \boldsymbol{\theta}_g)} \right\}' \right] dN_i(t), \\
\frac{\partial}{\partial \boldsymbol{\alpha}_g} U^B(\boldsymbol{\Theta}) &= \frac{\partial}{\partial \boldsymbol{\alpha}_g} \left(\sum_{h=1}^G \int_0^\infty \sum_{i:g(\mathcal{Z}_i(t))=h} Y_i(t) \left[\boldsymbol{W}_i^B(t) - \frac{E_h^B(t, \boldsymbol{\theta}_h)}{E_h(t, \boldsymbol{\theta}_h)} \right] dN_i(t) \right) \\
&= -\int_0^\infty \sum_{i:g(\mathcal{Z}_i(t))=g} Y_i(t) \left[\frac{E_g^{BA}(t, \boldsymbol{\theta}_g)}{E_g(t, \boldsymbol{\theta}_g)} - \frac{E_g^B(t, \boldsymbol{\theta}_g)}{E_g(t, \boldsymbol{\theta}_g)} \left\{ \frac{E_g^A(t, \boldsymbol{\theta}_g)}{E_g(t, \boldsymbol{\theta}_g)} \right\}' \right] dN_i(t) \\
&= \left(\frac{\partial}{\partial \boldsymbol{\beta}} U_g^A(\boldsymbol{\theta}_g) \right)'.
\end{aligned}$$

Recall that $\hat{\boldsymbol{\Theta}}$ and $\hat{\boldsymbol{\Psi}}(\boldsymbol{\Theta}_0)$ almost surely converge to $\boldsymbol{\Theta}_0$ and $\boldsymbol{\Psi}(\boldsymbol{\Theta}_0)$, respectively. That implies the almost sure convergence of $\hat{\boldsymbol{\Psi}}(\hat{\boldsymbol{\Theta}})$ to $\boldsymbol{\Psi}(\boldsymbol{\Theta}_0)$ due to the continuity of $\boldsymbol{\Psi}(\cdot)$. Furthermore, the result from Section 3.4.1 implies that $n^{-1/2}\boldsymbol{U}(\boldsymbol{\Theta}_0)$ weakly converges to a Gaussian process with mean zero and covariance function $\boldsymbol{\Phi}(\boldsymbol{\Theta}_0) \equiv \boldsymbol{\Phi}(\infty, \infty, \boldsymbol{\Theta}_0)$. Then,

by Slutsky's Theorem, $\sqrt{n}(\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0)$ converges in distribution to a normal random vector with mean $\mathbf{0}$ and variance $\boldsymbol{AV}(\boldsymbol{\Theta}_0) = \boldsymbol{\Psi}^{-1}(\boldsymbol{\Theta}_0)\boldsymbol{\Phi}(\boldsymbol{\Theta}_0)\boldsymbol{\Psi}^{-1}(\boldsymbol{\Theta}_0)$.

A natural estimator for $\boldsymbol{AV}(\boldsymbol{\Theta}_0)$ is

$$\widehat{\boldsymbol{AV}}(\hat{\boldsymbol{\Theta}}) = \hat{\boldsymbol{\Psi}}^{-1}(\hat{\boldsymbol{\Theta}})\hat{\boldsymbol{\Phi}}(\hat{\boldsymbol{\Theta}})\hat{\boldsymbol{\Psi}}^{-1}(\hat{\boldsymbol{\Theta}}),$$

where

$$\begin{aligned}
\hat{\boldsymbol{\Phi}}(\boldsymbol{\Theta}) &= \frac{1}{n}\sum_{i=1}^{n} \hat{\boldsymbol{\Omega}}_i(\hat{\boldsymbol{\Theta}})\hat{\boldsymbol{\Omega}}_i(\hat{\boldsymbol{\Theta}})', \\
\hat{\boldsymbol{\Omega}}_i(\boldsymbol{\Theta}) &= (\hat{\Omega}_{i1}^A(\boldsymbol{\theta}_1)', \cdots, \hat{\Omega}_{iG}^A(\boldsymbol{\theta}_G)', \hat{\Omega}_i^B(\boldsymbol{\Theta})')', \\
\hat{\Omega}_{ig}^A(\boldsymbol{\theta}) &= \int_0^\infty Y_i(t) I(g(\mathcal{Z}_i(t)) = g) \left[\boldsymbol{W}_i^A(t) - \frac{E_g^A(t, \boldsymbol{\theta})}{E_g(t, \boldsymbol{\theta})} \right] d\hat{M}_{ig}(t, \boldsymbol{\theta}), \text{ and} \\
\hat{\Omega}_i^B(\boldsymbol{\Theta}) &= \sum_{g=1}^{G} \int_0^\infty Y_i(t) I(g(\mathcal{Z}_i(t)) = g) \left[\boldsymbol{W}_i^B(t) - \frac{E_g^B(t, \boldsymbol{\theta}_g)}{E_g(t, \boldsymbol{\theta}_g)} \right] d\hat{M}_{ig}(t, \boldsymbol{\theta}_g)
\end{aligned}$$

with $\hat{M}_{ig}(t, \boldsymbol{\theta})$ equal to

$$N_i(t) I(g(\mathcal{Z}_i(t)) = g) - \int_0^t Y_i(u) I(g(\mathcal{Z}_i(u)) = g) \exp\{\boldsymbol{\theta}' \boldsymbol{W}_i(u)\} d\hat{\Lambda}_{0g}(u).$$

We show in Section 3.4.4 that $d\hat{\Lambda}_{0g}(\cdot)$ converges almost surely to $d\Lambda_{0g}(\cdot)$, so that $\hat{\boldsymbol{\Phi}}(\hat{\boldsymbol{\Theta}})$ converges almost surely to $\boldsymbol{\Phi}(\boldsymbol{\Theta}_0)$. Therefore, since $\hat{\boldsymbol{\Psi}}(\boldsymbol{\Theta}_0)$ almost surely converges to $\boldsymbol{\Psi}(\boldsymbol{\Theta}_0)$ by the uniform strong law of large numbers (Pollard 1990), and $\hat{\boldsymbol{\Theta}}$ almost surely converges to $\boldsymbol{\Theta}_0$ as shown in Section 3.4.2, we conclude that $\widehat{\boldsymbol{AV}}(\hat{\boldsymbol{\Theta}})$ converges almost surely to $\boldsymbol{AV}(\boldsymbol{\Theta}_0)$.

3.4.4 Consistency of Baseline Hazard Function Estimator

For a fixed $g \in \{1, \cdots, G\}$, we have by the uniform strong law of large numbers (Pollard 1990) that

$$\begin{aligned}
n^{-1} \sum_{i:g(\mathcal{Z}_i(t))=g} Y_i(t) dN_i(t) &\xrightarrow[n\to\infty]{} \mathbb{E}\{Y(t) dN(t)\} \\
n^{-1} \sum_{j:g(\mathcal{Z}_j(t))=g} Y_j(t) \exp\{\boldsymbol{\theta}_g' \boldsymbol{W}_j(t)\} &\xrightarrow[n\to\infty]{} \mathbb{E}\{Y(t) \exp\{\boldsymbol{\theta}_g' \boldsymbol{W}(t)\}\}
\end{aligned}$$

almost surely. This entails that as $n \to \infty$, $d\hat{\Lambda}_{0g}(t, \boldsymbol{\theta}_g)$ uniformly converges to

$$\frac{\mathbb{E}\{E_g(t, \boldsymbol{\theta}_{g0})\}}{\mathbb{E}\{E_g(t, \boldsymbol{\theta}_g)\}} d\Lambda_{0g}(t)$$

under (3.2). Since $\hat{\boldsymbol{\theta}}_g$ converges almost surely to $\boldsymbol{\theta}_{g0}$, this hence implies that $d\hat{\Lambda}_{0g}(t;\hat{\boldsymbol{\theta}}_g)$ converges almost surely to $d\Lambda_{0g}(t)$ uniformly in t.

3.4.5 Weak Convergence of Baseline Hazard Function Estimator

We start by establishing the weak convergence of $\sqrt{n}(\hat{\Lambda}_{0g}(t;\hat{\boldsymbol{\theta}}_g)-\Lambda_{0g}(t))$, where $\hat{\Lambda}_{0g}(t;\boldsymbol{\theta}) = \int_0^t d\hat{\Lambda}_{0g}(u;\boldsymbol{\theta})$. Let $\boldsymbol{\theta}_{g0} = (\boldsymbol{\alpha}'_{g0}, \beta'_0)'$ denote the true value $\boldsymbol{\theta}_g$ under (3.2), and

$$\begin{aligned}\boldsymbol{\Lambda}_0(t) &= (\Lambda_{01}(t), \cdots, \Lambda_{0G}(t))' \\ \hat{\boldsymbol{\Lambda}}_0(t,\boldsymbol{\Theta}) &= (\hat{\Lambda}_{01}(t;\boldsymbol{\theta}_1), \cdots, \hat{\Lambda}_{0G}(t,\boldsymbol{\theta}_g))'.\end{aligned}$$

Consider the gth element of $\sqrt{n}(\hat{\boldsymbol{\Lambda}}_0(t;\hat{\boldsymbol{\Theta}}) - \boldsymbol{\Lambda}_0(t))$, which can be expressed as

$$\begin{aligned}&\sqrt{n}(\hat{\Lambda}_{0g}(t;\hat{\boldsymbol{\theta}}_g) - \Lambda_{0g}(t)) \\ &= \sqrt{n}\left(\int_0^t \sum_{i:g(\mathcal{Z}_i(u))=g} \frac{Y_i(u)dN_i(u)}{E_g(u,\hat{\boldsymbol{\theta}}_g)} - \Lambda_{0g}(t)\right) \\ &= \sqrt{n}\left(\int_0^t \sum_{i:g(\mathcal{Z}_i(u))=g} \frac{Y_i(u)dN_i(u)}{E_g(u,\boldsymbol{\theta}_{g0})} - \Lambda_{0g}(t)\right) \\ &\quad + \sqrt{n}\left(\int_0^t \sum_{i:g(\mathcal{Z}_i(u))=g} \frac{Y_i(u)dN_i(u)}{E_g(u,\hat{\boldsymbol{\theta}}_g)} - \int_0^t \sum_{i:g(\mathcal{Z}_i(u))=g} \frac{Y_i(u)dN_i(u)}{E_g(u,\boldsymbol{\theta}_{g0})}\right) \\ &= \sqrt{n}Q_g(t) + \sqrt{n}R_g(t,\hat{\boldsymbol{\theta}}_g),\end{aligned}$$

where

$$\begin{aligned}Q_g(t) &= \left(\int_0^t \sum_{i:g(\mathcal{Z}_i(u))=g} \frac{Y_i(u)dN_i(u)}{E_g(u,\boldsymbol{\theta}_{g0})} - \Lambda_{0g}(t)\right) \\ R_g(t,\boldsymbol{\theta}) &= \left(\int_0^t \sum_{i:g(\mathcal{Z}_i(u))=g} \frac{Y_i(u)dN_i(u)}{E_g(u,\boldsymbol{\theta})} - \int_0^t \sum_{i:g(\mathcal{Z}_i(u))=g} \frac{Y_i(u)dN_i(u)}{E_g(u,\boldsymbol{\theta}_0)}\right).\end{aligned}$$

It can be shown for τ as specified in regularity condition (a) and $t \le \tau$,

$$Q_g(t) = n^{-1/2}\int_0^t \sum_{i:g(\mathcal{Z}_i(u))=g} \frac{Y_i(u)dM_{ig}(u)}{E_g(u,\boldsymbol{\theta}_{g0})},$$

where

$$M_{ig}(t) = N_i(t) - \int_0^t Y_i(u)\exp\{\boldsymbol{\theta}'_{g0}\boldsymbol{W}_i(u)\}d\Lambda_{0g}(u).$$

By the arguments presented in Section 3.4.1, $Q_g(t)$ is tight and uniformly converges to

$$Q_{g0}(t) = n^{-1/2} \int_0^t \sum_{i:g(\mathcal{Z}_i(u))=g} \frac{Y_i(u) dM_{ig}(u)}{\mathbb{E}\{E_g(u, \boldsymbol{\theta}_{g0})\}} + o_p(1).$$

Now consider $\boldsymbol{R}(t, \hat{\boldsymbol{\Theta}})$, where

$$\boldsymbol{R}(t, \boldsymbol{\Theta}) = (R_1(t, \boldsymbol{\theta}_1), \cdots, R_G(t, \boldsymbol{\theta}_g))'.$$

By taking a (first-order) Taylor expansion of $\boldsymbol{R}(t, \hat{\boldsymbol{\Theta}})$ around $\boldsymbol{\Theta}_0$, we have

$$\begin{aligned}
\boldsymbol{R}(t, \hat{\boldsymbol{\Theta}}) &= -\boldsymbol{\xi}(t, \boldsymbol{\Theta}^\dagger)' \sqrt{n}(\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0), \\
\boldsymbol{\xi}(t, \boldsymbol{\theta}) &= \begin{bmatrix}
\frac{\partial}{\partial \boldsymbol{\alpha}_1} R_1(t, \boldsymbol{\theta}_1) & \mathbf{0} & \cdots & \mathbf{0} \\
\mathbf{0} & \frac{\partial}{\partial \boldsymbol{\alpha}_2} R_2(t, \boldsymbol{\theta}_2) & \cdots & \mathbf{0} \\
\vdots & \vdots & \ddots & \vdots \\
\mathbf{0} & \mathbf{0} & \cdots & \frac{\partial}{\partial \boldsymbol{\alpha}_G} R_G(t, \boldsymbol{\theta}_g) \\
\frac{\partial}{\partial \boldsymbol{\beta}} R_1(t, \boldsymbol{\theta}_1) & \frac{\partial}{\partial \boldsymbol{\beta}} R_2(t, \boldsymbol{\theta}_2) & \cdots & \frac{\partial}{\partial \boldsymbol{\beta}} R_G(t, \boldsymbol{\theta}_g)
\end{bmatrix} \\
\frac{\partial}{\partial \boldsymbol{\alpha}_g} R_g(t, \boldsymbol{\theta}) &= \int_0^t \sum_{i:g(\mathcal{Z}_i(u))=g} \frac{E_g^A(u, \boldsymbol{\theta})}{(E_g(u, \boldsymbol{\theta}))^2} Y_i(u) dN_i(u), \\
\frac{\partial}{\partial \boldsymbol{\beta}} R_g(t, \boldsymbol{\theta}) &= \int_0^t \sum_{i:g(\mathcal{Z}_i(u))=g} \frac{E_g^B(u, \boldsymbol{\theta})}{(E_g(u, \boldsymbol{\theta}))^2} Y_i(u) dN_i(u),
\end{aligned}$$

where $\boldsymbol{\Theta}^\dagger$ is on the line segment between $\hat{\boldsymbol{\Theta}}$ and $\boldsymbol{\Theta}_0$. By considering a partition of the study period into finite intervals as in Section 3.4.1, we can apply the uniform strong law of large numbers (Pollard (1990) page 41) to show that $\boldsymbol{\xi}(t, \boldsymbol{\Theta}_0)$ converges almost surely to $\boldsymbol{\Xi}(t, \boldsymbol{\Theta}_0)$ uniformly on t, where

$$\begin{aligned}
\boldsymbol{\Xi}(t, \boldsymbol{\theta}) &= \begin{bmatrix}
S_1^A(t, \boldsymbol{\theta}_1) & \mathbf{0} & \cdots & \mathbf{0} \\
\mathbf{0} & S_2^A(t, \boldsymbol{\theta}_2) & \cdots & \mathbf{0} \\
\vdots & \vdots & \ddots & \vdots \\
\mathbf{0} & \mathbf{0} & \cdots & S_G^A(t, \boldsymbol{\theta}_g) \\
S_1^B(t, \boldsymbol{\theta}_1) & S_2^B(t, \boldsymbol{\theta}_2) & \cdots & S_G^B(t, \boldsymbol{\theta}_g)
\end{bmatrix} \\
S_g^A(t, \boldsymbol{\theta}) &= \int_0^t \frac{\mathbb{E}\{E_g^A(u, \boldsymbol{\theta})\}}{\mathbb{E}\{E_g(u, \boldsymbol{\theta})\}} d\Lambda_{0g}(u), \\
S_g^B(t, \boldsymbol{\theta}) &= \int_0^t \frac{\mathbb{E}\{E_g^B(u, \boldsymbol{\theta})\}}{\mathbb{E}\{E_g(u, \boldsymbol{\theta})\}} d\Lambda_{0g}(u).
\end{aligned}$$

Furthermore, we established in Section 3.4.3 that

$$\begin{aligned}
\sqrt{n}(\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0) &= \boldsymbol{\Psi}^{-1}(\boldsymbol{\Theta}_0)\sqrt{n}\sum_{i=1}^{n}\boldsymbol{\Omega}_i(\infty, \boldsymbol{\Theta}) + o_p(1), \\
\boldsymbol{\Omega}_i(t, \boldsymbol{\Theta}) &= (\Omega_{i1}^A(t, \boldsymbol{\theta}_1)', \cdots, \Omega_{iG}^A(t, \boldsymbol{\theta}_G)', \Omega_i^B(t, \boldsymbol{\Theta})')', \\
\Omega_{ig}^A(t, \boldsymbol{\theta}) &= \int_0^t Y_i(u) I(g(\mathcal{Z}_i(u)) = g) \left[\boldsymbol{W}_i^A(u) - \frac{\mathbb{E}\{E_g^A(u, \boldsymbol{\theta})\}}{\mathbb{E}\{E_g(u, \boldsymbol{\theta})\}} \right] dM_i(u), \text{ and} \\
\Omega_i^B(t, \boldsymbol{\Theta}) &= \sum_{g=1}^{G} \int_0^t Y_i(u) I(g(\mathcal{Z}_i(u)) = g) \left[\boldsymbol{W}_i^B(u) - \frac{\mathbb{E}\{E_g^B(u, \boldsymbol{\theta}_g)\}}{\mathbb{E}\{E_g(u, \boldsymbol{\theta}_g)\}} \right] dM_i(u),
\end{aligned}$$

where $\boldsymbol{\Psi}(\boldsymbol{\Theta}_0)$ is as defined in regularity condition (d). This implies that $\boldsymbol{R}(t, \boldsymbol{\Theta}_0)$ is tight, and

$$\boldsymbol{R}(t, \boldsymbol{\Theta}_0) = -\boldsymbol{\Xi}(t, \boldsymbol{\Theta}_0)'\boldsymbol{\Psi}^{-1}(\boldsymbol{\Theta}_0)\sqrt{n}\sum_{i=1}^{n}\boldsymbol{\Omega}_i(t, \boldsymbol{\Theta}_0) + o_p(1).$$

Hence, $\sqrt{n}(\hat{\boldsymbol{\Lambda}}_0(t; \hat{\boldsymbol{\Theta}}) - \boldsymbol{\Lambda}_0(t)) = n^{-1/2}\sum_{i=1}^{n}\boldsymbol{\Delta}_i(t) + o_p(1)$, where

$$\boldsymbol{\Delta}_i(t) = \boldsymbol{K}_i(t) - \boldsymbol{\Xi}(t, \boldsymbol{\Theta}_0)'\boldsymbol{\Psi}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{\Omega}_i(t, \boldsymbol{\Theta}_0) + o_p(1),$$

with the gth element of $\boldsymbol{K}_i(t)$ being

$$n^{-1/2}\int_0^t \frac{Y_i(u) I(g(\mathcal{Z}_i(u)) = g) dM_{ig}(u)}{\mathbb{E}\{E_g(u, \boldsymbol{\theta}_{g0})\}} + o_p(1).$$

That is, $\sqrt{n}(\hat{\boldsymbol{\Lambda}}_0(t; \hat{\boldsymbol{\Theta}}) - \boldsymbol{\Lambda}_0(t))$ converges weakly to a zero-mean Gaussian process with covariance function $\boldsymbol{\zeta}(s, t) = \mathbb{E}\{\boldsymbol{\Delta}(s)\boldsymbol{\Delta}(t)'\}$ for $0 < s \leq t$. A natural estimator for $\boldsymbol{\zeta}(s, t)$ is

$$\hat{\boldsymbol{\zeta}}(s, t) = \frac{1}{n}\sum_{i=1}^{n}\hat{\boldsymbol{\Delta}}_i(s, \hat{\boldsymbol{\Theta}})\hat{\boldsymbol{\Delta}}_i(t, \hat{\boldsymbol{\Theta}})',$$

where

$$\begin{aligned}
\hat{\boldsymbol{\Delta}}_i(t,\boldsymbol{\Theta}) &= \hat{\boldsymbol{K}}_i(t,\boldsymbol{\Theta}) - \boldsymbol{\xi}(t,\boldsymbol{\Theta})'\hat{\boldsymbol{\Psi}}^{-1}(\boldsymbol{\Theta})\hat{\boldsymbol{\Omega}}_i(t,\boldsymbol{\Theta}),\\
\hat{\boldsymbol{K}}_i(t,\boldsymbol{\Theta}) &= (\hat{K}_{i1}(t,\boldsymbol{\theta}_1),\cdots,\hat{K}_{iG}(t,\boldsymbol{\theta}_g))'\\
\hat{K}_{ig}(t,\boldsymbol{\theta}) &= \int_0^t \frac{Y_i(u)I(g(\mathcal{Z}_i(u))=g)d\hat{M}_{ig}(u,\boldsymbol{\theta})}{E_g(u,\boldsymbol{\theta})}\\
\hat{M}_{ig}(t,\boldsymbol{\theta}) &= N_i(t)I(g(\mathcal{Z}_i(t))=g) - \int_0^t Y_i(u)I(g(\mathcal{Z}_i(u))=g)\exp\{\boldsymbol{\theta}'\boldsymbol{W}_i(u)\}d\hat{\Lambda}_{0g}(u;\hat{\boldsymbol{\theta}})\\
\hat{\boldsymbol{\Psi}}(\boldsymbol{\Theta}) &= -\frac{1}{n}\frac{\partial}{\partial\boldsymbol{\Theta}}\boldsymbol{U}(\boldsymbol{\Theta}),\\
\hat{\boldsymbol{\Omega}}_i(t,\boldsymbol{\Theta}) &= (\hat{\Omega}^A_{i1}(t,\boldsymbol{\theta}_1)',\cdots,\hat{\Omega}^A_{iG}(t,\boldsymbol{\theta}_g)',\hat{\Omega}^B_i(t,\boldsymbol{\Theta})')',\\
\hat{\Omega}^A_{ig}(t,\boldsymbol{\theta}) &= \int_0^t Y_i(u)I(g(\mathcal{Z}_i(u))=g)\left[\boldsymbol{W}^A_i(u) - \frac{E^A_g(u,\boldsymbol{\theta})}{E_g(u,\boldsymbol{\theta})}\right]d\hat{M}_{ig}(u,\boldsymbol{\theta}),\ \text{and}\\
\hat{\Omega}^B_i(t,\boldsymbol{\Theta}) &= \sum_{g=1}^G\int_0^t Y_i(u)I(g(\mathcal{Z}_i(u))=g)\left[\boldsymbol{W}^B_i(u) - \frac{E^B_g(u,\boldsymbol{\theta}_g)}{E_g(u,\boldsymbol{\theta}_g)}\right]d\hat{M}_{ig}(u,\boldsymbol{\theta}_g).
\end{aligned}$$

Based on the asymptotic normality of $\hat{\Lambda}_{0g}(\cdot)$ and consistency of $\hat{\boldsymbol{\zeta}}(\cdot,\cdot)$, we can construct pointwise confidence intervals for $\Lambda_{0g}(t)$, for each $g=1,\cdots,G$. Since $\Lambda_{0g}(t)$ is non-negative, we consider the transformed random variable $\sqrt{n}(\log\hat{\Lambda}_{0g}(t;\hat{\boldsymbol{\theta}}) - \log\Lambda_{0g}(t))$, so that an approximate $1-\alpha^*$ pointwise confidence interval for $\Lambda_{0g}(t)$ is

$$\hat{\Lambda}_{0g}(t;\hat{\boldsymbol{\theta}}_g)\exp\left(\pm n^{-1/2}\phi_{\alpha^*/2}\frac{\sqrt{\hat{\zeta}_{(g,g)}(t,t)}}{\hat{\Lambda}_{0g}(t;\hat{\boldsymbol{\theta}}_g)}\right),$$

where $\phi_{\alpha^*/2}$ is the $(1-\alpha^*)$ critical point of the standard normal distribution, and $\hat{\zeta}_{(g,g)}(t,t)$ is the (g,g) element of $\hat{\boldsymbol{\zeta}}(t,t)$.

To construct simultaneous confidence bands for $\Lambda_{0g}(t)$ over a time interval of interest for $[t_1,t_2]$ $(0<t_1<t_2)$, we need to approximate the distribution of the supremum of $\sqrt{n}(\hat{\Lambda}_{0g}(t;\hat{\boldsymbol{\theta}}_g)-\Lambda_{0g}(t))$ for $t\in[t_1,t_2]$. As this distribution is challenging to evaluate analytically, we can approximate the distribution of $\sqrt{n}(\hat{\Lambda}_{0g}(t;\hat{\boldsymbol{\theta}}_g)-\Lambda_0(t))$ with $n^{-1/2}\sum_{i=1}^n\hat{\Delta}_{ig}(t,\hat{\boldsymbol{\theta}}_g)\phi_i$, where $(\phi_1,\cdots,\phi_n)$ are independent and identically distributed standard normal random variables, and $\hat{\Delta}_{ig}(t,\hat{\boldsymbol{\theta}}_g)$ is the gth element of $\hat{\boldsymbol{\Delta}}_i(t,\hat{\boldsymbol{\Theta}})$. We can then repeatedly generate random variables $(\phi_1,\cdots,\phi_n)$, and determine an approximate value of $\phi^*_{\alpha^*/2}$ which satisfies

$$P\left(\sup_{t_1\le t\le t_2}\left|\frac{n^{-1/2}\sum_{i=1}^n\hat{\Delta}_{ig}(t,\hat{\boldsymbol{\theta}}_g)\phi_i}{\sqrt{\hat{\zeta}_{(g,g)}(t,t)}}\right| < \phi^*_{\alpha^*/2}\right) = 1-\alpha^*.$$

Hence, an approximate $1-\alpha^*$ simultaneous confidence band of $\Lambda_{0g}(t)$ for $t \in [t_1, t_2]$ is

$$\hat{\Lambda}_{0g}(t;\hat{\boldsymbol{\theta}}_g)\exp\left(\pm n^{-1/2}\phi^*_{\alpha^*/2}\frac{\sqrt{\hat{\zeta}_{(g,g)}(t,t)}}{\hat{\Lambda}_{0g}(t;\hat{\boldsymbol{\theta}}_g)}\right).$$

By the multivariate central limit theorem, the process $n^{-1/2}\sum_{i=1}^{n}\hat{\Delta}_{ig}(t,\hat{\boldsymbol{\theta}}_g)\phi_i$ weakly converges to a zero-mean Gaussian process with covariance function $\hat{\zeta}_{(g,g)}(s,t)$. Since $\hat{\zeta}_{(g,g)}(s,t)$ converges to $\zeta_{(g,g)}(s,t)$ almost surely uniformly in s and t, then provided that $n^{-1/2}\sum_{i=1}^{n}\hat{\Delta}_{ig}(t,\hat{\boldsymbol{\theta}}_g)\phi_i$ is tight, it converges to a zero-mean Gaussian process with covariance function $\zeta_{(g,g)}(s,t)$. Thus, the weak convergence of $\sqrt{n}(\hat{\Lambda}_{0g}(t;\hat{\boldsymbol{\theta}}_g)-\Lambda_{0g}(t))$ establishes the weak convergence of $\sqrt{n}(d\hat{\Lambda}_{0g}(t;\hat{\boldsymbol{\theta}}_g)-d\Lambda_{0g}(t))$.

3.5 Analysis of the Provincial OAT Dispensation Records (I)

We started our analysis by fitting the extended Cox regression model (3.4) with the observed data. The same set of covariates from our preliminary analysis in Chapter 2 were included, except we omitted the *OAT dispensation rate* and *incarceration FPC score*, and included the following time-varying covariates: *OAT dispensation indicator* at time t, *incarceration status* at time t, and the *number of (observed) incarcerations* by time t. The parameter estimates are displayed in Table 3.1, in which we again considered both *time since first recorded OAT dispensation* and *age* as time scales. For the variables that are present in both our preliminary and current analysis, we see the estimates are quite similar. However, the effect pertaining to the OAT dispensation process is quite different, where our current analysis implies that individuals dispensed OAT have a significantly lower mortality risk (at time t) relative to those not dispensed OAT. To shed some light on this discrepancy, we note that 31.34% of individuals that died had an OAT dispensation covering their date of death. Since the average (and median) *OAT dispensation rate* reported in Table 2.4 is larger than 50%, we would anticipate the *OAT dispensation indicator* to be negatively associated with mortality risk. On the other hand, since Table 2.4 reveals *OAT dispensation rate* to have a positive association with mortality risk, this highlights individuals modifying their *OAT dispensation indicator* shortly before their date of death; a phenomenon further investigated by Pearce *et al.* (2020).

We proceeded to fit a stratified Cox regression model (3.2), in which we stratified individuals (at time t) according to their OAT episode number. An OAT episode number at time t can loosely be conceived as the number of (long-term) "not dispensed OAT" to "dispensed OAT" transitions by time t, as illustrated in Figure 3.1. We specified $G = 9$ levels for the stratification variable: (i) 1 OAT episode; (ii) 2-3 OAT episodes; (iii) 4-5 OAT episodes; (iv) 6-7 OAT episodes; (v) 8-10 OAT episodes; (vi) 11-15 OAT episodes; (vii) 16-20 OAT

episodes; (viii) 21-30 OAT episodes; and (ix) 31+ OAT episodes. The G levels were selected based on a combination of summary statistics for the number of OAT episodes individuals experienced by their end of follow-up date, as well as expert opinion. The estimates of $\boldsymbol{\Theta}$, in which we started by specifying each $\boldsymbol{\theta}_g = \boldsymbol{\alpha}_g$ when $g(\mathcal{Z}(t)) = g$, are illustrated in Figures 3.2 and 3.3, where the time scales are *time since first recorded OAT dispensation* and *age*, respectively. We can see that the *OAT dispensation indicator*, the *birth generation indicators*, and the *ever on PharmaCare plans C or G* indicator have varying effects across strata, whereas the other effects are static. This motivates us to update our modelling by specifying the variables with a grey background in Figures 3.2 and 3.3 to have a constant effect, $\boldsymbol{\beta}$. This reduces the number of parameters needed to estimate, and thus serves to improve the statistical efficiency. The updated parameter estimates illustrated in Figures 3.4 and 3.5 are upon specifying $\boldsymbol{\theta}_g = (\boldsymbol{\alpha}_g', \boldsymbol{\beta}')'$ when $g(\mathcal{Z}(t)) = g$. We can see that the estimates that vary across strata are quite similar to their corresponding estimates in Figures 3.2 and 3.3, and the estimates with a constant effect are similar to the estimates shown in Table 3.1. We also estimated the baseline hazard functions with (3.5), and illustrate the LOESS-smoothed estimates in the top row of Figure 3.6. Under the *time since first recorded OAT dispensation* time scale, Figure 3.6 shows us that individuals with higher OAT episodes have a higher mortality risk, whereas the estimates in the *age* time scale shows us that older individuals have a higher mortality risk compared to younger individuals, regardless of their OAT episode number. The fact that older individuals generally experienced fewer OAT episodes compared to younger individuals is a byproduct of older individuals having a higher *OAT dispensation rate*, as shown in Table 2.8.

We proceeded to conduct the Wald test, where the results of the eight tests are shown in Tables 3.2 and 3.3. For each test, we displayed the estimates of both $\boldsymbol{\alpha}_{g-1}$ and $\boldsymbol{\alpha}_g$, the test statistic J_g in (3.7), and the resulting p-value. By applying a Bonferonni correction for the multiple testing, we specified the type I error rate to be $\alpha^* = 0.05/(G-1)$. The results of the test under both *time since first recorded OAT dispensation* and *age* time scales reveals that the stratification variable should be updated to the following $H = 2$ levels: (i) 1-3 OAT episodes; and (ii) 4+ OAT episodes. The results with the updated stratification are illustrated in Figures 3.7 and 3.8, in which we see the *OAT dispensation indicator* effect has a higher protective effect against mortality for individuals with more OAT episodes, whereas the opposite is true for the *birth generation indicator* and *ever on PharmaCare plans C or G* effects. The bottom row of Figure 3.6 illustrates the LOESS-smoothed estimates of the baseline hazard function estimates (3.5), in which we see the two baseline hazard functions highly overlap when the time scale is *age*. This finding would give us reason to further improve the efficiency by updating our model to (3.3), where we specify the two baseline hazard functions to be the same.

3.6 Simulation Study

We conducted three simulation studies to examine the finite-sample performance of the proposed estimator and assess the performance of the Wald testing procedure. Specifically, we generated data based on the Cox regression model (3.4) in the first simulation study, generated data based on the stratified Cox regression model in (3.2) in the second simulation study, and assessed the robustness of the Wald test against model misspecification in the third simulation study. We also provide a discussion on the required sample size needed for the test for the Wald test to achieve a certain power.

3.6.1 Data Generation

We simulated a study with $n = 10{,}000$ independent units, where we generated right-censored observations of an event time, and an alternating binary process that affects the event time's hazard. For each study unit, we generated observations as follows:

(i) Generate two baseline covariates, X_1 and X_2, where $X_1 \sim \text{Uniform}(0, 1)$, and $X_2 \sim \text{Bernoulli}(0.5)$.

(ii) Generate a time-varying alternating binary indicator, $Z(t)$. To be consistent with our data application, we specified $Z(0) \equiv 1$, $g(\mathcal{Z}(0)) \equiv 1$, and $g(\mathcal{Z}(t))$ is determined by the number zero-to-one changes up to time t, as a way to mimic the role of OAT episode numbers. We specified $G = 10$, and keep units with more than 10 zero-to-one changes in group G. To generate $Z(t)$, we simulated the time an individual changes their binary indicator status from the exponential distribution, where ρ_{0g} and ρ_{1g} are the rates for the time $Z(t)$ transitions to zero and one, respectively, when $g(\mathcal{Z}(t)) = g$. Here, we specified

$$\rho_{0g} = \begin{cases} 20 & \text{if } g \in \{1, 2, 3\} \\ 30 & \text{if } g \in \{4, 5, 6\} \\ 50 & \text{if } g = 7 \\ 10 & \text{if } g \in \{8, 9, 10\} \end{cases} \quad \text{and} \quad \rho_{1g} = \begin{cases} 10 & \text{if } g \in \{1, 2, 3\} \\ 20 & \text{if } g \in \{4, 5, 6\} \\ 50 & \text{if } g = 7 \\ 5 & \text{if } g \in \{8, 9, 10\} \end{cases}.$$

(iii) With $\boldsymbol{W}(t) = (Z(t), X_1, X_2)'$ as the vector of covariates at time t, we specify group specific parameter vector $\boldsymbol{\theta}_g$, and baseline hazard λ_g in order to compute the event hazard

$$\lambda^{\dagger}(t; \mathcal{Z}(t), X_1, X_2) = \lambda_{0g} f(\boldsymbol{\theta}_g' \boldsymbol{W}(t)),$$

for some function $f(\cdot) > 0$ which depends on the simulation setting. We discretized the time interval for which $Z(t)$ is constant into subintervals of length $\Delta t = 0.0001$, and simulated the event occurrence at time t as a Bernoulli random variable with

success probability $\lambda^{\dagger}(t; \mathcal{Z}(t), X_1, X_2) \times \Delta t$. This procedure continues until we observe a success.

(iv) Generate censoring times from the exponential distribution with rate $\lambda_C \in \{0.5, 1.5\}$ to produce right-censored event times.

Overall, this data generation procedure produces the following independent observations

$$\{(T_i^*, \delta_i, \mathcal{Z}_i(T_i^*), X_{i1}, X_{i2}) : i = 1, \cdots, n\},$$

with $T_i^* = T_i \wedge C_i$, and $\delta_i = I(T_i \leq C_i)$. With a simulated dataset, we proceeded to estimate $\boldsymbol{\Theta}$ and conduct the Wald test. We replicated the data generation and inference procedure 150 times.

3.6.2 Simulation Outcome 1: Reduction to the Cox Regression Model

We first conducted a simulation study where we generated event times under the Cox regression model with the baseline hazard function specified as a constant over time. The data generation procedure is as presented in Section 3.6.1, but we take $\lambda_{0g} \equiv \lambda_0$, $\boldsymbol{\theta}_g \equiv \boldsymbol{\theta}$, and $f(x) = \exp\{x\}$. That is, we generated event times under the hazard model

$$\lambda^{\dagger}(t; \mathcal{Z}(t), X_1, X_2) = \lambda_0 \exp\{\boldsymbol{\theta}' \boldsymbol{W}(t)\},$$

with $\lambda_0 = 1.75$ and $\boldsymbol{\theta} = (-1, 0, -0.5)'$. Based on the two specifications of λ_C, this resulted in (on average) approximately 29% and 52% of event times being right-censored.

Table 3.4 summarizes the estimation results upon fitting the Cox regression model to the data, where we consider two parameter settings for λ_C: **Censoring Parameter Setting I** specifies $\lambda_C = 0.5$ and **Censoring Parameter Setting II** specifies $\lambda_C = 1.5$. We considered two approaches to estimate $\boldsymbol{\theta}$: (a) solve $\boldsymbol{U}(\boldsymbol{\theta}) = \boldsymbol{0}$ in which we fix $G \equiv 1$, and (ii) implement the `coxph` function in the `survival` (Therneau 2015) R package. As expected, we the estimates of $\boldsymbol{\theta}$ and $Var(\hat{\boldsymbol{\theta}})$ are practically identical across the two procedures. The similarity of the estimated standard errors between the two approaches is likely due to fitting the appropriate data generating model to the data, which is resulting in $\boldsymbol{\Psi}^{-1}(\boldsymbol{\theta}) \approx \boldsymbol{\Phi}(\boldsymbol{\theta})$.

We fit a stratified Cox regression model (3.2) to the simulated data, where we summarized the parameter estimates in Table 3.5. We see that the standard error estimates are larger relative to the standard error estimates under the Cox regression model, due to less information being used to estimate $\boldsymbol{\theta}_g$, for $g = 1, \cdots, G$. Moreover, the largest standard error corresponds to *Group 7*, which is attributed to the relatively large values specified for ρ_{07} and ρ_{17}. We also note that **Censoring Parameter Setting II** results in larger standard error estimates of $\hat{\boldsymbol{\Theta}}$, which is due to the larger censoring rate.

We then proceeded to conduct the Wald test from Section 3.3.3, and assess its performance in correctly recovering the Cox regression model. Table 3.6 presents a matrix for

each of the four parameter specifications, where the (g, g') element is the proportion that groups g and g' are classified to the same class, with $\alpha^* = 0.05$. The matrices inform us that there is approximately a 95% chance of group $g-1$ (correctly) being merged together with group g. To address the multiple testing issue, we propose a Bonferonni correction and specify $\alpha^* = 0.05/(G-1)$. Table 3.7 shows the results of the risk classes identified by the Wald test, with and without a Bonferonni correction. Overall, we see that the test can adequately recover the Cox regression model.

3.6.3 Simulation Outcome 2: Correctly Identifying the Number of Risk Classes in the Stratified Cox Model

We conducted a simulation study where we generated event times under the stratified Cox regression model with constant baseline hazards. That is, we generated event times under the hazard model

$$\lambda^{\dagger}(t; \mathcal{Z}(t), X_1, X_2) = \lambda_{0g} \exp\{\boldsymbol{\theta}_g' \boldsymbol{W}(t)\},$$

so that $f(x) = \exp\{x\}$, and we specified λ_{0g} and $\boldsymbol{\theta}_g$ as

$$\lambda_{0g} = \begin{cases} 0.5 & \text{if } g \in \{1,2,3\} \\ 1.75 & \text{if } g \in \{4,5,6\} \\ 3 & \text{if } g \in \{7,8,9,10\} \end{cases} \quad \text{and} \quad \boldsymbol{\theta}_g = \begin{cases} (-0.5,-2,-2)' & \text{if } g \in \{1,2,3\} \\ (-1,0,-0.5)' & \text{if } g \in \{4,5,6\} \\ (-2,2,1.5)' & \text{if } g \in \{7,8,9,10\} \end{cases}.$$

With the two specifications of λ_C, this resulted in (on average) approximately 38% and 73% of event times being right-censored.

We started by naively fitting the Cox regression model in (3.4), where Table 3.8 presents the estimation results. We can see the resulting estimates resemble a weighted average of the $\boldsymbol{\theta}_g$ parameters across the G groups. Furthermore, we see that our standard error estimates are generally larger relative to the standard error estimates reported by the `coxph` function, which is to be expected.

We proceeded to fit the true data generating model in (3.2) to the simulated data, where the parameter estimates are presented in Table 3.9. We can see the estimates $\hat{\boldsymbol{\Theta}}$ and $\widehat{\boldsymbol{AV}}(\hat{\boldsymbol{\Theta}})$ appear to consistently estimate $\boldsymbol{\Theta}$ and $\boldsymbol{AV}(\boldsymbol{\Theta})$, respectively. Similar to Section 3.6.2, we generally see the standard error estimates under **Censoring Parameter Setting I** to be smaller than the standard error estimates under **Censoring Parameter Setting II** due to the smaller censoring rate.

We then conducted the Wald test to assess its performance in correctly recovering the three data-generating treatment classes. Table 3.10 presents a matrix for each parameter specification, where the (g, g') element is the proportion that groups g and g' are classified to the same class, with $\alpha^* = 0.05$. The matrices inform us that there is approximately a 95% chance that group $g-1$ is merged together with group g if they belong to the same

class. We present the risk classes identified by the Wald test in Table 3.11. By applying a Bonferonni correction, we can see that the test performs adequately in recovering the three risk classes.

3.6.4 Simulation Outcome 3: Robustness to Model Misspecification

We conducted a simulation study where we generated event times under a misspecified stratified Cox regression model. Specifically, we generated event times under the hazard model

$$\lambda^{\dagger}(t; \mathcal{Z}(t), X_1, X_2) = \lambda_{0g} f(\boldsymbol{\theta}_g' \boldsymbol{W}(t)),$$

where $f(x) = (1+x)I(x \geq 0) + (1-x)I(x < 0)$, for $g = 1, \cdots, 10$. Upon specifying $\lambda_C = 0.5$ and $\lambda_C = 1.5$, this resulted in (on average) approximately 29% and 62% of individuals having right-censored death times, respectively. The purpose of this simulation study is to assess if the Wald test is robust to model misspecification.

By estimating the model parameters under the stratified Cox regression model (3.2), Table 3.12 shows that the Wald test is robust to model misspecification upon applying a Bonferonni correction.

3.7 On Sample Size Determination

Our simulation study demonstrates that we can adequately recover the true risk classes with our specification of $\boldsymbol{\Theta}$, where the type I error rate is $\alpha^* = 0.05/(G-1)$, and $n = 10{,}000$. We anticipate a larger sample size is required to correctly identify the correct risk classes if we instead specify difference between stratum-specific effects across successive groups to be "small". We now consider the minimum sample size needed for the Wald test to achieve a certain power, $1-\beta^*$ given α^*, and effect difference $\boldsymbol{\alpha}_g - \boldsymbol{\alpha}_{g-1}$. In other words, consider the following simple hypothesis for a fixed $g \in \{2, \cdots, G\}$:

$$H_0 : \boldsymbol{\alpha}_g - \boldsymbol{\alpha}_{g-1} = \mathbf{0} \quad \text{vs.} \quad H_a : \boldsymbol{\alpha}_g - \boldsymbol{\alpha}_{g-1} = \boldsymbol{\gamma}_g \neq 0. \tag{3.10}$$

Here, $\boldsymbol{\gamma}_g$ is the smallest value in which we view the difference in effects to be meaningful. In order for the Wald test to conduct the hypothesis test in (3.10) adequately, we want to reject H_0 with probability $1-\beta^*$ when H_a is true, and reject H_0 with probability α^* when H_0 is true. Here, β^* is referred to as the type II error rate, and $1-\beta^*$ is referred to as the power of the test. Recall in Section 3.3.3 that under H_0 in (3.10), $J_g \dot{\sim} F_{q_A}(\cdot)$, where $F_{q_A}(\cdot)$ denotes the χ^2-distribution function with q_A degrees of freedom. However, under H_a in (3.10), $J_g \dot{\sim} F_{q_A,\nu_g}(\cdot)$, where $F_{q_A,\nu_g}(\cdot)$ denotes the non-central χ^2-distribution with q_A

degrees of freedom with non-centrality parameter ν_g, with

$$\nu_g = \boldsymbol{\gamma}_g' \left[\frac{\text{Var}(\hat{\boldsymbol{\alpha}}_g - \hat{\boldsymbol{\alpha}}_{g-1})}{n} \right]^{-1} \boldsymbol{\gamma}_g. \tag{3.11}$$

Then for α^* and β^* given, we want to find ν_g such that

$$F_{q_A,\nu_g}(F_{q_A}^{-1}(1-\alpha^*)) = \beta^*.$$

Hence, we can proceed to solve for n in (3.11) as

$$n_g = \nu_g \left[\boldsymbol{\gamma}_g' \text{Var}(\hat{\boldsymbol{\alpha}}_g - \hat{\boldsymbol{\alpha}}_{g-1})^{-1} \boldsymbol{\gamma}_g \right]^{-1}.$$

But since we considered the test (3.10) for a fixed g, we iterate over g to obtain $n_2, \cdots, n_G$, and conclude that the sample size needed to correctly recover the correct group structure with type I error α^* and type II error β^* is

$$n = \max\{n_g : g = 2, \cdots, G\}.$$

3.8 Discussion and Outlook for Chapter 4

We proposed a generalized Cox regression model, under which we conducted time-dependent stratification, where the strata are defined in terms of the history of the OAT dispensation process. To accommodate for internal time-varying covariates, we adopted an estimating-equation based inference procedure that bypasses interpretation challenges brought on by constructing the likelihood function. Large sample properties of the proposed estimators were established, and our simulation study shows that we are able to consistently estimate both the regression parameters and their estimate's standard errors. To determine if the effects between groups are significant, we proposed a Wald test that sequentially tests if successive groups should be merged. Upon applying a Bonferonni correction, we showed through a simulation study that this test can correctly recover the true grouping structure in a satisfactory manner, and is robust to model misspecification. We applied the proposed methodology to provincial OAT dispensation records, in which two risk classes based on an individual's history of OAT use were identified. We summarized an individual's history of the OAT dispensation process with their OAT episode number, which can loosely be conceived as the number of observed long term "not dispensed OAT" to "dispensed OAT" transitions. Other summaries, such as proportion of time dispensed OAT, can seamlessly be used as an alternative stratification variable.

Note that the models specified in this chapter are for the hazard functions of the survival times conditioning on the internal covariate process at time t. Although we are able to estimate the mortality hazard function, the conventional relationship between the hazard

and survivor functions does not hold. The models presented in this chapter cannot provide survival predictions. This differs from the approach taken in Chapter 2, which specifies a model for $\lambda(t; \mathcal{Z}_i(\infty), \boldsymbol{\mathcal{X}}_{(}t))$ by summarizing the entire internal covariate process with a (time-independent) latent variable. This is the key difference between the modelling in Chapters 2 and 3. In fact, as this chapter models $\lambda(t; \mathcal{Z}_i(t), \boldsymbol{\mathcal{X}}_{(}t))$, there is generally no connection between these models with joint modelling.

Thus far, we have presented two different approaches to properly handle internal covariates within hazard regression modelling, where each method has their own set of advantages and disadvantages. To further establish desirable features of the approach presented in Chapter 2, we proceed to modify the inference procedure to produce consistent estimates in Chapter 4.

Figure 3.1: A multistate representation of the OAT dispensation process. The OAT episode of an individual at time t is the number of (long-term) *not dispensed OAT* to *dispensed OAT* transitions they experience by time t, where individuals initialize in OAT episode 1.

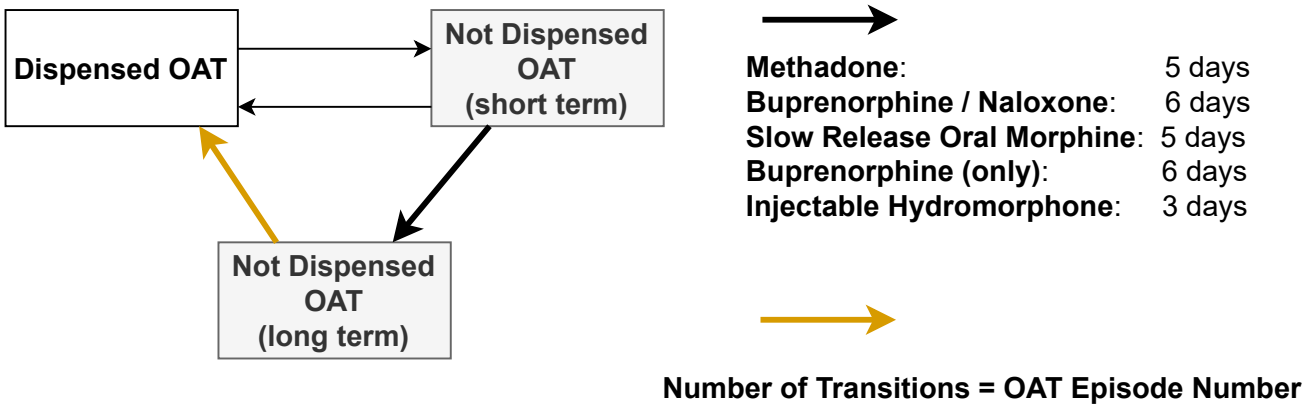


Figure 3.2: Estimates of regression coefficients under the stratified Cox regression model (3.2), where the time scale is *time since first recorded OAT dispensation*, and $\boldsymbol{\theta}_g = \boldsymbol{\alpha}_g$. Variables with a grey background appear to have a constant effect across strata. As a reference, we illustrate the estimated effect under the Cox model (3.4) with a red line.

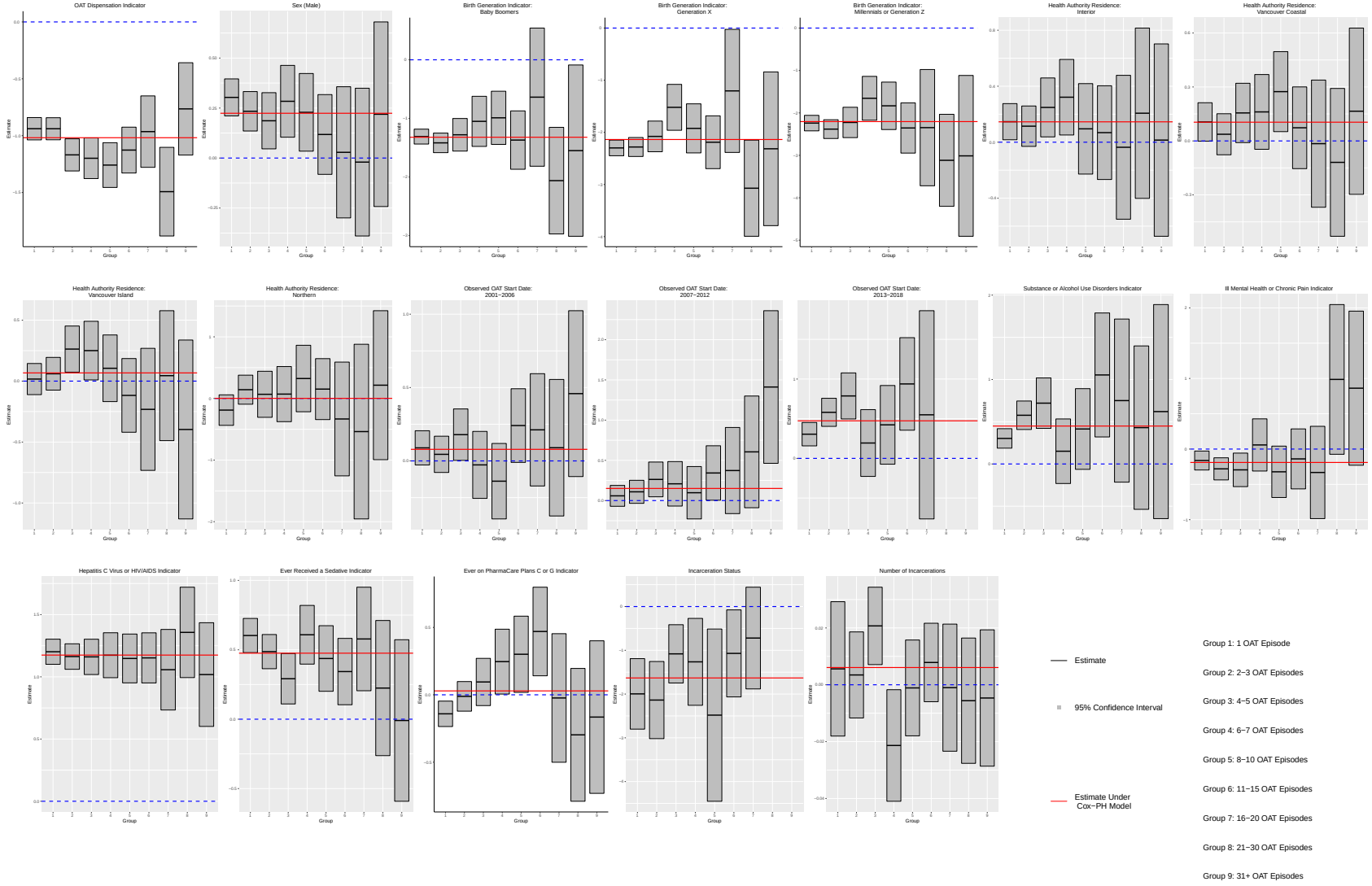


Figure 3.3: Estimates of regression coefficients under the stratified Cox regression model (3.2), where the time scale is *age*, and $\boldsymbol{\theta}_g = \boldsymbol{\alpha}_g$. Variables with a grey background appear to have a constant effect across strata. As a reference, we illustrate the estimated effect under the Cox model (3.4) with a red line.

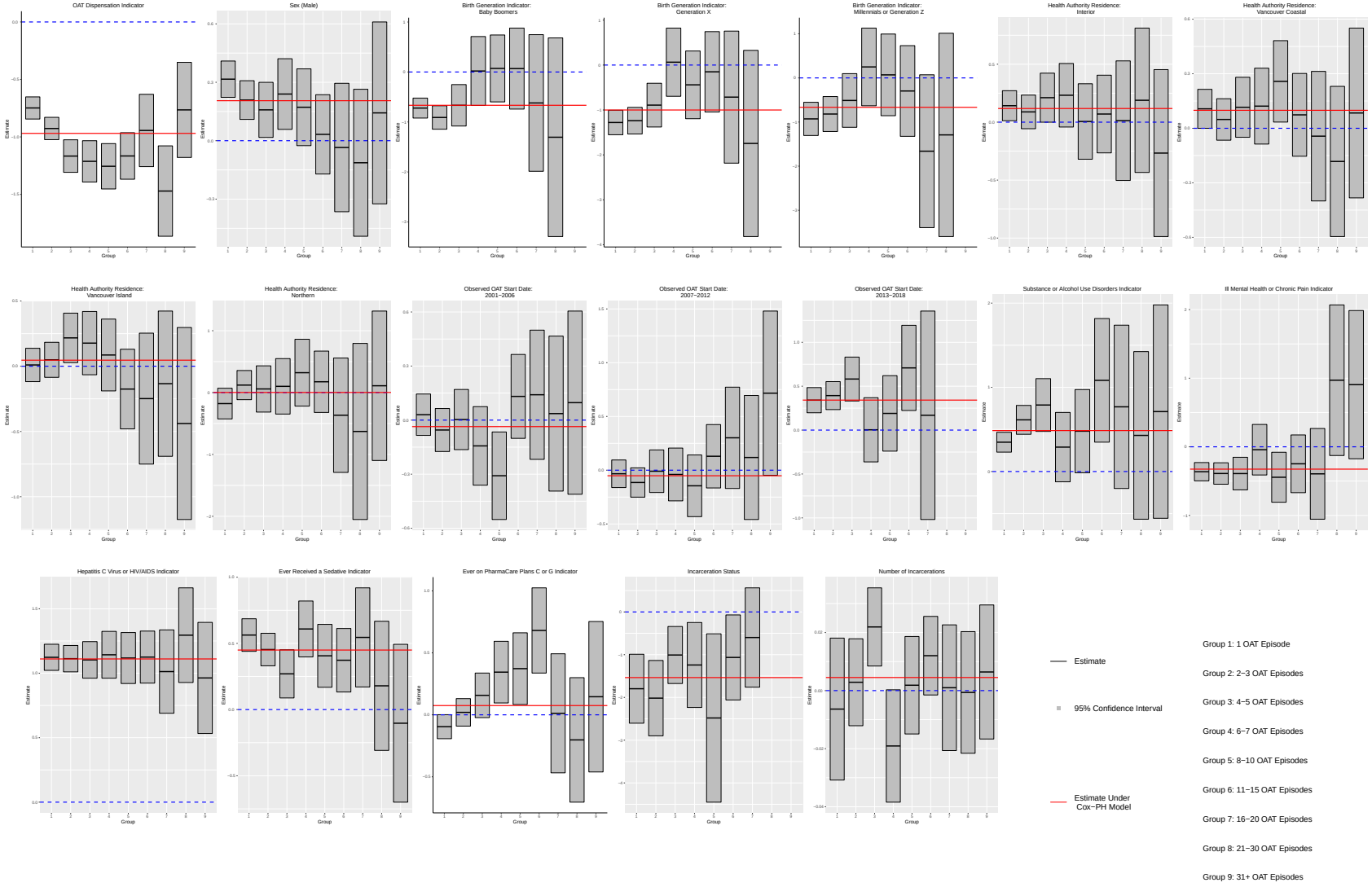
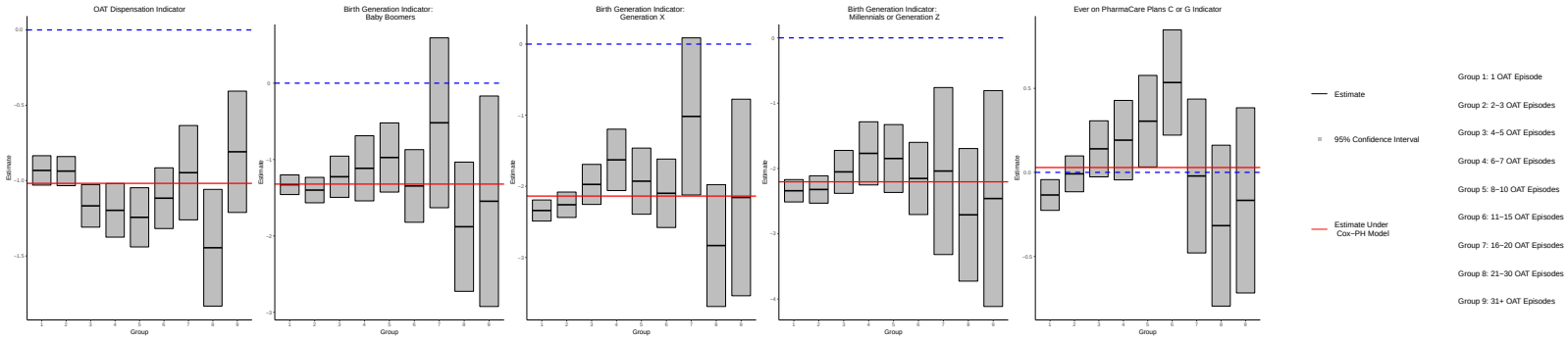
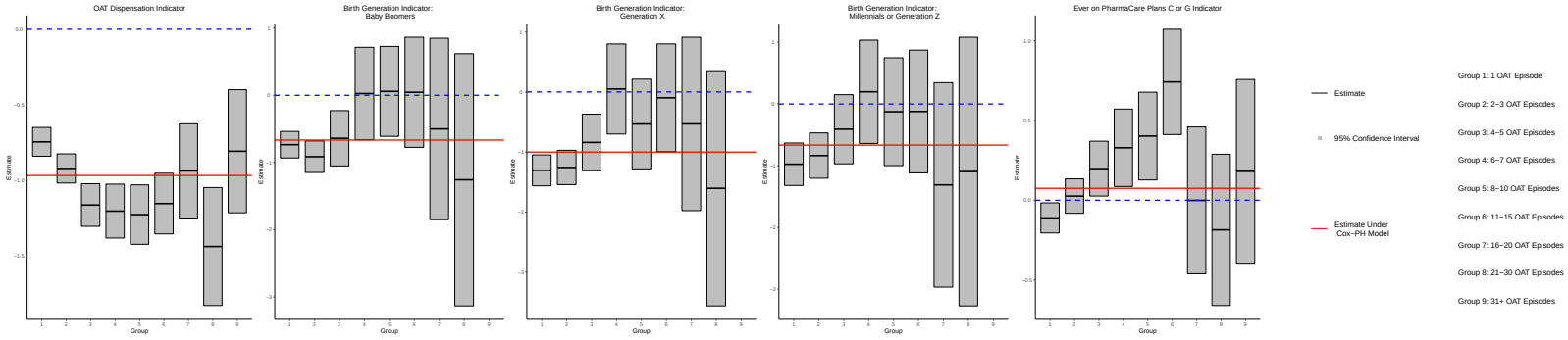


Figure 3.4: Estimates of regression coefficients under the stratified Cox regression model (3.2), where the time scale is *time since first recorded OAT dispensation*, and $\boldsymbol{\theta}_g = (\boldsymbol{\alpha}_g', \boldsymbol{\beta}')'$. Variables with effect $\boldsymbol{\alpha}_g$ and $\boldsymbol{\beta}$ are illustrated and tabulated below, respectively.



Covariate Name	Estimate	S.E.
Sex (vs. *Female*)	-	-
Male	**0.2340**	0.0260
Heath Authority (vs *Fraser Health*)	-	-
Interior	**0.1535**	0.0385
Vancouver Coastal	**0.1079**	0.0301
Vancouver Island	**0.0723**	0.0366
Northern	0.0114	0.0675
Year Category (vs. *1996-2000*)	-	-
2001-2006	**0.0791**	0.0338
2007-2012	**0.1430**	0.0387
2013-2018	**0.4619**	0.0498
Alcohol or Other Substance Use Disorders	**0.4329**	0.0447
Ill Mental Health or Chronic pain	**-0.1887**	0.0434
Hepatitis C Virus or HIV/AIDS	**1.1704**	0.0273
Ever Received a Sedative	**0.4791**	0.0335
Incarceration Status	**-1.6286**	0.1726
Number of Incarcerations	0.0014	0.0030

Figure 3.5: Estimates of regression coefficients under the stratified Cox regression model (3.2), where the time scale is *age*, and $\boldsymbol{\theta}_g = (\boldsymbol{\alpha}_g', \boldsymbol{\beta}')'$. Variables with effect $\boldsymbol{\alpha}_g$ and $\boldsymbol{\beta}$ are illustrated and tabulated below, respectively.



Covariate Name	Estimate	S.E.
Sex (vs. *Female*)	-	-
Male	**0.2071**	0.0261
Heath Authority (vs *Fraser Health*)	-	-
Interior	**0.1249**	0.0388
Vancouver Coastal	**0.0975**	0.0301
Vancouver Island	0.0467	0.0367
Northern	0.0100	0.0677
Year Category (vs. *1996-2000*)	-	-
2001-2006	-0.0259	0.0320
2007-2012	-0.0404	0.0380
2013-2018	**0.3658**	0.0462
Alcohol or Other Substance Use Disorders	**0.4770**	0.0443
Ill Mental Health or Chronic pain	**-0.3222**	0.0425
Hepatitis C Virus or HIV/AIDS	**1.1188**	0.0273
Ever Received a Sedative	**0.4534**	0.0336
Incarceration Status	**-1.5387**	0.1742
Number of Incarcerations	0.0034	0.0030

Figure 3.6: Smoothed estimates of $\lambda_{0g}(\cdot)$, where we stratify by the OAT episode number at time t, where the time scale is *time since first recorded OAT dispensation* and *age*.
Top Row: Estimates prior to the Wald test.
Bottom Row: Estimates following the Wald test.

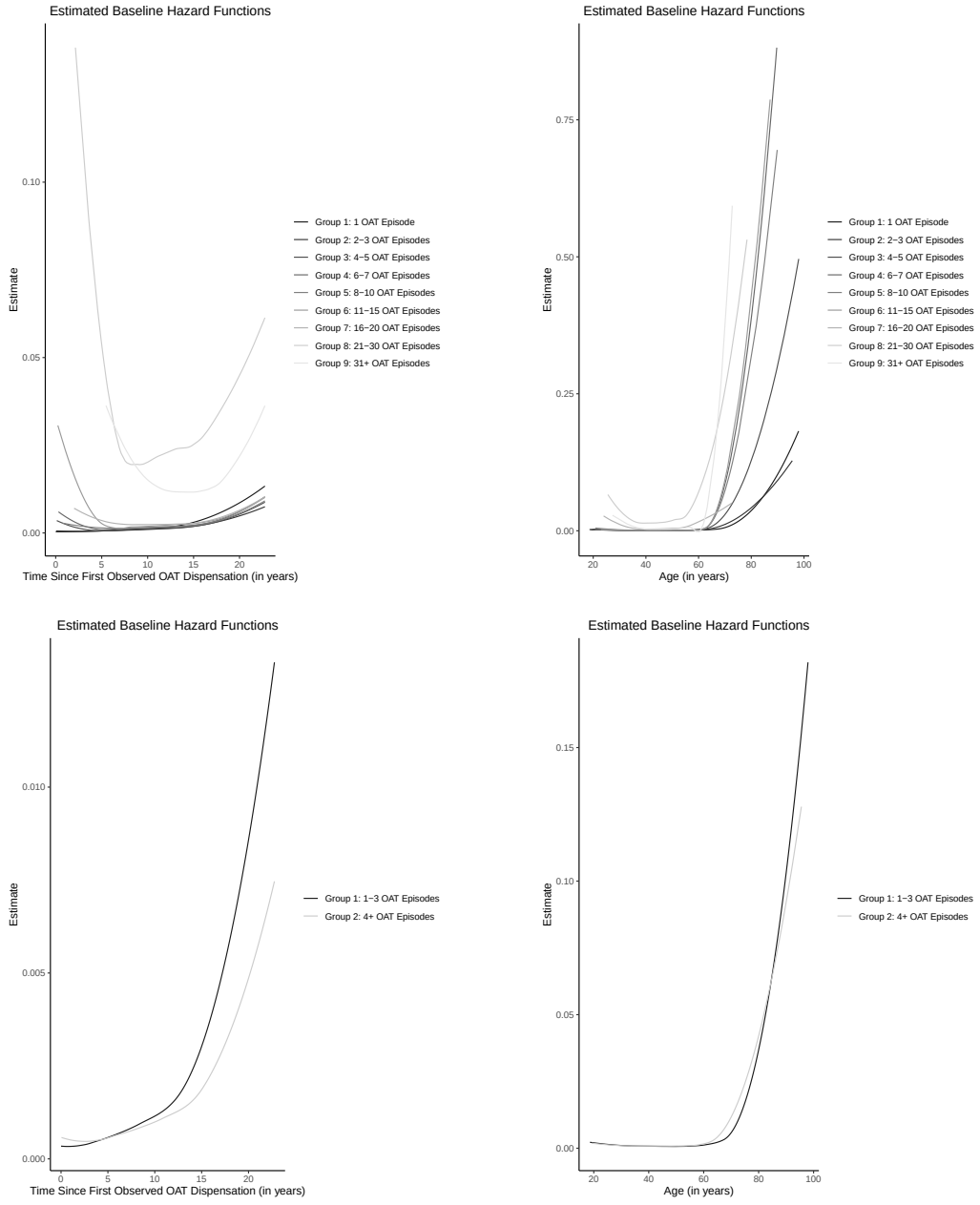
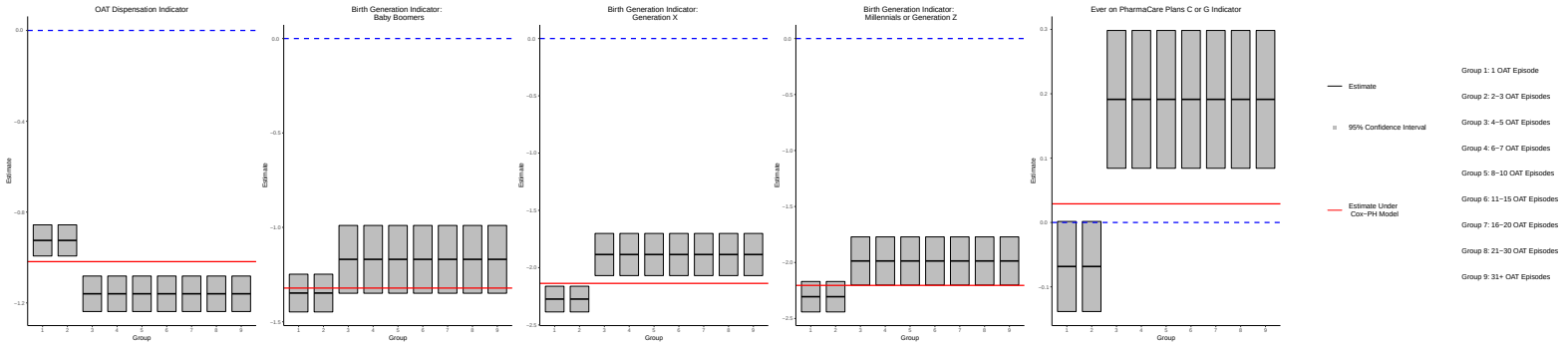
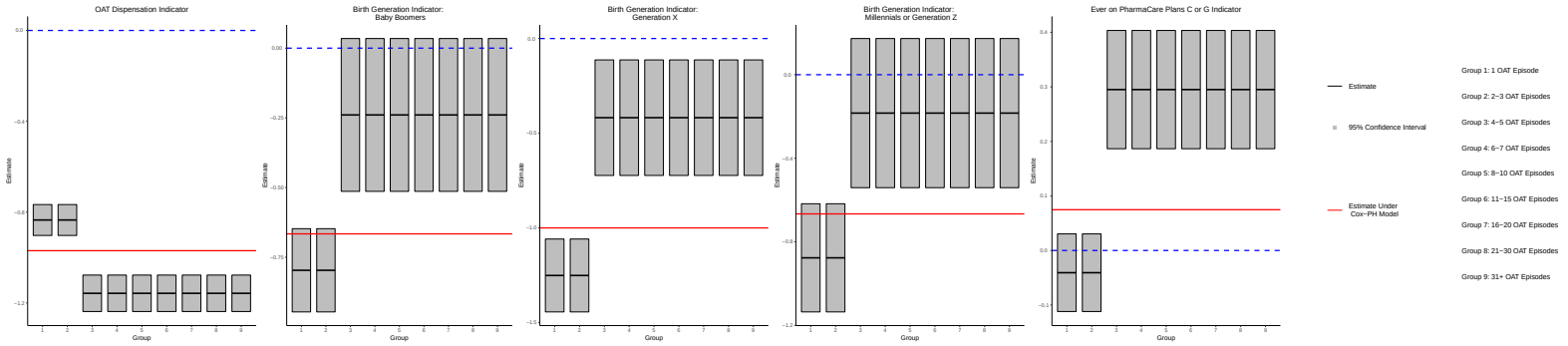


Figure 3.7: Estimates of regression coefficients under the stratified Cox regression model (3.2) following the Wald test, where the time scale is *time since first recorded OAT dispensation*, and $\boldsymbol{\theta}_g = (\boldsymbol{\alpha}_g', \boldsymbol{\beta}')'$. Variables with effect $\boldsymbol{\alpha}_g$ and $\boldsymbol{\beta}$ are illustrated and tabulated below, respectively.



Covariate Name	Estimate	S.E.
Sex (vs. *Female*)	-	-
Male	**0.2327**	0.0262
Heath Authority (vs *Fraser Health*)	-	-
Interior	**0.1517**	0.0387
Vancouver Coastal	**0.1062**	0.0302
Vancouver Island	0.0706	0.0362
Northern	0.0062	0.0676
Year Category (vs. *1996-2000*)	-	-
2001-2006	**0.0819**	0.0329
2007-2012	**0.1488**	0.0391
2013-2018	**0.4706**	0.0492
Alcohol or Other Substance Use Disorders	**0.4457**	0.0423
Ill Mental Health or Chronic pain	**-0.1904**	0.0430
Hepatitis C Virus or HIV/AIDS	**1.1725**	0.0273
Ever Received a Sedative	**0.4768**	0.0333
Incarceration Status	**-1.6244**	0.1793
Number of Incarcerations	0.0023	0.0029

Figure 3.8: Estimates of regression coefficients under the stratified Cox regression model (3.2) following the Wald test, where the time scale is *age*, and $\boldsymbol{\theta}_g = (\boldsymbol{\alpha}_g', \boldsymbol{\beta}')'$. Variables with effect $\boldsymbol{\alpha}_g$ and $\boldsymbol{\beta}$ are illustrated and tabulated below, respectively.



Covariate Name	Estimate	S.E.
Sex (vs. *Female*)	-	-
Male	**0.2089**	0.0263
Heath Authority (vs *Fraser Health*)	-	-
Interior	**0.1251**	0.0389
Vancouver Coastal	**0.0990**	0.0302
Vancouver Island	0.0471	0.0362
Northern	0.0089	0.0677
Year Category (vs. *1996-2000*)	-	-
2001-2006	-0.0267	0.0316
2007-2012	-0.0418	0.0366
2013-2018	**0.3606**	0.0447
Alcohol or Other Substance Use Disorders	**0.4851**	0.0435
Ill Mental Health or Chronic pain	**-0.3249**	0.0429
Hepatitis C Virus or HIV/AIDS	**1.1129**	0.0273
Ever Received a Sedative	**0.4515**	0.0332
Incarceration Status	**-1.5313**	0.1787
Number of Incarcerations	0.0031	0.0029

Table 3.1: Estimates of regression coefficients under the Cox regression model (3.4). The reported standard-error (S.E.) estimates of $\hat{\boldsymbol{\Theta}}$ correspond to the square-root of the diagonal elements of $\widehat{\boldsymbol{AV}}(\hat{\boldsymbol{\Theta}})$. The **bolded** estimates are statistically significant with the type 1 error rate set at $\alpha^* = 5\%$.

Time Scale	Time Since First Observed OAT Dispensation		Age	
Covariate Name	Estimate	S.E.	Estimate	S.E.
OAT Dispensation Indicator	**-1.0182**	0.0265	**-0.9696**	0.0264
Sex (vs. *Female*)	-	-	-	-
Male	**0.2239**	0.0261	**0.2055**	0.0263
Birth Generation (vs *Greatest & Silent Generations*)	-	-	-	-
Baby Boomers	**-1.3215**	0.0443	**-0.6661**	0.0662
Generation X	**-2.1372**	0.0478	**-1.0006**	0.0812
Millennials & Generation Z	**-2.2044**	0.0591	**-0.6663**	0.1066
Heath Authority (vs *Fraser Health*)	-	-	-	-
Interior	**0.1455**	0.0387	**0.1179**	0.0388
Vancouver Coastal	**0.1041**	0.0302	**0.0989**	0.0302
Vancouver Island	0.0674	0.0362	0.0465	0.0362
Northern	0.0015	0.0677	0.0018	0.0677
Year Category (vs. *1996-2000*)	-	-	-	-
2001-2006	**0.0789**	0.0330	-0.0360	0.0314
2007-2012	**0.1509**	0.0391	-0.0516	0.0361
2013-2018	**0.4733**	0.0492	**0.3416**	0.0435
Alcohol or Other Substance Use Disorders	**0.4492**	0.0424	**0.4864**	0.0434
Ill Mental Health or Chronic pain	**-0.1898**	0.0431	**-0.3255**	0.0428
Hepatitis C Virus or HIV/AIDS	**1.1748**	0.0272	**1.1109**	0.0272
Ever Received a Sedative	**0.4756**	0.0332	**0.4497**	0.0331
Ever on PharmaCare Plans C or G	0.0290	0.0299	**0.0748**	0.0305
Incarceration Status	**-1.6318**	0.1782	**-1.5358**	0.1784
Number of Incarcerations	**0.0061**	0.0028	0.0045	0.0028

Table 3.2: Results of the Wald test when we stratify by the number of OAT episode by time t, where the time scale is *time since first recorded OAT dispensation*, and we applied a Bonferonni correction so that the type 1 error rate is $\alpha^* = 0.05/(G-1)$. For each test, the estimates on the left- and right-hand side correspond to $\hat{\boldsymbol{\alpha}}_{g-1}$ and $\hat{\boldsymbol{\alpha}}_g$ and their estimated standard errors, respectively. **Bolded** test statistic(s) and p-value(s) indicate tests that rejected the null hypothesis in (3.6).

	Test 1: Group 1 vs. Group 2				**Test 2**: Groups 1,2 vs. Group 3				**Test 3**: Group 3 vs. Group 4			
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
OAT Dispensation Indicator	-0.9324	0.0498	-0.9369	0.0493	-0.9252	0.0350	-1.1679	0.0720	-1.1679	0.0720	-1.1974	0.0908
Birth Generation (vs *Greatest & Silent Generations*)	-	-	-	-	-	-	-	-	-	-	-	-
Baby Boomers	-1.3298	0.0656	-1.3999	0.0850	-1.3467	0.0510	-1.2268	0.1399	-1.2268	0.1399	-1.1164	0.2156
Generation X	-2.3404	0.0756	-2.2593	0.0911	-2.2719	0.0573	-1.9749	0.1454	-1.9749	0.1454	-1.6300	0.2188
Millennials & Generation Z	-2.3416	0.0873	-2.3219	0.1070	-2.3014	0.0695	-2.0543	0.1691	-2.0543	0.1691	-1.7708	0.2487
Ever on PharmaCare Plans C or G	-0.1345	0.0467	-0.0086	0.0544	-0.0684	0.0357	0.1378	0.0864	0.1378	0.0864	0.1892	0.1183
Test Statistic		6.9201				**20.8622**				4.9197		
p-value		0.2266				**0.0009**				0.4258		
	Test 4 Groups 3,4 vs. Group 5				**Test 5**: Groups 3,4,5 vs. Group 6				**Test 6**: Groups 3,4,5,6 vs. Group 7			
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
OAT Dispensation Indicator	-1.1757	0.0565	-1.2442	0.0999	-1.1914	0.0492	-1.1171	0.1025	-1.1802	0.0440	-0.9477	0.1595
Birth Generation (vs *Greatest & Silent Generations*)	-	-	-	-	-	-	-	-	-	-	-	-
Baby Boomers	-1.1922	0.1169	-0.9751	0.2286	-1.1446	0.1044	-1.3471	0.2510	-1.1691	0.0957	-0.5234	0.5948
Generation X	-1.8458	0.1204	-1.9299	0.2354	-1.8694	0.1081	-2.1007	0.2536	-1.8919	0.0985	-1.0243	0.5912
Millennials & Generation Z	-1.9514	0.1396	-1.8502	0.2663	-1.9287	0.1246	-2.1561	0.2873	-1.9478	0.1146	-2.0486	0.6715
Ever on PharmaCare Plans C or G	0.1537	0.0701	0.3017	0.1389	0.1882	0.0626	0.5313	0.1633	0.2397	0.0586	-0.0262	0.2342
Test Statistic		4.9197				4.7481				12.0358		
p-value		0.4258				0.4474				0.0343		
	Test 7: Groups 3,4,5,6,7 vs. Group 8				**Test 8**: Groups 3,4,5,6,7,8 vs. Group 9							
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.				
OAT Dispensation Indicator	-1.1629	0.0419	-1.4455	0.1962	-1.1740	0.0416	-0.8097	0.2062				
Birth Generation (vs *Greatest & Silent Generations*)	-	-	-	-	-	-	-	-				
Baby Boomers	-1.1505	0.0944	-1.8798	0.4469	-1.1642	0.0915	-1.5463	0.7347				
Generation X	-1.8571	0.0974	-2.8373	0.4522	-1.8852	0.0948	-2.1605	0.7350				
Millennials & Generation Z	-1.9653	0.1113	-2.7166	0.5223	-1.9821	0.1120	-2.4673	0.9264				
Ever on PharmaCare Plans C or G	0.2240	0.0554	-0.3222	0.2440	0.2035	0.0561	-0.1720	0.2796				
Test Statistic		13.3683				5.1710						
p-value		0.0202				0.3954						

Table 3.3: Results of the Wald test when we stratify by the number of OAT episode by time t, where the time scale is *age*, and we applied a Bonferonni correction so that the type 1 error rate is $\alpha^* = 0.05/(G-1)$. For each test, the estimates on the left- and right-hand side correspond to $\hat{\boldsymbol{\alpha}}_{g-1}$ and $\hat{\boldsymbol{\alpha}}_g$ and their estimated standard errors, respectively. **Bolded** test statistic(s) and p-value(s) indicate tests that rejected the null hypothesis in (3.6).

	Test 1: Group 1 vs. Group 2				**Test 2**: Groups 1,2 vs. Group 3				**Test 3**: Group 3 vs. Group 4			
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
OAT Dispensation Indicator	-0.7465	0.0488	-0.9227	0.0494	-0.8342	0.0347	-1.1654	0.0721	-1.1654	0.0721	-1.2059	0.0911
Birth Generation (vs *Greatest & Silent Generations*)	-	-	-	-	-	-	-	-	-	-	-	-
Baby Boomers	-0.7348	0.1008	-0.9134	0.1196	-0.7965	0.0762	-0.6379	0.2091	-0.6379	0.2091	0.0269	0.3486
Generation X	-1.3052	0.1313	-1.2567	0.1458	-1.2510	0.0985	-0.8369	0.2396	-0.8369	0.2396	0.0520	0.3799
Millennials & Generation Z	-0.9761	0.1755	-0.8363	0.1873	-0.8738	0.1324	-0.4016	0.2852	-0.4016	0.2852	0.2043	0.4243
Ever on PharmaCare Plans C or G	-0.1107	0.0482	0.0262	0.0550	-0.0415	0.0363	0.1972	0.0886	0.1972	0.0886	0.3280	0.1231
Test Statistic	14.2582				**26.2272**				6.2905			
p-value	0.0140				**0.0001**				0.2790			
	Test 4: Groups 3,4 vs. Group 5				**Test 5**: Group 3,4,5 vs. Group 6				**Test 6**: Groups 3,4,5,6 vs. Group 7			
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
OAT Dispensation Indicator	-1.1767	0.0565	-1.2292	0.1000	-1.1889	0.0491	-1.1557	0.1025	-1.1771	0.0440	-0.9389	0.1595
Birth Generation (vs *Greatest & Silent Generations*)	-	-	-	-	-	-	-	-	-	-	-	-
Baby Boomers	-0.4268	0.1790	0.0591	0.3389	-0.3108	0.1585	0.0453	0.4162	-0.2494	0.1476	-0.5003	0.6999
Generation X	-0.5340	0.2009	-0.5317	0.3784	-0.5296	0.1779	-0.0960	0.4562	-0.4537	0.1651	-0.5312	0.7468
Millennials & Generation Z	-0.2066	0.2345	-0.1200	0.4425	-0.1863	0.2086	-0.1183	0.5059	-0.1714	0.1927	-1.3081	0.8538
Ever on PharmaCare Plans C or G	0.2468	0.0720	0.4018	0.1415	0.2798	0.0642	0.7422	0.1709	0.3433	0.0593	-0.0025	0.2318
Test Statistic	9.2871				9.8890				10.5314			
p-value	0.0981				0.0784				0.0615			
	Test 7: Groups 3,4,5,6,7 vs. Group 8				**Test 8**: Groups 3,4,5,6,7,8 vs. Group 9							
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.				
OAT Dispensation Indicator	-1.1607	0.0421	-1.4405	0.1982	-1.1728	0.0416	-0.8083	0.2089				
Birth Generation (vs *Greatest & Silent Generations*)	-	-	-	-	-	-	-	-				
Baby Boomers	-0.2487	0.1427	-1.2555	0.9547	-0.2514	0.1419	NA	NA				
Generation X	-0.4374	0.1607	-1.6023	0.9987	-0.4518	0.1616	NA	NA				
Millennials & Generation Z	-0.2173	0.1874	-1.0918	1.1089	-0.2238	0.1900	NA	NA				
Ever on PharmaCare Plans C or G	0.3231	0.0586	-0.1875	0.2427	0.3039	0.0567	0.1821	0.2916				
Test Statistic	8.3051				2.9938							
p-value	0.1402				0.2238							

Table 3.4: Estimates of regression coefficients under (3.2) with $G \equiv 1$, and the extended Cox proportional hazards model via the `survival::coxph` R function, across 150 datasets with $n = 10{,}000$ independent units for Simulation Outcome 1. The data is generated under the model in (3.4).
Abbreviations:
SM: Sample mean;
SSE: Sample standard error;
SM_{SE}: Sample mean of the standard error estimator.

Censoring Parameter Setting I: $\lambda_C = 0.5$															
					$Z(\cdot)$ Effect Estimate				X_1 Effect Estimate				X_2 Effect Estimate			
						SM	SSE	SM_{SE}		SM	SSE	SM_{SE}		SM	SSE	SM_{SE}
Parameter	$Z(\cdot)$ Effect	X_1 Effect	X_2 Effect	λ_0	Our Method	-1.0174	0.0289	0.0285	Our Method	-0.0031	0.0364	0.0387	Our Method	-0.4951	0.0233	0.0228
All Groups	-1	0	-0.5	1.75	`coxph`	-1.0174	0.0289	0.0286	`coxph`	-0.0031	0.0364	0.0387	`coxph`	-0.4951	0.0233	0.0228
Censoring Parameter Setting II $\lambda_C = 1.5$																
					$Z(\cdot)$ Effect Estimate				X_1 Effect Estimate				X_2 Effect Estimate			
						SM	SSE	SM_{SE}		SM	SSE	SM_{SE}		SM	SSE	SM_{SE}
Parameter	$Z(\cdot)$ Effect	X_1 Effect	X_2 Effect	λ_0	Our Method	-1.0429	0.0339	0.0333	Our Method	-0.0034	0.0448	0.0451	Our Method	-0.4898	0.0264	0.0264
All Groups	-1	0	-0.5	1.75	`coxph`	-1.0429	0.0339	0.0333	`coxph`	-0.0034	0.0448	0.0451	`coxph`	-0.4898	0.0264	0.0264

Table 3.5: Estimates of regression coefficients and true values for 150 simulations for Simulation Outcome 1. Each dataset has $n = 10{,}000$ independent units, and the data is generated under the model in (3.4).
Abbreviations:
SM: Sample mean;
SSE: Sample standard error;
SM_{SE}: Sample mean of the standard error estimator.

Censoring Parameter Setting I: $\lambda_C = 0.5$													
					$Z(\cdot)$ Effect Estimate			X_1 Effect Estimate			X_2 Effect Estimate		
Parameter	$Z(\cdot)$ Effect	X_1 Effect	X_2 Effect	λ_0	SM	SSE	SM_{SE}	SM	SSE	SM_{SE}	SM	SSE	SM_{SE}
Group 1	-1	0	-0.5	1.75	-1.0129	0.0644	0.0618	-0.0130	0.0757	0.0734	-0.4959	0.0446	0.0433
Group 2	-1	0	-0.5	1.75	-1.0226	0.0659	0.0667	0.0036	0.0896	0.0868	-0.4933	0.0508	0.0507
Group 3	-1	0	-0.5	1.75	-1.0146	0.0794	0.0774	0.0010	0.1014	0.1025	-0.5013	0.0567	0.0596
Group 4	-1	0	-0.5	1.75	-1.0056	0.1161	0.1108	-0.0027	0.1507	0.1574	-0.4908	0.0948	0.0913
Group 5	-1	0	-0.5	1.75	-1.0239	0.1245	0.1218	0.0216	0.1868	0.1717	-0.5032	0.1009	0.0998
Group 6	-1	0	-0.5	1.75	-1.0170	0.1320	0.1334	-0.0080	0.1863	0.1890	-0.4938	0.0982	0.1095
Group 7	-1	0	-0.5	1.75	-1.0068	0.2148	0.1961	-0.0007	0.3153	0.3044	-0.5032	0.1629	0.1755
Group 8	-1	0	-0.5	1.75	-1.0323	0.0885	0.0921	-0.0007	0.1383	0.1282	-0.4857	0.0709	0.0750
Group 9	-1	0	-0.5	1.75	-1.0436	0.1233	0.1221	-0.0110	0.1724	0.1715	-0.5027	0.0996	0.1014
Group 10	-1	0	-0.5	1.75	-1.0149	0.0948	0.1035	-0.0047	0.1564	0.1466	-0.4902	0.0968	0.0939
Censoring Parameter Setting II: $\lambda_C = 1.5$													
					$Z(\cdot)$ Effect Estimate			X_1 Effect Estimate			X_2 Effect Estimate		
Parameter	$Z(\cdot)$ Effect	X_1 Effect	X_2 Effect	λ_0	SM	SSE	SM_{SE}	SM	SSE	SM_{SE}	SM	SSE	SM_{SE}
Group 1	-1	0	-0.5	1.75	-1.0427	0.0658	0.0620	-0.0129	0.0737	0.0748	-0.4924	0.0470	0.0442
Group 2	-1	0	-0.5	1.75	-1.0513	0.0715	0.0712	0.0006	0.1004	0.0939	-0.4897	0.0535	0.0549
Group 3	-1	0	-0.5	1.75	-1.0418	0.0905	0.0877	0.0022	0.1157	0.1180	-0.4946	0.0647	0.0685
Group 4	-1	0	-0.5	1.75	-1.0223	0.1378	0.1334	0.0018	0.1878	0.1910	-0.4904	0.1154	0.1107
Group 5	-1	0	-0.5	1.75	-1.0262	0.1509	0.1522	0.0256	0.2289	0.2172	-0.4966	0.1286	0.1260
Group 6	-1	0	-0.5	1.75	-1.0145	0.1672	0.1725	0.0001	0.2505	0.2480	-0.4954	0.1232	0.1432
Group 7	-1	0	-0.5	1.75	-0.9961	0.2691	0.2657	-0.0388	0.4251	0.4144	-0.4941	0.2361	0.2392
Group 8	-1	0	-0.5	1.75	-1.0700	0.1227	0.1273	0.0076	0.1825	0.1815	-0.4719	0.1032	0.1058
Group 9	-1	0	-0.5	1.75	-1.0984	0.2007	0.1883	-0.0237	0.2625	0.2713	-0.4742	0.1549	0.1587
Group 10	-1	0	-0.5	1.75	-1.0641	0.1932	0.1996	-0.0062	0.3467	0.2930	-0.4738	0.1788	0.1790

Table 3.6: Results of the Wald test pairing groups together with $\alpha^* = 0.05$ across 150 datasets and $n = 10{,}000$ for Simulation Outcome 1. The data is generated under the Cox regression model (3.4). The (g, g') element in each matrix is the proportion groups g and g' were merged to the same class.

Censoring Parameter Setting I: $\lambda_C = 0.5$

Parameter	$Z(\cdot)$ Effect	X_1 Effect	X_2 Effect	λ_0
All Groups	-1	0	-0.5	1.75

Group \ Group	1	2	3	4	5	6	7	8	9	10
1	1									
2	0.9533	1								
3	0.8733	0.9133	1							
4	0.8400	0.8733	0.9400	1						
5	0.7800	0.8067	0.8733	0.9200	1					
6	0.7400	0.7667	0.8200	0.8667	0.9400	1				
7	0.7133	0.7400	0.7933	0.8400	0.9067	0.9667	1			
8	0.6867	0.7000	0.7400	0.7867	0.8400	0.8933	0.8933	1		
9	0.6467	0.6533	0.6933	0.7400	0.7933	0.8333	0.8333	0.9067	1	
10	0.6267	0.6333	0.6733	0.7200	0.7733	0.8133	0.8133	0.8867	0.9733	1

Censoring Parameter Setting II: $\lambda_C = 1.5$

Parameter	$Z(\cdot)$ Effect	X_1 Effect	X_2 Effect	λ_0
All Groups	-1	0	-0.5	1.75

Group \ Group	1	2	3	4	5	6	7	8	9	10
1	1									
2	0.9467	1								
3	0.8933	0.9467	1							
4	0.8533	0.9067	0.9533	1						
5	0.8067	0.8533	0.8933	0.9200	1					
6	0.7600	0.8067	0.8467	0.8667	0.9333	1				
7	0.6933	0.7400	0.7800	0.8000	0.8667	0.9200	1			
8	0.6533	0.6800	0.7133	0.7267	0.7933	0.8400	0.8933	1		
9	0.5667	0.5933	0.6200	0.6267	0.6933	0.7400	0.7867	0.8867	1	
10	0.5133	0.5400	0.5667	0.5733	0.6400	0.6800	0.7267	0.8267	0.9333	1

Table 3.7: Results of the risk classes identified by the Wald test with $\alpha^* = 0.05$ and $\alpha^* = 0.05/(G-1)$ for Simulation Outcome 1. The data is generated under the Cox regression model (3.4), and $n =$ 10,000. The correct (data-generating) class is **bolded**.

$\alpha^* = 0.05$			
Censoring Parameter Setting I: $\lambda_C = 0.5$		**Censoring Parameter Setting II**: $\lambda_C = 1.5$	
Updated Groups	Proportion	Updated Groups	Proportion
1, 1, 1, 1, 1, 1, 1, 1, 1, 1	**0.6267**	**1, 1, 1, 1, 1, 1, 1, 1, 1, 1**	**0.5133**
1, 1, 2, 2, 2, 2, 2, 2, 2, 2	0.0400	1, 1, 1, 1, 1, 1, 1, 1, 2, 2	0.0800
1, 1, 1, 1, 1, 1, 1, 1, 2, 2	0.0400	1, 1, 1, 1, 1, 1, 1, 1, 1, 2	0.0533
1, 1, 1, 1, 2, 2, 2, 2, 2, 2	0.0333	1, 1, 1, 1, 2, 2, 2, 2, 2, 2	0.0400
1, 1, 1, 1, 1, 2, 2, 2, 2, 2	0.0267	1, 1, 1, 1, 1, 1, 1, 2, 2, 2	0.0333
1, 1, 1, 2, 2, 2, 2, 2, 2, 2	0.0267	1, 1, 1, 1, 1, 1, 2, 2, 2, 2	0.0333
1, 1, 1, 1, 1, 1, 2, 3, 3, 3	0.0267	1, 2, 2, 2, 2, 2, 2, 2, 2, 2	0.0267
1, 1, 2, 3, 3, 3, 3, 3, 3, 3	0.0200	1, 1, 2, 2, 2, 2, 2, 2, 2, 2	0.0267
1, 1, 1, 1, 1, 1, 1, 1, 1, 2	0.0200	1, 1, 1, 1, 1, 1, 2, 3, 3, 3	0.0267
1, 1, 1, 1, 1, 1, 1, 2, 3, 3	0.0200	1, 1, 1, 1, 1, 2, 2, 2, 2, 2	0.0267
1, 1, 2, 2, 2, 2, 2, 3, 3, 3	0.0133	1, 2, 2, 2, 2, 2, 2, 3, 3, 3	0.0200
1, 2, 2, 2, 2, 2, 2, 3, 3, 3	0.0067	1, 1, 1, 2, 3, 3, 3, 3, 3, 3	0.0200
1, 2, 3, 3, 3, 4, 4, 4, 4, 4	0.0067	1, 1, 1, 1, 1, 2, 3, 3, 3, 3	0.0133
1, 1, 1, 1, 2, 2, 2, 3, 3, 3	0.0067	1, 2, 2, 2, 3, 4, 4, 4, 4, 4	0.0067
1, 1, 1, 1, 1, 2, 2, 2, 3, 3	0.0067	1, 1, 2, 2, 2, 2, 2, 3, 3, 3	0.0067
Other	0.0798	Other	0.0733
$\alpha^* = 0.05/(G-1)$			
Censoring Parameter Setting I: $\lambda_C = 0.5$		**Censoring Parameter Setting II**: $\lambda_C = 1.5$	
Updated Groups	Proportion	Updated Groups	Proportion
1, 1, 1, 1, 1, 1, 1, 1, 1, 1	**0.9400**	**1, 1, 1, 1, 1, 1, 1, 1, 1, 1**	**0.9000**
1, 2, 2, 2, 2, 2, 2, 2, 2, 2	0.0133	1, 1, 1, 1, 1, 1, 1, 1, 2, 2	0.0467
1, 1, 1, 1, 1, 1, 1, 1, 2, 2	0.0133	1, 1, 1, 1, 1, 1, 1, 1, 1, 2	0.0200
1, 1, 1, 1, 2, 2, 2, 2, 2, 2	0.0133	1, 2, 2, 2, 2, 2, 2, 3, 3, 3	0.0067
1, 1, 1, 2, 2, 2, 2, 2, 2, 2	0.0067	1, 1, 1, 1, 1, 2, 2, 2, 2, 2	0.0067
1, 1, 1, 2, 3, 3, 3, 3, 3, 3	0.0067	1, 1, 1, 2, 2, 2, 2, 2, 2, 2	0.0067
1, 1, 1, 1, 1, 1, 2, 2, 2, 2	0.0067	1, 2, 2, 2, 2, 2, 2, 2, 2, 2	0.0067
		1, 1, 2, 2, 2, 2, 2, 2, 2, 2	0.0067

Table 3.8: Estimates of regression coefficients under (3.2) with $G \equiv 1$, and the extended Cox proportional hazards model via the `survival::coxph` R function, across 150 datasets with $n = 10{,}000$ independent units for Simulation Outcome 2. The data is generated under the model in (3.2).
Abbreviations:
SM: Sample mean;
SSE: Sample standard error;
SM_{SE}: Sample mean of the standard error estimator.

Censoring Parameter Setting I: $\lambda_C = 0.5$															
					$Z(\cdot)$ Effect Estimate				X_1 Effect Estimate				X_2 Effect Estimate			
						SM	SSE	SM_{SE}		SM	SSE	SM_{SE}		SM	SSE	SM_{SE}
Parameter	$Z(\cdot)$ Effect	X_1 Effect	X_2 Effect	λ_{0g}	Our Method	-0.8802	0.0320	0.0299	Our Method	0.4700	0.0457	0.0431	Our Method	0.2422	0.0230	0.0246
Group 1	-0.5	-2	-2	0.5	`coxph`	-0.8802	0.0320	0.0286	`coxph`	0.4700	0.0457	0.0417	`coxph`	0.2422	0.0230	0.0243
Group 2	-0.5	-2	-2	0.5												
Group 3	-0.5	-2	-2	0.5												
Group 4	-1	0	-0.5	1.75												
Group 5	-1	0	-0.5	1.75												
Group 6	-1	0	-0.5	1.75												
Group 7	-2	2	1.5	3												
Group 8	-2	2	1.5	3												
Group 9	-2	2	1.5	3												
Group 10	-2	2	1.5	3												
Censoring Parameter Setting II: $\lambda_C = 1.5$																
$\lambda_C = 1.5$					$Z(\cdot)$ Effect Estimate				X_1 Effect Estimate				X_2 Effect Estimate			
						SM	SSE	SM_{SE}		SM	SSE	SM_{SE}		SM	SSE	SM_{SE}
Parameter	$Z(\cdot)$ Effect	X_1 Effect	X_2 Effect	λ_{0g}	Our Method	-0.7935	0.0431	0.0404	Our Method	0.3245	0.0577	0.0587	Our Method	0.1202	0.0313	0.0334
Group 1	-0.5	-2	-2	0.5	`coxph`	-0.7935	0.0431	0.0388	`coxph`	0.3245	0.0577	0.0571	`coxph`	0.1202	0.0313	0.0331
Group 2	-0.5	-2	-2	0.5												
Group 3	-0.5	-2	-2	0.5												
Group 4	-1	0	-0.5	1.75												
Group 5	-1	0	-0.5	1.75												
Group 6	-1	0	-0.5	1.75												
Group 7	-2	2	1.5	3												
Group 8	-2	2	1.5	3												
Group 9	-2	2	1.5	3												
Group 10	-2	2	1.5	3												

Table 3.9: Estimates of regression coefficients and true values for 150 simulations, where each dataset has n = 10,000 independent units for Simulation Outcome 2. The data is generated under the model in (3.2).

Abbreviations:
SM: Sample mean;
SSE: Sample standard error;
SM_{SE}: Sample mean of the standard error estimator.

Censoring Parameter Setting I: $\lambda_C = 0.5$													
					$Z(\cdot)$ Effect Estimate			X_1 Effect Estimate			X_2 Effect Estimate		
Parameter	$Z(\cdot)$ Effect	X_1 Effect	X_2 Effect	λ_{0g}	SM	SSE	SM_{SE}	SM	SSE	SM_{SE}	SM	SSE	SM_{SE}
Group 1	-0.5	-2	-2	0.5	-0.5495	0.2323	0.2286	-2.0103	0.3074	0.3096	-2.0112	0.2676	0.2548
Group 2	-0.5	-2	-2	0.5	-0.5061	0.2094	0.2112	-2.0135	0.3764	0.3264	-2.0234	0.2737	0.2657
Group 3	-0.5	-2	-2	0.5	-0.5574	0.2275	0.2204	-1.9973	0.3386	0.3420	-2.0449	0.3312	0.2805
Group 4	-1	0	-0.5	1.75	-1.0070	0.0982	0.0996	0.0290	0.1502	0.1403	-0.5088	0.0768	0.0825
Group 5	-1	0	-0.5	1.75	-1.0002	0.0988	0.1063	0.0026	0.1729	0.1500	-0.4972	0.0854	0.0880
Group 6	-1	0	-0.5	1.75	-1.0004	0.1202	0.1134	0.0047	0.1576	0.1597	-0.4931	0.0943	0.0932
Group 7	-2	2	1.5	3	-1.9971	0.0679	0.0598	1.9972	0.0849	0.0870	1.4937	0.0563	0.0560
Group 8	-2	2	1.5	3	-1.9871	0.0460	0.0472	1.9728	0.0810	0.0738	1.4773	0.0449	0.0422
Group 9	-2	2	1.5	3	-2.0034	0.1184	0.1171	1.9718	0.1850	0.1774	1.4793	0.1164	0.1093
Group 10	-2	2	1.5	3	-2.0231	0.1972	0.1917	1.9968	0.3030	0.3021	1.4970	0.2530	0.2429
Censoring Parameter Setting II: $\lambda_C = 1.5$													
					$Z(\cdot)$ Effect Estimate			X_1 Effect Estimate			X_2 Effect Estimate		
Parameter	$Z(\cdot)$ Effect	X_1 Effect	X_2 Effect	λ_{0g}	SM	SSE	SM_{SE}	SM	SSE	SM_{SE}	SM	SSE	SM_{SE}
Group 1	-0.5	-2	-2	0.5	-0.5842	0.2365	0.2292	-1.9990	0.3082	0.3152	-2.0060	0.2715	0.2594
Group 2	-0.5	-2	-2	0.5	-0.5557	0.2238	0.2278	-2.0309	0.3989	0.3569	-2.0327	0.3129	0.2937
Group 3	-0.5	-2	-2	0.5	-0.6080	0.2467	0.2558	-1.9849	0.3761	0.4001	-2.0342	0.4042	0.3315
Group 4	-1	0	-0.5	1.75	-1.0226	0.1288	0.1229	0.0191	0.1789	0.1746	-0.5094	0.0938	0.1026
Group 5	-1	0	-0.5	1.75	-1.0180	0.1219	0.1367	0.0030	0.2067	0.1940	-0.4902	0.1099	0.1138
Group 6	-1	0	-0.5	1.75	-1.0078	0.1638	0.1510	-0.0301	0.2181	0.2147	-0.5081	0.1271	0.1255
Group 7	-2	2	1.5	3	-1.9890	0.0863	0.0826	1.9958	0.1248	0.1213	1.4867	0.0805	0.0777
Group 8	-2	2	1.5	3	-1.9594	0.0664	0.0666	1.9240	0.1142	0.1063	1.4363	0.0664	0.0607
Group 9	-2	2	1.5	3	-2.0034	0.1929	0.1790	1.8914	0.2819	0.2788	1.4369	0.1727	0.1694
Group 10	-2	2	1.5	3	-2.0553	0.3921	0.3451	1.9406	0.5482	0.5394	1.5485	0.5169	0.4369

Table 3.10: Results of the Wald test pairing groups together with $\alpha^* = 0.05$ across 150 datasets and $n = 10{,}000$ for Simulation Outcome 2. The data is generated under the stratified Cox regression model (3.2). The (g, g') element in each matrix is the proportion groups g and g' were merged to the same class.

Censoring Parameter Setting I: $\lambda_C = 0.5$

Parameter	$Z(\cdot)$ Effect	X_1 Effect	X_2 Effect	λ_{0g}
Group 1	-0.5	-2	-2	0.5
Group 2	-0.5	-2	-2	0.5
Group 3	-0.5	-2	-2	0.5
Group 4	-1	0	-0.5	1.75
Group 5	-1	0	-0.5	1.75
Group 6	-1	0	-0.5	1.75
Group 7	-2	2	1.5	3
Group 8	-2	2	1.5	3
Group 9	-2	2	1.5	3
Group 10	-2	2	1.5	3

Group \ Group	1	2	3	4	5	6	7	8	9	10
1	1									
2	0.9267	1								
3	0.8800	0.9467	1							
4	0	0	0	1						
5	0	0	0	0.9600	1					
6	0	0	0	0.9000	0.9400	1				
7	0	0	0	0	0	0	1			
8	0	0	0	0	0	0	0.9000	1		
9	0	0	0	0	0	0	0.8600	0.9533	1	
10	0	0	0	0	0	0	0.7867	0.8667	0.9000	1

Censoring Parameter Setting II: $\lambda_C = 1.5$

Parameter	$Z(\cdot)$ Effect	X_1 Effect	X_2 Effect	λ_{0g}
Group 1	-0.5	-2	-2	0.5
Group 2	-0.5	-2	-2	0.5
Group 3	-0.5	-2	-2	0.5
Group 4	-1	0	-0.5	1.75
Group 5	-1	0	-0.5	1.75
Group 6	-1	0	-0.5	1.75
Group 7	-2	2	1.5	3
Group 8	-2	2	1.5	3
Group 9	-2	2	1.5	3
Group 10	-2	2	1.5	3

Group \ Group	1	2	3	4	5	6	7	8	9	10
1	1									
2	0.9333	1								
3	0.8933	0.9467	1							
4	0	0	0	1						
5	0	0	0	0.9533	1					
6	0	0	0	0.9000	0.9467	1				
7	0	0	0	0	0	0	1			
8	0	0	0	0	0	0	0.9000	1		
9	0	0	0	0	0	0	0.8333	0.9333	1	
10	0	0	0	0	0	0	0.7133	0.8000	0.8533	1

Table 3.11: Results of the risk classes identified by the Wald test with $\alpha^* = 0.05$ and $\alpha^* = 0.05/(G-1)$ for Simulation Outcome 2. The data is generated under the stratified Cox regression model (3.2), and $n = 10{,}000$. The correct (data-generating) class is **bolded**.

$\alpha^* = 0.05$			
Censoring Parameter Setting I: $\lambda_C = 0.5$		**Censoring Parameter Setting II**: $\lambda_C = 1.5$	
Updated Groups	Proportion	Updated Groups	Proportion
1, 1, 1, 2, 2, 2, 3, 3, 3, 3	**0.6200**	**1, 1, 1, 2, 2, 2, 3, 3, 3, 3**	**0.5733**
1, 1, 1, 2, 2, 2, 3, 4, 4, 4	0.0667	1, 1, 1, 2, 2, 2, 3, 3, 3, 4	0.1000
1, 1, 1, 2, 2, 2, 3, 3, 3, 4	0.0533	1, 1, 1, 2, 2, 2, 3, 4, 4, 4	0.0733
1, 2, 2, 3, 3, 3, 4, 4, 4, 4	0.0467	1, 1, 1, 2, 2, 3, 4, 4, 4, 4	0.0333
1, 1, 2, 3, 3, 3, 4, 4, 4, 4	0.0400	1, 1, 1, 2, 3, 3, 4, 4, 4, 4	0.0333
1, 1, 1, 2, 2, 3, 4, 4, 4, 4	0.0400	1, 1, 2, 3, 3, 3, 4, 4, 4, 4	0.0333
1, 1, 1, 2, 3, 3, 4, 4, 4, 4	0.0267	1, 1, 1, 2, 2, 2, 3, 3, 4, 4	0.0267
1, 1, 1, 2, 2, 2, 3, 3, 4, 4	0.0200	1, 2, 2, 3, 3, 3, 4, 4, 4, 4	0.0200
1, 1, 1, 2, 2, 2, 3, 4, 4, 5	0.0133	1, 1, 1, 2, 2, 2, 3, 4, 4, 5	0.0133
1, 1, 1, 2, 2, 3, 4, 5, 5, 5	0.0133	1, 2, 2, 3, 3, 3, 4, 4, 5, 5	0.0133
1, 1, 1, 2, 3, 3, 4, 4, 4, 5	0.0133	1, 1, 1, 2, 2, 2, 3, 3, 4, 5	0.0133
1, 1, 1, 2, 2, 2, 3, 3, 4, 5	0.0067	1, 2, 3, 4, 4, 4, 5, 5, 5, 5	0.0133
1, 2, 2, 3, 3, 4, 5, 5, 5, 5	0.0067	1, 2, 2, 3, 3, 3, 4, 5, 5, 5	0.0067
1, 2, 3, 4, 4, 4, 5, 5, 5, 5	0.0067	1, 1, 1, 2, 2, 3, 4, 4, 4, 5	0.0067
1, 1, 2, 3, 3, 3, 4, 4, 5, 5	0.0067	1, 1, 1, 2, 2, 3, 4, 5, 5, 5	0.0067
Other	0.0199	Other	0.0335
$\alpha^* = 0.05/(G-1)$			
Censoring Parameter Setting I: $\lambda_C = 0.5$		**Censoring Parameter Setting II**: $\lambda_C = 1.5$	
Updated Groups	Proportion	Updated Groups	Proportion
1, 1, 1, 2, 2, 2, 3, 3, 3, 3	**0.9533**	**1, 1, 1, 2, 2, 2, 3, 3, 3, 3**	**0.9333**
1, 1, 1, 2, 2, 2, 3, 4, 4, 4	0.0133	1, 1, 1, 2, 2, 2, 3, 3, 3, 4	0.0267
1, 1, 1, 2, 3, 3, 4, 4, 4, 4	0.0133	1, 1, 1, 2, 2, 2, 3, 4, 4, 4	0.0133
1, 1, 1, 2, 2, 2, 3, 3, 3, 4	0.0067	1, 1, 1, 2, 2, 2, 3, 3, 4, 4	0.0133
1, 1, 1, 2, 2, 3, 4, 4, 4, 4	0.0067	1, 1, 1, 2, 2, 2, 3, 3, 4, 5	0.0067
1, 1, 1, 2, 2, 2, 3, 3, 4, 4	0.0067	1, 1, 1, 2, 3, 3, 4, 4, 4, 4	0.0067

Table 3.12: Results of the risk classes identified by the Wald test with $\alpha^* = 0.05$ and $\alpha^* = 0.05/(G-1)$ for Simulation Outcome 3. The data is generated under a misspecified stratified Cox regression model, and $n = 10{,}000$. The correct (data-generating) class is **bolded**.

$\alpha^* = 0.05$			
Censoring Parameter Setting I: $\lambda_C = 0.5$		**Censoring Parameter Setting II**: $\lambda_C = 1.5$	
Updated Groups	Proportion	Updated Groups	Proportion
1, 1, 1, 2, 2, 2, 3, 3, 3, 3	**0.6667**	**1, 1, 1, 2, 2, 2, 3, 3, 3, 3**	**0.6333**
1, 1, 1, 2, 2, 2, 3, 3, 3, 4	0.1067	1, 1, 1, 2, 2, 2, 3, 3, 3, 4	0.0733
1, 1, 1, 2, 2, 2, 3, 3, 4, 4	0.0467	1, 1, 1, 2, 2, 2, 3, 4, 4, 4	0.0600
1, 1, 1, 2, 2, 2, 3, 4, 4, 4	0.0467	1, 1, 2, 3, 3, 3, 4, 4, 4, 4	0.0600
1, 2, 2, 3, 3, 3, 4, 4, 4, 4	0.0267	1, 1, 1, 2, 2, 2, 3, 3, 4, 4	0.0533
1, 1, 1, 2, 2, 2, 3, 3, 4, 5	0.0200	1, 1, 1, 2, 3, 3, 4, 4, 4, 4	0.0333
1, 1, 1, 2, 3, 3, 4, 4, 4, 4	0.0200	1, 2, 2, 3, 3, 3, 4, 4, 4, 4	0.0200
1, 1, 1, 2, 2, 2, 3, 4, 5, 5	0.0133	1, 1, 1, 2, 2, 3, 4, 4, 4, 5	0.0133
1, 2, 2, 3, 3, 3, 4, 4, 4, 5	0.0067	1, 2, 3, 4, 4, 4, 5, 5, 5, 5	0.0133
1, 1, 1, 2, 2, 3, 4, 4, 4, 4	0.0067	1, 1, 2, 3, 3, 3, 4, 5, 5, 5	0.0067
1, 1, 1, 2, 3, 3, 4, 4, 4, 5	0.0067	1, 1, 1, 2, 3, 3, 4, 4, 4, 5	0.0067
1, 1, 2, 3, 3, 3, 4, 4, 4, 4	0.0067	1, 2, 2, 3, 3, 4, 5, 6, 6, 6	0.0067
1, 1, 2, 3, 3, 3, 4, 5, 5, 5	0.0067	1, 1, 2, 3, 4, 4, 5, 5, 5, 5	0.0067
1, 1, 2, 3, 3, 4, 5, 6, 6, 6	0.0067	1, 1, 1, 2, 3, 4, 5, 5, 5, 5	0.0067
1, 1, 1, 2, 3, 4, 5, 5, 5, 5	0.0067	1, 1, 1, 2, 2, 3, 4, 4, 4, 4	0.0067
Other	0.0063		
$\alpha^* = 0.05/(G-1)$			
Censoring Parameter Setting I: $\lambda_C = 0.5$		**Censoring Parameter Setting II**: $\lambda_C = 1.5$	
Updated Groups	Proportion	Updated Groups	Proportion
1, 1, 1, 2, 2, 2, 3, 3, 3, 3	**0.9400**	**1, 1, 1, 2, 2, 2, 3, 3, 3, 3**	**0.9533**
1, 1, 1, 2, 2, 2, 3, 3, 4, 4	0.0200	1, 1, 1, 2, 2, 2, 3, 3, 3, 4	0.0333
1, 1, 1, 2, 2, 2, 3, 3, 3, 4	0.0133	1, 1, 1, 2, 2, 2, 2, 3, 3, 3	0.0067
1, 1, 1, 2, 3, 3, 4, 4, 4, 5	0.0067	1, 2, 3, 4, 4, 4, 5, 5, 5, 5	0.0067
1, 1, 1, 2, 3, 3, 4, 4, 4, 4	0.0067		
1, 1, 1, 2, 2, 2, 3, 4, 4, 4	0.0067		
1, 2, 3, 4, 4, 4, 5, 5, 5, 5	0.0067		
1, 1, 1, 2, 2, 2, 3, 3, 3, 4	0.0333		

Chapter 4

Developing a Predictive Survival Model with Administrative Service Utilization Records

4.1 Introduction

To best target their patient's specific needs, many clinicians utilize survival predictions as a tool to find personalized treatment regimens. The challenge however is that the corresponding survival probabilities are meaningless whenever internal covariates are included in the hazard regression model. A popular strategy to overcome such challenges is to conduct joint modelling, but implementing such an analysis can be challenging due to its heavy computation costs. As shown in Chapter 2, projecting the distribution of random effects to the available data space and simplifying the distribution of the random effect greatly reduces the computational costs, but it places a very restrictive parametric assumption. Unless the assumption is satisfied, which is highly unlikely in practice, the resulting inference can be misleading. In this chapter, we revisit our modelling from Chapter 2, and avoid placing any parametric assumptions on the random effect's distribution. We extend the conditional score approach of Tsiatis and Davidian (2001) by allowing successive observations in the longitudinal sub-model in the joint model to be correlated. This procedure retains the computational simplicity from the approach presented in Chapter 2, and also provides consistent estimates of model parameters. It naturally produces survival predictions by summarizing an individual's service utilization, such as OAT usage in our application.

This chapter is organized as follows. Section 4.2 presents the notation that will be utilized, and presents the longitudinal and hazard sub-models that comprise the joint modelling. Section 4.3 outlines the conditional score approach of Tsiatis and Davidian (2001), and its extension for correlated observations. We also estimate correlation parameters, summarize our proposed method through a two-stage inference procedure, and outline how to obtain survival predictions. We also outline a promising inference procedure that can be

more efficient than the proposed method in Section 4.3.6. The proposed conditional score estimator is applied to an analysis of the provincial administrative database in Section 4.4. We conduct a simulation study in Section 4.5 to verify our findings from our data analysis, and to show the performance of our proposed estimator relative over naive approaches. We conclude this chapter with a summary and discussion in Section 4.6.

4.2 Jointly Modelling OAT Dispensation and Mortality Risk Processes

Adopting the notation used in Chapter 2, let T denote an individual's survival time (measured in *days*) since their first recorded OAT dispensation record, which is subject to a right-censoring time, C. The available information on T is (T^*, δ), where $T^* = T \wedge C$ is the minimum between T and C, and $\delta = I(T \leq C)$. Let $Z(t) \in \{0, 1\}$ denote an individual's OAT dispensation indicator at time $t \geq 0$, which is obtained from daily pharmaceutical dispensation records, and $\mathcal{Z}(t) = \{Z(u) : 0 \leq u \leq t\}$. We additionally let $\boldsymbol{X}(t) = (X_1(t), \cdots, X_q(t))'$ denote external time-varying covariates and time-independent characteristics of an individual, and $\boldsymbol{\mathcal{X}}(t) = \{\boldsymbol{X}(u) : 0 \leq u \leq t\}$. Here, the covariate $X_k(t)$ is time-independent if $X_k(0) \equiv X_k(t)$, for all $t > 0$, and $k = 1, \cdots, q$. Our statistical goal is to estimate the conditional hazard function of T (at time t), given the processes $\mathcal{Z}(t)$ and $\boldsymbol{\mathcal{X}}(t)$:

$$\lambda(t; \mathcal{Z}(t), \boldsymbol{\mathcal{X}}(t)) = \lim_{dt \to 0^+} \frac{1}{dt} P(t \leq T < t + \Delta t | T \geq t, \mathcal{Z}(t), \boldsymbol{\mathcal{X}}(t)). \tag{4.1}$$

If we specify a model for (4.1) that is capable of providing survival predictions, we cannot directly use $\mathcal{Z}(t)$, since this would obstruct the conventional relationship between the hazard and survivor functions. The standard approach to overcome the problem is to model the joint distribution of T and $\mathcal{Z}(t)$, given $\boldsymbol{\mathcal{X}}(t)$. This is done by linking two sub-models together: one sub-model is for the process $\mathcal{Z}(\cdot)$, and the other sub-model is for (4.1). Since this eliminates the direct use of internal covariates within the modelling, a likelihood based estimating procedure can then be used to estimate the parameters of interest in the hazard sub-model.

To summarize the OAT dispensation process for individual $i = 1, \cdots, n$, let $R_i(t) = \int_0^t Z_i(u)du/t$ denote the proportion of time individual i is dispensed OAT over $[0, t]$. In practice, one may choose to use $\{R_i(t_{ij}) : j = 1, \cdots, m_i\}$, where $0 \leq t_{i,m_i} \leq T_i^*$ in their analysis. Using the simplified notation $R_{ij} = R_i(t_{ij})$, we specify the following model for the OAT dispensation history of individual i:

$$h(R_{ij}) = \nu_i + \varepsilon_{ij}, \tag{4.2}$$

for some function $h(\cdot)$, where ν_i is a subject-specific value that summarizes individual i's OAT dispensation history, and ε_{ij} captures the deviation between ν_i and the transformed measurements $h(R_{ij})$, with $\mathbb{E}(\varepsilon_{ij}) = 0$ and $Cov(\varepsilon_{ij}, \varepsilon_{i,j-k}) = \rho^k\sigma^2/(1-\rho^2)$. That is, $\{\varepsilon_{ij} : j = 1, \cdots, m_i\}$ follows an $AR(1)$ process. This specification will be justified in Section 4.4.

We model an individual's mortality hazard function with

$$\lambda(t; \mathcal{Z}_i(t), \boldsymbol{\mathcal{X}}_i(t)) = \lambda_0(t) \exp\{\gamma\nu_i + \boldsymbol{\theta}'\boldsymbol{X}_i(t)\} \quad (i = 1, \cdots, n), \tag{4.3}$$

where $\lambda_0(t)$ is an unspecified baseline hazard function, and γ and $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_q)'$ are unknown regression parameters. Note that the model in (4.3) explicitly assumes that $T_i \perp\!\!\!\perp \mathcal{Z}_i(\infty)|\{\nu_i, \boldsymbol{\mathcal{X}}_i(t)\}$, meaning that we assume ν_i serves as an adequate summary of the process $\mathcal{Z}_i(\cdot)$ in the presence of $\boldsymbol{\mathcal{X}}_i(t)$. As we summarize $\mathcal{Z}_i(\cdot)$ with ν_i in (4.3), this preserves the conventional relationship between the hazard and survivor functions. Following the arguments presented in Section 1.1.3, we can show that

$$P(T_i > t|\mathcal{Z}_i(t), \boldsymbol{\mathcal{X}}_i(t)) = \exp\left\{-\int_0^t \lambda_0(u) \exp\{\gamma\nu_i + \boldsymbol{\theta}'\boldsymbol{X}_i(u)\}du\right\}.$$

A likelihood based procedure can be used to infer the unknown regression parameters. Here, ν_i links the models (4.2) and (4.3) together. The challenge of inferring parameters in (4.3) is that ν_i is an unknown quantity defined by (4.2). We proceed to outline the proposed inference procedure under the model assumptions.

4.3 Inference Procedure

4.3.1 Conditional Score Based Approach

With reference to the likelihood function (2.4),Tsiatis and Davidian (2001) specified the random effect distribution as $F(\nu_i) = I(\nu_i \geq \hat{\nu}_i)$, where $\hat{\nu}_i = \sum_{j=1}^{m_i} h(R_{ij})/m_i$ is an unbiased estimator of ν_i under (4.2). Since $\hat{\nu}_i$ can be viewed as a "noisy measurement" of ν_i, directly replacing ν_i with $\hat{\nu}_i$ will generally produce biased estimates (Prentice 1982). To correct for this bias, Tsiatis and Davidian (2001) extended the conditional score approach for generalized linear models (Stefanski and Carroll 1987) to the extended Cox proportional hazards model, as in (4.3). This was done by deriving a *sufficient statistic* of ν_i, which depends on either known or estimable quantities. Their derivation assumes that the error terms in (4.2) are independent and identically distributed, which is an assumption likely violated in practice. We extend the approach by allowing for the error terms to follow an $AR(1)$ process.

By letting $n_i(t) = I(T_i \in [t, t+dt))$ and $Y_i(t) = I(T_i^* \geq t)$, the conditional likelihood of $\{n_i(t), \hat{\nu}_i\}$ given $\{Y_i(t) = 1, \nu_i, \boldsymbol{\mathcal{X}}_i(t)\}$ can be expressed up to order dt as

$$\exp\{\frac{\nu_i}{V_i(\sigma^2,\rho)}(\hat{\nu}_i + \gamma V_i(\sigma^2,\rho) n_i(t))\} \frac{(\lambda_0(t)\exp\{\boldsymbol{\theta}'\boldsymbol{X}_i(t)\}dt)^{n_i(t)}}{\sqrt{2\pi V_i(\sigma^2,\rho)}} \exp\{\frac{-\hat{\nu}_i^2 - \nu_i^2}{2V_i(\sigma^2,\rho)}\}, \quad (4.4)$$

where

$$V_i(\sigma^2,\rho) = Var(\hat{\nu}_i) = \frac{1}{m_i^2}\frac{\sigma^2}{1-\rho^2}\left(m_i + \frac{2\rho}{1-\rho}\sum_{j=1}^{m_i-1}(1-\rho^{m_i-j})\right).$$

Here, it is assumed that (i) $n_i(t) \perp\!\!\!\perp \hat{\nu}_i | \{Y_i(t) = 1, \boldsymbol{\mathcal{X}}_i(t), \nu_i\}$, and (ii) $\hat{\nu}_i \perp\!\!\!\perp \{Y_i(t) = 1, \boldsymbol{\mathcal{X}}_i(t)\} | \nu_i$. The first assumption is essentially the "nondifferential measurement error mechanism" assumption in the measurement error literature (Carroll *et al.* 2006, Yi 2017), and the second assumption results from our modelling assumption in (4.2). We recognize (4.4) to be a member of the exponential family of probability distributions. Thus

$$S_i(t) \equiv S_i(t,\gamma,\sigma^2,\rho) = \hat{\nu}_i + \gamma V_i(\sigma^2,\rho) n_i(t) \quad (4.5)$$

is a sufficient statistic for ν_i. Following the arguments presented in Tsiatis and Davidian (2001), we can show that

$$\begin{aligned}\lambda^\dagger(t; \mathcal{S}_i(t), \boldsymbol{\mathcal{X}}_i(t)) &= \lim_{\Delta t \to 0^+} \frac{1}{\Delta t} P(n_i(t) = 1 | Y_i(t) = 1, \mathcal{S}_i(t), \boldsymbol{\mathcal{X}}_i(t)) \\ &= \lambda_0(t)\exp\{\gamma S_i(t) - \frac{\gamma^2}{2}V_i(\sigma^2,\rho) + \boldsymbol{\theta}'\boldsymbol{X}_i(t)\}, \end{aligned} \quad (4.6)$$

where $\mathcal{S}_i(t) = \{S_i(u) : 0 \leq u \leq t\}$. Note that (4.6) involves only known or estimable quantities. Furthermore, the parameters in (4.6) are the same as in (4.3). This motivates us to base our estimation of $\lambda_0(\cdot)$, γ, and $\boldsymbol{\theta}$ on (4.6).

4.3.2 Estimating Model Parameters

Let $N_i(t) = I(T_i \leq t)$, $Y_i(t) = I(T_i^* \geq t)$, and $E_r(t,\gamma,\boldsymbol{\theta},\sigma^2,\rho) = \sum_{j=1}^n E_{rj}(t,\gamma,\boldsymbol{\theta},\sigma^2,\rho)$ with

$$E_{rj}(t,\gamma,\boldsymbol{\theta},\sigma^2,\rho) = Y_j(t)\exp\{\gamma S_j(t) - \frac{\gamma^2}{2}V_j(\sigma^2,\rho) + \boldsymbol{\theta}'\boldsymbol{X}_j(t)\}(S_j(t), \boldsymbol{X}_j(t)')'^{\otimes r},$$

where for a column-vector $\boldsymbol{a}$, $\boldsymbol{a}^{\otimes 0} = 1$ and $\boldsymbol{a}^{\otimes 1} = \boldsymbol{a}$. For the time being, assume that σ^2 and ρ are known (see Section 4.3.3), so that we can compute the sufficient statistic in (4.5). Under the model in (4.6), the following estimating equations are (conditionally) unbiased

for all $t > 0$:

$$\sum_{i=1}^{n}(Y_i(t)dN_i(t) - E_{0i}(t,\gamma,\boldsymbol{\theta},\sigma^2,\rho)d\Lambda_0(t)) = 0 \text{ and} \tag{4.7}$$

$$\sum_{i=1}^{n}\int_0^{\infty}(S_i(t),\boldsymbol{X}_i(t)')'(Y_i(t)dN_i(t) - E_{0i}(t,\gamma,\boldsymbol{\theta},\sigma^2,\rho)d\Lambda_0(t)) = \mathbf{0}, \tag{4.8}$$

where $d\Lambda_0(t) = \lambda_0(t)\ dt$. Solving for $d\Lambda_0(t)$ in (4.7) results in the following Breslow-like estimator

$$d\hat{\Lambda}_0(t;\gamma,\boldsymbol{\theta},\sigma^2,\rho) = \sum_{i=1}^{n}\frac{Y_i(t)dN_i(t)}{E_0(t,\gamma,\boldsymbol{\theta},\sigma^2,\rho)}. \tag{4.9}$$

Inserting (4.9) in place of $d\Lambda_0(t)$ in (4.8) yields

$$U_1(\gamma,\boldsymbol{\theta};\sigma^2,\rho) = \sum_{i=1}^{n}\int_0^{\infty}Y_i(t)\left[(S_i(t),\boldsymbol{X}_i(t)')' - \frac{E_1(t,\gamma,\boldsymbol{\theta},\sigma^2,\rho)}{E_0(t,\gamma,\boldsymbol{\theta},\sigma^2,\rho)}\right]dN_i(t).$$

We propose to estimate $(\gamma,\boldsymbol{\theta}')'$ with $(\hat{\gamma},\hat{\boldsymbol{\theta}}')'$, the solution to $U_1(\gamma,\boldsymbol{\theta};\sigma^2,\rho) = \mathbf{0}$. To estimate $d\Lambda_0(t)$, we can then use (4.9) after replacing γ and $\boldsymbol{\theta}$ with $\hat{\gamma}$ and $\hat{\boldsymbol{\theta}}$, respectively.

4.3.3 Estimating Additional Parameters

By assuming that $\{\varepsilon_{ij} : j = 1,\cdots,m_i\}$ follows an $AR(1)$ process for fixed i in (4.2), we have

$$h(R_{ij}) - \nu_i = \rho[h(R_{i,j-1}) - \nu_i] + e_{ij}. \tag{4.10}$$

where $\mathbb{E}(e_{ij}) = 0$ and $Var(e_{ij}) = \sigma^2$, and $e_{ij} \perp\!\!\!\perp \varepsilon_{i,j-1}$. Since ν_i is unavailable, we estimate it with $\hat{\nu}_i$, and proceed to estimate ρ with the least squares estimator, under (4.10), and estimate σ^2 based on the residual sum of squares under (4.10), upon pooling all the individual's information together. Specifically, $\hat{\rho}$ is the solution to the following (unbiased) estimating equation

$$U_2(\rho) = \sum_{i=1}^{n}\sum_{j=2}^{m_i}\{h(R_{ij}) - \hat{\nu}_i - \rho[h(R_{i,j-1}) - \hat{\nu}_i][h(R_{i,j-1}) - \hat{\nu}_i]\} = 0, \tag{4.11}$$

and $\hat{\sigma}^2$ is the solution to the following (unbiased) estimating equation

$$U_3(\sigma^2,\rho) = \sigma^2 - \frac{1}{|\mathcal{I}|}\sum_{i\in\mathcal{I}}\frac{1}{m_i-2}\sum_{j=2}^{m_i}\{h(R_{ij}) - \hat{\nu}_i - \rho[h(R_{i,j-1}) - \hat{\nu}_i]\}^2 = 0, \tag{4.12}$$

where $\mathcal{I} = \{i : m_i > 2\}$ and $|\mathcal{I}|$ is the size of $\mathcal{I}$. Essentially, (4.11) is proportional to the derivative of $\sum_{i=1}^{n}\sum_{j=2}^{m_i} e_{ij}^2$ with respect to ρ, where e_{ij} is from (4.10). Solving (4.12) for σ^2 results in an estimator that averages the conventional regression error variance estimator among all individuals with at least two observations.

4.3.4 Two-Stage Estimation Procedure and Variance Estimation

By stacking the estimating functions, let $\boldsymbol{U}(\gamma, \boldsymbol{\theta}, \sigma^2, \rho) = (U_1(\gamma, \boldsymbol{\theta}, ; \sigma^2, \rho)', U_{23}(\sigma^2, \rho)')'$ with $U_{23}(\sigma^2, \rho) = (U_2(\rho), U_3(\sigma^2, \rho))'$, where we can express each estimating function as the sum of each individual's contribution

$$U_1(\gamma, \boldsymbol{\theta}; \sigma^2, \rho) = \sum_{i=1}^{n} U_{i1}(\gamma, \boldsymbol{\theta}; \sigma^2, \rho)$$
$$U_{23}(\sigma^2, \rho) = \sum_{i=1}^{n} U_{i,23}(\sigma^2, \rho).$$

Since $\hat{\sigma}^2$ and $\hat{\rho}$ have analytic forms, we infer all of the parameters with the following two-stage procedure:

Step 1: Under the model $h(R_{ij}) = \nu_i + \varepsilon_{ij}$, estimate ν_i with $\hat{\nu}_i = \sum_{j=1}^{m_i} h(R_{ij})/m_i$, and proceed to estimate ρ and σ^2 by solving the equation $U_{23}(\sigma^2, \rho) = \mathbf{0}$.

Step 2: Estimate γ and $\boldsymbol{\theta}$ with the solutions to $U_1(\gamma, \boldsymbol{\theta}; \hat{\sigma}^2, \hat{\rho}) = \mathbf{0}$. Also, estimate $d\Lambda_0(t)$ with $d\hat{\Lambda}_0(t; \hat{\gamma}, \hat{\boldsymbol{\theta}}, \hat{\sigma}^2, \hat{\rho})$.

Letting $\boldsymbol{\Omega} = (\gamma, \boldsymbol{\theta}')'$ and $\hat{\boldsymbol{\Omega}} = (\hat{\gamma}, \hat{\boldsymbol{\theta}}')'$, the resulting estimate of $\boldsymbol{\Omega}$ under our two-stage inference procedure, it follows that $Var(\hat{\boldsymbol{\Omega}}) = \boldsymbol{\Psi}^{-1}\boldsymbol{\Phi}\boldsymbol{\Psi}^{-1}$ (c.f. Chapter 1 of Yi (2017)), where $\boldsymbol{\Psi} = \mathbb{E}\left(\frac{\partial}{\partial \boldsymbol{\Omega}} U_{11}(\boldsymbol{\Omega}; \sigma^2, \rho)\right)$ and $\boldsymbol{\Phi} = \mathbb{E}(Q_1(\boldsymbol{\Omega}, \sigma^2, \rho) Q_1(\boldsymbol{\Omega}, \sigma^2, \rho)')$ with

$$Q_i(\boldsymbol{\Omega}, \sigma^2, \rho) = U_{i1}(\boldsymbol{\Omega}; \sigma^2, \rho) - \mathbb{E}\left(\frac{\partial}{\partial(\sigma^2, \rho)'} U_{i1}(\boldsymbol{\Omega}; \sigma^2, \rho)\right)\left[\mathbb{E}\left(\frac{\partial}{\partial(\sigma^2, \rho)'} U_{i,23}(\sigma^2, \rho)\right)\right]^{-1} U_{i,23}(\sigma^2, \rho).$$

By replacing $\boldsymbol{\Psi}$ and $\boldsymbol{\Phi}$ with their empirical means: $\hat{\boldsymbol{\Psi}} = n^{-1}\sum_{i=1}^{n} \frac{\partial}{\partial \boldsymbol{\Omega}} U_{i1}(\boldsymbol{\Omega}; \hat{\sigma}^2, \hat{\rho})\Big|_{\boldsymbol{\Omega}=\hat{\boldsymbol{\Omega}}}$ and $\hat{\boldsymbol{\Phi}} = n^{-1}\sum_{i=1}^{n} \hat{Q}_i(\hat{\boldsymbol{\Omega}}, \hat{\sigma}^2, \hat{\rho})\hat{Q}_i(\hat{\boldsymbol{\Omega}}, \hat{\sigma}^2, \hat{\rho})'$ with

$$\hat{Q}_i(\boldsymbol{\Omega}, \sigma^2, \rho) = U_{i1}(\boldsymbol{\Omega}; \sigma^2, \rho) - \left[\frac{\partial}{\partial(\sigma^2, \rho)'} U_1(\boldsymbol{\Omega}; \sigma^2, \rho)\right]\left[\frac{\partial}{\partial(\sigma^2, \rho)'} U_{23}(\sigma^2, \rho)\right]^{-1} U_{i,23}(\sigma^2, \rho),$$

we have $\widehat{Var}(\hat{\boldsymbol{\Omega}}) = \hat{\boldsymbol{\Psi}}^{-1}\hat{\boldsymbol{\Phi}}\hat{\boldsymbol{\Psi}}^{-1}$, a consistent estimator for $Var(\hat{\boldsymbol{\Omega}})$.

4.3.5 Survival Prediction

Similar to (1.22), we can predict an individual's survival risk based on the external covariate $\boldsymbol{X}_i(\cdot)$, and the observed process $\mathcal{S}_i(C_i)$. Note that this individual does not necessarily need to be a study subject. We can use (4.6) to show for any $t > C_i$ that

$$\begin{aligned}\pi(t; \mathcal{S}_i(C_i), \boldsymbol{\mathcal{X}}_i(t)) &= P(T_i > t | T_i > C_i, \mathcal{S}_i(C_i), \boldsymbol{\mathcal{X}}_i(t)) \\ &= \exp\{-\int_{C_i}^{t} \lambda^{\dagger}(u; \mathcal{S}_i(C_i), \boldsymbol{\mathcal{X}}_i(u)) du\}, \end{aligned} \tag{4.13}$$

so that the model-based survival risk predictions are obtainable upon estimating $\lambda_0(\cdot)$, $\boldsymbol{\Omega} = (\gamma, \boldsymbol{\theta}')'$, σ^2, and ρ. Specifically, a point and interval estimate for individual i's survival probability at time $t > C_i$ can be obtained with the following algorithm:

Step 1: Based on the observed information over $[0, C_i]$, obtain $\hat{\nu}_i$.

Step 2: Letting $\hat{S}_i(t) = S_i(t, \hat{\gamma}, \hat{\sigma}^2, \hat{\rho})$ and $\hat{\mathcal{S}}_i(t) = \{\hat{S}_i(u) : 0 \leq u \leq t\}$, obtain survival probabilities with (4.13), where we replace $\mathcal{S}_i(t)$ with $\hat{\mathcal{S}}_i(t)$, and the parameters in (4.6) with their corresponding estimate.

Step 3: A $100(1-\alpha^*)\%$ confidence interval for $\pi(t; \mathcal{S}_i(C_i), \boldsymbol{\mathcal{X}}_i(t))$ can be obtained from the $\alpha^*/2$ and $1-\alpha^*/2$ percentiles of $\{\hat{\pi}^{(j)}(t; \hat{\mathcal{S}}_i^{(j)}(C_i), \boldsymbol{\mathcal{X}}_i(t)) : j = 1, \cdots, M\}$ for some $M >> 1$, where

$$\hat{\boldsymbol{\Omega}}^{(j)} = (\hat{\gamma}^{(j)}, \hat{\boldsymbol{\theta}}^{(j)'})' \sim \mathcal{N}(\hat{\boldsymbol{\Omega}}, \widehat{Var}(\hat{\boldsymbol{\Omega}})),$$

the baseline hazard function is estimated with (4.9) and $\boldsymbol{\Omega}$ replaced with $\hat{\boldsymbol{\Omega}}^{(j)}$, $\hat{S}_i^{(j)}(t) = S_i(t, \hat{\gamma}^{(j)}, \hat{\sigma}^2, \hat{\rho})$, $\hat{\mathcal{S}}_i^{(j)}(t) = \{\hat{S}_i^{(j)}(u) : 0 \leq u \leq t\}$, and $\hat{\pi}^{(j)}(t; \hat{\mathcal{S}}_i^{(j)}(C_i), \boldsymbol{\mathcal{X}}_i(t))$ is the jth estimate of (4.13). The asymptotic normality of the estimator $\hat{\boldsymbol{\Omega}}$ carries over to our setting from the results derived by Tsiatis and Davidian (2001).

Similar to so-called "dynamic predictions" (Rizopoulos 2011), the survival probability predictions can be updated once more information from individual i becomes available at time $C_i^* > C_i$, in which we simply repeat Steps 1-3 with C_i^* replacing C_i.

In addition to the survival probability predictions, we can directly predict when the event will occur with

$$\begin{aligned}\mathbb{E}(T_i | T_i > C_i, \mathcal{S}_i(C_i), \boldsymbol{\mathcal{X}}_i(\infty)) &= C_i + \int_{C_i}^{\infty} \frac{P(T_i > t | \mathcal{S}_i(C_i), \boldsymbol{\mathcal{X}}_i(t))\ dt}{P(T_i > C_i | \mathcal{S}_i(C_i), \boldsymbol{\mathcal{X}}_i(C_i))} \\ &= C_i + \int_{C_i}^{\infty} \exp\{-\int_{C_i}^{t} \lambda^{\dagger}(u; \mathcal{S}_i(C_i), \boldsymbol{\mathcal{X}}_i(u)) du\}\ dt, \end{aligned} \tag{4.14}$$

where the second term in (4.14) is the remaining life expectancy for individual i provided that they survived up to time C_i (Zhou and Sun 2022). Similarly to the survival probability

predictions, we can obtain a point and interval estimate for T_i with the following algorithm:

Step 1: Based on the observed information over $[0, C_i]$, obtain $\hat{\nu}_i$.

Step 2: Letting $\hat{S}_i(t) = S_i(t, \hat{\gamma}, \hat{\sigma}^2, \hat{\rho})$ and $\hat{\mathcal{S}}_i(t) = \{\hat{S}_i(u) : 0 \leq u \leq t\}$, estimate (4.14) by replacing $\mathcal{S}_i(t)$ with $\hat{\mathcal{S}}_i(t)$, and the parameters in (4.6) with their corresponding estimate.

Step 3: A $100(1-\alpha^*)\%$ prediction interval for T_i can be obtained from the $\alpha^*/2$ and $1-\alpha^*/2$ percentiles of $\{\hat{\mathbb{E}}^{(j)}(T_i|T_i > C_i, \hat{\mathcal{S}}_i^{(j)}(C_i), \boldsymbol{\mathcal{X}}_i(\infty)) : j = 1, \cdots, M\}$ for some $M >> 1$, where

$$\hat{\boldsymbol{\Omega}}^{(j)} = (\hat{\gamma}^{(j)}, \hat{\boldsymbol{\theta}}^{(j)\prime})' \sim \mathcal{N}(\hat{\boldsymbol{\Omega}}, \widehat{Var}(\hat{\boldsymbol{\Omega}})),$$

the baseline hazard function is estimated with (4.9) and $\boldsymbol{\Omega}$ replaced with $\hat{\boldsymbol{\Omega}}^{(j)}$, $\hat{S}_i^{(j)}(t) = S_i(t, \hat{\gamma}^{(j)}, \hat{\sigma}^2, \hat{\rho})$, $\hat{\mathcal{S}}_i^{(j)}(t) = \{\hat{S}_i^{(j)}(u) : 0 \leq u \leq t\}$, and $\hat{\mathbb{E}}^{(j)}(T_i|T_i > C_i, \hat{\mathcal{S}}_i^{(j)}(C_i), \boldsymbol{\mathcal{X}}_i(\infty))$ is the jth estimate of (4.14).

4.3.6 Alternative Procedure by Predicting the Random Effect

A drawback with estimating ν_i with $\hat{\nu}_i$ is that it only utilizes individual i's information, which can be problematic when m_i is small. An alternative approach that pools all individual's information together would hence be desirable.

In vector notation, we can express (4.2) as

$$\boldsymbol{h}_i = \mathbf{1}_i \nu_i + \boldsymbol{\varepsilon}_i = \mathbf{1}_i \mu + \mathbf{1}_i \eta_i + \boldsymbol{\varepsilon}_i, \tag{4.15}$$

where $\boldsymbol{h}_i = (h_i(R_{i1}), \cdots, h_i(R_{i,m_i}))'$, $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \cdots, \varepsilon_{i,m_i})'$, and $\nu_i = \mu + \eta_i$. Suppose we are willing to make the assumption that $\nu_i \sim \mathcal{N}(\mu, \sigma_\nu^2)$, and $\boldsymbol{\varepsilon}_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{C}_i(\sigma^2, \rho))$, where the (j,k) element of $\boldsymbol{C}_i(\sigma^2, \rho)$ is $Cov(\varepsilon_{ij}, \varepsilon_{ik}) = \rho^{|j-k|}\sigma^2/(1-\rho^2)$. Since (4.15) is a linear mixed effect model (Laird and Ware 1982), we can estimate μ with its maximum likelihood estimator

$$\hat{\mu}(\sigma^2, \rho, \sigma_\nu^2) = \left(\sum_{i=1}^{n} \mathbf{1}_i' \boldsymbol{\Sigma}_i^{-1}(\sigma^2, \rho, \sigma_\nu^2) \mathbf{1}_i \right)^{-1} \left(\sum_{i=1}^{n} \mathbf{1}_i' \boldsymbol{\Sigma}_i^{-1}(\sigma^2, \rho, \sigma_\nu^2) \boldsymbol{h}_i \right),$$

where $\boldsymbol{\Sigma}_i(\sigma^2, \rho, \sigma_\nu^2) = \sigma_\nu^2 \mathbf{1}_i \mathbf{1}_i' + \boldsymbol{C}_i(\sigma^2, \rho)$ and estimate $(\sigma^2, \rho, \sigma_\nu^2)$ via restricted maximum likelihood estimation:

$$
\begin{aligned}
(\hat{\sigma}^2, \hat{\rho}, \hat{\sigma}_\nu^2) &= \underset{(\sigma,\rho,\sigma_\nu^2)}{\operatorname{argmax}} \{\ell_R(\sigma^2, \rho, \sigma_\nu^2)\}, \\
\ell_R(\sigma^2, \rho, \sigma_\nu^2) &= -\frac{1}{2}\sum_{i=1}^{n} \log\det(\boldsymbol{\Sigma}_i(\sigma^2, \rho, \sigma_\nu^2)) - \frac{1}{2}\sum_{i=1}^{n} \log\det(\mathbf{1}_i' \boldsymbol{\Sigma}_i^{-1}(\sigma^2, \rho, \sigma_\nu^2)\mathbf{1}_i) \\
&\quad - \frac{1}{2}\sum_{i=1}^{n} [\boldsymbol{h}_i - \mathbf{1}_i \hat{\mu}(\sigma^2, \rho, \sigma_\nu^2)]' \boldsymbol{\Sigma}_i^{-1}(\sigma^2, \rho, \sigma_\nu^2)[\boldsymbol{h}_i - \mathbf{1}_i \hat{\mu}(\sigma^2, \rho, \sigma_\nu^2)].
\end{aligned}
$$

We can predict η_i with its best linear unbiased predictor

$$
\hat{\eta}_i = \hat{\sigma}_\nu^2 \mathbf{1}_i' \boldsymbol{\Sigma}_i^{-1}(\hat{\sigma}^2, \hat{\rho}, \hat{\sigma}_\nu^2)[\boldsymbol{h}_i - \mathbf{1}_i \hat{\mu}(\hat{\sigma}^2, \hat{\rho}, \hat{\sigma}_\nu^2)].
$$

This results in predicting ν_i with $\tilde{\nu}_i = \hat{\mu} + \hat{\eta}_i$. Since $\tilde{\nu}_i$ uses information from other individuals through $\hat{\sigma}^2$, $\hat{\rho}$, and $\hat{\sigma}_\nu^2$, we anticipate $\tilde{\nu}_i$ to be more efficient compared to $\hat{\nu}_i$, especially when m_i is small. We can then proceed with the conditional score inference procedure to estimate the regression parameters in (4.3) by updating the sufficient statistic in (4.5) into

$$
S_i^*(t) \equiv S_i^*(t, \gamma, \sigma^2, \rho, \sigma_\nu^2) = \hat{\nu}_i + \gamma V_i(\sigma^2, \rho, \sigma_\nu^2) n_i(t),
$$

where

$$
\begin{aligned}
V_i(\sigma^2, \rho, \sigma_\nu^2) &= Var(\eta_i) - Var(\hat{\eta}_i) \\
&= \sigma_\nu^2 (1 - \sigma_\nu^2 \mathbf{1}_i' \boldsymbol{\Sigma}_i(\sigma^2, \rho, \sigma_\nu^2)\mathbf{1}_i)
\end{aligned}
$$

is the extra variability induced by replacing η_i with $\hat{\eta}_i$. The two-stage inference procedure and variance estimation procedure can be updated using the stacked estimating functions

$$
\boldsymbol{U}(\gamma, \boldsymbol{\theta}, \sigma^2, \rho, \sigma_\nu^2) = (U_1(\gamma, \boldsymbol{\theta}; \sigma^2, \rho, \sigma_\nu^2)', U_2(\sigma^2, \rho, \sigma_\nu^2)')',
$$

where

$$
\begin{aligned}
U_1(\gamma, \boldsymbol{\theta}; \sigma^2, \rho, \sigma_\nu^2) &= \sum_{i=1}^{n} U_{i1}(\gamma, \boldsymbol{\theta}; \sigma^2, \rho, \sigma_\nu^2) \\
&= \sum_{i=1}^{n} \int_0^\infty Y_i(t) \left((S_i^*(t), \boldsymbol{X}_i(t)')' - \frac{E_1(t, \gamma, \boldsymbol{\theta}, \sigma^2, \rho, \sigma_\nu^2)}{E_0(t, \gamma, \boldsymbol{\theta}, \sigma^2, \rho, \sigma_\nu^2)} \right) dN_i(t), \\
E_r(t, \gamma, \boldsymbol{\theta}, \sigma^2, \rho, \sigma_\nu^2) &= \sum_{j=1}^{n} E_{rj}(t, \gamma, \boldsymbol{\theta}, \sigma^2, \rho, \sigma_\nu^2) \\
&= \sum_{j=1}^{n} Y_j(t) \exp\{\gamma S_j^*(t) - \frac{\gamma^2}{2} V_j(\sigma^2, \rho, \sigma_\nu^2) + \boldsymbol{\theta}' \boldsymbol{X}_j(t)\} (S_j^*(t), \boldsymbol{X}_j(t)')'^{\otimes r} \text{ for all } t > 0, \\
U_2(\sigma^2, \rho, \sigma_\nu^2) &= \frac{\partial}{\partial(\sigma^2, \rho, \sigma_\nu^2)} \ell_R(\sigma^2, \rho, \sigma_\nu^2) = \sum_{i=1}^{n} U_{i2}(\sigma^2, \rho, \sigma_\nu^2) \\
\hat{\boldsymbol{\Psi}} &= \frac{1}{n} \sum_{i=1}^{n} \frac{\partial}{\partial \boldsymbol{\Omega}} U_{i1}(\boldsymbol{\Omega}; \hat{\sigma}^2, \hat{\rho}, \hat{\sigma}_\nu^2) \Big|_{\boldsymbol{\Omega} = \hat{\boldsymbol{\Omega}}}, \text{ and} \\
\hat{\boldsymbol{\Phi}} &= \frac{1}{n} \sum_{i=1}^{n} \hat{Q}_i(\hat{\boldsymbol{\Omega}}, \hat{\sigma}^2, \hat{\rho}, \hat{\sigma}_\nu^2) \hat{Q}_i(\hat{\boldsymbol{\Omega}}, \hat{\sigma}^2, \hat{\rho}, \hat{\sigma}_\nu^2)'
\end{aligned}
$$

with $\hat{Q}_i(\boldsymbol{\Omega}, \sigma^2, \rho, \sigma_\nu^2)$ defined as

$$
U_{i1}(\boldsymbol{\Omega}; \sigma^2, \rho) - \left[\frac{\partial}{\partial(\sigma^2, \rho, \sigma_\nu^2)'} U_1(\boldsymbol{\Omega}; \sigma^2, \rho, \sigma_\nu^2) \right] \left[\frac{\partial}{\partial(\sigma^2, \rho, \sigma_\nu^2)'} U_2(\sigma^2, \rho, \sigma_\nu^2) \right]^{-1} U_{i2}(\sigma^2, \rho, \sigma_\nu^2).
$$

Not only can this approach result in improving the statistical efficiency by estimating $\boldsymbol{\Omega}$, but this framework allows us to consider alternative correlation structures, such as $AR(p)$ with $p > 1$, or exponential.

4.4 Analysis of the Provincial OAT Dispensation Records (II)

We applied the proposed estimation procedure outlined in Section 4.3 to the provincial administrative database pertaining to OAT dispensation records. We include the same set of risk factors as in the analysis of Chapter 2.

We specified $h(x) = logit(x)$ so that both sides of the equation in (4.2) is unconstrained, and began our analysis by obtaining $\hat{\nu}_i$ under model (4.2) for each individual. Some individuals had OAT dispensations cover every day they were in the study, in which we defined $logit(1) \coloneqq 20$. To explore the potential autocorrelation within the error terms in (4.2), Figure 4.1 illustrates the sample autocorrelation and partial autocorrelation functions of the residuals $\hat{e}_{ij} = logit(R_{ij}) - \hat{\nu}_i$ for four randomly selected individuals. The time unit for measurements in (4.2) was in *months*, as this resulted in the time series processes to be approximately stationary. The findings from Figure 4.1 motivated the $AR(1)$ specification

in (4.2). Table 4.1 presents the resulting estimates of γ and $\boldsymbol{\theta}$, where we considered *time since first recorded OAT dispensation* and *age* as time scales. The four sets of estimates correspond to the following procedures:

(a) Estimates corresponding to maximizing the joint likelihood function (2.4) upon specifying $\nu_i \sim N(\mu, \sigma_\nu^2)$.

(b) Estimates obtained by directly replacing ν_i with $\hat{\nu}_i$ in (4.3).

(c) Estimates obtained from the conditional score approach as described in Section 4.3 with $\rho \equiv 0$.

(d) Estimates obtained from the conditional score approach as described in Section 4.3 with $\rho \neq 0$.

For each of the two time scales, we can see the estimated effect of the average OAT dispensation rate by using approach (d) is much larger than by approach (b), and the results of using approach (a) are similar to those for approach (c). We remark that using the popular R package `JM` (Rizopoulos 2010) to obtain estimates under approach (a) required notably longer computation time relative to the other approaches. Obtaining estimates by approach (a) can be computationally infeasible if either a more complex model than (4.2) is considered or if there are several internal covariate processes to model.

We see the effect of the average OAT dispensation rate to be both positive and significant, which is similar to our results from Chapter 2. This suggests that individuals with higher OAT usage proportions have higher risks of mortality. We again took the approach from Chapter 2 and stratified individuals with respect to their corresponding birth generation, where Table 4.2 displays the summary statistics for $\hat{\nu}_i$'s under (4.2). Just as in Table 2.8, we see differing OAT usage patterns between survival groups varies across birth generation. In particular, survivors have lower OAT dispensation rates within the older generations, whereas nonsurvivors have lower OAT dispensation rates for the younger generations. This finding indicates that the results in Table 4.1 are likely confounded by age. Upon stratifying individuals into their corresponding birth generations, we re-estimated the model parameters in (4.3), where the estimates are presented in Tables 4.3-4.6. We see that the estimated effect of ν decreases with birth generation, which corroborates with Table 4.2 and our preliminary analysis from Chapter 2.

Since the estimates in Table 4.1 were obtained by pooling all individuals together, we can view the estimates in Table 4.1 as weighted averages of the estimates from Tables 4.3-4.6, where the weights partially depend on the follow-up time. Figure 4.2 illustrates both the distribution of first recorded OAT dispensation date and follow-up time (i.e. T^*) across birth generations, in which the majority of Millennials & Generation Z individuals had their first recorded OAT dispensation record near the end of the data extraction window, and hence contribute less information relative to individuals from the earlier birth generations.

This result can explain why we see the estimated effect of the average OAT dispensation rate to be positive on the hazard: larger weights are placed on older individuals due to their relative longer follow-up times.

To obtain an overall effect estimate for our population, we weight the estimates from Tables 4.3-4.6 based on the number of individuals within each of the four birth generations, that is, the size of each birth generation stratum. Letting n_g denote the number of individuals that belong to the gth birth generation and $(\hat{\gamma}_g, \hat{\boldsymbol{\theta}}_g)$ denote the estimates from Tables 4.3-4.6 ($g = 1, 2, 3, 4$), we computed the weighted estimates of γ and $\boldsymbol{\theta}$ as $\tilde{\gamma} = \sum_{g=1}^{4} \frac{n_g}{n} \hat{\gamma}_g$, and $\tilde{\boldsymbol{\theta}} = \sum_{g=1}^{4} \frac{n_g}{n} \hat{\boldsymbol{\theta}}_g$. The corresponding variances of $\tilde{\gamma}$ and $\tilde{\boldsymbol{\theta}}$ are

$$Var(\tilde{\gamma}) = Var\left(\sum_{g=1}^{G} \frac{n_g}{n} \hat{\gamma}_g\right) \approx \sum_{g=1}^{G} \left(\frac{n_g}{n}\right)^2 Var(\hat{\gamma}_g), \text{ and}$$
$$Var(\tilde{\boldsymbol{\theta}}) = Var\left(\sum_{g=1}^{G} \frac{n_g}{n} \hat{\boldsymbol{\theta}}_g\right) \approx \sum_{g=1}^{G} \left(\frac{n_g}{n}\right)^2 Var(\hat{\boldsymbol{\theta}}_g).$$

Table 4.7 presents $\tilde{\gamma}$ and $\tilde{\boldsymbol{\theta}}$, as well as conditional score approach with $AR(1)$ errors from Table 4.1 as a reference. By up-weighting the *Millennials & Generation Z* estimates, the resulting analysis indicates that individuals with a higher OAT dispensation rate have a lower mortality risk.

4.5 Simulation Study

To verify the findings from our data analysis in Section 4.4, we conducted a simulation study to assess the relative performance of the naive and conditional score inference procedures presented in Table 4.1, compared to the "ideal" approach of fitting (4.3) with known ν_i.

4.5.1 Data Generation

We started by simulating $n \in \{5000, 20000, 50000\}$ event times under (4.3) with $h(x) = logit(x)$. The true parameters were $\boldsymbol{\theta} = (-2, -1.75, -2.5)'$, $\gamma \in \{-0.5, 0, 0.5\}$, and $\lambda_0(t) \equiv \lambda_0 = 1$. The different specifications for γ are to reflect the varying signs of $\hat{\gamma}$ in Tables 4.3-4.6, and also examine the performance of each method when the effect is insignificant. We generated three baseline covariates for each sample unit $\boldsymbol{X}_i = (X_{i1}, X_{i2}, X_{i3})'$, where $X_{i1} \sim$ Normal$(0, 1)$, $X_{i2} \sim$ Uniform$(0, 1)$, and $X_{i3} \sim$ Bernoulli(0.5). Upon generating $\nu_i \sim$ Normal$(0, 5)$, where the specification of the variance was based on findings from our data application, event times were then generated as $T_i \sim$ Exponential(λ_i), where $\lambda_i = \lambda_0 \exp\{\gamma\nu_i + \boldsymbol{\theta}'\boldsymbol{X}_i\}$. Simulating right-censored event times $C_i \sim$ Exponential(3), we

obtained observations $\{(T_i^*, \delta_i, \boldsymbol{X}_i) : i = 1, \cdots, n\}$, in which $\sum_{i=1}^{n} \delta_i \approx 14\%$ to roughly match our data application.

We then generated longitudinal observations under (4.2), where we discretized $[0, T_i^*]$ into time units of length 0.003, and compute $m_i = \lfloor T_i^*/0.003 \rfloor$. We specified 0.003 as the time unit so that the distribution of m_i across individuals is roughly similar to the observed distribution from our data application. As it is possible to generate $m_i = 0$ (this pertains to $\approx 1\%$ of the simulated sample units), we omit these units to be consistent with the analysis in Section 4.4. We then generated observations under (4.2) upon simulating ε_{ij} from a normal distribution with $\mathbb{E}(\varepsilon_{ij}) = 0$, and $Cov(\varepsilon_{ij}, \varepsilon_{i,j-k}) = \rho^k \sigma^2/(1-\rho^2)$, in which $\rho \in \{0, 0.5, 0.9\}$ to reflect varying levels of correlation within the error terms, and $\sigma^2/(1-\rho^2) \equiv 1.4$, as motivated by its estimate in Table 4.1. The entire data generation procedure was replicated 1,000 times.

4.5.2 Simulation Outcome 1: Bias Assessment

We summarized the average bias across the 1,000 data replicates for the "ideal", naive, and conditional score inference procedures in Table 4.8, across different specifications of n, γ, and ρ. Key findings from Table 4.8 include the following:

- When $\gamma \in \{-0.5, 0.5\}$ (the first and third columns of Table 4.8), we see that directly replacing ν_i with $\hat{\nu}$ in (4.3) produces biased estimates, and is especially noticeable when $\rho > 0$. Also, the conditional score approach with $\rho \equiv 0$ only performs well when the data is generated with $\rho = 0$. This is because the sufficient statistic requires an estimate for $Var(\hat{\nu}_i)$, and is underestimated when $\rho > 0$ if one assumes the errors are IID. Our proposed approach that allows for $AR(1)$ errors is able to consistently estimate γ and $\boldsymbol{\theta}$.

- When $\gamma = 0$ (the second column in Table 4.8), all four methods are able to consistently estimate the model parameters. This indicates that under this setting, the parameter estimates are not only robust to misspecification of the correlation structure in the error terms, but are robust to the differences between $\hat{\nu}$ and ν themselves. This latter result is due to γ generally being underestimated when ν is directly replaced with $\hat{\nu}$, in the sense of being biased towards zero. When $\gamma = 0$ however, this seems to serve as an advantage.

4.5.3 Simulation Outcome 2: Alternative Rate Modelling

For each simulated dataset, we also estimated the model parameters in (4.6) by linking it to (4.2), except we specify (i) $h(x) = x$; and (ii) $h(x) = -\log x$. Our objective is to assess the implications on inferring parameters in (4.3), when using these alternative specifications of $h(\cdot)$, since the left hand side of $h(R_{ij})$ is constrained to $(0, 1)$ and $(0, \infty)$ when $h(x) = x$

and $h(x) = -\log x$, respectively. Figure 4.3 and Figure 4.4 illustrates the estimates of γ for various levels of ρ and data-generating values of γ with $h(x) = x$ and $h(x) = -\log x$, respectively, where we fixed $n = 50{,}000$. We see that

- Under $h(x) = x$, larger values of $\hat{\nu}_i$ would correspond to larger dispensation rates. The results in Figure 4.3 indicate that this approach is robust to model misspecification when $\gamma = 0$, and is able to identify the correct sign of γ when $\gamma \in \{-0.5, 0.5\}$.
- Under $h(x) = -\log x$, larger values of $\hat{\nu}_i$ corresponds to smaller dispensation rates. We obtain similar findings as before, in the sense that this approach is able to properly identify whether the effect is positive or negative.

4.6 Discussion and Outlook for Chapter 5

Motivated by predicting an individual's mortality risk given their OAT dispensation history, we used a latent variable in the survival models to summarize an individual's OAT dispensation process. We predict the latent variable by projecting it to the available data space, and extend the conditional score approach of Tsiatis and Davidian (2001) to account for within-subject correlation in estimating parameters of an extended Cox regression model. Results from our data application showcases the estimated effect of the OAT dispensation rate varies with respect to time of birth in our population, in which the OAT protective effect against mortality is strongest for Millennials and Generation Z individuals. This finding suggests that more personalized approaches may be needed to effectively manage OAT use within different age groups.

As older individuals generally have longer follow-up times relative to younger individuals, this disparity induces a bias in the results if we conventionally pool all individuals together. To overcome that, we implemented a weighted average of birth generation stratified estimates. This produced an overall estimate that demonstrates an overall OAT protective effect for the population. Our simulation study verifies the findings from the data application, where in particular, the proposed conditional score estimator is able to consistently estimate model parameters, whereas the naive approaches fail.

Observe that the model in (4.2) explicitly assumes that $h(R_{ij})$ does not depend on any risk factors or time. We can overcome this limitation by extending (4.2) to

$$h(R_{ij}) = \nu_i(j, \boldsymbol{\mathcal{X}}_i(t_{ij}); \boldsymbol{\alpha}_i) + \varepsilon_{ij},$$

where $\nu_i(j, \boldsymbol{\mathcal{X}}_i(t_{ij}); \boldsymbol{\alpha}_i)$ is a parametric function that depends on both time and other risk factors $\boldsymbol{X}_i(\cdot)$, and is parameterized by $\boldsymbol{\alpha}_i$. Figure 4.5 illustrates $logit(R_{ij})$ over time for the same four randomly selected individuals from Figure 4.1, and appears to be reasonably modelled by a quadratic model of time. Accounting for this relationship would likely provide more accurate survival predictions than the ones obtained under (4.2). This extra step can

be realized under our current setup if assume $T_i \perp\!\!\!\perp \mathcal{Z}_i(\infty)|\boldsymbol{\alpha}_i$, and use $\nu_i(j, \boldsymbol{\mathcal{X}}_i(t_{ij}); \hat{\boldsymbol{\alpha}}_i)$ as a "noisy measurement" of $\nu_i(\cdot)$, where $\hat{\boldsymbol{\alpha}}_i$ is the least squares estimate of $\boldsymbol{\alpha}_i$. We could then essentially proceed as before to estimate parameters in the hazard sub-model.

To this point, we have developed methodology to conduct statistical inference when an internal covariate is present in a hazard regression model. To simplify our task and better understand the OAT dispensation process within our population, we neglected the type of OAT, as well as its dosage. We are now in a position to lift this self-imposed restriction, and extend the methods developed in Chapters 3 and 4 to include additional factors pertaining to OAT usage.

Figure 4.1: Sample autocorrelation and partial autocorrelation of the residuals under the model $logit(R_{ij}) = \nu_i + \varepsilon_{ij}$, for four randomly selected individuals, where the time unit (t_{ij}) is in *months*.

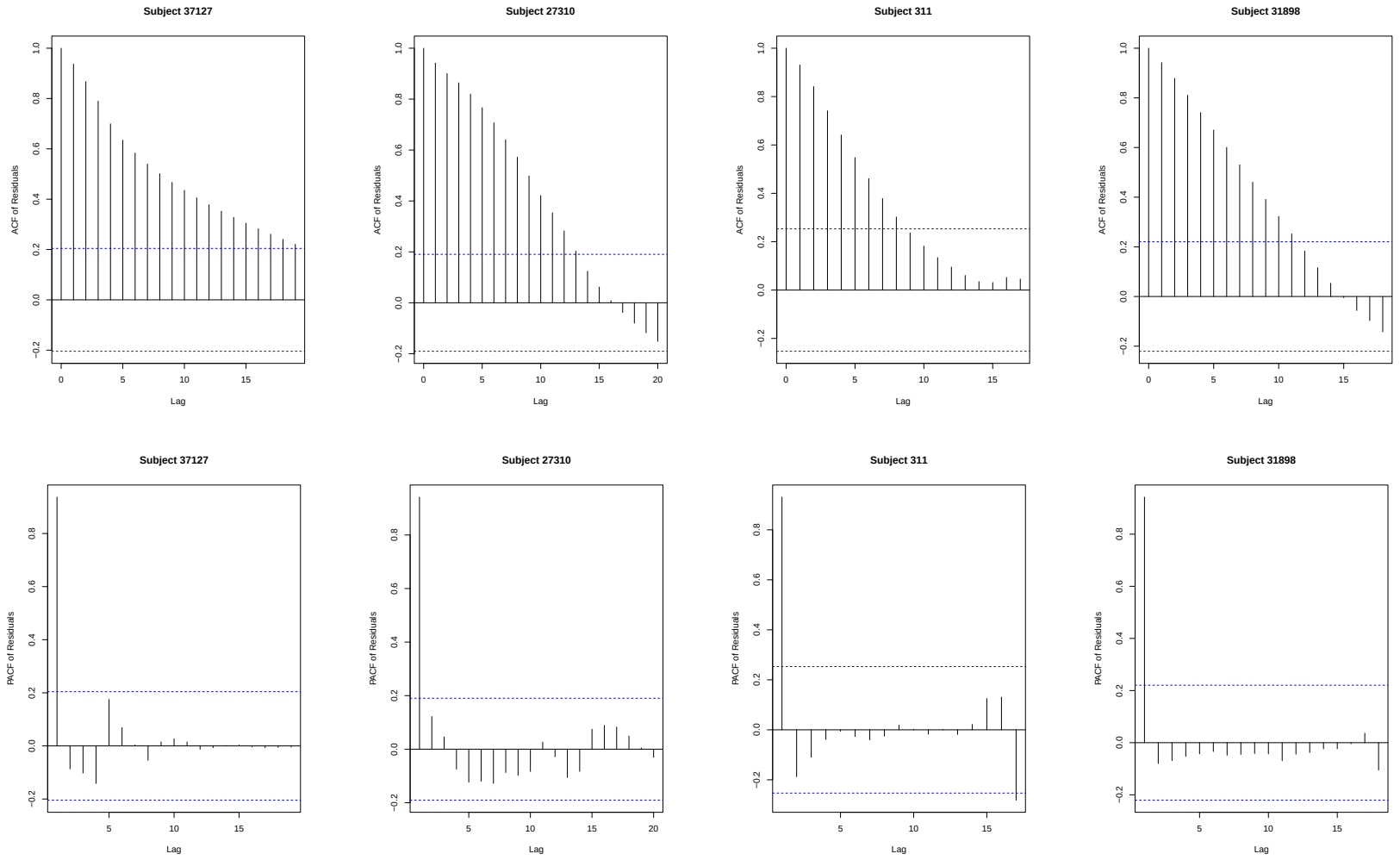


Figure 4.2: Illustration of the distributions for date of first recorded OAT dispensations, and T^* (in years), across birth generation levels.

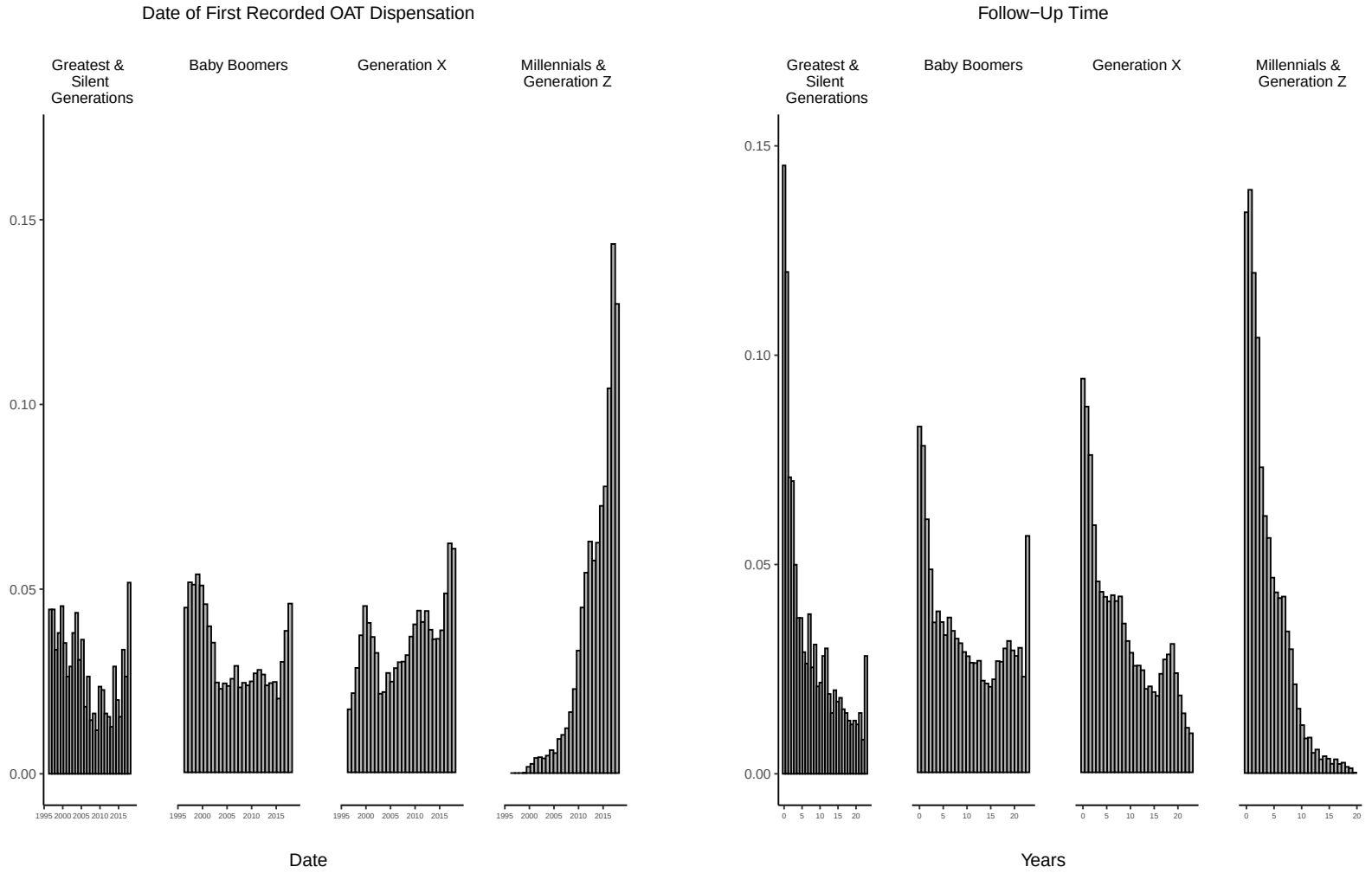


Figure 4.3: Histograms of $\hat{\gamma}$ from 1,000 simulated data replications upon specifying $h(x) = x$, for different specifications of γ and ρ. Here, the number of study units in each simulated dataset is fixed at $n = 50{,}000$.

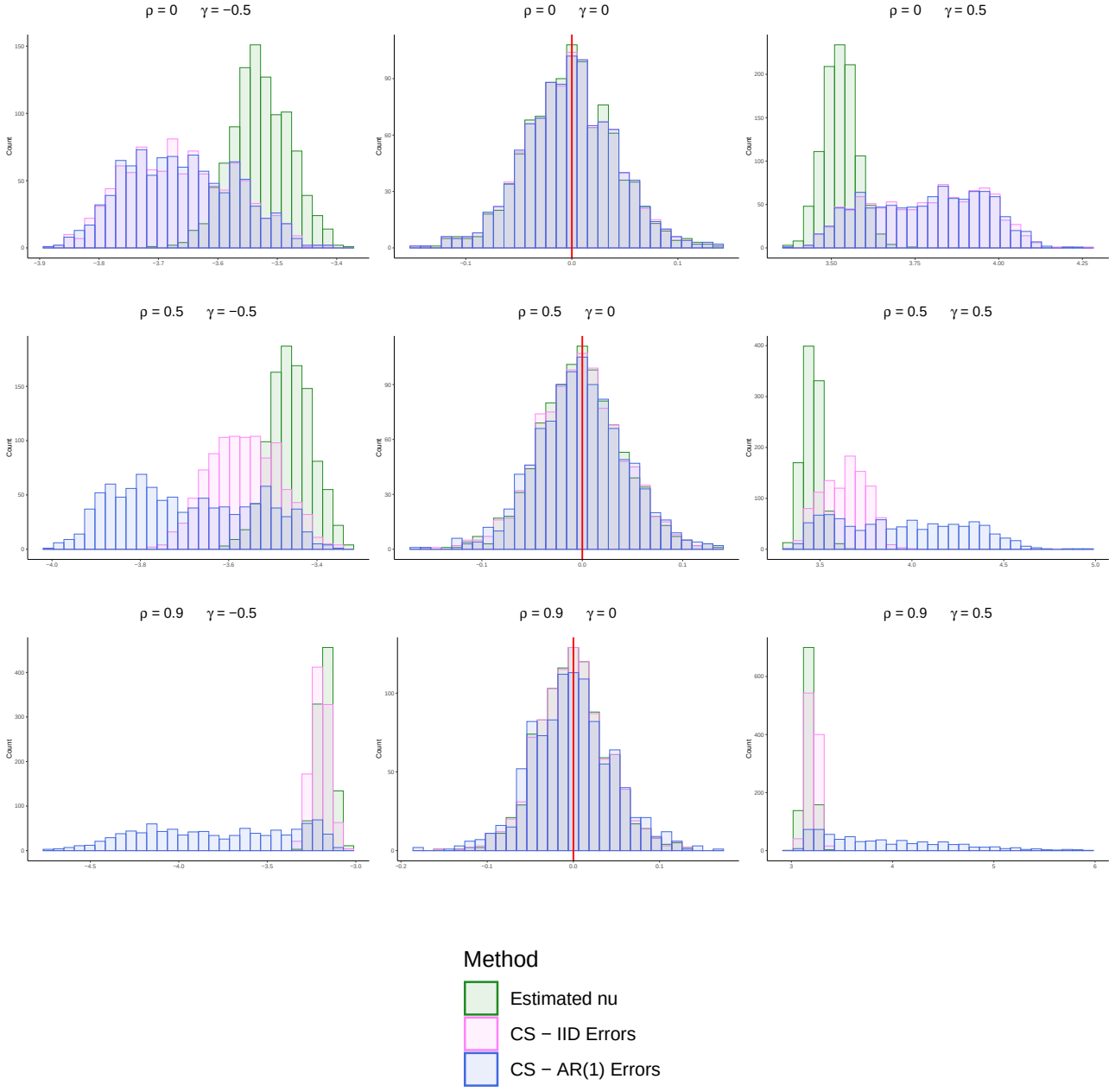


Figure 4.4: Histograms of $\hat{\gamma}$ from 1,000 simulated data replications upon specifying $h(x) = -\log x$, for different specifications of γ and ρ. Here, the number of study units in each simulated dataset is fixed at $n = 50{,}000$.

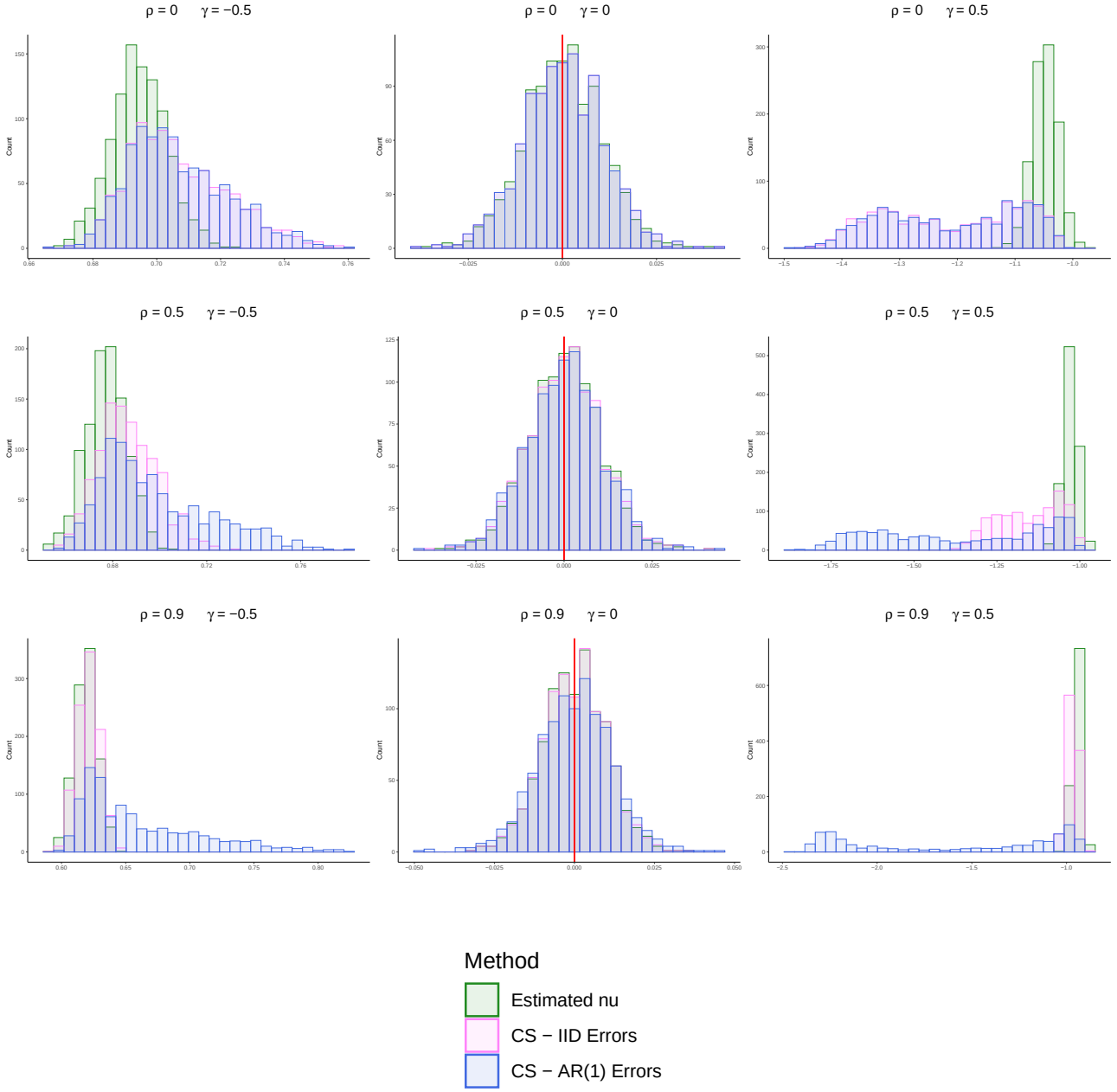


Figure 4.5: Illustration of $logit(R_{ij})$ vs. j or the four randomly selected individuals from Figure 4.1.
The red line is $\hat{\nu}_i$ under (4.2).
The blue curve is the LOESS fit.
The green curve is the fit under the model $logit(R_i(t)) = \nu_{i0} + \nu_{i1}t + \nu_{i2}t^2 + \varepsilon_i(t)$.

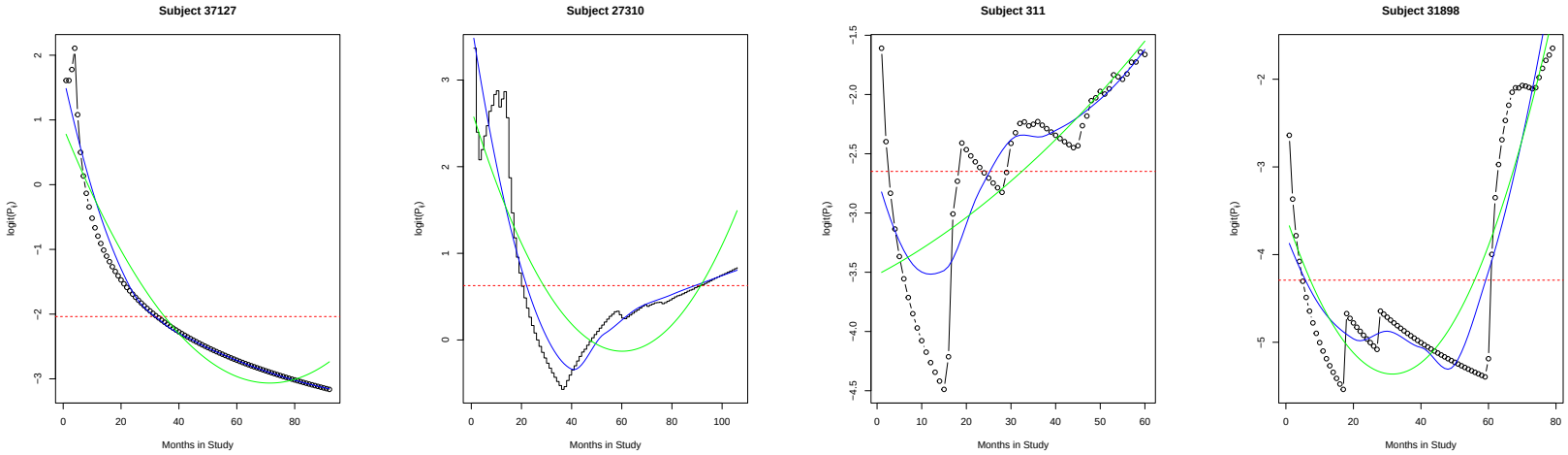


Table 4.1: Estimates of regression coefficients under (a) a joint model where we specify $\nu_i \sim N(\mu, \sigma_\nu^2)$, (b) the Cox model that directly replaces ν_i with $\hat{\nu}_i$, (c) the conditional score method with independent errors, and (d) the conditional score method with AR(1) errors. The **bolded** estimates are statistically significant with the type 1 error rate set at 5%.

Time Since First Recorded OAT Dispensation								
	Frailty Model		**Cox Model**		**Conditional Score IID Errors**		**Conditional Score AR(1) Errors**	
	$\hat{\mu} = 1.3711$		$\sigma^2/(1-\rho^2) \equiv 0$		$\hat{\sigma}^2/(1-\rho^2) = 1.3852$		$\hat{\sigma}^2/(1-\hat{\rho}^2) = 1.3852$	
	$\hat{\sigma}_\nu^2 = 4.3374$		$\rho \equiv 0$		$\rho \equiv 0$		$\hat{\rho} = 0.7108$	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
ν	**0.0515**	0.0039	**0.0499**	0.0033	**0.0510**	0.0039	**0.0539**	0.0044
Incarceration FPCS	**0.0310**	0.0071	**0.0310**	0.0071	**0.0310**	0.0071	**0.0311**	0.0071
Sex (vs. *Female*)	-	-	-	-	-	-	-	-
Male	**0.1778**	0.0265	**0.1774**	0.0265	**0.1771**	0.0265	**0.1763**	0.0265
Birth Generation (vs *Greatest & Silent Generations*)	-	-	-	-	-	-	-	-
Baby Boomers	**-1.2456**	0.0459	**-1.2448**	0.0459	**-1.2440**	0.0459	**-1.2419**	0.0459
Generation X	**-1.9527**	0.0495	**-1.9507**	0.0495	**-1.9486**	0.0495	**-1.9431**	0.0495
Millennials & Generation Z	**-2.0275**	0.0613	**-2.0250**	0.0613	**-2.0228**	0.0613	**-2.0166**	0.0613
Heath Authority (vs *Fraser Health*)	-	-	-	-	-	-	-	-
Interior	**0.2105**	0.0397	**0.2109**	0.0397	**0.2114**	0.0397	**0.2127**	0.0397
Vancouver Coastal	**0.1130**	0.0307	**0.1132**	0.0307	**0.1134**	0.0307	**0.1138**	0.0307
Vancouver Island	**0.0810**	0.0368	**0.0809**	0.0368	**0.0809**	0.0368	**0.0809**	0.0368
Northern	0.0533	0.0692	0.0535	0.0692	0.0538	0.0692	0.0546	0.0692
Year Category (vs. *1996-2000*)	-	-	-	-	-	-	-	-
2001-2006	**0.1518**	0.0337	**0.1523**	0.0337	**0.1528**	0.0337	**0.1541**	0.0337
2007-2012	**0.2443**	0.0399	**0.2443**	0.0399	**0.2444**	0.0399	**0.2446**	0.0399
2013-2018	**0.6137**	0.0510	**0.6117**	0.0510	**0.6108**	0.0510	**0.6084**	0.0511
Alcohol or Other Substance Use Disorders	**0.4235**	0.0448	**0.4236**	0.0448	**0.4234**	0.0448	**0.4229**	0.0448
Ill Mental Health or Chronic pain	**-0.1907**	0.0448	**-0.1906**	0.0448	**-0.1905**	0.0448	**-0.1904**	0.0448
Hepatitis C Virus or HIV/AIDS	**1.1611**	0.0275	**1.1611**	0.0275	**1.1611**	0.0275	**1.1612**	0.0276
Ever Received a Sedative	**0.4802**	0.0335	**0.4802**	0.0335	**0.4802**	0.0335	**0.4803**	0.0335
Ever on PharmaCare Plans C or G	**-0.2222**	0.0306	**-0.2238**	0.0306	**-0.2253**	0.0306	**-0.2296**	0.0306
Age								
	Frailty Model		**Cox Model**		**Conditional Score IID Errors**		**Conditional Score AR(1) Errors**	
	$\hat{\mu} = 1.3711$		$\sigma^2/(1-\rho^2) \equiv 0$		$\hat{\sigma}^2/(1-\rho^2) = 1.3852$		$\hat{\sigma}^2/(1-\hat{\rho}^2) = 1.3852$	
	$\hat{\sigma}_\nu^2 = 4.3374$		$\rho \equiv 0$		$\rho \equiv 0$		$\hat{\rho} = 0.7108$	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
ν	**0.0489**	0.0043	**0.0472**	0.0033	**0.0482**	0.0039	**0.0509**	0.0046
Incarceration FPCS	**0.0326**	0.0071	**0.0326**	0.0071	**0.0325**	0.0071	**0.0326**	0.0071
Sex (vs. *Female*)	-	-	-	-	-	-	-	-
Male	**0.1629**	0.0266	**0.1626**	0.0266	**0.1626**	0.0266	**0.1620**	0.0266
Birth Generation (vs *Greatest & Silent Generations*)	-	-	-	-	-	-	-	-
Baby Boomers	**-0.4444**	0.0685	**-0.4435**	0.0685	**-0.4423**	0.0685	**-0.4401**	0.0684
Generation X	**-0.6298**	0.0833	**-0.6283**	0.0833	**-0.6265**	0.0833	**-0.6223**	0.0833
Millennials & Generation Z	**-0.2779**	0.1087	**-0.2763**	0.1087	**-0.2749**	0.1087	**-0.2715**	0.1087
Heath Authority (vs *Fraser Health*)	-	-	-	-	-	-	-	-
Interior	**0.1712**	0.0398	**0.1715**	0.0398	**0.1720**	0.0398	**0.1731**	0.0398
Vancouver Coastal	**0.0966**	0.0307	**0.0967**	0.0307	**0.0964**	0.0307	**0.0967**	0.0307
Vancouver Island	0.0577	0.0368	0.0576	0.0368	0.0576	0.0368	0.0576	0.0368
Northern	0.0476	0.0692	0.0477	0.0692	0.0485	0.0692	0.0493	0.0692
Year Category (vs. *1996-2000*)	-	-	-	-	-	-	-	-
2001-2006	0.0093	0.0320	0.0097	0.0320	0.0097	0.0320	0.0109	0.0320
2007-2012	-0.0330	0.0365	-0.0332	0.0365	-0.0344	0.0365	-0.0343	0.0365
2013-2018	**0.2425**	0.0449	**0.2398**	0.0449	**0.2357**	0.0449	**0.2328**	0.0449
Alcohol or Other Substance Use Disorders	**0.5125**	0.0459	**0.5127**	0.0459	**0.5116**	0.0459	**0.5110**	0.0458
Ill Mental Health or Chronic pain	**-0.2742**	0.0446	**-0.2736**	0.0446	**-0.2732**	0.0446	**-0.2726**	0.0446
Hepatitis C Virus or HIV/AIDS	**1.1088**	0.0275	**1.1090**	0.0275	**1.1088**	0.0275	**1.1092**	0.0275
Ever Received a Sedative	**0.4599**	0.0335	**0.4600**	0.0335	**0.4600**	0.0335	**0.4601**	0.0335
Ever on PharmaCare Plans C or G	**-0.1433**	0.0313	**-0.1448**	0.0313	**-0.1471**	0.0313	**-0.1510**	0.0313

Table 4.2: Summary statistics of $\hat{\nu}$ across birth generations and survival status.
(*): Among individuals with $R_{ij} \neq 1$ for all $j = 1, \cdots, m_i$.

Greatest & Silent Generations: 1901-1945				**Baby Boomers: 1946-1964**			
	Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total
Minimum	-7.9979	-7.8148	-7.9979	Minimum	-7.9904	-7.3155	-7.9904
1st Quartile	-2.8862	-1.6070	-2.2005	1st Quartile	-1.1312	-0.2740	-0.9036
Median	-0.4048	1.8379	0.9485	Median	1.2243	2.0777	1.4494
3rd Quartile	3.2339	5.5032	4.6510	3rd Quartile	3.7305	4.3637	3.9229
Maximum*	16.8276	19.4207	19.4207	Maximum*	19.7516	19.5262	19.7516
N	447	654	1,101	N	9,904	3,575	13,479
Generation X: 1965-1980				**Millennials & Generation Z: 1981+**			
	Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total
Minimum	-7.9126	-6.6852	-7.9126	Minimum	-7.8369	-7.0285	-7.8369
1st Quartile	-1.2177	-1.2753	-1.2265	1st Quartile	-1.4127	-2.3128	-1.4576
Median	0.6233	0.4265	0.6080	Median	0.4660	-0.4127	0.4264
3rd Quartile	2.6975	2.2999	2.6571	3rd Quartile	2.7066	1.4970	2.6618
Maximum*	19.8185	18.0020	19.8185	Maximum*	19.7073	18.5611	19.7073
N	18,861	1,998	20,859	N	18,519	781	19,300

Table 4.3: Estimates of regression coefficients under (a) a joint model where we specify $\nu_i \sim N(\mu, \sigma_\nu^2)$, (b) the Cox model that directly replaces ν_i with $\hat{\nu}_i$, (c) the conditional score method with independent errors, and (d) the conditional score method with AR(1) errors; for subjects born between **1901-1945**. The **bolded** estimates are statistically significant with the type 1 error rate set at 5%.

Time Since First Recorded OAT Dispensation								
	Frailty Model		**Cox Model**		**Conditional Score IID Errors**		**Conditional Score AR(1) Errors**	
	$\hat{\mu} = 1.8438$		$\sigma^2/(1-\rho^2) \equiv 0$		$\hat{\sigma}^2/(1-\rho^2) = 2.9872$		$\hat{\sigma}^2/(1-\hat{\rho}^2) = 2.9872$	
	$\hat{\sigma}_\nu^2 = 5.4214$		$\rho \equiv 0$		$\rho \equiv 0$		$\hat{\rho} = 0.6306$	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
ν	**0.0962**	0.0095	**0.0922**	0.0087	**0.0958**	0.0101	**0.1023**	0.0112
Incarceration FPCS	**0.3387**	0.1407	**0.3388**	0.1407	**0.3393**	0.1408	**0.3405**	0.1410
Sex (vs. *Female*)	-	-	-	-	-	-	-	-
Male	-0.0454	0.0904	-0.0475	0.0904	-0.0535	0.0904	-0.0645	0.0905
Heath Authority (vs *Fraser Health*)	-	-	-	-	-	-	-	-
Interior	**0.3273**	0.1319	**0.3289**	0.1319	**0.3365**	0.1320	**0.3499**	0.1321
Vancouver Coastal	**0.2459**	0.1048	**0.2460**	0.1048	**0.2468**	0.1048	**0.2483**	0.1048
Vancouver Island	0.1417	0.1324	0.1420	0.1324	0.1447	0.1324	0.1490	0.1325
Northern	-0.2293	0.2935	-0.2270	0.2935	-0.2169	0.2936	-0.1992	0.2938
Year Category (vs. *1996-2000*)	-	-	-	-	-	-	-	-
2001-2006	**0.4452**	0.1118	**0.4464**	0.1117	**0.4553**	0.1118	**0.4717**	0.1120
2007-2012	0.1275	0.1538	0.1281	0.1539	0.1292	0.1539	0.1316	0.1539
2013-2018	0.1346	0.1839	0.1355	0.1839	0.1410	0.1840	0.1513	0.1841
Alcohol or Other Substance Use Disorders	**-0.2370**	0.0994	**-0.2375**	0.0994	**-0.2412**	0.0993	**-0.2476**	0.0992
Ill Mental Health or Chronic pain	0.1859	0.1532	0.1872	0.1532	0.1869	0.1533	0.1862	0.1535
Hepatitis C Virus or HIV/AIDS	**1.0347**	0.1014	**1.0351**	0.1014	**1.0346**	0.1015	**1.0343**	0.1016
Ever Received a Sedative	**0.3553**	0.1543	**0.3547**	0.1543	**0.3550**	0.1544	**0.3572**	0.1544
Ever on PharmaCare Plans C or G	**-0.4774**	0.1001	**-0.4789**	0.1000	**-0.4890**	0.1000	**-0.5068**	0.1000
Age								
	Frailty Model		**Cox Model**		**Conditional Score IID Errors**		**Conditional Score AR(1) Errors**	
	$\hat{\mu} = 1.8438$		$\sigma^2/(1-\rho^2) \equiv 0$		$\hat{\sigma}^2/(1-\rho^2) = 2.9872$		$\hat{\sigma}^2/(1-\hat{\rho}^2) = 2.9872$	
	$\hat{\sigma}_\nu^2 = 5.4214$		$\rho \equiv 0$		$\rho \equiv 0$		$\hat{\rho} = 0.6306$	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
ν	**0.1039**	0.0091	**0.0996**	0.0086	**0.1032**	0.0106	**0.1098**	0.0118
Incarceration FPCS	**0.4018**	0.1414	**0.4016**	0.1415	**0.4016**	0.1415	**0.4011**	0.1416
Sex (vs. *Female*)	-	-	-	-	-	-	-	-
Male	0.0237	0.0929	0.0214	0.0930	0.0155	0.0930	0.0041	0.0932
Heath Authority (vs *Fraser Health*)	-	-	-	-	-	-	-	-
Interior	0.1839	0.1344	0.1855	0.1344	0.1916	0.1345	0.2037	0.1346
Vancouver Coastal	0.1718	0.1056	0.1716	0.1056	0.1713	0.1056	0.1707	0.1056
Vancouver Island	0.1201	0.1327	0.1201	0.1327	0.1231	0.1327	0.1269	0.1328
Northern	-0.0988	0.2938	-0.0978	0.2939	-0.0863	0.2939	-0.0675	0.2942
Year Category (vs. *1996-2000*)	-	-	-	-	-	-	-	-
2001-2006	0.3085	0.1100	**0.3075**	0.1100	**0.3137**	0.1100	**0.3263**	0.1102
2007-2012	-0.1233	0.1529	-0.1244	0.1529	-0.1304	0.1530	-0.1359	0.1533
2013-2018	-0.0319	0.1765	-0.0352	0.1766	-0.0420	0.1768	-0.0419	0.1771
Alcohol or Other Substance Use Disorders	-0.0987	0.1041	-0.0981	0.1041	-0.1027	0.1039	**-0.1089**	0.1037
Ill Mental Health or Chronic pain	-0.0133	0.1507	-0.0097	0.1507	-0.0097	0.1508	-0.0093	0.1509
Hepatitis C Virus or HIV/AIDS	**0.9984**	0.0997	**1.0001**	0.0998	**1.0015**	0.0997	**1.0050**	0.0998
Ever Received a Sedative	**0.3273**	0.1544	**0.3255**	0.1545	**0.3236**	0.1546	**0.3241**	0.1548
Ever on PharmaCare Plans C or G	**-0.2718**	0.1062	**-0.2720**	0.1061	**-0.2793**	0.1060	**-0.2938**	0.1060

Table 4.4: Estimates of regression coefficients under (a) a joint model where we specify $\nu_i \sim N(\mu, \sigma_\nu^2)$, (b) the Cox model that directly replaces ν_i with $\hat{\nu}_i$, (c) the conditional score method with independent errors, and (d) the conditional score method with AR(1) errors; for subjects born between **1946-1964**. The **bolded** estimates are statistically significant with the type 1 error rate set at 5%.

Time Since First Recorded OAT Dispensation								
	Frailty Model		**Cox Model**		**Conditional Score IID Errors**		**Conditional Score AR(1) Errors**	
	$\hat{\mu} = 1.9122$		$\sigma^2/(1-\rho^2) \equiv 0$		$\hat{\sigma}^2/(1-\rho^2) = 2.5269$		$\hat{\sigma}^2/(1-\hat{\rho}^2) = 2.5269$	
	$\hat{\sigma}_\nu^2 = 4.4554$		$\rho \equiv 0$		$\rho \equiv 0$		$\hat{\rho} = 0.7359$	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
ν	**0.0731**	0.0061	**0.0717**	0.0045	**0.0738**	0.0059	**0.0808**	0.0066
Incarceration FPCS	**0.0250**	0.0118	**0.0250**	0.0118	**0.0250**	0.0118	**0.0252**	0.0118
Sex (vs. *Female*)	-	-	-	-	-	-	-	-
Male	**0.1484**	0.0375	**0.1480**	0.0375	**0.1474**	0.0375	**0.1454**	0.0375
Heath Authority (vs *Fraser Health*)	-	-	-	-	-	-	-	-
Interior	**0.1771**	0.0580	**0.1778**	0.0579	**0.1795**	0.0580	**0.1849**	0.0580
Vancouver Coastal	0.0648	0.0431	0.0651	0.0431	0.0655	0.0432	0.0669	0.0432
Vancouver Island	0.0952	0.0506	0.0953	0.0506	0.0958	0.0506	0.0973	0.0506
Northern	0.1529	0.0958	0.1527	0.0958	0.1529	0.0958	0.1536	0.0959
Year Category (vs. *1996-2000*)	-	-	-	-	-	-	-	-
2001-2006	**0.1675**	0.0444	**0.1679**	0.0444	**0.1689**	0.0444	**0.1720**	0.0444
2007-2012	**0.1814**	0.0577	**0.1816**	0.0577	**0.1821**	0.0577	**0.1837**	0.0577
2013-2018	**0.3886**	0.0844	**0.3832**	0.0845	**0.3792**	0.0846	**0.3661**	0.0847
Alcohol or Other Substance Use Disorders	**0.3333**	0.0656	**0.3338**	0.0656	**0.3338**	0.0656	**0.3340**	0.0656
Ill Mental Health or Chronic pain	**-0.2160**	0.0635	**-0.2159**	0.0635	**-0.2158**	0.0635	**-0.2156**	0.0635
Hepatitis C Virus or HIV/AIDS	**1.2607**	0.0362	**1.2606**	0.0362	**1.2605**	0.0362	**1.2605**	0.0362
Ever Received a Sedative	**0.3457**	0.0471	**0.3456**	0.0471	**0.3454**	0.0471	**0.3447**	0.0471
Ever on PharmaCare Plans C or G	**-0.2047**	0.0455	**-0.2066**	0.0455	**-0.2099**	0.0455	**-0.2206**	0.0455
Age								
	Frailty Model		**Cox Model**		**Conditional Score IID Errors**		**Conditional Score AR(1) Errors**	
	$\hat{\mu} = 1.9122$		$\sigma^2/(1-\rho^2) \equiv 0$		$\hat{\sigma}^2/(1-\rho^2) = 2.5269$		$\hat{\sigma}^2/(1-\hat{\rho}^2) = 2.5269$	
	$\hat{\sigma}_\nu^2 = 4.4554$		$\rho \equiv 0$		$\rho \equiv 0$		$\hat{\rho} = 0.7359$	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
ν	**0.0651**	0.0060	**0.0637**	0.0045	**0.0657**	0.0059	**0.0718**	0.0066
Incarceration FPCS	**0.0269**	0.0118	**0.0269**	0.0118	**0.0264**	0.0117	**0.0265**	0.0118
Sex (vs. *Female*)	-	-	-	-	-	-	-	-
Male	**0.1019**	0.0375	**0.1016**	0.0375	**0.1012**	0.0375	**0.0997**	0.0376
Heath Authority (vs *Fraser Health*)	-	-	-	-	-	-	-	-
Interior	**0.1532**	0.0579	**0.1540**	0.0579	**0.1556**	0.0579	**0.1604**	0.0579
Vancouver Coastal	0.0569	0.0431	0.0571	0.0431	0.0572	0.0431	0.0579	0.0431
Vancouver Island	0.0669	0.0506	0.0671	0.0506	0.0676	0.0506	0.0693	0.0506
Northern	0.1671	0.0958	0.1668	0.0958	0.1676	0.0958	0.1683	0.0958
Year Category (vs. *1996-2000*)	-	-	-	-	-	-	-	-
2001-2006	0.0454	0.0423	0.0458	0.0423	0.0464	0.0423	0.0495	0.0423
2007-2012	**-0.1225**	0.0539	**-0.1224**	0.0539	**-0.1227**	0.0539	**-0.1205**	0.0539
2013-2018	-0.0560	0.0777	-0.0618	0.0778	-0.0671	0.0778	-0.0773	0.0779
Alcohol or Other Substance Use Disorders	**0.3521**	0.0655	**0.3525**	0.0655	**0.3519**	0.0655	**0.3517**	0.0655
Ill Mental Health or Chronic pain	**-0.3041**	0.0633	**-0.3034**	0.0633	**-0.3022**	0.0633	**-0.3000**	0.0633
Hepatitis C Virus or HIV/AIDS	**1.2063**	0.0360	**1.2064**	0.0360	**1.2062**	0.0360	**1.2069**	0.0360
Ever Received a Sedative	**0.3293**	0.0470	**0.3293**	0.0470	**0.3294**	0.0470	**0.3289**	0.0470
Ever on PharmaCare Plans C or G	**-0.1074**	0.0460	**-0.1094**	0.0460	**-0.1133**	0.0460	**-0.1237**	0.0460

Table 4.5: Estimates of regression coefficients under (a) a joint model where we specify $\nu_i \sim N(\mu, \sigma_\nu^2)$, (b) the Cox model that directly replaces ν_i with $\hat{\nu}_i$, (c) the conditional score method with independent errors, and (d) the conditional score method with AR(1) errors; for subjects born between **1956-1980**. The **bolded** estimates are statistically significant with the type 1 error rate set at 5%.

Time Since First Recorded OAT Dispensation								
	Frailty Model		**Cox Model**		**Conditional Score IID Errors**		**Conditional Score AR(1) Errors**	
	$\hat{\mu} = 1.1519$		$\sigma^2/(1-\rho^2) \equiv 0$		$\hat{\sigma}^2/(1-\rho^2) = 1.4208$		$\hat{\sigma}^2/(1-\hat{\rho}^2) = 1.4208$	
	$\hat{\sigma}_\nu^2 = 4.1112$		$\rho \equiv 0$		$\rho \equiv 0$		$\hat{\rho} = 0.7396$	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
ν	**0.0338**	0.0075	**0.0312**	0.0074	**0.0321**	0.0075	**0.0353**	0.0078
Incarceration FPCS	0.0284	0.0196	0.0285	0.0196	0.0285	0.0196	0.0287	0.0196
Sex (vs. *Female*)	-	-	-	-	-	-	-	-
Male	**0.1769**	0.0486	**0.1770**	0.0486	**0.1770**	0.0486	**0.1772**	0.0486
Heath Authority (vs *Fraser Health*)	-	-	-	-	-	-	-	-
Interior	**0.2285**	0.0737	**0.2286**	0.0737	**0.2286**	0.0737	**0.2287**	0.0737
Vancouver Coastal	**0.1471**	0.0563	**0.1472**	0.0563	**0.1473**	0.0563	**0.1474**	0.0563
Vancouver Island	0.0407	0.0696	0.0403	0.0696	0.0400	0.0696	0.0389	0.0696
Northern	-0.0776	0.1280	-0.0774	0.1280	-0.0772	0.1280	-0.0767	0.1280
Year Category (vs. *1996-2000*)	-	-	-	-	-	-	-	-
2001-2006	0.0685	0.0621	0.0681	0.0621	0.0679	0.0621	0.0670	0.0621
2007-2012	**0.3804**	0.0690	**0.3797**	0.0690	**0.3791**	0.0690	**0.3771**	0.0691
2013-2018	**0.8072**	0.0921	**0.8050**	0.0921	**0.8038**	0.0921	**0.7992**	0.0922
Alcohol or Other Substance Use Disorders	**1.0907**	0.1181	**1.0912**	0.1181	**1.0915**	0.1181	**1.0925**	0.1181
Ill Mental Health or Chronic pain	**-0.3457**	0.0825	**-0.3462**	0.0825	**-0.3465**	0.0825	**-0.3477**	0.0825
Hepatitis C Virus or HIV/AIDS	**1.0698**	0.0519	**1.0700**	0.0519	**1.0700**	0.0519	**1.0704**	0.0519
Ever Received a Sedative	**0.6463**	0.0585	**0.6465**	0.0585	**0.6466**	0.0585	**0.6470**	0.0585
Ever on PharmaCare Plans C or G	-0.1081	0.0589	-0.1093	0.0589	-0.1101	0.0589	-0.1131	0.0590
Age								
	Frailty Model		**Cox Model**		**Conditional Score IID Errors**		**Conditional Score AR(1) Errors**	
	$\hat{\mu} = 1.1519$		$\sigma^2/(1-\rho^2) \equiv 0$		$\hat{\sigma}^2/(1-\rho^2) = 1.4208$		$\hat{\sigma}^2/(1-\hat{\rho}^2) = 1.4208$	
	$\hat{\sigma}_\nu^2 = 4.1112$		$\rho \equiv 0$		$\rho \equiv 0$		$\hat{\rho} = 0.7396$	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
ν	**0.0305**	0.0078	**0.0278**	0.0074	**0.0287**	0.0075	**0.0315**	0.0078
Incarceration FPCS	0.0301	0.0196	0.0302	0.0196	0.0303	0.0196	0.0304	0.0196
Sex (vs. *Female*)	-	-	-	-	-	-	-	-
Male	**0.1455**	0.0487	**0.1456**	0.0487	**0.1462**	0.0487	**0.1464**	0.0487
Heath Authority (vs *Fraser Health*)	-	-	-	-	-	-	-	-
Interior	**0.2130**	0.0737	**0.2131**	0.0737	**0.2130**	0.0737	**0.2131**	0.0737
Vancouver Coastal	**0.1358**	0.0563	**0.1359**	0.0563	**0.1351**	0.0563	**0.1353**	0.0563
Vancouver Island	0.0399	0.0696	0.0395	0.0696	0.0390	0.0696	0.0380	0.0696
Northern	-0.0884	0.1279	-0.0883	0.1279	-0.0875	0.1280	-0.0869	0.1280
Year Category (vs. *1996-2000*)	-	-	-	-	-	-	-	-
2001-2006	-0.0262	0.0591	-0.0265	0.0591	-0.0276	0.0591	-0.0283	0.0591
2007-2012	**0.1797**	0.0619	**0.1791**	0.0619	**0.1766**	0.0619	**0.1750**	0.0619
2013-2018	**0.4409**	0.0807	**0.4387**	0.0808	**0.4332**	0.0808	**0.4291**	0.0809
Alcohol or Other Substance Use Disorders	**1.1135**	0.1182	**1.1141**	0.1182	**1.1138**	0.1182	**1.1147**	0.1182
Ill Mental Health or Chronic pain	**-0.3785**	0.0825	**-0.3787**	0.0825	**-0.3791**	0.0825	**-0.3799**	0.0825
Hepatitis C Virus or HIV/AIDS	**1.0186**	0.0523	**1.0188**	0.0523	**1.0183**	0.0523	**1.0187**	0.0524
Ever Received a Sedative	**0.6240**	0.0585	**0.6242**	0.0585	**0.6247**	0.0585	**0.6252**	0.0585
Ever on PharmaCare Plans C or G	-0.1012	0.0589	-0.1023	0.0589	-0.1046	0.0589	-0.1073	0.0590

Table 4.6: Estimates of regression coefficients under (a) a joint model where we specify $\nu_i \sim N(\mu, \sigma_\nu^2)$, (b) the Cox model that directly replaces ν_i with $\hat{\nu}_i$, (c) the conditional score method with independent errors, and (d) the conditional score method with AR(1) errors; for subjects born after **1980**. The **bolded** estimates are statistically significant with the type 1 error rate set at 5%.

Time Since First Recorded OAT Dispensation								
	Frailty Model		**Cox Model**		**Conditional Score IID Errors**		**Conditional Score AR(1) Errors**	
	$\hat{\mu} = 1.0722$		$\sigma^2/(1-\rho^2) \equiv 0$		$\hat{\sigma}^2/(1-\rho^2) = 0.9634$		$\hat{\sigma}^2/(1-\hat{\rho}^2) = 0.9634$	
	$\hat{\sigma}_\nu^2 = 4.5476$		$\rho \equiv 0$		$\rho \equiv 0$		$\hat{\rho} = 0.6646$	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
ν	**-0.0458**	0.0121	**-0.0438**	0.0120	**-0.0452**	0.0122	**-0.0490**	0.0128
Incarceration FPCS	**0.0439**	0.0145	**0.0439**	0.0145	**0.0439**	0.0145	**0.0439**	0.0145
Sex (vs. *Female*)	-	-	-	-	-	-	-	-
Male	**0.5045**	0.0808	**0.5042**	0.0808	**0.5044**	0.0808	**0.5050**	0.0808
Heath Authority (vs *Fraser Health*)	-	-	-	-	-	-	-	-
Interior	**0.3233**	0.1053	**0.3221**	0.1053	**0.3232**	0.1053	**0.3260**	0.1054
Vancouver Coastal	**0.2044**	0.0966	**0.2045**	0.0966	**0.2043**	0.0966	**0.2038**	0.0966
Vancouver Island	0.1647	0.1116	0.1636	0.1116	0.1645	0.1116	0.1669	0.1117
Northern	0.0295	0.1973	0.0291	0.1973	0.0296	0.1973	0.0311	0.1973
Year Category (vs. *1996-2000*)	-	-	-	-	-	-	-	-
2001-2006	0.7590	0.4704	0.7585	0.4704	0.7589	0.4704	0.7601	0.4704
2007-2012	**0.9305**	0.4711	**0.9303**	0.4711	**0.9305**	0.4711	**0.9309**	0.4711
2013-2018	**1.5618**	0.4744	**1.5615**	0.4744	**1.5618**	0.4744	**1.5628**	0.4744
Alcohol or Other Substance Use Disorders	**1.2165**	0.1605	**1.2167**	0.1605	**1.2164**	0.1605	**1.2157**	0.1605
Ill Mental Health or Chronic pain	**-0.3629**	0.1302	**-0.3630**	0.1302	**-0.3627**	0.1302	**-0.3622**	0.1302
Hepatitis C Virus or HIV/AIDS	**0.6425**	0.1418	**0.6428**	0.1418	**0.6423**	0.1418	**0.6411**	0.1418
Ever Received a Sedative	**0.7458**	0.0993	**0.7462**	0.0993	**0.7459**	0.0993	**0.7452**	0.0993
Ever on PharmaCare Plans C or G	**-0.3532**	0.0807	**-0.3547**	0.0807	**-0.3534**	0.0807	**-0.3499**	0.0808
Age								
	Frailty Model		**Cox Model**		**Conditional Score IID Errors**		**Conditional Score AR(1) Errors**	
	$\hat{\mu} = 1.0722$		$\sigma^2/(1-\rho^2) \equiv 0$		$\hat{\sigma}^2/(1-\rho^2) = 0.9634$		$\hat{\sigma}^2/(1-\hat{\rho}^2) = 0.9634$	
	$\hat{\sigma}_\nu^2 = 4.5476$		$\rho \equiv 0$		$\rho \equiv 0$		$\hat{\rho} = 0.6646$	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
ν	**-0.0500**	0.0121	**-0.0480**	0.0119	**-0.0496**	0.0122	**-0.0537**	0.0127
Incarceration FPCS	**0.0419**	0.0147	**0.0419**	0.0147	**0.0420**	0.0147	**0.0420**	0.0147
Sex (vs. *Female*)	-	-	-	-	-	-	-	-
Male	**0.4751**	0.0810	**0.4749**	0.0810	**0.4757**	0.0810	**0.4761**	0.0810
Heath Authority (vs *Fraser Health*)	-	-	-	-	-	-	-	-
Interior	**0.2992**	0.1053	**0.2981**	0.1053	**0.2992**	0.1053	**0.3020**	0.1053
Vancouver Coastal	**0.1932**	0.0966	**0.1933**	0.0966	**0.1914**	0.0966	**0.1908**	0.0966
Vancouver Island	0.1356	0.1116	0.1346	0.1116	0.1354	0.1116	0.1378	0.1116
Northern	-0.0187	0.1973	-0.0190	0.1973	-0.0175	0.1973	-0.0158	0.1973
Year Category (vs. *1996-2000*)	-	-	-	-	-	-	-	-
2001-2006	0.7477	0.4628	0.7472	0.4628	0.7457	0.4628	0.7468	0.4628
2007-2012	0.8416	0.4532	0.8416	0.4532	0.8386	0.4532	0.8386	0.4532
2013-2018	**1.3431**	0.4525	**1.3430**	0.4525	**1.3378**	0.4525	**1.3388**	0.4525
Alcohol or Other Substance Use Disorders	**1.2683**	0.1607	**1.2684**	0.1607	**1.2674**	0.1607	**1.2664**	0.1607
Ill Mental Health or Chronic pain	**-0.3443**	0.1303	**-0.3443**	0.1303	**-0.3438**	0.1303	**-0.3435**	0.1303
Hepatitis C Virus or HIV/AIDS	**0.6174**	0.1418	**0.6177**	0.1418	**0.6168**	0.1418	**0.6152**	0.1418
Ever Received a Sedative	**0.7484**	0.0992	**0.7488**	0.0992	**0.7485**	0.0992	**0.7476**	0.0992
Ever on PharmaCare Plans C or G	**-0.3372**	0.0807	**-0.3388**	0.0806	**-0.3404**	0.0807	**-0.3367**	0.0808

Table 4.7: Estimates of the conditional score approach with $AR(1)$ errors from Table 4.1, and weighted estimates $\tilde{\gamma}$ and $\tilde{\boldsymbol{\theta}}$.

Time Since First Recorded OAT Dispensation				
	Conditional Score AR(1) Errors		**Weighted Average (Birth Generation)**	
Covariate Name	Estimate	S.E.	Estimate	S.E.
ν	**0.0539**	0.0044	**-0.0130**	0.0055
Incarceration FPCS	**0.0311**	0.0071	**0.0395**	0.0099
Sex (vs. *Female*)	-	-	-	-
Male	**0.1763**	0.0265	**0.2801**	0.0352
Birth Generation (vs *Greatest & Silent Generations*)	-	-	-	-
Baby Boomers	**-1.2419**	0.0459	-	-
Generation X	**-1.9431**	0.0495	-	-
Millennials & Generation Z	**-2.0166**	0.0613	-	-
Heath Authority (vs *Fraser Health*)	-	-	-	-
Interior	**0.2127**	0.0397	**0.2546**	0.0488
Vancouver Coastal	**0.1138**	0.0307	**0.1495**	0.0417
Vancouver Island	**0.0809**	0.0368	**0.1006**	0.0491
Northern	0.0546	0.0692	0.0156	0.0884
Year Category (vs. *1996-2000*)	-	-	-	-
2001-2006	**0.1541**	0.0337	**0.3454**	0.1679
2007-2012	**0.2446**	0.0399	**0.5198**	0.1688
2013-2018	**0.6084**	0.0511	**0.9488**	0.1722
Alcohol or Other Substance Use Disorders	**0.4229**	0.0448	**0.9222**	0.0741
Ill Mental Health or Chronic pain	**-0.1904**	0.0448	**-0.3096**	0.0579
Hepatitis C Virus or HIV/AIDS	**1.1612**	0.0276	**0.9651**	0.0546
Ever Received a Sedative	**0.4803**	0.0335	**0.6013**	0.0432
Ever on PharmaCare Plans C or G	**-0.2296**	0.0306	**-0.2310**	0.0380
Age				
	Conditional Score AR(1) Errors		**Weighted Average (Birth Generation)**	
Covariate Name	Estimate	S.E.	Estimate	S.E.
ν	**0.0509**	0.0046	**-0.0181**	0.0055
Incarceration FPCS	**0.0326**	0.0071	**0.0410**	0.0100
Sex (vs. *Female*)	-	-	-	-
Male	**0.1620**	0.0266	**0.2483**	0.0353
Birth Generation (vs *Greatest & Silent Generations*)	-	-	-	-
Baby Boomers	**-0.4401**	0.0684	-	-
Generation X	**-0.6223**	0.0833	-	-
Millennials & Generation Z	**-0.2715**	0.1087	-	-
Heath Authority (vs *Fraser Health*)	-	-	-	-
Interior	**0.1731**	0.0398	**0.2313**	0.0488
Vancouver Coastal	**0.0967**	0.0307	**0.1365**	0.0417
Vancouver Island	0.0576	0.0368	0.0827	0.0491
Northern	0.0493	0.0692	0.0014	0.0884
Year Category (vs. *1996-2000*)	-	-	-	-
2001-2006	0.0109	0.0320	0.2713	0.1651
2007-2012	-0.0343	0.0365	**0.3300**	0.1621
2013-2018	**0.2328**	0.0449	**0.6157**	0.1637
Alcohol or Other Substance Use Disorders	**0.5110**	0.0458	**0.9557**	0.0742
Ill Mental Health or Chronic pain	**-0.2726**	0.0446	**-0.3399**	0.0579
Hepatitis C Virus or HIV/AIDS	**1.1092**	0.0275	**0.9225**	0.0546
Ever Received a Sedative	**0.4601**	0.0335	**0.5893**	0.0432
Ever on PharmaCare Plans C or G	**-0.1510**	0.0313	**-0.1959**	0.0381

Table 4.8: Averaged bias for parameter Ω: $\sum_{j=1}^{1000}(\hat{\Omega}_j - \Omega)/1000$, where $\hat{\Omega}_j$ is an estimate of Ω with the jth data replicate ($j = 1, \cdots, 1000$).

Abbreviation: CS = Conditional Score.

$n = 5000$; $\gamma = -0.5$; $\rho = 0$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	-0.0009	0.0073	-0.0041	-0.0041
$\hat{\theta}_1$	-0.0035	0.0091	0.0008	0.0008
$\hat{\theta}_2$	0.0030	0.0144	0.0070	0.0070
$\hat{\theta}_3$	-0.0033	0.0113	0.0001	< 0.0001

$n = 5000$; $\gamma = 0$; $\rho = 0$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	0.0004	0.0004	0.0004	0.0004
$\hat{\theta}_1$	-0.0067	-0.0067	-0.0067	-0.0067
$\hat{\theta}_2$	-0.0013	-0.0013	-0.0013	-0.0013
$\hat{\theta}_3$	-0.0061	-0.0061	-0.0061	-0.0061

$n = 5000$; $\gamma = 0.5$; $\rho = 0$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	0.0008	-0.0073	0.0053	0.0053
$\hat{\theta}_1$	-0.0052	0.0075	-0.0049	-0.0049
$\hat{\theta}_2$	-0.0017	0.0094	-0.0016	-0.0016
$\hat{\theta}_3$	-0.0018	0.0131	-0.0025	-0.0025

$n = 5000$; $\gamma = -0.5$; $\rho = 0.5$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	-0.0009	0.0176	0.0112	-0.0022
$\hat{\theta}_1$	-0.0035	0.0237	0.0188	0.0090
$\hat{\theta}_2$	0.0030	0.0268	0.0225	0.0139
$\hat{\theta}_3$	-0.0033	0.0280	0.0215	0.0085

$n = 5000$; $\gamma = 0$; $\rho = 0.5$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	0.0004	0.0005	0.0005	0.0005
$\hat{\theta}_1$	-0.0067	-0.0067	-0.0067	-0.0067
$\hat{\theta}_2$	-0.0013	-0.0013	-0.0013	-0.0013
$\hat{\theta}_3$	-0.0061	-0.0061	-0.0061	-0.0061

$n = 5000$; $\gamma = 0.5$; $\rho = 0.5$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	0.0008	-0.0172	-0.0103	0.0048
$\hat{\theta}_1$	-0.0052	0.0215	0.0148	-0.0007
$\hat{\theta}_2$	-0.0017	0.0209	0.0149	0.0012
$\hat{\theta}_3$	-0.0018	0.0297	0.0212	0.0019

$n = 5000$; $\gamma = -0.5$; $\rho = 0.9$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	-0.0009	0.0546	0.0542	-0.0005
$\hat{\theta}_1$	-0.0035	0.0700	0.0697	0.0278
$\hat{\theta}_2$	0.0030	0.0672	0.0669	0.0302
$\hat{\theta}_3$	-0.0033	0.0842	0.0838	0.0301

$n = 5000$; $\gamma = 0$; $\rho = 0.9$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	0.0004	0.0001	0.0001	0.0001
$\hat{\theta}_1$	-0.0067	-0.0067	-0.0067	-0.0068
$\hat{\theta}_2$	-0.0013	-0.0012	-0.0012	-0.0012
$\hat{\theta}_3$	-0.0061	-0.0061	-0.0061	-0.0062

$n = 5000$; $\gamma = 0.5$; $\rho = 0.9$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	0.0008	-0.0550	-0.0546	0.0082
$\hat{\theta}_1$	-0.0052	0.0685	0.0682	0.0021
$\hat{\theta}_2$	-0.0017	0.0621	0.0618	0.0042
$\hat{\theta}_3$	-0.0018	0.0860	0.0855	0.0042

$n = 20000$; $\gamma = -0.5$; $\rho = 0$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	-0.0004	0.0079	-0.0033	-0.0033
$\hat{\theta}_1$	-0.0006	0.0122	0.0039	0.0039
$\hat{\theta}_2$	0.0005	0.0120	0.0047	0.0047
$\hat{\theta}_3$	0.0002	0.0152	0.0039	0.0039

$n = 20000$; $\gamma = 0$; $\rho = 0$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	< 0.0001	< 0.0001	< 0.0001	< 0.0001
$\hat{\theta}_1$	-0.0006	-0.0006	-0.0006	-0.0006
$\hat{\theta}_2$	0.0007	0.0007	0.0007	0.0007
$\hat{\theta}_3$	-0.0008	-0.0008	-0.0009	-0.0009

$n = 20000$; $\gamma = 0.5$; $\rho = 0$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	0.0003	-0.0081	0.0046	0.0047
$\hat{\theta}_1$	< 0.0001	0.0129	0.0006	0.0006
$\hat{\theta}_2$	-0.0004	0.0110	0.0001	0.0001
$\hat{\theta}_3$	-0.0016	0.0133	-0.0023	-0.0023

$n = 20000$; $\gamma = -0.5$; $\rho = 0.5$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	-0.0004	0.0180	0.0116	-0.0016
$\hat{\theta}_1$	-0.0006	0.0266	0.0217	0.0120
$\hat{\theta}_2$	0.0005	0.0244	0.0201	0.0116
$\hat{\theta}_3$	0.0002	0.0322	0.0256	0.0127

$n = 20000$; $\gamma = 0$; $\rho = 0.5$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	< 0.0001	0.0001	0.0001	0.0001
$\hat{\theta}_1$	-0.0006	-0.0006	-0.0006	-0.0006
$\hat{\theta}_2$	0.0007	0.0007	0.0007	0.0007
$\hat{\theta}_3$	-0.0008	-0.0008	-0.0008	-0.0008

$n = 20000$; $\gamma = 0.5$; $\rho = 0.5$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	0.0003	-0.0180	-0.0111	0.0040
$\hat{\theta}_1$	< 0.0001	0.0271	0.0204	0.0051
$\hat{\theta}_2$	-0.0004	0.0235	0.0176	0.0041
$\hat{\theta}_3$	-0.0016	0.0302	0.0218	0.0027

$n = 20000$; $\gamma = -0.5$; $\rho = 0.9$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	-0.0004	0.0549	0.0546	< 0.0001
$\hat{\theta}_1$	-0.0006	0.0731	0.0728	0.0304
$\hat{\theta}_2$	0.0005	0.0644	0.0642	0.0269
$\hat{\theta}_3$	0.0002	0.0882	0.0878	0.0334

$n = 20000$; $\gamma = 0$; $\rho = 0.9$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	< 0.0001	0.0001	0.0001	0.0001
$\hat{\theta}_1$	-0.0006	-0.0006	-0.0006	-0.0006
$\hat{\theta}_2$	0.0007	0.0007	0.0007	0.0007
$\hat{\theta}_3$	-0.0008	-0.0009	-0.0009	-0.0009

$n = 20000$; $\gamma = 0.5$; $\rho = 0.9$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	0.0003	-0.0553	-0.0549	0.0070
$\hat{\theta}_1$	< 0.0001	0.0731	0.0728	0.0086
$\hat{\theta}_2$	-0.0004	0.0632	0.0629	0.0065
$\hat{\theta}_3$	-0.0016	0.0860	0.0855	0.0067

$n = 50000$; $\gamma = -0.5$; $\rho = 0$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	-0.0002	0.0081	-0.0032	-0.0032
$\hat{\theta}_1$	-0.0013	0.0118	0.0035	0.0035
$\hat{\theta}_2$	0.0008	0.0121	0.0047	0.0047
$\hat{\theta}_3$	-0.0008	0.0143	0.0031	0.0031

$n = 50000$; $\gamma = 0$; $\rho = 0$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	-0.0002	-0.0002	-0.0002	-0.0002
$\hat{\theta}_1$	-0.0008	-0.0008	-0.0008	-0.0008
$\hat{\theta}_2$	-0.0006	-0.0006	-0.0006	-0.0006
$\hat{\theta}_3$	-0.0021	-0.0022	-0.0022	-0.0022

$n = 50000$; $\gamma = 0.5$; $\rho = 0$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	-0.0001	-0.0084	0.0042	0.0042
$\hat{\theta}_1$	-0.0007	0.0124	0.0001	0.0001
$\hat{\theta}_2$	-0.0020	0.0095	-0.0014	-0.0014
$\hat{\theta}_3$	-0.0013	0.0139	-0.0016	-0.0017

$n = 50000$; $\gamma = -0.5$; $\rho = 0.5$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	-0.0002	0.0183	0.0119	-0.0013
$\hat{\theta}_1$	-0.0013	0.0262	0.0213	0.0118
$\hat{\theta}_2$	0.0008	0.0247	0.0204	0.0119
$\hat{\theta}_3$	-0.0008	0.0314	0.0248	0.0120

$n = 50000$; $\alpha = 0$; $\rho = 0.5$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	-0.0002	-0.0002	-0.0002	-0.0002
$\hat{\theta}_1$	-0.0008	-0.0008	-0.0008	-0.0008
$\hat{\theta}_2$	-0.0006	-0.0006	-0.0006	-0.0006
$\hat{\theta}_3$	-0.0021	-0.0021	-0.0021	-0.0022

$n = 50000$; $\gamma = 0.5$; $\rho = 0.5$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	-0.0001	-0.0186	-0.0117	0.0033
$\hat{\theta}_1$	-0.0007	0.0268	0.0201	0.0049
$\hat{\theta}_2$	-0.0020	0.0223	0.0164	0.0030
$\hat{\theta}_3$	-0.0013	0.0310	0.0226	0.0035

$n = 50000$; $\gamma = -0.5$; $\rho = 0.9$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	-0.0002	0.0552	0.0549	-0.0013
$\hat{\theta}_1$	-0.0013	0.0725	0.0723	0.0288
$\hat{\theta}_2$	0.0008	0.0650	0.0647	0.0263
$\hat{\theta}_3$	-0.0008	0.0870	0.0867	0.0309

$n = 50000$; $\gamma = 0$; $\rho = 0.9$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	-0.0002	-0.0001	-0.0001	-0.0001
$\hat{\theta}_1$	-0.0008	-0.0008	-0.0008	-0.0008
$\hat{\theta}_2$	-0.0006	-0.0006	-0.0006	-0.0006
$\hat{\theta}_3$	-0.0021	-0.0021	-0.0021	-0.0022

$n = 50000$; $\gamma = 0.5$; $\rho = 0.9$				
	True ν	$\hat{\nu}$	CS (IID Errors)	CS (AR(1) Errors)
$\hat{\gamma}$	-0.0001	-0.0554	-0.0550	0.0059
$\hat{\theta}_1$	-0.0007	0.0731	0.0727	0.0094
$\hat{\theta}_2$	-0.0020	0.0622	0.0618	0.0064
$\hat{\theta}_3$	-0.0013	0.0869	0.0864	0.0084

Chapter 5

Strategies to Handle a Multivariate Internal Covariate Process in Survival Analysis

5.1 Introduction

Chapters 2-4 considered the OAT dispensation indicator as the primary internal covariate of interest, and omitted other factors pertaining to OAT usage. To better align with the National Academies of Sciences, Engineering, Medicine (2019) study report, we incorporate OAT type and dose into our modelling. Since the internal covariate is now multivariate, we proceed to extend the methodologies developed in Chapters 3 and 4 to the current setting.

This chapter is organized as follows. Section 5.2 presents the notation that will be utilized, and extend the modelling from the previous chapters to include a multivariate internal covariate. In particular, we extend the $AR(1)$ specification from (4.2) to a vector autoregressive process of order 1 ($VAR(1)$), which accounts for potential correlation between OAT usage factors. Section 5.3 extends the approaches presented in Sections 3.3 and 4.3. The extended approaches are applied to analyzing the provincial administrative database in Section 5.4. A summary and discussion is presented in Section 5.5.

5.2 Notation and Modelling

Let T denote an individual's survival time (measured in *days*) since their first recorded OAT dispensation record, on which the observations are subject to a right-censoring time, denoted by C. The available information on T is (T^*, δ), where $T^* = T \wedge C$ is the minimum between T and C, and $\delta = I(T \leq C)$. Let $\boldsymbol{Z}(t)$ represent a vector of relevant features pertaining to an individual's OAT dispensation at time $t \geq 0$, and $\boldsymbol{\mathcal{Z}}(t) = \{\boldsymbol{Z}(u) : 0 \leq u \leq t\}$. We additionally let $\boldsymbol{X}(t) = (X_1(t), \cdots, X_q(t))'$ denote external time-varying covariates and time-independent characteristics of an individual, and $\boldsymbol{\mathcal{X}}(t) = \{\boldsymbol{X}(u) : 0 \leq u \leq t\}$. Here,

the covariate $X_k(t)$ is time-independent if $X_k(0) \equiv X_k(t)$, for all $t > 0$, and $k = 1, \cdots, q$. Our objective is to estimate the conditional hazard function of T (at time t), given the processes $\boldsymbol{\mathcal{Z}}(t)$ and $\boldsymbol{\mathcal{X}}(t)$:

$$\lambda(t; \boldsymbol{\mathcal{Z}}(t), \boldsymbol{\mathcal{X}}(t)) = \lim_{\Delta t \to 0^+} \frac{1}{\Delta t} P(t \leq T < t + \Delta t | T \geq t, \boldsymbol{\mathcal{Z}}(t), \boldsymbol{\mathcal{X}}(t)). \tag{5.1}$$

In terms of OAT dispensation factors, we let $\boldsymbol{Z}(t) = (Z_1(t), Z_2(t), Z_3(t))'$, where $Z_1(t)$ denotes an individual's OAT dispensation indicator at time t, $Z_2(t)$ is a (time-varying) categorical variable pertaining to the OAT type dispensed at time t, and $Z_3(t)$ is the dosage dispensed at time t. Note that the values of $Z_2(t)$ and $Z_3(t)$ are only relevant if a person is dispensed OAT at time t. Based on this information, we specify $K + 1$ OAT dispensation classes. Thus we can reflect the information in $\boldsymbol{Z}(t)$ with the following categorical variable:

$$Z^*(t) = \begin{cases} 1 & \text{if an individual belongs to OAT dispensation class 1 at time } t \\ \cdots & \cdots \\ K & \text{if an individual belongs to OAT dispensation class } K \text{ at time } t \\ 0 & \text{if an individual belongs to OAT dispensation class } K+1 \text{ at time } t \end{cases}.$$

For $k = 1, \cdots, K$, let $Z_k^*(t) = I(Z^*(t) = k)$ denote K *dummy variables*, where the $(K+1)$th category is the reference category in our analyses.

5.2.1 Time-Varying Stratified Cox Regression Modelling

Let $\boldsymbol{W}(t) = (Z_1^*(t), \cdots, Z_K^*(t), \boldsymbol{X}(t)')'$ denote all covariates at time t. The most natural extension of (3.1) is the following generalized Cox regression model for the hazard function of T: for $t > 0$,

$$\lambda(t; \boldsymbol{\mathcal{Z}}(t), \boldsymbol{\mathcal{X}}(t)) = \lambda_0(t; \boldsymbol{\mathcal{Z}}(t)) \exp\{\boldsymbol{\theta}(\boldsymbol{\mathcal{Z}}(t))' \boldsymbol{W}(t)\}, \tag{5.2}$$

where $\lambda_0(t; \boldsymbol{\mathcal{Z}}(t))$ is an arbitrary baseline hazard function, and $\boldsymbol{\theta}(\boldsymbol{\mathcal{Z}}(t))$ is a known function up to finite dimensional parameters with the same dimension as $\boldsymbol{W}(t)$. We again let $g(\boldsymbol{\mathcal{Z}}(t)) \in \{1, 2, \cdots, G\}$ denote a stratification variable that is fully determined by an individual's dispensation history up to time $t > 0$, where $G < \infty$ is known. We assume $\lambda_0(t; \boldsymbol{\mathcal{Z}}(t)) = \lambda_{0g}(t)$ and $\boldsymbol{\theta}(\boldsymbol{\mathcal{Z}}(t)) = \boldsymbol{\theta}_g = (\boldsymbol{\alpha}_g', \boldsymbol{\beta}')'$ when $g(\boldsymbol{\mathcal{Z}}(t)) = g$, where $\boldsymbol{\theta}_g$ is a vector of regression parameters. Here, $\boldsymbol{\alpha}_g$ is a q_A-dimensional vector of stratum-specific effects, and $\boldsymbol{\beta}$ is a q_B-dimensional vector of shared effects across strata. Without loss of generality, we partition the covariates as $\boldsymbol{W}(t) = (\boldsymbol{W}^A(t)', \boldsymbol{W}^B(t)')'$, where $\boldsymbol{W}^A(t)$ and $\boldsymbol{W}^B(t)$ have the same dimensions as $\boldsymbol{\alpha}_g$ and $\boldsymbol{\beta}$, respectively. The model in (5.2) then becomes

$$\lambda(t; \boldsymbol{\mathcal{Z}}(t), \boldsymbol{\mathcal{X}}(t)) = \lambda_{0g}(t) \exp\{\boldsymbol{\theta}_g' \boldsymbol{W}(t)\} \tag{5.3}$$

when $g(\boldsymbol{\mathcal{Z}}_i(t)) = g$. That is, (5.3) is a direct extension to the model (3.2), where we have now included multiple binary alternating covariates. By fixing $G \equiv 1$, we can simplify the model in (5.3) into

$$\lambda(t; \boldsymbol{\mathcal{Z}}(t), \boldsymbol{\mathcal{X}}(t)) = \lambda_0(t) \exp\{\boldsymbol{\theta}' \boldsymbol{W}(t)\}, \tag{5.4}$$

which is similar to the model (3.4), without considering stratification.

5.2.2 Jointly Modelling OAT Dispensation Process and Mortality Risk

To extend our modelling from Chapter 4, let $R_{ik}(t) = \int_0^t Z_{ik}^*(u)du/t$ denote the proportion of time individual i is in OAT dispensation class k during the study period up to time t, for $k = 1, \cdots, K$. Letting $R_{ijk} = R_{ik}(t_{ij})$, where $\{t_{ij} : j = 1, \cdots, m_i\}$ is the set of chosen discretization observation times, with $0 \leq t_{i,m_i} \leq T_i^*$. It is natural to extend (4.2) by modelling each R_{ijk} $(k = 1, \cdots, K)$. We consider the transformation R_{ijk} to $Y_{ijk} = \log(R_{ijk}/(1 - \sum_{k=1}^{K} R_{ijk}))$ so that $Y_{ijk} \in (-\infty, \infty)$. To account for the correlation between OAT factors, consider the following longitudinal sub-model

$$\boldsymbol{Y}_{ij} = \boldsymbol{\nu}_i + \boldsymbol{\varepsilon}_{ij}, \tag{5.5}$$

where $\boldsymbol{Y}_{ij} = (Y_{ij1}, \cdots, Y_{ijK})'$, $\boldsymbol{\nu}_i = (\nu_{i1}, \cdots, \nu_{iK})'$, $\boldsymbol{\varepsilon}_{ij} = (\varepsilon_{ij1}, \cdots, \varepsilon_{ijK})'$. Motivated by our previous analysis in Chapter 4, $\{\boldsymbol{\varepsilon}_{ij} : j = 1, \cdots, m_i\}$ is assumed to follow a mean zero $VAR(1)$ process. Specifically,

$$\boldsymbol{\varepsilon}_{ij} = \boldsymbol{P}\boldsymbol{\epsilon}_{i,j-1} + \boldsymbol{e}_{ij},$$

where $\boldsymbol{P}$ is the $K \times K$ matrix with (k, ℓ) element $\rho_{k\ell}$, and $\boldsymbol{e}_{ij} = (e_{ij1}, \cdots, e_{ijK})'$ satisfies with $\mathbb{E}(\boldsymbol{e}_{ij}) = \boldsymbol{0}$, $Var(\boldsymbol{e}_{ij}) = \boldsymbol{\Xi}$ is a diagonal matrix with σ_k^2 along the main diagonal, and $\boldsymbol{e}_{ij} \perp\!\!\!\perp \boldsymbol{\varepsilon}_{i,j-1}$. We assume that the $VAR(1)$ process is stable, meaning that the eigenvalues of $\boldsymbol{P}$ have modulus less than one. We can show that $\mathbb{E}(\boldsymbol{\epsilon}_{ij}\boldsymbol{\epsilon}_{i,j-\ell}') = \boldsymbol{P}^\ell$, and $vec(Var(\boldsymbol{\epsilon}_{ij})) = (\boldsymbol{I}_{K^2 \times K^2} - \boldsymbol{P} \otimes \boldsymbol{P})^{-1} vec(\boldsymbol{\Xi})$, where $vec(\boldsymbol{A})$ is a vector that stacks the columns of the matrix $\boldsymbol{A}$, $\boldsymbol{I}_{P \times P}$ is the $P \times P$ identity matrix, and $\boldsymbol{A} \otimes \boldsymbol{B}$ is the kronecker product between matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ (Lütkepohl 2005).

By summarizing the OAT dispensation process with $\boldsymbol{\nu}_i$, that is, assuming $T_i \perp\!\!\!\perp \boldsymbol{\mathcal{Z}}_i(\infty) | \{\boldsymbol{\nu}_i, \boldsymbol{\mathcal{X}}_i(t)\}$, we specify model (5.1) into

$$\lambda(t; \boldsymbol{\mathcal{Z}}_i(t), \boldsymbol{\mathcal{X}}_i(t)) = \lambda_0(t) \exp\{\boldsymbol{\gamma}' \boldsymbol{\nu}_i + \boldsymbol{\theta}' \boldsymbol{X}_i(t)\} \quad \text{for } i = 1, \cdots, n, \tag{5.6}$$

where $\lambda_0(t)$ is an unspecified baseline hazard function, and $\boldsymbol{\gamma} = (\gamma_1, \cdots, \gamma_K)'$ and $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_q)'$ are unknown regression parameters.

5.3 Estimation Procedures

For the models presented in Sections 5.2.1 and 5.2.2, we proceed to extend the estimation procedures developed in Chapters 3 and 4, respectively.

5.3.1 Estimating Equation Under the Time-Varying Stratified Cox Model (5.3)

As in Chapter 3, let $N_i(t) = I(T_i \leq t)$, $Y_i(t) = I(T_i^* \geq t)$, and $\boldsymbol{\Theta} = (\boldsymbol{\alpha}_1', \cdots, \boldsymbol{\alpha}_G', \boldsymbol{\beta}')'$ be all the regression parameters in model (5.3), with $\boldsymbol{\Theta}_0$ denoting the true value of $\boldsymbol{\Theta}$. Under (5.3), consider the following estimating function for $\boldsymbol{\Theta}$: $\boldsymbol{U}(\boldsymbol{\Theta}) = (U_1^A(\boldsymbol{\theta}_1)', \cdots, U_G^A(\boldsymbol{\theta}_G)', U^B(\boldsymbol{\Theta})')'$, where

$$
\begin{aligned}
U_g^A(\boldsymbol{\theta}_g) &= \int_0^\infty \sum_{i:g(\boldsymbol{\mathcal{Z}}_i(t))=g} Y_i(t) \left[\boldsymbol{W}_i^A(t) - \frac{E_g^A(t, \boldsymbol{\theta}_g)}{E_g(t, \boldsymbol{\theta}_g)} \right] dN_i(t), \quad g = 1, \cdots, G, \\
U^B(\boldsymbol{\Theta}) &= \sum_{g=1}^G \int_0^\infty \sum_{i:g(\boldsymbol{\mathcal{Z}}_i(t))=g} Y_i(t) \left[\boldsymbol{W}_i^B(t) - \frac{E_g^B(t, \boldsymbol{\theta}_g)}{E_g(t, \boldsymbol{\theta}_g)} \right] dN_i(t), \\
E_g(t, \boldsymbol{\theta}) &= \sum_{i:g(\boldsymbol{\mathcal{Z}}_i(t))=g} Y_i(t) \exp\{\boldsymbol{\theta}' \boldsymbol{W}_i(t)\}, \\
E_g^A(t, \boldsymbol{\theta}) &= \sum_{i:g(\boldsymbol{\mathcal{Z}}_i(t))=g} Y_i(t) \exp\{\boldsymbol{\theta}' \boldsymbol{W}_i(t)\} \boldsymbol{W}_i^A(t), \text{ and} \\
E_g^B(t, \boldsymbol{\theta}) &= \sum_{i:g(\boldsymbol{\mathcal{Z}}_i(t))=g} Y_i(t) \exp\{\boldsymbol{\theta}' \boldsymbol{W}_i(t)\} \boldsymbol{W}_i^B(t).
\end{aligned}
$$

We can again view $\boldsymbol{U}(\boldsymbol{\Theta})$ as an estimating function for $\boldsymbol{\Theta}$, and by extending the derivation in Section 3.4.1 we can show that $\boldsymbol{U}(\boldsymbol{\Theta}_0)$ is centred at zero asymptotically under (5.3), so that a consistent estimate for $\boldsymbol{\Theta}$ is therefore the solution to the equation $\boldsymbol{U}(\boldsymbol{\Theta}) = \boldsymbol{0}$. Furthermore, we can extend the derivation in Section 3.4.3 to show that $\sqrt{n}(\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0)$ converges in distribution to a zero-mean normal random variable, in which the corresponding asymptotic variance of $\sqrt{n}(\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0)$ can be consistently estimated with $\widehat{\boldsymbol{AV}}(\hat{\boldsymbol{\Theta}}) = \hat{\boldsymbol{\Psi}}^{-1}(\hat{\boldsymbol{\Theta}}) \hat{\boldsymbol{\Phi}}(\hat{\boldsymbol{\Theta}}) \hat{\boldsymbol{\Psi}}^{-1}(\hat{\boldsymbol{\Theta}})$, where $\hat{\boldsymbol{\Psi}}(\boldsymbol{\Theta}) = -\frac{1}{n} \frac{\partial}{\partial \boldsymbol{\Theta}} \boldsymbol{U}(\boldsymbol{\Theta})$, $\hat{\boldsymbol{\Phi}}(\boldsymbol{\Theta}) = \frac{1}{n} \sum_{i=1}^n \hat{\boldsymbol{\Omega}}_i(\boldsymbol{\Theta}) \hat{\boldsymbol{\Omega}}_i(\boldsymbol{\Theta})'$,

$$
\begin{aligned}
\hat{\boldsymbol{\Omega}}_i(\boldsymbol{\Theta}) &= (\hat{\Omega}_{i1}^A(\boldsymbol{\theta}_1)', \cdots, \hat{\Omega}_{iG}^A(\boldsymbol{\theta}_G)', \hat{\Omega}_i^B(\boldsymbol{\Theta})')', \\
\hat{\Omega}_{ig}^A(\boldsymbol{\theta}) &= \int_0^\infty Y_i(t) I(g(\boldsymbol{\mathcal{Z}}_i(t)) = g) \left[\boldsymbol{W}_i^A(t) - \frac{E_g^A(t, \boldsymbol{\theta})}{E_g(t, \boldsymbol{\theta})} \right] d\hat{M}_{ig}(t, \boldsymbol{\theta}), \\
\hat{\Omega}_i^B(\boldsymbol{\Theta}) &= \sum_{g=1}^G \int_0^\infty Y_i(t) I(g(\boldsymbol{\mathcal{Z}}_i(t)) = g) \left[\boldsymbol{W}_i^B(t) - \frac{E_g^B(t, \boldsymbol{\theta}_g)}{E_g(t, \boldsymbol{\theta}_g)} \right] d\hat{M}_{ig}(t, \boldsymbol{\theta}_g), \\
\hat{M}_{ig}(t, \boldsymbol{\theta}) &= N_i(t) I(g(\boldsymbol{\mathcal{Z}}_i(t)) = g) - \int_0^t Y_i(u) I(g(\boldsymbol{\mathcal{Z}}_i(u)) = g) \exp\{\boldsymbol{\theta}' \boldsymbol{W}_i(u)\} d\hat{\Lambda}_{0g}(u),
\end{aligned}
$$

and

$$d\hat{\Lambda}_{0g}(t;\boldsymbol{\theta}_g) = \sum_{i:g(\boldsymbol{\mathcal{Z}}_i(t))=g} \frac{Y_i(t)dN_i(t)}{\sum\limits_{j:g(\boldsymbol{\mathcal{Z}}_j(t))=g} Y_j(t)\exp\{\boldsymbol{\theta}_g'\boldsymbol{W}_j(t)\}}. \tag{5.7}$$

Here, we take the convention that $0/0 = 0$. By replacing the unknown $\boldsymbol{\theta}_g$ with its corresponding estimate, $\hat{\boldsymbol{\theta}}_g = (\hat{\boldsymbol{\alpha}}_g', \hat{\boldsymbol{\beta}}')'$, the baseline hazard function is estimated with a Breslow-type estimator $d\hat{\Lambda}_{0g}(t;\hat{\boldsymbol{\theta}}_g)$. We can extend the argument in Section (3.4.4) to show that $d\hat{\Lambda}_{0g}(t;\hat{\boldsymbol{\theta}}_g)$ is a consistent estimator for $d\Lambda_{0g}(t)$. Moreover, we can adopt the Wald test from Section 3.3.3.

5.3.2 Estimating Parameters in the Joint Model

5.3.2.1 Multivariate Conditional Score Approach

We take the approach from Chapter 4 and estimate $\boldsymbol{\nu}_i$ under (5.5) with

$$\hat{\boldsymbol{\nu}}_i = \frac{1}{m_i}\sum_{j=1}^{m_i} \boldsymbol{Y}_{ij}.$$

Viewing $\hat{\boldsymbol{\nu}}_i$ as a "noisy measurement" of $\boldsymbol{\nu}_i$, Song *et al.* (2002) extended the conditional score approach of Tsiatis and Davidian (2001) to the settings where multiple longitudinal covariates are measured with errors. The idea is the same as if that we derive a sufficient statistic of $\boldsymbol{\nu}_i$, which depends on either known or estimable quantities. Although Song *et al.* (2002) allowed for possible correlation between multiple covariates, they assumed that successive observations over time are independent. As shown by our data analysis in Chapter 4, this assumption is clearly violated. That motivated us to extend their derivation to allow for autocorrelated errors.

By letting $n_i(t) = I(T_i \in [t, t+dt))$ and $Y_i(t) = I(T_i^* \geq t)$, and

$$\boldsymbol{\Sigma}_i = \boldsymbol{\Sigma}_i(\boldsymbol{\sigma}, \boldsymbol{P}) = Var(\hat{\boldsymbol{\nu}}_i) = \frac{1}{m_i^2}\left(m_i Var(\boldsymbol{\varepsilon}_{ij}) + 2\sum_{j=1}^{m_i-1}\sum_{\ell=j+1}^{m_i} \mathbb{E}(\boldsymbol{\varepsilon}_{ij}\boldsymbol{\varepsilon}_{i\ell}')\right),$$

the conditional likelihood of $\{n_i(t), \hat{\boldsymbol{\nu}}_i\}$ given $\{Y_i(t) = 1, \boldsymbol{\nu}_i, \boldsymbol{\mathcal{X}}_i(t)\}$ can be expressed up to order dt as

$$\begin{aligned}
&[n_i(t), \hat{\boldsymbol{\nu}}_i | Y_i(t) = 1, \boldsymbol{\mathcal{X}}_i(t), \boldsymbol{\nu}_i] \\
&= \exp\{\boldsymbol{S}_i(t, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_i)'\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\nu}_i\}\frac{(\lambda_0(t)\exp\{\boldsymbol{\theta}'\boldsymbol{X}_i(t)\}dt)^{n_i(t)}}{(2\pi)^{k/2}\det(\boldsymbol{\Sigma}_i)^{1/2}}\exp\{-\hat{\boldsymbol{\nu}}_i'\boldsymbol{\Sigma}_i^{-1}\hat{\boldsymbol{\nu}}_i/2 - \boldsymbol{\nu}_i'\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\nu}_i/2\},
\end{aligned} \tag{5.8}$$

where $\boldsymbol{\sigma} = (\sigma_1^2, \cdots, \sigma_K^2)$, $\boldsymbol{\Sigma}_i$ depends on i only through m_i, and we assume
(i) $n_i(t) \perp\!\!\!\perp \hat{\boldsymbol{\nu}}_i | \{Y_i(t) = 1, \boldsymbol{\mathcal{X}}_i(t), \boldsymbol{\nu}_i\}$, and (ii) $\hat{\boldsymbol{\nu}}_i \perp\!\!\!\perp \{Y_i(t) = 1, \boldsymbol{\mathcal{X}}_i(t)\} | \boldsymbol{\nu}_i$. The first assumption is essentially the "nondifferential measurement error mechanism" assumption within the measurement error literature (Carroll *et al.* 2006, Yi 2017), and the second assumption follows from our model specification in (5.5). We recognize (5.8) to be a member of the exponential family of probability distributions, so that

$$\boldsymbol{S}_i(t) \equiv \boldsymbol{S}_i(t, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_i) = \hat{\boldsymbol{\nu}}_i + n_i(t)\boldsymbol{\Sigma}_i\boldsymbol{\gamma} \tag{5.9}$$

is a sufficient statistic for $\boldsymbol{\nu}_i$. Following along the arguments presented in Song *et al.* (2002), we can show that

$$\begin{aligned}\lambda^{\dagger}(t; \boldsymbol{\mathcal{S}}_i(t), \boldsymbol{\mathcal{X}}_i(t)) &= \lim_{\Delta t \to 0^+} \frac{1}{\Delta t} P(n_i(t) = 1 | Y_i(t) = 1, \boldsymbol{\mathcal{S}}_i(t), \boldsymbol{\mathcal{X}}_i(t)) \\ &= \lambda_0(t) \exp\{\boldsymbol{\gamma}'\boldsymbol{S}_i(t) - \frac{1}{2}\boldsymbol{\gamma}'\boldsymbol{\Sigma}_i\boldsymbol{\gamma} + \boldsymbol{\theta}'\boldsymbol{X}_i(t)\},\end{aligned} \tag{5.10}$$

where $\boldsymbol{\mathcal{S}}_i(t) = \{\boldsymbol{S}_i(u) : 0 \leq u \leq t\}$. Since (5.10) involves known or estimable quantities, and the parameters in (4.6) are the same as in (5.6). This motivates us to base our inference around (5.10).

5.3.2.2 Estimating Model Parameters

Let $N_i(t) = I(T_i \leq t)$, $Y_i(t) = I(T_i^* \geq t)$, and $E_r(t, \boldsymbol{\gamma}, \boldsymbol{\theta}, \boldsymbol{\sigma}, \boldsymbol{P}) = \sum_{j=1}^{n} E_{rj}(t, \boldsymbol{\gamma}, \boldsymbol{\theta}, \boldsymbol{\sigma}, \boldsymbol{P})$ with

$$E_{rj}(t, \boldsymbol{\gamma}, \boldsymbol{\theta}, \boldsymbol{\sigma}, \boldsymbol{P}) = Y_j(t) \exp\{\boldsymbol{\gamma}'\boldsymbol{S}_j(t) - \frac{1}{2}\boldsymbol{\gamma}'\boldsymbol{\Sigma}_j\boldsymbol{\gamma} + \boldsymbol{\theta}'\boldsymbol{X}_j(t)\}(\boldsymbol{S}_j(t)', \boldsymbol{X}_j(t)')^{\prime \otimes r}.$$

For the time being, assume that $\boldsymbol{\sigma}$ and $\boldsymbol{P}$ are known, so that we can compute the sufficient statistic in (5.9). Under the model in (5.10), the following estimating equations are (conditionally) unbiased:

$$\sum_{i=1}^{n} (Y_i(t)dN_i(t) - E_{0i}(t, \boldsymbol{\gamma}, \boldsymbol{\theta}, \boldsymbol{\sigma}, \boldsymbol{P})d\Lambda_0(t)) = 0 \text{ for } t > 0 \tag{5.11}$$

$$\sum_{i=1}^{n} \int_0^{\infty} (\boldsymbol{S}_i(t)', \boldsymbol{X}_i(t)')'(Y_i(t)dN_i(t) - E_{0i}(t, \boldsymbol{\gamma}, \boldsymbol{\theta}, \boldsymbol{\sigma}, \boldsymbol{P})d\Lambda_0(t)) = \boldsymbol{0}, \tag{5.12}$$

where $d\Lambda_0(t) = \lambda_0(t)\ dt$. Solving for $d\Lambda_0(t)$ in (5.11) results in the following Breslow-type estimator

$$d\hat{\Lambda}_0(t; \boldsymbol{\gamma}, \boldsymbol{\theta}, \boldsymbol{\sigma}, \boldsymbol{P}) = \sum_{i=1}^{n} \frac{Y_i(t)dN_i(t)}{\sum_{j=1}^{n} E_{0j}(t, \boldsymbol{\gamma}, \boldsymbol{\theta}, \boldsymbol{\sigma}, \boldsymbol{P})}. \tag{5.13}$$

Inserting (5.7) in place of $d\Lambda_0(t)$ in (5.12) yields

$$U_1(\boldsymbol{\gamma}, \boldsymbol{\theta}; \boldsymbol{\sigma}, \boldsymbol{P}) = \sum_{i=1}^{n} \int_0^{\infty} Y_i(t) \left((\boldsymbol{S}_i(t)', \boldsymbol{X}_i(t)')' - \frac{E_1(t, \boldsymbol{\gamma}, \boldsymbol{\theta}, \boldsymbol{\sigma}, \boldsymbol{P})}{E_0(t, \boldsymbol{\gamma}, \boldsymbol{\theta}, \boldsymbol{\sigma}, \boldsymbol{P})} \right) dN_i(t)$$

We propose to estimate $(\boldsymbol{\gamma}', \boldsymbol{\theta}')'$ with $(\hat{\boldsymbol{\gamma}}', \hat{\boldsymbol{\theta}}')'$, which is the solution to $U_1(\boldsymbol{\gamma}, \boldsymbol{\theta}; \boldsymbol{\sigma}, \boldsymbol{P}) = \mathbf{0}$. To estimate $d\Lambda_0(t)$, we can plug-in $\hat{\gamma}$ and $\hat{\boldsymbol{\theta}}$ in (5.13).

5.3.2.3 Estimating Additional Parameters

Letting $\boldsymbol{\mathcal{E}}_i^y = (\boldsymbol{\varepsilon}_{i2}, \cdots, \boldsymbol{\varepsilon}_{i,m_i})$, $\boldsymbol{\mathcal{E}}_i^x = (\boldsymbol{\varepsilon}_{i1}, \cdots, \boldsymbol{\varepsilon}_{i,m_i-1})$, and $\boldsymbol{e}_i = (\boldsymbol{e}_{i2}, \cdots, \boldsymbol{e}_{i,m_i})$, we can express the $VAR(1)$ model from (5.5) in matrix notation as

$$\begin{aligned} \boldsymbol{\mathcal{E}}_i^y &= \boldsymbol{P}\boldsymbol{\mathcal{E}}_i^x + \boldsymbol{e}_i, \\ vec(\boldsymbol{\mathcal{E}}_i^y) &= vec(\boldsymbol{P}\boldsymbol{\mathcal{E}}_i^x) + vec(\boldsymbol{e}_i) \\ &= ((\boldsymbol{\mathcal{E}}_i^x)' \otimes \boldsymbol{I}_{K\times K}) vec(\boldsymbol{P}) + vec(\boldsymbol{e}_i). \end{aligned}$$

With known $\boldsymbol{\mathcal{E}}_i^y$ and $\boldsymbol{\mathcal{E}}_i^x$, we could then estimate $vec(\boldsymbol{P})$ with

$$\begin{aligned} \widehat{vec(\boldsymbol{P})} &= \underset{vec(\boldsymbol{P})}{\operatorname{argmin}} \left\{ \sum_{i=1}^{n} vec(\boldsymbol{e}_i)' vec(\boldsymbol{e}_i) \right\} \\ &= \left(\sum_{i=1}^{n} \{\boldsymbol{\mathcal{E}}_i^x (\boldsymbol{\mathcal{E}}_i^x)' \otimes \boldsymbol{I}_{K\times K}\} \right)^{-1} \left(\sum_{i=1}^{n} (\boldsymbol{\mathcal{E}}_i^x \otimes \boldsymbol{I}_{K\times K}) vec(\boldsymbol{\mathcal{E}}_i^y) \right). \end{aligned}$$

With $\boldsymbol{\mathcal{E}}_i^y$ and $\boldsymbol{\mathcal{E}}_i^x$ unknown, we replace them with $\hat{\boldsymbol{\mathcal{E}}}_i^y = (\hat{\boldsymbol{\varepsilon}}_{i2}, \cdots, \hat{\boldsymbol{\varepsilon}}_{i,m_i})$ and $\hat{\boldsymbol{\mathcal{E}}}_i^x = (\hat{\boldsymbol{\varepsilon}}_{i1}, \cdots, \hat{\boldsymbol{\varepsilon}}_{i,m_i-1})$, where $\hat{\boldsymbol{\varepsilon}}_{ij} = \boldsymbol{Y}_{ij} - \hat{\boldsymbol{\nu}}_i$. Specifically, note that $\widehat{vec(\boldsymbol{P})}$ is the solution to $U_2(\boldsymbol{P}) = \mathbf{0}$, where

$$U_2(\boldsymbol{P}) = \sum_{i=1}^{n} \{\hat{\boldsymbol{\mathcal{E}}}_i^x \otimes \boldsymbol{I}_{K\times K}\} \{((\hat{\boldsymbol{\mathcal{E}}}_i^x)' \otimes \boldsymbol{I}_{K\times K}) vec(\boldsymbol{P}) - vec(\hat{\boldsymbol{\mathcal{E}}}_i^y)\}.$$

Furthermore, it is natural to estimate $\boldsymbol{\sigma}$ under the following model

$$vec(\boldsymbol{\mathcal{E}}_i^y) = ((\boldsymbol{\mathcal{E}}_i^x)' \otimes \boldsymbol{I}_{K^2\times K^2}) vec(\boldsymbol{P}) + vec(\boldsymbol{e}_i).$$

By pooling all of the individuals together, we estimate $\boldsymbol{\sigma}$ with

$$\hat{\boldsymbol{\sigma}} = \frac{1}{|\mathcal{I}|} \sum_{i\in\mathcal{I}} \frac{1}{m_i - 1} \hat{\boldsymbol{e}}_{ij}' \hat{\boldsymbol{e}}_{ij},$$

where $\mathcal{I} = \{i : m_i > 1\}$, and $\hat{\boldsymbol{e}}_{ij} = \hat{\boldsymbol{\varepsilon}}_{ij} - \hat{\boldsymbol{P}}\hat{\boldsymbol{\varepsilon}}_{i,j-1}$. Note that $\hat{\boldsymbol{\sigma}}$ is the solution to $U_3(\boldsymbol{\sigma}, \boldsymbol{P}) = \boldsymbol{0}$, where

$$U_3(\boldsymbol{\sigma}, \boldsymbol{P}) = \sum_{i \in \mathcal{I}} \left(\boldsymbol{\sigma} - \frac{1}{m_i - 1} (\hat{\boldsymbol{\varepsilon}}_{ij} - \boldsymbol{P}\hat{\boldsymbol{\varepsilon}}_{i,j-1})'(\hat{\boldsymbol{\varepsilon}}_{ij} - \boldsymbol{P}\hat{\boldsymbol{\varepsilon}}_{i,j-1}) \right).$$

5.3.2.4 Two-Stage Inference Procedure & Variance Estimation

Let $\boldsymbol{U}(\boldsymbol{\gamma}, \boldsymbol{\theta}, \boldsymbol{\sigma}, \boldsymbol{P}) = (U_1(\boldsymbol{\gamma}, \boldsymbol{\theta}, ; \boldsymbol{\sigma}, \boldsymbol{P})', U_{23}(\boldsymbol{\sigma}, \boldsymbol{P})')'$ with $U_{23}(\boldsymbol{\sigma}, \boldsymbol{P}) = (U_2(\boldsymbol{P})', U_3(\boldsymbol{\sigma}, \boldsymbol{P})')'$, where each estimating function as the sum of each individual's contribution: $U_1(\boldsymbol{\gamma}, \boldsymbol{\theta}; \boldsymbol{\sigma}, \boldsymbol{P}) = \sum_{i=1}^{n} U_{i1}(\boldsymbol{\gamma}, \boldsymbol{\theta}; \boldsymbol{\sigma}, \boldsymbol{P})$ and $U_{23}(\boldsymbol{\sigma}, \boldsymbol{P}) = \sum_{i=1}^{n} U_{i,23}(\boldsymbol{\sigma}, \boldsymbol{P})$. Note that $\hat{\boldsymbol{\sigma}}$ and $\widehat{vec(\boldsymbol{P})}$ have analytic forms. We infer all of the parameters with the following two-stage procedure:

Step 1: Under the model $\boldsymbol{Y}_{ij} = \boldsymbol{\nu}_i + \boldsymbol{\varepsilon}_{ij}$, estimate $\boldsymbol{\nu}_i$ with $\hat{\boldsymbol{\nu}}_i = \sum_{j=1}^{m_i} \boldsymbol{Y}_{ij}/m_i$, and proceed to estimate $vec(\boldsymbol{P})$ and $\boldsymbol{\sigma}$ with $\widehat{vec(\boldsymbol{P})}$ and $\hat{\boldsymbol{\sigma}}$, respectively.

Step 2: Estimate $\boldsymbol{\gamma}$ and $\boldsymbol{\theta}$ with the solutions to $U_1(\boldsymbol{\gamma}, \boldsymbol{\theta}; \hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{P}}) = \boldsymbol{0}$. Also, estimate $d\Lambda_0(t)$ with $d\hat{\Lambda}_0(t; \hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{P}})$.

Letting $\boldsymbol{\Omega} = (\boldsymbol{\gamma}', \boldsymbol{\theta}')'$ and $\hat{\boldsymbol{\Omega}} = (\hat{\boldsymbol{\gamma}}', \hat{\boldsymbol{\theta}}')'$ denote the resulting estimate of $\boldsymbol{\Omega}$ under our two-stage inference procedure, it follows that $Var(\hat{\boldsymbol{\Omega}}) = \boldsymbol{\Psi}^{-1}\boldsymbol{\Phi}\boldsymbol{\Psi}^{-1}$ (c.f. Chapter 1 of Yi (2017)), where $\boldsymbol{\Psi} = \mathbb{E}\left(\frac{\partial}{\partial \boldsymbol{\Omega}} U_{11}(\boldsymbol{\Omega}; \boldsymbol{\sigma}, \boldsymbol{P})\right)$ and $\boldsymbol{\Phi} = \mathbb{E}(Q_1(\boldsymbol{\Omega}, \boldsymbol{\sigma}, \boldsymbol{P}) Q_1(\boldsymbol{\Omega}, \boldsymbol{\sigma}, \boldsymbol{P})')$ with

$$Q_i(\boldsymbol{\Omega}, \boldsymbol{\sigma}, \boldsymbol{P}) = U_{i1}(\boldsymbol{\Omega}; \boldsymbol{\sigma}, \boldsymbol{P}) - \mathbb{E}\left(\frac{\partial}{\partial(\boldsymbol{\sigma}, \boldsymbol{P})'} U_{i1}(\boldsymbol{\Omega}; \boldsymbol{\sigma}, \boldsymbol{P})\right) \left[\mathbb{E}\left(\frac{\partial}{\partial(\boldsymbol{\sigma}, \boldsymbol{P})'} U_{i,23}(\boldsymbol{\sigma}, \boldsymbol{P})\right)\right]^{-1} U_{i,23}(\boldsymbol{\sigma}, \boldsymbol{P}).$$

By replacing $\boldsymbol{\Psi}$ and $\boldsymbol{\Phi}$ with their empirical means: $\hat{\boldsymbol{\Psi}} = \frac{1}{n} \sum_{i=1}^{n} \frac{\partial}{\partial \boldsymbol{\Omega}} U_{i1}(\boldsymbol{\Omega}; \hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{P}}) \Big|_{\boldsymbol{\Omega} = \hat{\boldsymbol{\Omega}}}$ and $\hat{\boldsymbol{\Phi}} = \frac{1}{n} \sum_{i=1}^{n} \hat{Q}_i(\hat{\boldsymbol{\Omega}}, \hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{P}}) \hat{Q}_i(\hat{\boldsymbol{\Omega}}, \hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{P}})'$ with

$$\hat{Q}_i(\boldsymbol{\Omega}, \boldsymbol{\sigma}, \boldsymbol{P}) = U_{i1}(\boldsymbol{\Omega}; \boldsymbol{\sigma}, \boldsymbol{P}) - \left[\frac{\partial}{\partial(\boldsymbol{\sigma}, \boldsymbol{P})'} U_1(\boldsymbol{\Omega}; \boldsymbol{\sigma}, \boldsymbol{P})\right] \left[\frac{\partial}{\partial(\boldsymbol{\sigma}, \boldsymbol{P})'} U_{23}(\boldsymbol{\sigma}, \boldsymbol{P})\right]^{-1} U_{i,23}(\boldsymbol{\sigma}, \boldsymbol{P}),$$

we have $\widehat{Var}(\hat{\boldsymbol{\Omega}}) = \hat{\boldsymbol{\Psi}}^{-1}\hat{\boldsymbol{\Phi}}\hat{\boldsymbol{\Psi}}^{-1}$ is consistent for $Var(\hat{\boldsymbol{\Omega}})$.

5.4 Analysis of the Provincial OAT Dispensation Records (III)

We proceed to apply the proposed inference procedures to the provincial OAT dispensation records. We specify eight OAT dispensation classes, pertaining to OAT type and dosage levels. In particular, we follow Piske *et al.* (2021) with the following dose levels for each OAT type:

Dispensed Low Dose of Methadone: OAT type is methadone and the dose is ≤ 40 mg.

Dispensed Medium Dose of Methadone: OAT type is methadone and the dose is in $(40, 85)$ mg.

Dispensed High Dose of Methadone: OAT type is methadone and the dose is ≥ 85 mg.

Dispensed Low Dose of Buprenorphine: OAT type is buprenorphine and the dose is ≤ 6 mg.

Dispensed Medium Dose of Buprenorphine: OAT type is buprenorphine and the dose is in $(6, 16)$ mg.

Dispensed High Dose of Buprenorphine: OAT type is buprenorphine and the dose is ≥ 16 mg.

Dispensed Low Dose of Slow Release Oral Morphine: OAT type is slow release oral morphine and the dose is ≤ 350 mg.

Dispensed Medium Dose of Slow Release Oral Morphine: OAT type is slow release oral morphine and the dose is in $(350, 750)$ mg.

Dispensed High Dose of Slow Release Oral Morphine: OAT type is slow release oral morphine and the dose is ≥ 750 mg.

Dispensed Low Dose of Hydromorphone: OAT type is injectable hydromorphone and the dose is ≤ 70 mg.

Dispensed Medium Dose of Hydromorphone: OAT type is injectable hydromorphone and the dose is in $(70, 150)$ mg.

Dispensed High Dose of Hydromorphone: OAT type is injectable hydromorphone and the dose is ≥ 150 mg.

However, due to slow release oral morphine and injectable hydromorphone being unavailable throughout most of the data extraction window, the amount of information pertaining to these OAT types is rather limited. This motivated us to specify the following $K = 8$ OAT dispensation classes:

(i) *Dispensed Low Dose of Methadone*

(ii) *Dispensed Medium Dose of Methadone*

(iii) *Dispensed High Dose of Methadone*

(iv) *Dispensed Low Dose of Buprenorphine*

(v) *Dispensed Medium Dose of Buprenorphine*

(vi) *Dispensed High Dose of Buprenorphine*

(vii) *Dispensed Low Dose of Other OAT Type*

(viii) *Dispensed Medium or High Dose of Other OAT Type*

Here, "Other OAT Type" refers to slow release oral morphine and injectable hydromorphone.

5.4.1 Analysis Under Time-Varying Stratified Cox Regression Modelling (5.3)

We started our analysis by fitting the extended Cox regression model (5.4) to the observed data, where aside from the OAT dispensation process variables, we used the same set of covariates from the analysis in Chapter 3. Table 5.1 displays the corresponding estimates in both *time since first recorded OAT dispensation* and *age* time scales. We see that the OAT dispensation indicator estimate in Table 3.1 resembles a weighted average of the OAT dispensation class estimates in Table 5.1, and the other estimates in Table 5.1 are similar to those in Table 3.1.

We then proceeded to fit a stratified Cox regression model (5.3), where we again stratified individuals (at time t) according to their OAT episode number. As in Chapter 3, we initially specified the following $G = 9$ levels for the stratification variable: (i) 1 OAT episode; (ii) 2-3 OAT episodes; (iii) 4-5 OAT episodes; (iv) 6-7 OAT episodes; (v) 8-10 OAT episodes; (vi) 11-15 OAT episodes; (vii) 16-20 OAT episodes; (viii) 21-30 OAT episodes; and (ix) 31+ OAT episodes. By using the results from Chapter 3, we specified the *OAT dispensation class* variables, the *birth generation indicators*, and the *ever on PharmaCare plans C or G* indicator to have varying effects, and specify the effects for other variables to be constant across strata. Figures 5.1 and 5.2 displays the estimates of $\boldsymbol{\theta}_g$ for $g = 1, \cdots, G$. We also estimated the baseline hazard functions with (5.7), and illustrate the LOESS-smoothed

estimates in the top row in Figure 5.3. The same conclusions apply from interpreting the top row of Figure 3.6.

We conducted the Wald test, where the results of the eight tests are shown in Tables 5.2 and 5.3. For each test, we displayed the estimates of both $\boldsymbol{\alpha}_{g-1}$ and $\boldsymbol{\alpha}_g$, the test statistic J_g in (3.7), and the resulting p-value. The difference from our Chapter 3 analysis comes from *age* time scale, where the test reveals the following $H = 3$ levels for the stratification variable (i) 1 OAT episode; (ii) 2-10 OAT episodes; (iii) 11+ OAT episodes. The results with the updated stratification are illustrated in Figures 5.4 and 5.5. We can see the protective effect of *buprenorphine* against mortality increases as individuals increase their OAT episode number. The bottom row of Figure 5.3 illustrates the LOESS-smoothed estimates of the baseline hazard function estimates (5.7), in which we see the baseline hazard functions do not overlap, and promotes our model specification of having strata-specific baseline hazard functions in (5.3).

5.4.2 Joint Modelling of OAT Dispensation (5.5) and Mortality Risk Processes (5.6)

We began our analysis by obtaining $\hat{\boldsymbol{\nu}}_i$ under our model (5.5), for each individual. Due to the *non-methadone* OATs being introduced relatively late during the data extraction period, there is a limited amount of information pertaining to the usage of these OAT types, as revealed in Tables 5.4-5.8. We proceeded to instead consider $K = 4$ OAT usage classes, where we merged all of the *non-methadone* OAT classes together. Furthermore, to overcome some computational challenges, we defined $R_{ijk} = 1.0 \times 10^{-10}$ if $R_{ijk} = 0$ and $R_{ijk} = 1 - 1.0 \times 10^{-10}$ if $R_{ijk} = 1$. To be consistent with our analysis in Chapter 4, we again specified the time unit for measurements in (5.5) to be in *months*. Compared to the descriptive results in Chapter 4 however, it is less apparent that *age* is a confounding variable, and difficult to assess how the OAT classes are associated with mortality risk.

Table 5.9 presents the resulting estimates of $\boldsymbol{\gamma}$ and $\boldsymbol{\theta}$, where we considered *time since first recorded OAT dispensation* and *age* as time scales. The three sets of estimates correspond to: (a) estimates obtained by directly replacing $\boldsymbol{\nu}_i$ with $\hat{\boldsymbol{\nu}}_i$ in (5.6), (b) estimates obtained from the conditional score approach as described in Section 5.3 with $\rho_{k\ell} = 0$ when $k \neq \ell$, and (c) estimates obtained from the conditional score approach as described in Section 5.3. We see that the estimates pertaining to the OAT dispensation process from (b) & (c) varies from the estimates in (a), which is to be expected. The fact that the estimates in (b) to (c) is due to the correlation between processes, and based on the results from Chapter 4, we anticipate the estimates in (b) to underestimate the parameters. The same conclusions apply to Tables 5.10-5.13, where we stratified individuals according to their birth generation.

Letting n_g denote the number of individuals that belong to the gth birth generation and $(\hat{\boldsymbol{\gamma}}_g, \hat{\boldsymbol{\theta}}_g)$ denote the estimates from Tables 5.10-5.13 ($g = 1, 2, 3, 4$), we computed the

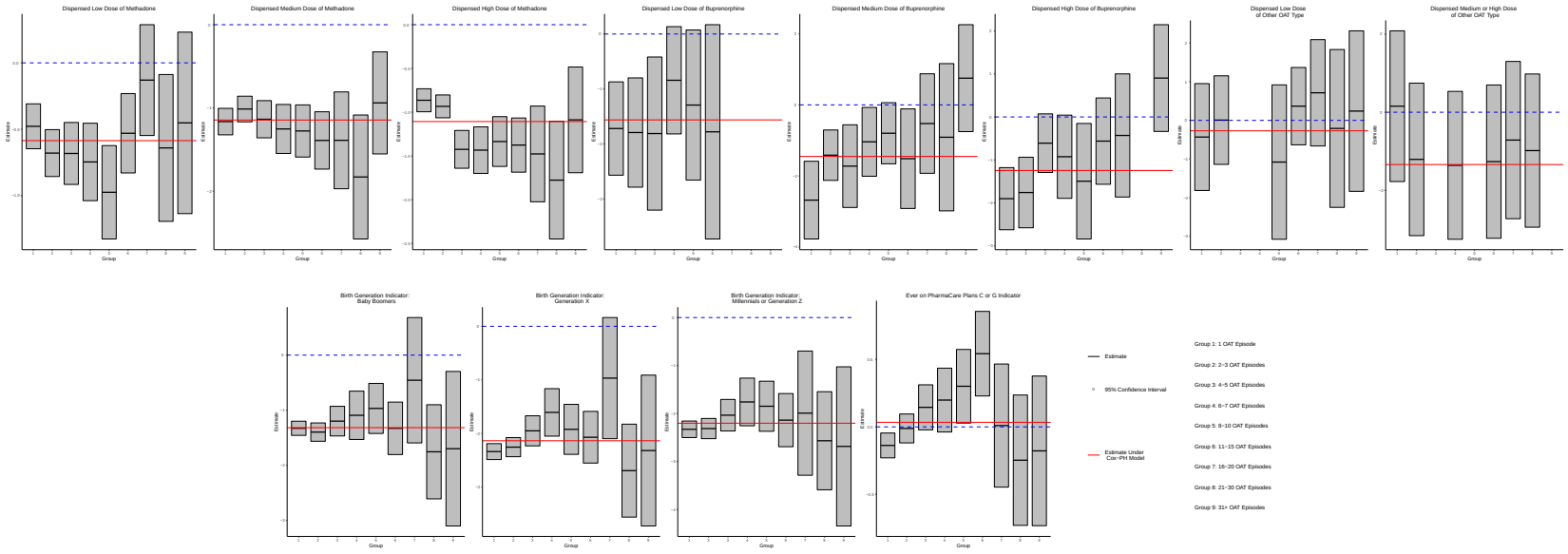
weighted estimates of $\boldsymbol{\gamma}$ and $\boldsymbol{\theta}$ as $\tilde{\gamma} = \sum_{g=1}^{4} \frac{n_g}{n} \hat{\gamma}_g$ and $\tilde{\boldsymbol{\theta}} = \sum_{g=1}^{4} \frac{n_g}{n} \hat{\boldsymbol{\theta}}_g$. The corresponding variances of $\tilde{\gamma}$ and $\tilde{\boldsymbol{\theta}}$ are $Var(\tilde{\gamma}) = Var\left(\sum_{g=1}^{G} \frac{N_g}{N} \hat{\gamma}_g\right) \approx \sum_{g=1}^{G} \left(\frac{N_g}{N}\right)^2 Var(\hat{\gamma}_g)$ and $Var(\tilde{\boldsymbol{\theta}}) = Var\left(\sum_{g=1}^{G} \frac{N_g}{N} \hat{\boldsymbol{\theta}}_g\right) \approx \sum_{g=1}^{G} \left(\frac{N_g}{N}\right)^2 Var(\hat{\boldsymbol{\theta}}_g)$. Table 5.14 presents $\tilde{\gamma}$ and $\tilde{\boldsymbol{\theta}}$, as well as conditional score estimates from Table 5.9 as a reference. Since *age* no longer appears to be a confounder, the weighted average is similar to the conditional score estimates.

5.5 Discussion

To accommodate additional features present in the OAT dispensation process, we extend the inference procedures developed in Chapters 3 and 4 to accommodate a multivariate internal covariate process. Specifically, we introduced a set of dummy variables pertaining to different combinations of OAT type and dosage levels. By directly estimating their effect on the hazard function, our time-dependent stratified Cox regression model shows buprenorphine to have the largest protective effect against mortality. This observation is less obvious when we jointly model the OAT dispensation and mortality risk processes, but the proposed extension accommodates for the correlation between the OAT dispensation processes. Using the results from Chapter 4, omitting this correlation would underestimate the parameters. Moreover, the results from the conditional score method can be used to provide survival predictions.

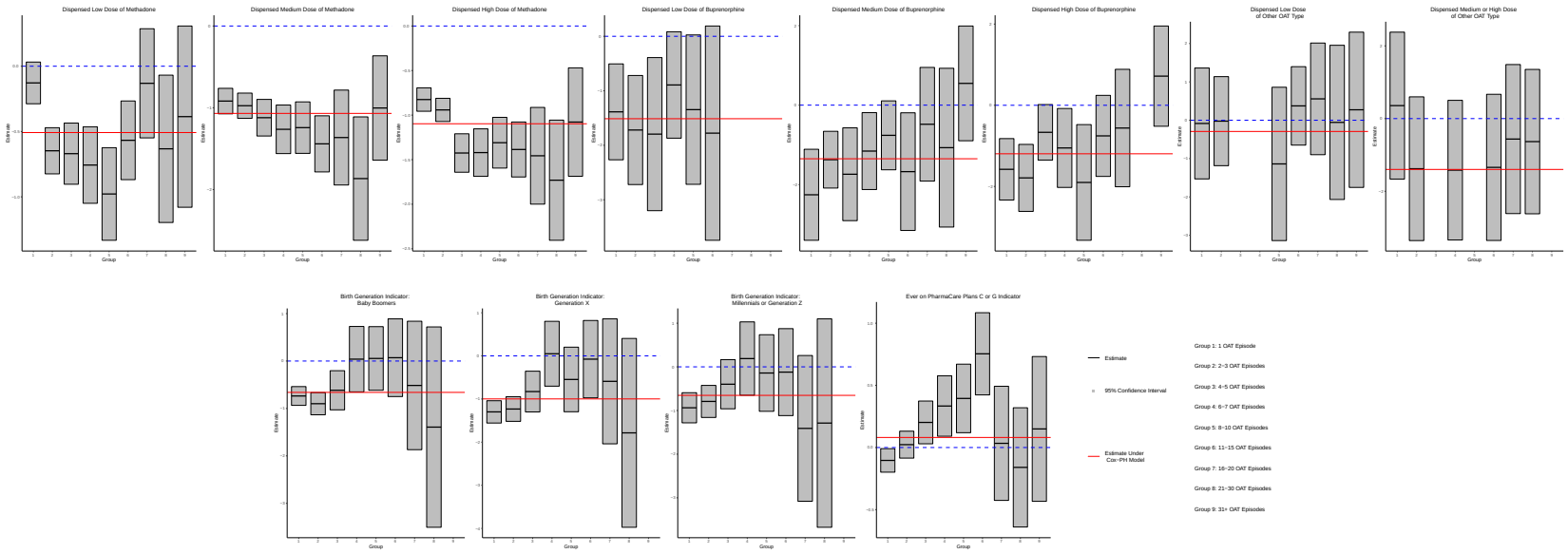
Although our data application pertains to OAT dispensation records, we anticipate our modelling to have broad applications where the effect of a high-dimensional internal covariate process on mortality risk is of interest. For example, longitudinal metabolomics studies can in principle adopt both modelling approaches, but the joint modelling approach might be more appealing in order to accommodate for potential measurement error. Provided that the researcher can adequately specify $\boldsymbol{\Sigma}_i$ in (5.10) and has collected enough data to estimate it, we anticipate the proposed two-stage inference procedure from Section 5.3.2.4 to be computationally feasible.

Figure 5.1: Estimates of regression coefficients under the stratified Cox regression model (5.3), where the time scale is *time since first recorded OAT dispensation*, and $\boldsymbol{\theta}_g = (\boldsymbol{\alpha}_g', \boldsymbol{\beta}')'$. Variables with effect $\boldsymbol{\alpha}_g$ and $\boldsymbol{\beta}$ are illustrated and tabulated below, respectively.



Covariate Name	Estimate	S.E.
Sex (vs. *Female*)	-	-
Male	**0.2314**	0.0260
Heath Authority (vs *Fraser Health*)	-	-
Interior	**0.1611**	0.0385
Vancouver Coastal	**0.1056**	0.0301
Vancouver Island	**0.0794**	0.0366
Northern	0.0137	0.0680
Year Category (vs. *1996-2000*)	-	-
2001-2006	0.0635	0.0337
2007-2012	**0.1270**	0.0386
2013-2018	**0.4784**	0.0505
Alcohol or Other Substance Use Disorders	**0.4320**	0.0448
Ill Mental Health or Chronic pain	**-0.1893**	0.0428
Hepatitis C Virus or HIV/AIDS	**1.1690**	0.0273
Ever Received a Sedative	**0.4783**	0.0335
Incarceration Status	**-1.6111**	0.1729
Number of Incarcerations	0.0022	0.0030

Figure 5.2: Estimates of regression coefficients under the stratified Cox regression model (5.3), where the time scale is *age*, and $\boldsymbol{\theta}_g = (\boldsymbol{\alpha}_g', \boldsymbol{\beta}')'$. Variables with effect $\boldsymbol{\alpha}_g$ and $\boldsymbol{\beta}$ are illustrated and tabulated below, respectively.



Covariate Name	Estimate	S.E.
Sex (vs. *Female*)	-	-
Male	**0.2053**	0.0261
Heath Authority (vs *Fraser Health*)	-	-
Interior	**0.1342**	0.0389
Vancouver Coastal	**0.0943**	0.0301
Vancouver Island	0.0550	0.0367
Northern	0.0133	0.0680
Year Category (vs. *1996-2000*)	-	-
2001-2006	-0.0399	0.0319
2007-2012	-0.0569	0.0369
2013-2018	**0.3749**	0.0466
Alcohol or Other Substance Use Disorders	**0.4779**	0.0461
Ill Mental Health or Chronic pain	**-0.3208**	0.0425
Hepatitis C Virus or HIV/AIDS	**1.1195**	0.0273
Ever Received a Sedative	**0.4537**	0.0336
Incarceration Status	**-1.5239**	0.1744
Number of Incarcerations	0.0043	0.0030

Figure 5.3: Smoothed estimates of $\lambda_{0g}(\cdot)$, where we stratify by the OAT episode number at time t, where the time scale is *time since first recorded OAT dispensation* and *age*.
Top Row: Estimates prior to the Wald test.
Bottom Row: Estimates following the Wald test.

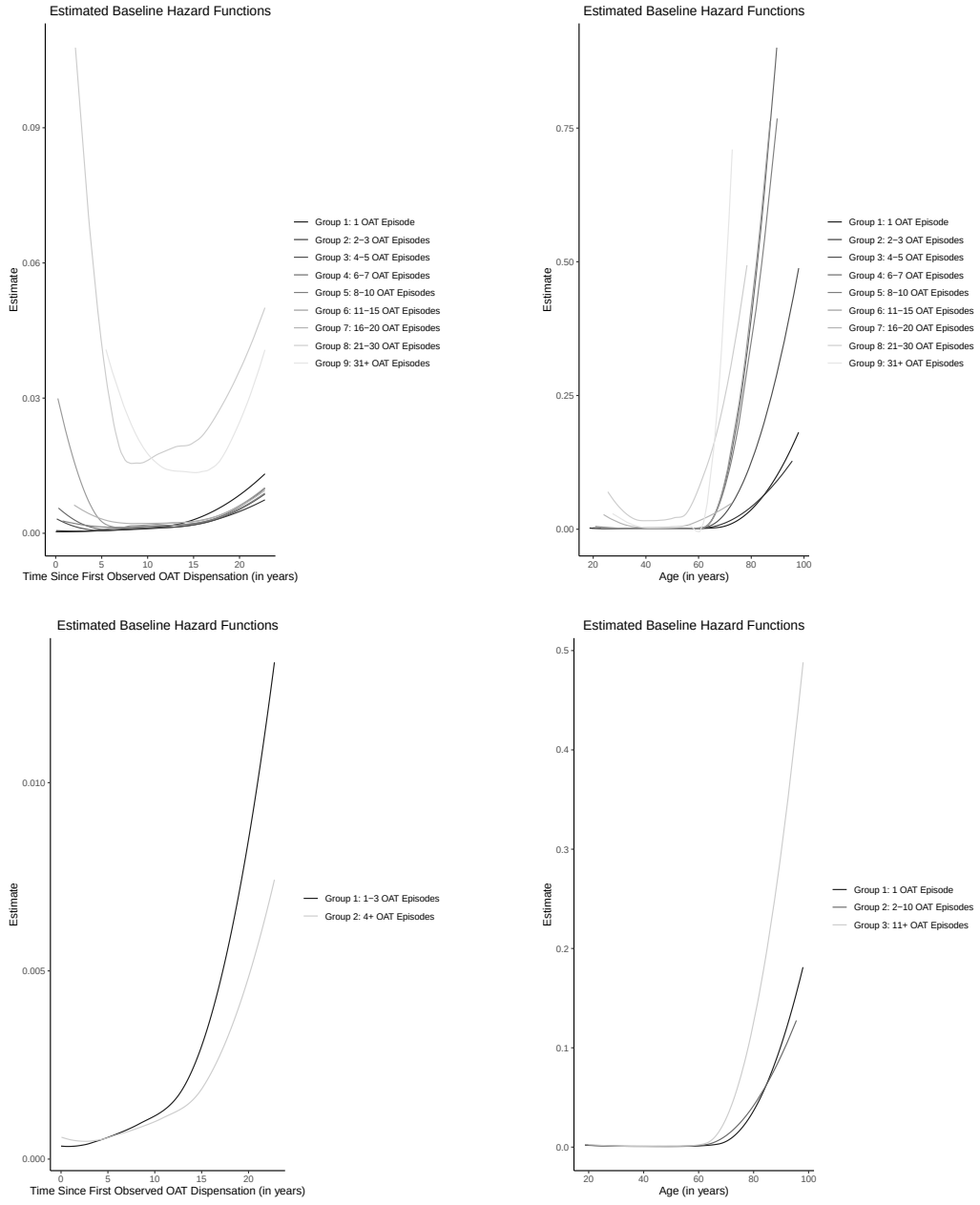
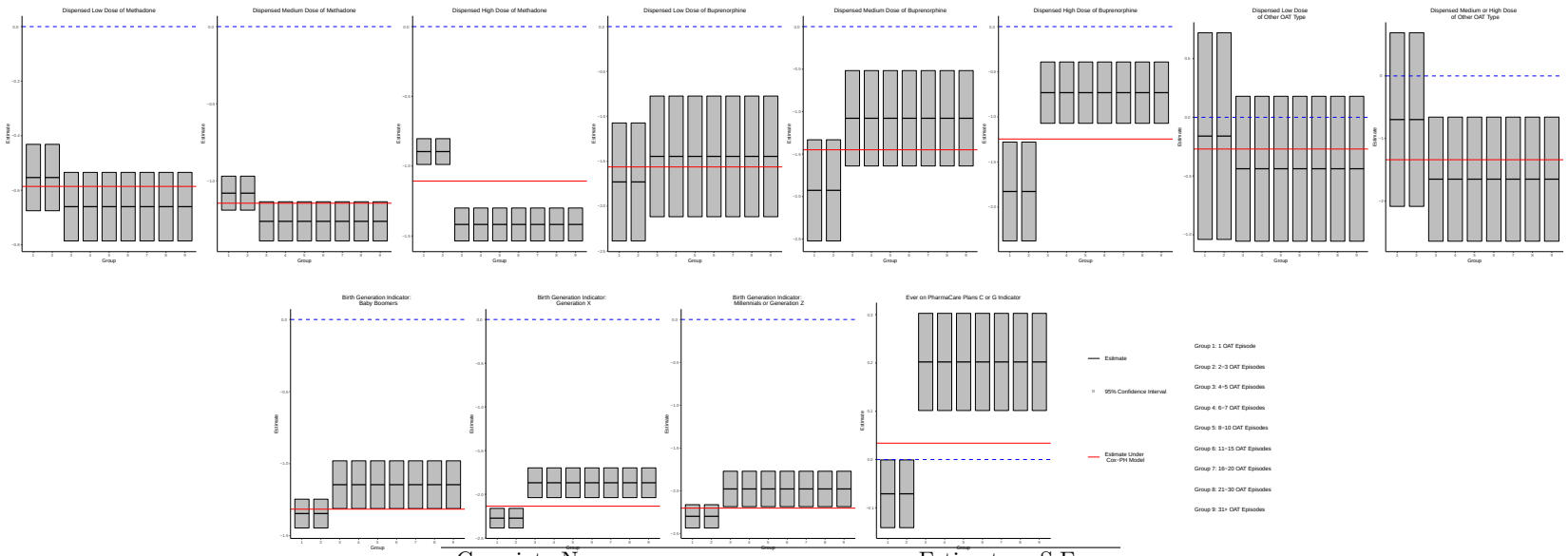
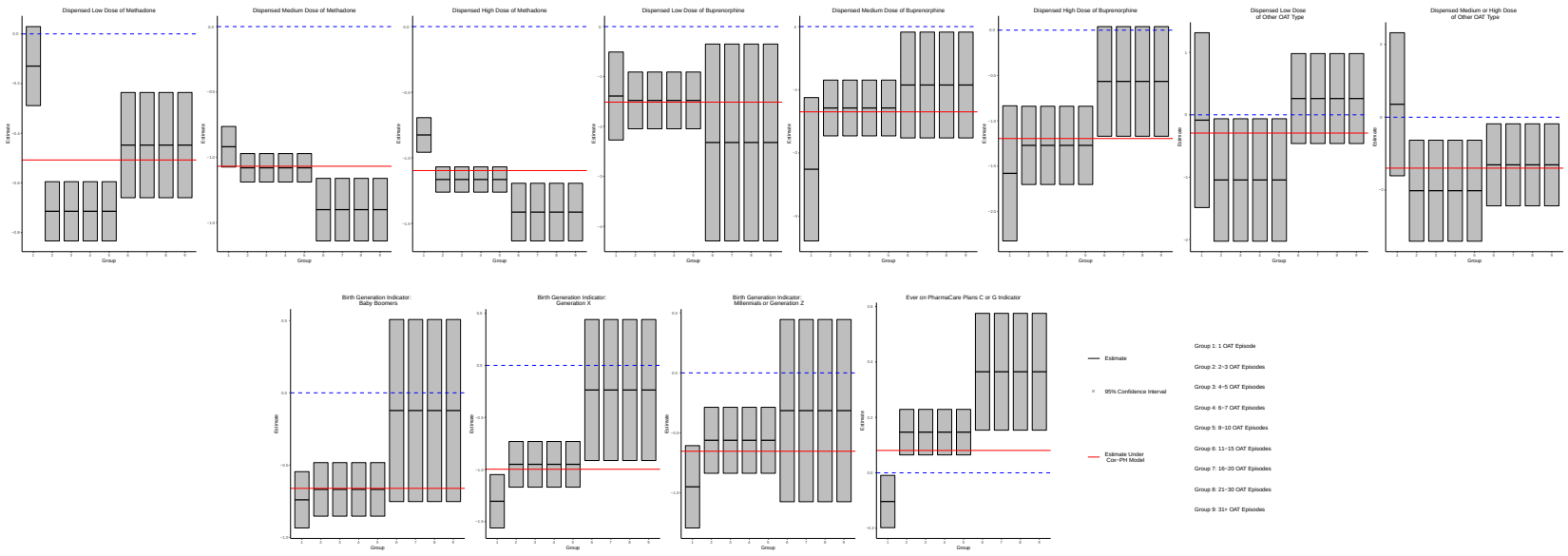


Figure 5.4: Estimates of regression coefficients under the stratified Cox regression model (5.3) following the Wald test, where the time scale is *time since first recorded OAT dispensation*, and $\boldsymbol{\theta}_g = (\boldsymbol{\alpha}_g', \boldsymbol{\beta}')'$. Variables with effect $\boldsymbol{\alpha}_g$ and $\boldsymbol{\beta}$ are illustrated and tabulated below, respectively.



Covariate Name	Estimate	S.E.
Sex (vs. *Female*)	-	-
Male	**0.2311**	0.0262
Heath Authority (vs *Fraser Health*)	-	-
Interior	**0.1592**	0.0387
Vancouver Coastal	**0.1057**	0.0302
Vancouver Island	**0.0776**	0.0362
Northern	0.0094	0.0676
Year Category (vs. *1996-2000*)	-	-
2001-2006	**0.0706**	0.0330
2007-2012	**0.1393**	0.0392
2013-2018	**0.4923**	0.0497
Alcohol or Other Substance Use Disorders	**0.4457**	0.0423
Ill Mental Health or Chronic pain	**-0.1910**	0.0431
Hepatitis C Virus or HIV/AIDS	**1.1741**	0.0273
Ever Received a Sedative	**0.4784**	0.0333
Incarceration Status	**-1.6067**	0.1787
Number of Incarcerations	0.0031	0.0029

Figure 5.5: Estimates of regression coefficients under the stratified Cox regression model (5.3) following the Wald test, where the time scale is *age*, and $\boldsymbol{\theta}_g = (\boldsymbol{\alpha}_g', \boldsymbol{\beta}')'$. Variables with effect $\boldsymbol{\alpha}_g$ and $\boldsymbol{\beta}$ are illustrated and tabulated below, respectively.



Covariate Name	Estimate	S.E.
Sex (vs. *Female*)	-	-
Male	**0.2071**	0.0264
Heath Authority (vs *Fraser Health*)	-	-
Interior	**0.1312**	0.0389
Vancouver Coastal	**0.1014**	0.0302
Vancouver Island	0.0526	0.0363
Northern	0.0080	0.0677
Year Category (vs. *1996-2000*)	-	-
2001-2006	-0.0379	0.0316
2007-2012	-0.0464	0.0367
2013-2018	**0.3822**	0.0452
Alcohol or Other Substance Use Disorders	**0.4838**	0.0436
Ill Mental Health or Chronic pain	**-0.3195**	0.0429
Hepatitis C Virus or HIV/AIDS	**1.1131**	0.0273
Ever Received a Sedative	**0.4513**	0.0332
Incarceration Status	**-1.5132**	0.1782
Number of Incarcerations	0.0036	0.0029

Table 5.1: Estimates of regression coefficients under the Cox regression model (5.4). The reported standard-error (S.E.) estimates of $\hat{\boldsymbol{\Theta}}$ correspond to the square-root of the diagonal elements of $\widehat{\boldsymbol{AV}}(\hat{\boldsymbol{\Theta}})$. The **bolded** estimates are statistically significant with the type 1 error rate set at $\alpha^* = 5\%$.

Time Scale	Time Since First Observed OAT Dispensation		Age	
Covariate Name	Estimate	S.E.	Estimate	S.E.
Dispensed Low Dose of Methadone	**-0.5859**	0.0448	**-0.5079**	0.0442
Dispensed Medium Dose of Methadone	**-1.1450**	0.0430	**-1.0689**	0.0428
Dispensed High Dose of Methadone	**-1.1072**	0.0375	**-1.0971**	0.0375
Dispensed Low Dose of Buprenorphine	**-1.5628**	0.2369	**-1.5126**	0.2367
Dispensed Medium Dose of Buprenorphine	**-1.4500**	0.1843	**-1.3492**	0.1839
Dispensed High Dose of Buprenorphine	**-1.2459**	0.1663	**-1.1953**	0.1661
Dispensed Low Dose of Other OAT Type	-0.2703	0.2685	-0.2928	0.2685
Dispensed Medium or High Dose of Other OAT Type	**-1.3395**	0.4091	**-1.3991**	0.4090
Sex (vs. *Female*)	-	-	-	-
Male	**0.2242**	0.0261	**0.2058**	0.0263
Birth Generation (vs *Greatest & Silent Generations*)	-	-	-	-
Baby Boomers	**-1.3159**	0.0444	**-0.6610**	0.0663
Generation X	**-2.1358**	0.0478	**-0.9975**	0.0813
Millennials & Generation Z	**-2.2016**	0.0591	**-0.6531**	0.1067
Heath Authority (vs *Fraser Health*)	-	-	-	-
Interior	**0.1517**	0.0387	**0.1251**	0.0388
Vancouver Coastal	**0.1031**	0.0302	**0.0980**	0.0302
Vancouver Island	**0.0726**	0.0362	0.0524	0.0362
Northern	0.0040	0.0677	0.0047	0.0677
Year Category (vs. *1996-2000*)	-	-	-	-
2001-2006	**0.0701**	0.0330	-0.0451	0.0314
2007-2012	**0.1452**	0.0392	-0.0584	0.0361
2013-2018	**0.4863**	0.0497	**0.3506**	0.0439
Alcohol or Other Substance Use Disorders	**0.4491**	0.0424	**0.4857**	0.0434
Ill Mental Health or Chronic pain	**-0.1903**	0.0431	**-0.3239**	0.0428
Hepatitis C Virus or HIV/AIDS	**1.1764**	0.0273	**1.1141**	0.0272
Ever Received a Sedative	**0.4788**	0.0332	**0.4542**	0.0332
Ever on PharmaCare Plans C or G	0.0339	0.0299	**0.0808**	0.0305
Incarceration Status	**-1.6159**	0.1783	**-1.5226**	0.1784
Number of Incarcerations	**0.0064**	0.0028	0.0049	0.0028

Table 5.2: Results of the Wald test when we stratify by the number of OAT episode by time t, where the time scale is *time since first recorded OAT dispensation*, and we applied a Bonferonni correction so that the type 1 error rate is $\alpha^* = 0.05/(G-1)$. For each test, the estimates on the left- and right-hand side correspond to $\hat{\boldsymbol{\alpha}}_{g-1}$ and $\hat{\boldsymbol{\alpha}}_g$ and their estimated standard errors, respectively. **Bolded** test statistic(s) and p-value(s) indicate tests that rejected the null hypothesis in (3.6).

	Test 1: Group 1 vs. Group 2				**Test 2**: Groups 1,2 vs. Group 3				**Test 3**: Group 3 vs. Group 4			
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Dispensed Low Dose of Methadone	-0.4779	0.0863	-0.6797	0.0901	-0.5536	0.0626	-0.6824	0.1192	-0.6824	0.1192	-0.7460	0.1479
Dispensed Medium Dose of Methadone	-1.1641	0.0812	-1.0130	0.0787	-1.0816	0.0563	-1.1356	0.1139	-1.1356	0.1139	-1.2521	0.1493
Dispensed High Dose of Methadone	-0.8638	0.0671	-0.9337	0.0666	-0.8962	0.0472	-1.4256	0.1091	-1.4256	0.1091	-1.4347	0.1356
Dispensed Low Dose of Buprenorphine	-1.7198	0.4332	-1.7925	0.5072	-1.7275	0.3390	-1.8113	0.7088	-1.8113	0.7088	-0.8417	0.5032
Dispensed Medium Dose of Buprenorphine	-2.6802	0.5578	-1.4153	0.3618	-1.9252	0.3020	-1.7239	0.5799	-1.7239	0.5799	-1.0398	0.5037
Dispensed High Dose of Buprenorphine	-1.9022	0.3689	-1.7586	0.4190	-1.8258	0.2799	-0.6109	0.3372	-0.6109	0.3372	-0.9263	0.5038
Dispensed Low Dose of Other OAT Type	-0.4309	0.7054	0.0074	0.5844	-0.1560	0.4515	NA	NA	NA	NA	NA	NA
Dispensed Medium or High Dose of Other OAT Type	0.1548	0.9845	-1.2073	0.9963	-0.6926	0.7068	NA	NA	NA	NA	-1.3620	1.0029
Birth Generation (vs *Greatest & Silent Generations*)	-	-	-	-	-	-	-	-	-	-	-	-
Baby Boomers	-1.3271	0.0656	-1.3965	0.0848	-1.3445	0.0506	-1.1991	0.1398	-1.1991	0.1398	-1.0932	0.2156
Generation X	-2.3379	0.0756	-2.2553	0.0911	-2.2698	0.0570	-1.9530	0.1454	-1.9530	0.1454	-1.6091	0.2188
Millennials & Generation Z	-2.3283	0.0879	-2.3130	0.1074	-2.2918	0.0692	-2.0368	0.1693	-2.0368	0.1693	-1.7605	0.2490
Ever on PharmaCare Plans C or G	-0.1376	0.0469	-0.0117	0.0546	-0.0709	0.0358	0.1425	0.0863	0.1425	0.0863	0.1960	0.1182
Test Statistic	17.4119				**40.6295**				7.8344			
p-value	0.1347				**< 0.0001**				0.6450			
	Test 4: Groups 3,4 vs. Group 5				**Test 5**: Groups 3,4,5 vs. Group 6				**Test 6**: Groups 3,4,5,6 vs. Group 7			
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Dispensed Low Dose of Methadone	-0.7087	0.0930	-0.9746	0.1781	-0.7773	0.0844	-0.5304	0.1524	-0.7239	0.0702	-0.1304	0.2125
Dispensed Medium Dose of Methadone	-1.1816	0.0909	-1.2765	0.1582	-1.2061	0.0791	-1.3904	0.1756	-1.2411	0.0734	-1.3885	0.2939
Dispensed High Dose of Methadone	-1.4282	0.0850	-1.3365	0.1434	-1.4138	0.0718	-1.3768	0.1567	-1.4120	0.0656	-1.4779	0.2772
Dispensed Low Dose of Buprenorphine	-1.2636	0.4166	-1.2926	0.7102	-1.2775	0.3546	-1.7810	1.0023	-1.3470	0.3294	NA	NA
Dispensed Medium Dose of Buprenorphine	-1.3923	0.3773	-0.7980	0.4517	-1.1828	0.2784	-1.5154	0.7102	-1.2432	0.2710	-0.5264	0.7162
Dispensed High Dose of Buprenorphine	-0.7110	0.2867	-1.4975	0.7101	-0.8659	0.2609	-0.5640	0.5046	-0.8096	0.2254	-0.4340	0.7162
Dispensed Low Dose of Other OAT Type	NA	NA	-1.0814	1.0050	-2.0775	1.1092	0.3639	0.5069	-0.7600	0.4672	0.7113	0.7188
Dispensed Medium or High Dose of Other OAT Type	-2.0191	0.9891	NA	NA	-2.4051	0.8442	-1.2677	1.0032	-1.9837	0.7132	-0.7159	1.0111
Birth Generation (vs *Greatest & Silent Generations*)	-	-	-	-	-	-	-	-	-	-	-	-
Baby Boomers	-1.1678	0.1186	-0.9684	0.2285	-1.1233	0.1055	-1.3289	0.2510	-1.1472	0.0940	-0.4608	0.5950
Generation X	-1.8274	0.1219	-1.9259	0.2353	-1.8539	0.1093	-2.0737	0.2538	-1.8740	0.0969	-0.9758	0.5914
Millennials & Generation Z	-1.9400	0.1413	-1.8512	0.2665	-1.9206	0.1258	-2.1422	0.2878	-1.9381	0.1131	-2.0039	0.6720
Ever on PharmaCare Plans C or G	0.1624	0.0700	0.2976	0.1388	0.1968	0.0632	0.5383	0.1634	0.2483	0.0599	0.0040	0.2338
Test Statistic	11.7750				12.7953				22.2579			
p-value	0.3004				0.3841				0.0224			
	Test 7: Groups 3,4,5,6,7 vs. Group 8				**Test 8**: Groups 3,4,5,6,7,8 vs. Group 9							
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.				
Dispensed Low Dose of Methadone	-0.6720	0.0630	-0.6410	0.2768	-0.6674	0.0647	-0.4513	0.3427				
Dispensed Medium Dose of Methadone	-1.2497	0.0723	-1.8284	0.3703	-1.2749	0.0672	-0.9407	0.3185				
Dispensed High Dose of Methadone	-1.4142	0.0659	-1.7767	0.3342	-1.4293	0.0644	-1.0890	0.3120				
Dispensed Low Dose of Buprenorphine	-1.3862	0.3348	NA	NA	-1.4325	0.3197	NA	NA				
Dispensed Medium Dose of Buprenorphine	-1.1765	0.2315	-0.9104	1.0071	-1.1709	0.2299	0.7520	0.7257				
Dispensed High Dose of Buprenorphine	-0.7795	0.2095	NA	NA	-0.8384	0.2110	0.9031	0.5989				
Dispensed Low Dose of Other OAT Type	-0.5221	0.4079	-0.2193	1.0121	-0.4972	0.3338	0.2365	1.0181				
Dispensed Medium or High Dose of Other OAT Type	-1.7362	0.5412	-0.9864	1.0125	-1.5807	0.5008	NA	NA				
Birth Generation (vs *Greatest & Silent Generations*)	-	-	-	-	-	-	-	-				
Baby Boomers	-1.1266	0.0943	-1.7527	0.4543	-1.1396	0.0947	-1.6993	0.7399				
Generation X	-1.8382	0.0961	-2.6970	0.4608	-1.8653	0.0980	-2.3245	0.7403				
Millennials & Generation Z	-1.9545	0.1105	-2.5740	0.5303	-1.9697	0.1148	-2.6947	0.9326				
Ever on PharmaCare Plans C or G	0.2334	0.0552	-0.2525	0.2469	0.2133	0.0527	-0.1840	0.2807				
Test Statistic	14.3567				16.4283							
p-value	0.1573				0.0880							

Table 5.3: Results of the Wald test when we stratify by the number of OAT episode by time t, where the time scale is *time since first recorded OAT dispensation*, and we applied a Bonferonni correction so that the type 1 error rate is $\alpha^* = 0.05/(G-1)$. For each test, the estimates on the left- and right-hand side correspond to $\hat{\boldsymbol{\alpha}}_{g-1}$ and $\hat{\boldsymbol{\alpha}}_g$ and their estimated standard errors, respectively. **Bolded** test statistic(s) and p-value(s) indicate tests that rejected the null hypothesis in (3.6).

	Test 1: Group 1 vs. Group 2				**Test 2**: Group 2 vs. Group 3				**Test 3**: Groups 2,3 vs. Group 4			
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Dispensed Low Dose of Methadone	-0.1279	0.0815	-0.6475	0.0902	-0.6475	0.0902	-0.6696	0.1195	-0.6578	0.0717	-0.7569	0.1483
Dispensed Medium Dose of Methadone	-0.9182	0.0796	-0.9749	0.0787	-0.9749	0.0787	-1.1214	0.1141	-1.0258	0.0649	-1.2633	0.1508
Dispensed High Dose of Methadone	-0.8255	0.0670	-0.9418	0.0664	-0.9418	0.0664	-1.4254	0.1100	-1.0917	0.0582	-1.4199	0.1353
Dispensed Low Dose of Buprenorphine	-1.3866	0.4489	-1.7198	0.5107	-1.7198	0.5107	-1.7953	0.7167	-1.7679	0.4057	-0.8944	0.5040
Dispensed Medium Dose of Buprenorphine	-2.2552	0.5842	-1.3723	0.3642	-1.3723	0.3642	-1.7385	0.5957	-1.5021	0.3097	-1.1575	0.5045
Dispensed High Dose of Buprenorphine	-1.5775	0.3859	-1.7891	0.4210	-1.7891	0.4210	-0.6661	0.3489	-1.2819	0.2518	-1.0530	0.5052
Dispensed Low Dose of Other OAT Type	-0.0843	0.7377	-0.0241	0.5924	-0.0241	0.5924	NA	NA	-0.7329	0.6337	NA	NA
Dispensed Medium or High Dose of Other OAT Type	0.3570	1.0313	-1.3792	1.0090	-1.3792	1.0090	NA	NA	-2.0498	0.8503	-1.4200	1.0046
Birth Generation (vs *Greatest & Silent Generations*)	-	-	-	-	-	-	-	-	-	-	-	-
Baby Boomers	-0.7368	0.1010	-0.8996	0.1192	-0.8996	0.1192	-0.6163	0.2105	-0.8316	0.1047	0.0385	0.3482
Generation X	-1.2983	0.1320	-1.2320	0.1457	-1.2320	0.1457	-0.8272	0.2410	-1.1262	0.1256	0.0470	0.3798
Millennials & Generation Z	-0.9375	0.1761	-0.7921	0.1873	-0.7921	0.1873	-0.3980	0.2871	-0.6904	0.1590	0.1833	0.4251
Ever on PharmaCare Plans C or G	-0.1039	0.0483	0.0230	0.0553	0.0230	0.0553	0.2019	0.0878	0.0778	0.0472	0.3344	0.1231
Test Statistic		**29.5934**				23.7381				24.0573		
p-value		**0.0032**				0.0083				0.0125		
	Test 4: Groups 2,3,4 vs. Group 5				**Test 5**: Groups 2,3,4,5 vs. Group 6				**Test 6**: Group 6 vs. Group 7			
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Dispensed Low Dose of Methadone	-0.6747	0.0651	-0.9791	0.1796	-0.7143	0.0621	-0.5677	0.1536	-0.5677	0.1536	-0.1310	0.2131
Dispensed Medium Dose of Methadone	-1.0581	0.0610	-1.2428	0.1585	-1.0808	0.0564	-1.4414	0.1758	-1.4414	0.1758	-1.3645	0.2943
Dispensed High Dose of Methadone	-1.1435	0.0535	-1.3096	0.1429	-1.1658	0.0494	-1.3848	0.1561	-1.3848	0.1561	-1.4557	0.2769
Dispensed Low Dose of Buprenorphine	-1.5086	0.3098	-1.3465	0.7113	-1.4769	0.2818	-1.7786	1.0045	-1.7786	1.0045	NA	NA
Dispensed Medium Dose of Buprenorphine	-1.4086	0.2512	-0.7608	0.4538	-1.2867	0.2213	-1.6758	0.7304	-1.6758	0.7304	-0.4823	0.7170
Dispensed High Dose of Buprenorphine	-1.2232	0.2224	-1.9015	0.7498	-1.2694	0.2157	-0.7543	0.5059	-0.7543	0.5059	-0.5607	0.7180
Dispensed Low Dose of Other OAT Type	-1.0516	0.5891	-1.1402	1.0059	-1.0441	0.5054	0.3723	0.5099	0.3723	0.5099	0.5560	0.7325
Dispensed Medium or High Dose of Other OAT Type	-1.7674	0.6798	NA	NA	-2.0183	0.7241	-1.3440	1.0053	-1.3440	1.0053	-0.5636	1.0121
Birth Generation (vs *Greatest & Silent Generations*)	-	-	-	-	-	-	-	-	-	-	-	-
Baby Boomers	-0.7456	0.0976	0.0547	0.3390	-0.6680	0.0959	0.0692	0.4166	0.0692	0.4166	-0.5139	0.6995
Generation X	-0.9844	0.1143	-0.5465	0.3786	-0.9488	0.1116	-0.0765	0.4568	-0.0765	0.4568	-0.5902	0.7471
Millennials & Generation Z	-0.5954	0.1433	-0.1432	0.4435	-0.5572	0.1390	-0.1248	0.5071	-0.1248	0.5071	-1.4126	0.8556
Ever on PharmaCare Plans C or G	0.1152	0.0447	0.3947	0.1414	0.1456	0.0430	0.7551	0.1712	0.7551	0.1712	0.0343	0.2321
Test Statistic		19.4901				**30.2368**				13.6488		
p-value		0.0528				**0.0026**				0.2530		
	Test 7: Groups 6,7 vs. Group 8				**Test 8**: Groups 6,7,8 vs. Group 9							
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.				
Dispensed Low Dose of Methadone	-0.4188	0.1212	-0.6328	0.2828	-0.4460	0.1153	-0.3863	0.3448				
Dispensed Medium Dose of Methadone	-1.4044	0.1514	-1.8659	0.3751	-1.4929	0.1395	-1.0012	0.3320				
Dispensed High Dose of Methadone	-1.4008	0.1379	-1.7295	0.3342	-1.4608	0.1303	-1.0773	0.3133				
Dispensed Low Dose of Buprenorphine	-2.0370	1.0136	NA	NA	-2.2816	0.9830	NA	NA				
Dispensed Medium Dose of Buprenorphine	-1.1988	0.4857	-1.0722	1.0113	-1.1846	0.4347	0.5434	0.7318				
Dispensed High Dose of Buprenorphine	-0.6982	0.4108	NA	NA	-0.8868	0.3818	0.7163	0.6109				
Dispensed Low Dose of Other OAT Type	0.4286	0.4160	-0.0565	1.0252	0.2809	0.3997	0.2700	1.0242				
Dispensed Medium or High Dose of Other OAT Type	-1.3482	0.7472	-0.6312	1.0134	-1.1894	0.5509	NA	NA				
Birth Generation (vs *Greatest & Silent Generations*)	-	-	-	-	-	-	-	-				
Baby Boomers	0.0161	0.3407	-1.3881	1.0502	-0.1122	0.3350	NA	NA				
Generation X	-0.1010	0.3743	-1.7803	1.0914	-0.2831	0.3689	NA	NA				
Millennials & Generation Z	-0.3224	0.4229	-1.2869	1.1937	-0.4004	0.4147	NA	NA				
Ever on PharmaCare Plans C or G	0.5295	0.1315	-0.1575	0.2442	0.4021	0.1122	0.1527	0.2942				
Test Statistic		13.7540				10.8011						
p-value		0.1845				0.1475						

Table 5.4: Summary statistics of $\hat{\boldsymbol{\nu}}$ for all individuals.
Abbreviations:
S.D.: Sample standard deviation.

Methadone: Low				**Methadone: Medium**				**Methadone: High**			
	Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total
1st Quartile	-13.1900	-8.1302	-12.2352	1st Quartile	-23.0259	-6.4733	-23.0259	1st Quartile	-23.0259	-23.0259	-23.0259
Median	-2.6456	-2.6359	-2.6439	Median	-2.7015	-0.7465	-2.3253	Median	-23.0259	-11.9474	-21.4880
3rd Quartile	-0.2898	-0.2251	-0.2808	3rd Quartile	0.5725	1.9642	0.7697	3rd Quartile	-2.3585	-0.7344	-2.1096
Mean	-6.3726	-5.5573	-6.2646	Mean	-7.8501	-4.1490	-7.3598	Mean	-13.4739	-11.3988	-13.1990
S.D.	10.1611	9.4681	10.0757	S.D.	11.4674	10.6042	11.4258	S.D.	11.0445	11.5089	11.1293
N	44,227	6,754	50,981	N	44,227	6,754	50,981	N	44,227	6,754	50,981
Buprenorphine: Low Dose				**Buprenorphine: Medium Dose**				**Buprenorphine: Medium Dose**			
	Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total
1st Quartile	-23.0259	-23.0259	-23.0259	1st Quartile	-23.0259	-23.0259	-23.0259	1st Quartile	-23.0259	-23.0259	-23.0259
Median	-23.0259	-23.0259	-23.0259	Median	-23.0259	-23.0259	-23.0259	Median	-23.0259	-23.0259	-23.0259
3rd Quartile	-23.0259	-23.0259	-23.0259	3rd Quartile	-13.9601	-23.0259	-17.1870	3rd Quartile	-23.0259	-23.0259	-23.0259
Mean	-19.2809	-21.9246	-19.6311	Mean	-17.3070	-21.5129	-17.8642	Mean	-19.5458	-22.2000	-19.8974
S.D.	7.9108	4.4828	7.5997	S.D.	9.7487	5.3307	9.3938	S.D.	8.1749	4.0953	7.8106
N	44,227	6,754	50,981	N	44,227	6,754	50,981	N	44,227	6,754	50,981
Other OAT Types: Low Dose				**Non-Methadone**							
	Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total				
1st Quartile	-23.0259	-23.0259	-23.0259	1st Quartile	-23.0259	-23.0259	-23.0259				
Median	-23.0259	-23.0259	-23.0259	Median	-23.0259	-23.0259	-23.0259				
3rd Quartile	-23.0259	-23.0259	-23.0259	3rd Quartile	-4.7037	-23.0259	-7.0567				
Mean	-22.8057	-22.9820	-22.8290	Mean	-15.4648	-20.9387	-16.1900				
S.D.	1.8438	0.8258	1.7445	S.D.	10.8276	6.2716	10.5052				
N	44,227	6,754	50,981	N	44,227	6,754	50,981				

Table 5.5: Summary statistics of $\hat{\nu}$ for individuals born between **1901-1945**.
Abbreviations:
S.D.: Sample standard deviation.

Methadone: Low				**Methadone: Medium**				**Methadone: High**			
	Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total
1st Quartile	-23.0259	-23.0259	-23.0259	1st Quartile	-23.0259	-23.0259	-23.0259	1st Quartile	-23.0259	-23.0259	-23.0259
Median	-4.0921	-3.4138	-3.6518	Median	-23.0259	-2.0692	-8.3101	Median	-23.0259	-23.0259	-23.0259
3rd Quartile	-0.6917	< 0.0001	-0.2906	3rd Quartile	-2.1526	3.2269	1.7492	3rd Quartile	-23.0259	-3.4852	-6.1631
Mean	-9.0202	-7.3077	-8.0205	Mean	-14.5360	-6.8765	-10.0648	Mean	-18.1257	-14.1370	-15.7973
S.D.	11.6014	11.8782	11.7883	S.D.	11.6678	13.5637	13.3483	S.D.	9.5471	11.6699	11.0091
N	420	589	1,009	N	420	589	1,009	N	420	589	1,009
Buprenorphine: Low Dose				**Buprenorphine: Medium Dose**				**Buprenorphine: Medium Dose**			
	Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total
1st Quartile	-23.0259	-23.0259	-23.0259	1st Quartile	-23.0259	-23.0259	-23.0259	1st Quartile	-23.0259	-23.0259	-23.0259
Median	-23.0259	-23.0259	-23.0259	Median	-23.0259	-23.0259	-23.0259	Median	-23.0259	-23.0259	-23.0259
3rd Quartile	-23.0259	-23.0259	-23.0259	3rd Quartile	-23.0259	-23.0259	-23.0259	3rd Quartile	-23.0259	-23.0259	-23.0259
Mean	-18.6140	-22.5730	-20.9251	Mean	-19.9792	-22.6067	-21.5130	Mean	-21.0623	-22.6925	-22.0139
S.D.	9.1697	3.1057	6.6626	S.D.	8.2743	2.9560	5.9359	S.D.	6.8538	2.8701	4.9978
N	420	589	1,009	N	420	589	1,009	N	420	589	1,009
Other OAT Types: Low Dose				**Non-Methadone**							
	Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total				
1st Quartile	-23.0259	-23.0259	-23.0259	1st Quartile	-23.0259	-23.0259	-23.0259				
Median	-23.0259	-23.0259	-23.0259	Median	-23.0259	-23.0259	-23.0259				
3rd Quartile	-23.0259	-23.0259	-23.0259	3rd Quartile	-5.6688	-23.0259	-23.0259				
Mean	-22.6470	-23.0259	-22.8681	Mean	-16.5394	-22.2522	-19.8742				
S.D.	2.9319	< 0.0001	1.8995	S.D.	11.0719	4.2089	8.3203				
N	420	589	1,009	N	420	589	1,009				

Table 5.6: Summary statistics of $\hat{\nu}$ for individuals born between **1946-1964**.
Abbreviations:
S.D.: Sample standard deviation.

Methadone: Low				**Methadone: Medium**				**Methadone: High**			
	Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total
1st Quartile	-12.0065	-10.9269	-11.5050	1st Quartile	-23.0259	-2.8832	-10.4448	1st Quartile	-23.0259	-23.0259	-23.0259
Median	-2.8033	-2.8027	-2.8033	Median	-1.3032	0.2598	-0.8427	Median	-12.2703	-6.2756	-10.1425
3rd Quartile	-0.1337	-0.0312	-0.0977	3rd Quartile	1.5464	2.9144	1.9642	3rd Quartile	-0.6584	0.5643	-0.3281
Mean	-6.1864	-5.8754	-6.1019	Mean	-5.7265	-1.9727	-4.7069	Mean	-11.4929	-9.4699	-10.9434
S.D.	10.2567	9.8046	10.1365	S.D.	11.3023	9.7239	11.0231	S.D.	11.6287	11.6631	11.6723
N	9,311	3,472	12,783	*N*	9,311	3,472	12,783	*N*	9,311	3,472	12,783
Buprenorphine: Low Dose				**Buprenorphine: Medium Dose**				**Buprenorphine: Medium Dose**			
	Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total
1st Quartile	-23.0259	-23.0259	-23.0259	1st Quartile	-23.0259	-23.0259	-23.0259	1st Quartile	-23.0259	-23.0259	-23.0259
Median	-23.0259	-23.0259	-23.0259	Median	-23.0259	-23.0259	-23.0259	Median	-23.0259	-23.0259	-23.0259
3rd Quartile	-23.0259	-23.0259	-23.0259	3rd Quartile	-23.0259	-23.0259	-23.0259	3rd Quartile	-23.0259	-23.0259	-23.0259
Mean	-20.8109	-22.6606	-21.3133	Mean	-20.1921	-22.4929	-20.8170	Mean	-21.1738	-22.7343	-21.5976
S.D.	6.5601	2.5857	5.8170	S.D.	7.5471	3.2098	6.7329	S.D.	6.3964	2.5095	5.6561
N	9,311	3,472	12,783	*N*	9,311	3,472	12,783	*N*	9,311	3,472	12,783
Other OAT Types: Low Dose				**Non-Methadone**							
	Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total				
1st Quartile	-23.0259	-23.0259	-23.0259	1st Quartile	-23.0259	-23.0259	-23.0259				
Median	-23.0259	-23.0259	-23.0259	Median	-23.0259	-23.0259	-23.0259				
3rd Quartile	-23.0259	-23.0259	-23.0259	3rd Quartile	-21.7579	-23.0259	-23.0259				
Mean	-22.7282	-22.9808	-22.7968	Mean	-18.6578	-22.2504	-19.6336				
S.D.	2.4107	0.9404	2.1180	S.D.	9.2056	3.8913	8.2698				
N	9,311	3,472	12,783	*N*	9,311	3,472	12,783				

Table 5.7: Summary statistics of $\hat{\nu}$ for individuals born between **1965-1980**.
Abbreviations:
S.D.: Sample standard deviation.

Methadone: Low				**Methadone: Medium**				**Methadone: High**			
	Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total
1st Quartile	-5.7789	-4.6060	-5.5760	1st Quartile	-20.4573	-5.3104	-17.9748	1st Quartile	-23.0259	-23.0259	-23.0259
Median	-2.2889	-2.2990	-2.2924	Median	-1.8698	-1.2757	-1.7886	Median	-14.8418	-11.7609	-14.4446
3rd Quartile	-0.2379	-0.3145	-0.2452	3rd Quartile	0.6603	0.7697	0.6761	3rd Quartile	-1.5745	-1.5055	-1.5668
Mean	-5.2472	-4.0890	-5.1320	Mean	-6.3717	-4.5557	-6.1910	Mean	-12.2738	-11.8100	-12.2277
S.D.	9.3465	7.7190	9.2038	S.D.	10.7596	9.4156	10.6472	S.D.	11.0532	10.8494	11.0337
N	17,659	1,951	19,610	N	17,659	1,951	19,610	N	17,659	1,951	19,610
Buprenorphine: Low Dose				**Buprenorphine: Medium Dose**				**Buprenorphine: Medium Dose**			
	Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total
1st Quartile	-23.0259	-23.0259	-23.0259	1st Quartile	-23.0259	-23.0259	-23.0259	1st Quartile	-23.0259	-23.0259	-23.0259
Median	-23.0259	-23.0259	-23.0259	Median	-23.0259	-23.0259	-23.0259	Median	-23.0259	-23.0259	-23.0259
3rd Quartile	-23.0259	-23.0259	-23.0259	3rd Quartile	-20.4650	-23.0259	-21.3859	3rd Quartile	-23.0259	-23.0259	-23.0259
Mean	-20.3045	-21.7602	-20.4493	Mean	-18.7638	-21.2956	-19.0157	Mean	-20.3359	-22.0126	-20.5027
S.D.	6.8387	4.7161	6.6720	S.D.	8.6517	5.6256	8.4336	S.D.	7.3471	4.5482	7.1357
N	17,659	1,951	19,610	N	17,659	1,951	19,610	N	17,659	1,951	19,610
Other OAT Types: Low Dose				**Non-Methadone**							
	Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total				
1st Quartile	-23.0259	-23.0259	-23.0259	1st Quartile	-23.0259	-23.0259	-23.0259				
Median	-23.0259	-23.0259	-23.0259	Median	-23.0259	-23.0259	-23.0259				
3rd Quartile	-23.0259	-23.0259	-23.0259	3rd Quartile	-15.4539	-23.0259	-16.7728				
Mean	-22.8215	-22.9799	-22.8373	Mean	-17.3053	-20.6613	-17.6392				
S.D.	1.7069	0.7512	1.6376	S.D.	9.8146	6.5483	9.5925				
N	17,659	1,951	19,610	N	17,659	1,951	19,610				

Table 5.8: Summary statistics of $\hat{\nu}$ for individuals born after **1980**.
Abbreviations:
S.D.: Sample standard deviation.

Methadone: Low				**Methadone: Medium**				**Methadone: High**			
	Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total
1st Quartile	-23.0259	-10.4705	-23.0259	1st Quartile	-23.0259	-23.0259	-23.0259	1st Quartile	-23.0259	-23.0259	-23.0259
Median	-3.0010	-2.8117	-2.9907	Median	-8.1542	-8.8909	-8.1720	Median	-23.0259	-23.0259	-23.0259
3rd Quartile	-0.4347	-0.7971	-0.4487	3rd Quartile	-0.2645	-0.7264	-0.2836	3rd Quartile	-6.1038	-12.0887	-6.2314
Mean	-7.5899	-6.5398	-7.5455	Mean	-10.4082	-11.0982	-10.4373	Mean	-15.7119	-17.1699	-15.7735
S.D.	10.7271	9.4001	10.6763	S.D.	11.7278	11.2117	11.7070	S.D.	10.2964	9.7427	10.2775
N	16,837	742	17,579	N	16,837	742	17,579	N	16,837	742	17,579
Buprenorphine: Low Dose				**Buprenorphine: Medium Dose**				**Buprenorphine: Medium Dose**			
	Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total
1st Quartile	-23.0259	-23.0259	-23.0259	1st Quartile	-23.0259	-23.0259	-23.0259	1st Quartile	-23.0259	-23.0259	-23.0259
Median	-23.0259	-23.0259	-23.0259	Median	-23.0259	-23.0259	-23.0259	Median	-23.0259	-23.0259	-23.0259
3rd Quartile	-11.3718	-17.0240	-11.5626	3rd Quartile	-3.2324	-8.1029	-3.3266	3rd Quartile	-16.6664	-23.0259	-17.0766
Mean	-17.3777	-18.3990	-17.4208	Mean	-14.1170	-16.6308	-14.2231	Mean	-17.7790	-19.8019	-17.8643
S.D.	9.1337	8.3810	9.1053	S.D.	10.9891	9.5421	10.9433	S.D.	9.4785	7.4153	9.4092
N	16,837	742	17,579	N	16,837	742	17,579	N	16,837	742	17,579
Other OAT Types: Low Dose				**Non-Methadone**							
	Survivors	Non-Survivors	Total		Survivors	Non-Survivors	Total				
1st Quartile	-23.0259	-23.0259	-23.0259	1st Quartile	-23.0259	-23.0259	-23.0259				
Median	-23.0259	-23.0259	-23.0259	Median	-14.7063	-23.0259	-15.0702				
3rd Quartile	-23.0259	-23.0259	-23.0259	3rd Quartile	-1.4259	-3.6129	-1.5064				
Mean	-22.8358	-22.9587	-22.8410	Mean	-11.7419	-14.4871	-11.8577				
S.D.	1.5628	0.7647	1.5377	S.D.	11.5735	10.4550	11.5414				
N	16,837	742	17,579	N	16,837	742	17,579				

Table 5.9: Estimates of regression coefficients where the time unit is *months*, under (a) the Cox model that directly replaces $\boldsymbol{\nu}_i$ with $\hat{\boldsymbol{\nu}}_i$, (b) the conditional score method with K (univariate) $AR(1)$ models are specified for the error term in (5.5), and (c) the conditional score method upon fitting a $VAR(1)$ model for the errors in (5.5); for **all** subjects. The **bolded** estimates are statistically significant with the type 1 error rate set at 5%.

Time Since First Recorded OAT Dispensation						
	Cox Model		**Conditional Score AR(1) Error Models**		**Conditional Score VAR(1) Error Model**	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Dispensed Low Dose of Methadone Rate	**-0.0198**	0.0017	**-0.0231**	0.0019	**-0.0248**	0.0020
Dispensed Medium Dose of Methadone Rate	0.0013	0.0017	0.0016	0.0019	0.0017	0.0020
Dispensed High Dose of Methadone Rate	**-0.0214**	0.0013	**-0.0247**	0.0014	**-0.0269**	0.0016
Dispensed Other OAT Rate	**-0.0335**	0.0028	**-0.0390**	0.0031	**-0.0422**	0.0033
Incarceration FPCS	**0.0289**	0.0070	**0.0289**	0.0070	**0.0289**	0.0070
Sex (vs. *Female*)	-	-	-	-	-	-
Male	**0.2130**	0.0265	**0.2131**	0.0265	**0.2131**	0.0265
Birth Generation (vs *Greatest & Silent Generations*)	-	-	-	-	-	-
Baby Boomers	**-1.1951**	0.0462	**-1.1943**	0.0462	**-1.1946**	0.0462
Generation X	**-1.9226**	0.0496	**-1.9213**	0.0496	**-1.9219**	0.0496
Millennials & Generation Z	**-1.9581**	0.0616	**-1.9562**	0.0616	**-1.9571**	0.0616
Heath Authority (vs *Fraser Health*)	-	-	-	-	-	-
Interior	**0.2275**	0.0398	**0.2277**	0.0398	**0.2277**	0.0398
Vancouver Coastal	**0.1167**	0.0307	**0.1169**	0.0307	**0.1168**	0.0307
Vancouver Island	**0.1189**	0.0368	**0.1191**	0.0368	**0.1191**	0.0368
Northern	0.0138	0.0693	0.0135	0.0693	0.0137	0.0693
Year Category (vs. *1996-2000*)	-	-	-	-	-	-
2001-2006	**0.1781**	0.0344	**0.1791**	0.0344	**0.1784**	0.0344
2007-2012	**0.3272**	0.0412	**0.3288**	0.0412	**0.3278**	0.0412
2013-2018	**0.7807**	0.0552	**0.7832**	0.0553	**0.7820**	0.0552
Alcohol or Other Substance Use Disorders	**0.5341**	0.0453	**0.5348**	0.0453	**0.5346**	0.0453
Ill Mental Health or Chronic pain	**-0.1579**	0.0448	**-0.1574**	0.0448	**-0.1576**	0.0448
Hepatitis C Virus or HIV/AIDS	**1.1902**	0.0276	**1.1903**	0.0276	**1.1903**	0.0276
Ever Received a Sedative	**0.5123**	0.0335	**0.5125**	0.0335	**0.5125**	0.0335
Ever on PharmaCare Plans C or G	-0.0410	0.0313	-0.0406	0.0313	-0.0406	0.0313
Age						
	Cox Model		**Conditional Score AR(1) Error Models**		**Conditional Score VAR(1) Error Model**	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Dispensed Low Dose of Methadone Rate	**-0.0200**	0.0017	**-0.0233**	0.0018	**-0.0251**	0.0020
Dispensed Medium Dose of Methadone Rate	0.0026	0.0017	0.0031	0.0019	0.0033	0.0020
Dispensed High Dose of Methadone Rate	**-0.0214**	0.0013	**-0.0247**	0.0014	**-0.0269**	0.0016
Dispensed Other OAT Rate	**-0.0327**	0.0027	**-0.0380**	0.0030	**-0.0411**	0.0033
Incarceration FPCS	**0.0309**	0.0071	**0.0309**	0.0071	**0.0309**	0.0071
Sex (vs. *Female*)	-	-	-	-	-	-
Male	**0.1949**	0.0266	**0.1950**	0.0266	**0.1950**	0.0266
Birth Generation (vs *Greatest & Silent Generations*)	-	-	-	-	-	-
Baby Boomers	**-0.4136**	0.0688	**-0.4127**	0.0688	**-0.4131**	0.0688
Generation X	**-0.6035**	0.0835	**-0.6019**	0.0835	**-0.6026**	0.0835
Millennials & Generation Z	-0.1861	0.1090	-0.1840	0.1091	-0.1848	0.1090
Heath Authority (vs *Fraser Health*)	-	-	-	-	-	-
Interior	**0.1908**	0.0399	**0.1910**	0.0399	**0.1910**	0.0399
Vancouver Coastal	**0.1044**	0.0307	**0.1046**	0.0307	**0.1045**	0.0307
Vancouver Island	**0.0936**	0.0368	**0.0938**	0.0368	**0.0938**	0.0368
Northern	0.0080	0.0692	0.0077	0.0692	0.0079	0.0692
Year Category (vs. *1996-2000*)	-	-	-	-	-	-
2001-2006	0.0431	0.0328	0.0441	0.0329	0.0434	0.0328
2007-2012	0.0536	0.0380	0.0552	0.0380	0.0542	0.0380
2013-2018	**0.4209**	0.0499	**0.4234**	**0.0499**	**0.4222**	0.0499
Alcohol or Other Substance Use Disorders	**0.6186**	0.0465	**0.6192**	**0.0465**	**0.6190**	0.0465
Ill Mental Health or Chronic pain	**-0.2479**	0.0447	**-0.2475**	**0.0447**	**-0.2478**	0.0447
Hepatitis C Virus or HIV/AIDS	**1.1358**	0.0275	**1.1358**	**0.0275**	**1.1358**	0.0275
Ever Received a Sedative	**0.4922**	0.0335	**0.4923**	**0.0335**	**0.4923**	0.0335
Ever on PharmaCare Plans C or G	0.0250	0.0319	0.0254	0.0319	0.0254	0.0319

Table 5.10: Estimates of regression coefficients where the time unit is *months*, under (a) the Cox model that directly replaces $\boldsymbol{\nu}_i$ with $\hat{\boldsymbol{\nu}}_i$, (b) the conditional score method with K (univariate) $AR(1)$ models are specified for the error term in (5.5), and (c) the conditional score method upon fitting a $VAR(1)$ model for the errors in (5.5); for individuals born between **1901-1945**. The **bolded** estimates are statistically significant with the type 1 error rate set at 5%.

Time Since First Recorded OAT Dispensation						
	Cox Model		**Conditional Score AR(1) Error Models**		**Conditional Score VAR(1) Error Model**	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Dispensed Low Dose of Methadone Rate	-0.0062	0.0047	-0.0073	0.0050	-0.0084	0.0052
Dispensed Medium Dose of Methadone Rate	0.0064	0.0046	0.0064	0.0047	0.0069	0.0053
Dispensed High Dose of Methadone Rate	**-0.0089**	0.0044	**-0.0093**	0.0045	**-0.0102**	0.0046
Dispensed Other OAT Rate	-0.0243	0.0129	-0.0274	0.0138	**-0.0350**	0.0165
Incarceration FPCS	**0.3763**	0.1408	**0.3785**	0.1409	**0.3814**	0.1411
Sex (vs. *Female*)	-	-	-	-	-	-
Male	0.0771	0.0911	0.0790	0.0912	0.0792	0.0916
Heath Authority (vs *Fraser Health*)	-	-	-	-	-	-
Interior	0.1981	0.1316	0.2000	0.1316	0.2054	0.1319
Vancouver Coastal	0.2009	0.1055	0.1995	0.1055	0.1975	0.1057
Vancouver Island	0.0912	0.1322	0.0925	0.1322	0.0955	0.1322
Northern	-0.3994	0.2932	-0.3969	0.2932	-0.3906	0.2935
Year Category (vs. *1996-2000*)	-	-	-	-	-	-
2001-2006	**0.2752**	0.1159	**0.2799**	0.1164	**0.2867**	0.1172
2007-2012	0.1133	0.1596	0.1218	0.1601	0.1369	0.1608
2013-2018	0.2336	0.2090	0.2532	0.2100	0.3048	0.2128
Alcohol or Other Substance Use Disorders	-0.1037	0.1040	-0.1002	0.1042	-0.0950	0.1053
Ill Mental Health or Chronic pain	0.2170	0.1526	0.2200	0.1527	0.2236	0.1527
Hepatitis C Virus or HIV/AIDS	**1.0625**	0.1009	**1.0636**	0.1009	**1.0651**	0.1011
Ever Received a Sedative	**0.3372**	0.1541	**0.3391**	0.1542	**0.3417**	0.1542
Ever on PharmaCare Plans C or G	**-0.2224**	0.1029	**-0.2218**	0.1031	**-0.2191**	0.1041
Age						
	Cox Model		**Conditional Score AR(1) Error Models**		**Conditional Score VAR(1) Error Model**	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Dispensed Low Dose of Methadone Rate	-0.0072	0.0047	-0.0084	0.0050	-0.0089	0.0053
Dispensed Medium Dose of Methadone Rate	**0.0107**	0.0047	**0.0110**	0.0048	**0.0130**	0.0055
Dispensed High Dose of Methadone Rate	-0.0067	0.0045	-0.0070	0.0045	-0.0078	0.0047
Dispensed Other OAT Rate	-0.0192	0.0128	-0.0218	0.0137	-0.0262	0.0161
Incarceration FPCS	**0.4506**	0.1416	**0.4525**	0.1417	**0.4530**	0.1418
Sex (vs. *Female*)	-	-	-	-	-	-
Male	0.1387	0.0933	0.1402	0.0934	0.1353	0.0938
Heath Authority (vs *Fraser Health*)	-	-	-	-	-	-
Interior	0.0800	0.1336	0.0825	0.1337	0.0914	0.1340
Vancouver Coastal	0.1561	0.1061	0.1557	0.1061	0.1570	0.1062
Vancouver Island	0.0706	0.1324	0.0720	0.1324	0.0758	0.1325
Northern	-0.2690	0.2942	-0.2648	0.2943	-0.2497	0.2947
Year Category (vs. *1996-2000*)	-	-	-	-	-	-
2001-2006	0.2127	0.1140	0.2193	0.1144	**0.2339**	0.1154
2007-2012	-0.0277	0.1559	-0.0191	0.1563	-0.0020	0.1569
2013-2018	0.2053	0.2016	0.2236	0.2029	0.2727	0.2066
Alcohol or Other Substance Use Disorders	0.0277	0.1097	0.0305	0.1099	0.0274	0.1109
Ill Mental Health or Chronic pain	0.0056	0.1504	0.0086	0.1505	0.0098	0.1506
Hepatitis C Virus or HIV/AIDS	**0.9895**	0.1006	**0.9906**	0.1007	**0.9898**	0.1008
Ever Received a Sedative	**0.3172**	0.1535	**0.3194**	0.1536	**0.3207**	0.1536
Ever on PharmaCare Plans C or G	-0.0719	0.1083	-0.0730	0.1085	-0.0784	0.1090

Table 5.11: Estimates of regression coefficients where the time unit is *months*, under (a) the Cox model that directly replaces $\boldsymbol{\nu}_i$ with $\hat{\boldsymbol{\nu}}_i$, (b) the conditional score method with K (univariate) $AR(1)$ models are specified for the error term in (5.5), and (c) the conditional score method upon fitting a $VAR(1)$ model for the errors in (5.5); for individuals born between **1946-1964**. The **bolded** estimates are statistically significant with the type 1 error rate set at 5%.

Time Since First Recorded OAT Dispensation						
	Cox Model		**Conditional Score AR(1) Error Models**		**Conditional Score VAR(1) Error Model**	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Dispensed Low Dose of Methadone Rate	**-0.0202**	0.0021	**-0.0243**	0.0024	**-0.0257**	0.0026
Dispensed Medium Dose of Methadone Rate	**0.0125**	0.0025	**0.0151**	0.0029	**0.0159**	0.0031
Dispensed High Dose of Methadone Rate	**-0.0206**	0.0017	**-0.0242**	0.0019	**-0.0262**	0.0021
Dispensed Other OAT Rate	**-0.0534**	0.0059	**-0.0633**	0.0067	**-0.0686**	0.0073
Incarceration FPCS	0.0222	0.0118	0.0222	0.0118	0.0222	0.0118
Sex (vs. *Female*)	-	-	-	-	-	-
Male	**0.1795**	0.0375	**0.1798**	0.0375	**0.1797**	0.0375
Heath Authority (vs *Fraser Health*)	-	-	-	-	-	-
Interior	**0.2055**	0.0583	**0.2079**	0.0583	**0.2068**	0.0583
Vancouver Coastal	**0.0866**	0.0431	**0.0877**	0.0431	**0.0873**	0.0431
Vancouver Island	**0.1377**	0.0507	**0.1390**	0.0507	**0.1386**	0.0507
Northern	0.1012	0.0959	0.0994	0.0959	0.1003	0.0959
Year Category (vs. *1996-2000*)	-	-	-	-	-	-
2001-2006	**0.2155**	0.0456	**0.2202**	0.0456	**0.2169**	0.0456
2007-2012	**0.3172**	0.0599	**0.3246**	0.0599	**0.3203**	0.0599
2013-2018	**0.8188**	0.0891	**0.8310**	0.0891	**0.8261**	0.0891
Alcohol or Other Substance Use Disorders	**0.4262**	0.0662	**0.4284**	0.0662	**0.4276**	0.0662
Ill Mental Health or Chronic pain	**-0.1814**	0.0636	**-0.1798**	0.0636	**-0.1806**	0.0636
Hepatitis C Virus or HIV/AIDS	**1.2892**	0.0364	**1.2898**	0.0364	**1.2896**	0.0364
Ever Received a Sedative	**0.3980**	0.0472	**0.3992**	0.0472	**0.3988**	0.0472
Ever on PharmaCare Plans C or G	-0.0414	0.0463	-0.0415	0.0463	-0.0407	0.0463
Age						
	Cox Model		**Conditional Score AR(1) Error Models**		**Conditional Score VAR(1) Error Model**	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Dispensed Low Dose of Methadone Rate	**-0.0191**	0.0021	**-0.0229**	0.0024	**-0.0242**	0.0026
Dispensed Medium Dose of Methadone Rate	**0.0118**	0.0025	**0.0143**	0.0028	**0.0149**	0.0030
Dispensed High Dose of Methadone Rate	**-0.0204**	0.0017	**-0.0239**	0.0019	**-0.0259**	0.0021
Dispensed Other OAT Rate	**-0.0525**	0.0059	**-0.0622**	0.0066	**-0.0673**	0.0072
Incarceration FPCS	**0.0250**	0.0113	**0.0244**	0.0117	**0.0244**	0.0117
Sex (vs. *Female*)	-	-	-	-	-	-
Male	**0.1341**	0.0375	**0.1346**	0.0375	**0.1344**	0.0375
Heath Authority (vs *Fraser Health*)	-	-	-	-	-	-
Interior	**0.1819**	0.0582	**0.1840**	0.0582	**0.1831**	0.0582
Vancouver Coastal	0.0828	0.0430	0.0839	0.0430	0.0835	0.0430
Vancouver Island	**0.1051**	0.0507	**0.1063**	0.0507	**0.1059**	0.0507
Northern	0.1156	0.0959	0.1139	0.0959	0.1148	0.0959
Year Category (vs. *1996-2000*)	-	-	-	-	-	-
2001-2006	**0.0931**	0.0437	**0.0977**	0.0438	**0.0944**	0.0437
2007-2012	0.0032	0.0563	0.0103	0.0564	0.0060	0.0564
2013-2018	**0.3539**	0.0835	**0.3660**	0.0836	**0.3611**	0.0836
Alcohol or Other Substance Use Disorders	**0.4431**	0.0660	**0.4451**	0.0660	**0.4445**	0.0660
Ill Mental Health or Chronic pain	**-0.2868**	0.0634	**-0.2855**	0.0634	**-0.2862**	0.0634
Hepatitis C Virus or HIV/AIDS	**1.2298**	0.0361	**1.2305**	0.0361	**1.2303**	0.0361
Ever Received a Sedative	**0.3818**	0.0472	**0.3831**	0.0472	**0.3827**	0.0472
Ever on PharmaCare Plans C or G	0.0528	0.0468	0.0528	0.0468	0.0535	0.0468

Table 5.12: Estimates of regression coefficients where the time unit is *months*, under (a) the Cox model that directly replaces $\boldsymbol{\nu}_i$ with $\hat{\boldsymbol{\nu}}_i$, (b) the conditional score method with K (univariate) $AR(1)$ models are specified for the error term in (5.5), and (c) the conditional score method upon fitting a $VAR(1)$ model for the errors in (5.5); for individuals born between **1965-1980**. The **bolded** estimates are statistically significant with the type 1 error rate set at 5%.

Time Since First Recorded OAT Dispensation						
	Cox Model		**Conditional Score AR(1) Error Models**		**Conditional Score VAR(1) Error Model**	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Dispensed Low Dose of Methadone Rate	**-0.0245**	0.0041	**-0.0323**	0.0048	**-0.0328**	0.0051
Dispensed Medium Dose of Methadone Rate	-0.0006	0.0035	-0.0006	0.0040	-0.0007	0.0043
Dispensed High Dose of Methadone Rate	**-0.0237**	0.0026	**-0.0280**	0.0028	**-0.0301**	0.0031
Dispensed Other OAT Rate	**-0.0392**	0.0050	**-0.0477**	0.0056	**-0.0504**	0.0061
Incarceration FPCS	0.0285	0.0200	0.0289	0.0200	0.0289	0.0200
Sex (vs. *Female*)	-	-	-	-	-	-
Male	**0.1725**	0.0485	**0.1819**	0.0486	**0.1724**	0.0485
Heath Authority (vs *Fraser Health*)	-	-	-	-	-	-
Interior	**0.2494**	0.0740	**0.2471**	0.0740	**0.2493**	0.0740
Vancouver Coastal	**0.1545**	0.0563	**0.1549**	0.0563	**0.1547**	0.0563
Vancouver Island	0.0866	0.0697	0.0852	0.0698	0.0868	0.0697
Northern	-0.1062	0.1281	-0.1116	0.1282	-0.1075	0.1281
Year Category (vs. *1996-2000*)	-	-	-	-	-	-
2001-2006	**0.1500**	0.0627	**0.1579**	0.0628	**0.1523**	0.0628
2007-2012	**0.5263**	0.0712	**0.5399**	0.0714	**0.5305**	0.0713
2013-2018	**1.0196**	0.0989	**1.0368**	0.0992	**1.0262**	0.0990
Alcohol or Other Substance Use Disorders	**1.1951**	0.1188	**1.2051**	0.1189	**1.1984**	0.1189
Ill Mental Health or Chronic pain	**-0.2917**	0.0826	**-0.2892**	0.0827	**-0.2907**	0.0826
Hepatitis C Virus or HIV/AIDS	**1.0954**	0.0519	**1.0950**	0.0519	**1.0956**	0.0519
Ever Received a Sedative	**0.6584**	0.0586	**0.6570**	0.0586	**0.6583**	0.0586
Ever on PharmaCare Plans C or G	0.0584	0.0601	0.0639	0.0601	0.0609	0.0601
Age						
	Cox Model		**Conditional Score AR(1) Error Models**		**Conditional Score VAR(1) Error Model**	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Dispensed Low Dose of Methadone Rate	**-0.0232**	0.0041	**-0.0306**	0.0047	**-0.0312**	0.0051
Dispensed Medium Dose of Methadone Rate	-0.0006	0.0035	-0.0007	0.0040	-0.0007	0.0043
Dispensed High Dose of Methadone Rate	**-0.0240**	0.0025	**-0.0283**	0.0028	**-0.0305**	0.0031
Dispensed Other OAT Rate	**-0.0390**	0.0050	**-0.0473**	0.0056	**-0.0500**	0.0061
Incarceration FPCS	0.0306	0.0201	0.0310	0.0201	0.0308	0.0201
Sex (vs. *Female*)	-	-	-	-	-	-
Male	**0.1406**	0.0486	**0.1402**	0.0486	**0.1405**	0.0486
Heath Authority (vs *Fraser Health*)	-	-	-	-	-	-
Interior	**0.2347**	0.0740	**0.2326**	0.0740	**0.2347**	0.0740
Vancouver Coastal	**0.1414**	0.0563	**0.1418**	0.0563	**0.1416**	0.0563
Vancouver Island	0.0847	0.0697	0.0835	0.0697	0.0850	0.0697
Northern	-0.1190	0.1281	-0.1242	0.1281	-0.1202	0.1281
Year Category (vs. *1996-2000*)	-	-	-	-	-	-
2001-2006	0.0457	0.0597	0.0532	0.0598	0.0478	0.0598
2007-2012	**0.3063**	0.0643	**0.3190**	0.0645	**0.3100**	0.0644
2013-2018	**0.6318**	0.0892	**0.6481**	0.0895	**0.6380**	0.0893
Alcohol or Other Substance Use Disorders	**1.2131**	0.1189	**1.2223**	0.1190	**1.2160**	0.1189
Ill Mental Health or Chronic pain	**-0.3300**	0.0826	**-0.3274**	0.0827	**-0.3291**	0.0826
Hepatitis C Virus or HIV/AIDS	**1.0448**	0.0523	**1.0449**	0.0523	**1.0452**	0.0523
Ever Received a Sedative	**0.6375**	0.0586	**0.6363**	0.0586	**0.6375**	0.0586
Ever on PharmaCare Plans C or G	0.0604	0.0601	0.0658	0.0601	0.0628	0.0601

Table 5.13: Estimates of regression coefficients where the time unit is *months*, under (a) the Cox model that directly replaces $\boldsymbol{\nu}_i$ with $\hat{\boldsymbol{\nu}}_i$, (b) the conditional score method with K (univariate) $AR(1)$ models are specified for the error term in (5.5), and (c) the conditional score method upon fitting a $VAR(1)$ model for the errors in (5.5); for individuals born after **1980**. The **bolded** estimates are statistically significant with the type 1 error rate set at 5%.

Time Since First Recorded OAT Dispensation						
	Cox Model		**Conditional Score AR(1) Error Models**		**Conditional Score VAR(1) Error Model**	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Dispensed Low Dose of Methadone Rate	**-0.0189**	0.0067	**-0.0225**	0.0075	**-0.0239**	0.0082
Dispensed Medium Dose of Methadone Rate	**-0.0200**	0.0051	**-0.0229**	0.0057	**-0.0248**	0.0062
Dispensed High Dose of Methadone Rate	**-0.0323**	0.0050	**-0.0374**	0.0055	**-0.0408**	0.0061
Dispensed Other OAT Rate	**-0.0318**	0.0056	**-0.0374**	0.0062	**-0.0402**	0.0067
Incarceration FPCS	**0.0417**	0.0148	**0.0417**	0.0148	**0.0417**	0.0148
Sex (vs. *Female*)	-	-	-	-	-	-
Male	**0.5325**	0.0812	**0.5329**	0.0812	**0.5329**	0.0812
Heath Authority (vs *Fraser Health*)	-	-	-	-	-	-
Interior	**0.3567**	0.1054	**0.3564**	0.1054	**0.3571**	0.1054
Vancouver Coastal	**0.2078**	0.0966	**0.2078**	0.0966	**0.2079**	0.0966
Vancouver Island	0.2155	0.1121	0.2154	0.1121	0.2159	0.1121
Northern	0.0299	0.1975	0.0290	0.1975	0.0296	0.1975
Year Category (vs. *1996-2000*)	-	-	-	-	-	-
2001-2006	0.7537	0.4699	0.7549	0.4699	0.7541	0.4699
2007-2012	**0.9326**	0.4713	**0.9349**	0.4713	**0.9336**	0.4713
2013-2018	**1.4621**	0.4762	**1.4651**	0.4762	**1.4633**	0.4762
Alcohol or Other Substance Use Disorders	**1.2853**	0.1607	**1.2871**	0.1608	**1.2860**	0.1608
Ill Mental Health or Chronic pain	**-0.3141**	0.1302	**-0.3131**	0.1302	**-0.3136**	0.1302
Hepatitis C Virus or HIV/AIDS	**0.6830**	0.1419	**0.6832**	0.1419	**0.6833**	0.1419
Ever Received a Sedative	**0.7788**	0.0993	**0.7791**	0.0993	**0.7791**	0.0993
Ever on PharmaCare Plans C or G	-0.1566	0.0828	-0.1543	0.0829	-0.1552	0.0828
Age						
	Cox Model		**Conditional Score AR(1) Error Models**		**Conditional Score VAR(1) Error Model**	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Dispensed Low Dose of Methadone Rate	**-0.0171**	0.0067	**-0.0204**	0.0075	**-0.0216**	0.0081
Dispensed Medium Dose of Methadone Rate	**-0.0199**	0.0051	**-0.0228**	0.0056	**-0.0247**	0.0061
Dispensed High Dose of Methadone Rate	**-0.0324**	0.0050	**-0.0375**	0.0055	**-0.0410**	0.0060
Dispensed Other OAT Rate	**-0.0324**	0.0056	**-0.0380**	0.0062	**-0.0409**	0.0067
Incarceration FPCS	**0.0398**	0.0149	**0.0398**	0.0149	**0.0398**	0.0149
Sex (vs. *Female*)	-	-	-	-	-	-
Male	**0.5031**	0.0814	**0.5034**	0.0814	**0.5034**	0.0814
Heath Authority (vs *Fraser Health*)	-	-	-	-	-	-
Interior	**0.3258**	0.1053	**0.3256**	0.1053	**0.3262**	0.1053
Vancouver Coastal	**0.1930**	0.0966	**0.1930**	0.0966	**0.1931**	0.0966
Vancouver Island	0.1828	0.1121	0.1829	0.1121	0.1834	0.1121
Northern	-0.0270	0.1975	-0.0278	0.1975	-0.0272	0.1975
Year Category (vs. *1996-2000*)	-	-	-	-	-	-
2001-2006	0.7275	0.4630	0.7284	0.4630	0.7276	0.4630
2007-2012	0.8211	0.4543	0.8231	0.4543	0.8216	0.4543
2013-2018	**1.1975**	0.4557	**1.2003**	0.4557	**1.1984**	0.4557
Alcohol or Other Substance Use Disorders	**1.3418**	0.1608	**1.3433**	0.1609	**1.3423**	0.1608
Ill Mental Health or Chronic pain	**-0.2893**	0.1302	**-0.2886**	0.1302	**-0.2890**	0.1302
Hepatitis C Virus or HIV/AIDS	**0.6663**	0.1418	**0.6665**	0.1418	**0.6665**	0.1418
Ever Received a Sedative	**0.7827**	0.0993	**0.7828**	0.0993	**0.7828**	0.0993
Ever on PharmaCare Plans C or G	-0.1589	0.0826	-0.1568	0.0826	-0.1575	0.0826

Table 5.14: Weighted estimates $\tilde{\boldsymbol{\alpha}}$ and $\tilde{\boldsymbol{\theta}}$, where the time unit is *months*.

Time Since First Recorded OAT Dispensation								
	Conditional Score AR(1) Error Models		**Weighted Average**		**Conditional Score VAR(1) Error Model**		**Weighted Average**	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Dispensed Low Dose of Methadone Rate	**-0.0231**	0.0019	**-0.0264**	0.0033	**-0.0248**	0.0020	**-0.0270**	0.0035
Dispensed Medium Dose of Methadone Rate	0.0016	0.0019	-0.0040	0.0026	0.0017	0.0020	-0.0049	0.0028
Dispensed High Dose of Methadone Rate	**-0.0247**	0.0014	**-0.0300**	0.0023	**-0.0269**	0.0016	**-0.0325**	0.0025
Dispensed Other OAT Rate	**-0.0390**	0.0031	**-0.0475**	0.0035	**-0.0422**	0.0033	**-0.0510**	0.0038
Incarceration FPCS	**0.0289**	0.0070	**0.0394**	0.0100	**0.0289**	0.0070	**0.0388**	0.0101
Sex (vs. *Female*)	-	-	-	-	-	-	-	-
Male	**0.2131**	0.0265	**0.2993**	0.0353	**0.2131**	0.0265	**0.2995**	0.0353
Birth Generation (vs *Greatest & Silent Generations*)	-	-	-	-	-	-	-	-
Baby Boomers	**-1.1943**	0.0462	-	-	**-1.1946**	0.0462	-	-
Generation X	**-1.9213**	0.0496	-	-	**-1.9219**	0.0496	-	-
Millennials & Generation Z	**-1.9562**	0.0616	-	-	**-1.9571**	0.0616	-	-
Heath Authority (vs *Fraser Health*)	**-**	-	-	-	-	-	-	-
Interior	**0.2277**	0.0398	**0.2750**	0.0489	**0.2277**	0.0398	**0.2760**	0.0489
Vancouver Coastal	**0.1169**	0.0307	**0.1579**	0.0417	**0.1168**	0.0307	**0.1577**	0.0417
Vancouver Island	**0.1191**	0.0368	**0.1445**	0.0493	**0.1191**	0.0368	**0.1453**	0.0493
Northern	0.0135	0.0693	-0.0158	0.0885	0.0137	0.0693	-0.0137	0.0885
Year Category (vs. *1996-2000*)	-	-	-	-	-	-	-	-
2001-2006	**0.1791**	0.0344	**0.3862**	0.1678	**0.1784**	0.0344	**0.3831**	0.1678
2007-2012	**0.3288**	0.0412	**0.6177**	0.1691	**0.3278**	0.0412	**0.6130**	0.1690
2013-2018	**0.7832**	0.0553	**1.1214**	0.1736	**0.7820**	0.0552	**1.1165**	0.1735
Alcohol or Other Substance Use Disorders	**0.5348**	0.0453	**1.0165**	0.0744	**0.5346**	0.0453	**1.0135**	0.0744
Ill Mental Health or Chronic pain	**-0.1574**	0.0448	**-0.2605**	0.0579	**-0.1576**	0.0448	**-0.2613**	0.0579
Hepatitis C Virus or HIV/AIDS	**1.1903**	0.0276	**0.9971**	0.0546	**1.1903**	0.0276	**0.9974**	0.0546
Ever Received a Sedative	**0.5125**	0.0335	**0.6302**	0.0432	**0.5125**	0.0335	**0.6306**	0.0432
Ever on PharmaCare Plans C or G	-0.0406	0.0313	-0.0447	0.0389	-0.0406	0.0313	-0.0459	0.0389
Age								
	Conditional Score AR(1) Error Models		**Weighted Average**		**Conditional Score VAR(1) Error Model**		**Weighted Average**	
Covariate Name	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Dispensed Low Dose of Methadone Rate	**-0.0233**	0.0018	**-0.0247**	0.0033	**-0.0251**	0.0020	**-0.0252**	0.0035
Dispensed Medium Dose of Methadone Rate	0.0031	0.0019	-0.0042	0.0026	0.0033	0.0020	-0.0050	0.0028
Dispensed High Dose of Methadone Rate	**-0.0247**	0.0014	**-0.0300**	0.0023	**-0.0269**	0.0016	**-0.0326**	0.0025
Dispensed Other OAT Rate	**-0.0380**	0.0030	**-0.0472**	0.0035	**-0.0411**	0.0033	**-0.0505**	0.0038
Incarceration FPCS	**0.0309**	0.0071	**0.0410**	0.0101	**0.0309**	0.0071	**0.0409**	0.0101
Sex (vs. *Female*)	-	-	-	-	-	-	-	-
Male	**0.1950**	0.0266	**0.2668**	0.0355	**0.1950**	0.0266	**0.2669**	0.0355
Birth Generation (vs *Greatest & Silent Generations*)	-	-	-	-	-	-	-	-
Baby Boomers	**-0.4127**	0.0688	-	-	**-0.4131**	0.0688	-	-
Generation X	**-0.6019**	0.0835	-	-	**-0.6026**	0.0835	-	-
Millennials & Generation Z	-0.1840	0.1091	-	-	-0.1848	0.1090	-	-
Heath Authority (vs *Fraser Health*)	-	-	-	-	-	-	-	-
Interior	**0.1910**	0.0399	**0.2504**	0.0489	**0.1910**	0.0399	**0.2514**	0.0489
Vancouver Coastal	**0.1046**	0.0307	**0.1459**	0.0417	**0.1045**	0.0307	**0.1458**	0.0417
Vancouver Island	**0.0938**	0.0368	**0.1239**	0.0493	**0.0938**	0.0368	**0.1247**	0.0493
Northern	0.0077	0.0692	-0.0344	0.0885	0.0079	0.0692	-0.0322	0.0885
Year Category (vs. *1996-2000*)	-	-	-	-	-	-	-	-
2001-2006	0.0441	0.0329	0.3056	0.1652	0.0434	0.0328	0.3027	0.1652
2007-2012	0.0552	0.0380	**0.4139**	0.1627	0.0542	0.0380	**0.4093**	0.1627
2013-2018	**0.4234**	**0.0499**	**0.7648**	**0.1656**	**0.4222**	0.0499	**0.7601**	0.1656
Alcohol or Other Substance Use Disorders	**0.6192**	**0.0465**	**1.0496**	**0.0744**	**0.6190**	0.0465	**1.0467**	0.0744
Ill Mental Health or Chronic pain	**-0.2475**	**0.0447**	**-0.2966**	**0.0579**	**-0.2478**	0.0447	**-0.2976**	0.0579
Hepatitis C Virus or HIV/AIDS	**1.1358**	**0.0275**	**0.9561**	**0.0546**	**1.1358**	0.0275	**0.9561**	0.0546
Ever Received a Sedative	**0.4923**	**0.0335**	**0.6192**	**0.0432**	**0.4923**	0.0335	**0.6196**	0.0432
Ever on PharmaCare Plans C or G	0.0254	0.0319	-0.0187	0.0389	0.0254	0.0319	-0.0200	0.0389

Chapter 6

Final Discussion

Motivated by a real world health crisis, this dissertation develops methodology to overcome challenges brought on by internal covariates in a survival analysis. In particular, we developed procedures that can directly estimate a covariate's effect on the hazard function, as well as methods that are capable of providing survival predictions. The proposed methods were illustrated with administrative service utilization records pertaining to opioid use disorder management. However, the proposed methodology can be adapted to any setting involving an event time and internal covariate(s). Furthermore, the proposed methods can naturally be extended to situations with competing risks, or recurrent events.

6.1 Summary of Contributions

We started with a preliminary analysis in Chapter 2, in which we summarized the OAT dispensation process with an OAT dispensation rate and other processes with either a binary one-jump process or functional principle scores. This served to greatly reduce the computational complexity that plagues joint modelling methods. Our results led us to identify age as a confounding variable, and the protective effect of OAT against mortality is most prevalent for Millennials and Generation Z individuals.

To directly estimate the OAT dispensation process effect on the mortality hazard function, Chapter 3 considered a generalized Cox regression model, under which we conducted time-dependent stratification. An estimating equation inference procedure was implemented in order to overcome the inherent difficulties brought on by internal covariates. The estimating function proposed takes the familiar form of the partial score function when external covariates are present in the model, and we used modern empirical process theory to establish large sample properties. Since the time-dependent stratification variable is an ordinal variable, we utilized the asymptotic normality of the proposed estimator to update the strata levels with a testing procedure. Our results revealed two risk classes: individuals experiencing (i) 1-3 OAT episodes; and (ii) 4+ OAT episodes; in which the protective effect against mortality for the OAT dispensation indicator is stronger in the second group. To

the best of our knowledge, this is the first application to reveal a dynamic effect pertaining to OAT usage.

In Chapter 4, we revisited our modelling from Chapter 2 to address two key issues: (a) correct for the apparent bias induced by directly replacing the OAT dispensation rate with an estimate, and (b) modify our estimates to account for confounding. For the first issue, we adopted the conditional score approach to conduct our statistical inference, where we accounted for the autocorrelation present within the OAT dispensation process. Our simulation study, where the simulated data mimics the observed data, showed that the proposed inference procedure is able to produce consistent estimates. We addressed the second issue by weighting birth generation specific estimates by the relative size in each group, which served to upweight the contributions made by younger individuals. This procedure was able to reveal a significant protective effect against mortality pertaining to the OAT dispensation rate, which corroborates with our Chapter 3 analysis.

We then considered a multivariate internal covariate process in Chapter 5, by additionally including OAT type and dosage from the OAT dispensation process. We expanded upon the methodology presented in Chapters 3 and 4. While implementing the former was relatively straightforward, the latter procedure required us to address for the apparent correlation between OAT types and dosage levels.

6.2 Future Investigation

We list a few possibilities for future investigations.

6.2.1 Alternatives to the Wald Test

An alternative test procedure based on the estimating function $\boldsymbol{U}(\boldsymbol{\Theta})$ worth exploring. Here, one would consider two models: groups $g-1$ and g are merged, and groups $g-1$ and g separated; and proceed to estimate $\boldsymbol{\Theta}$ by solving $\boldsymbol{U}(\boldsymbol{\Theta})$. Letting $\boldsymbol{U}_M(\boldsymbol{\Theta})$, $\hat{\boldsymbol{\Theta}}_M$ and $\boldsymbol{U}_S(\boldsymbol{\Theta})$, $\hat{\boldsymbol{\Theta}}_S$ denoting the estimating function $\boldsymbol{U}(\boldsymbol{\Theta})$ and $\hat{\boldsymbol{\Theta}}$ under the models when groups $g-1$ and g are merged and separated, respectively, we have that $\hat{\boldsymbol{\Theta}}_M$ is a consistent estimator for $\boldsymbol{\Theta}$ under the null hypothesis. Therefore, $\boldsymbol{U}_S(\hat{\boldsymbol{\Theta}}_M)$ and $\boldsymbol{U}_S(\boldsymbol{\Theta}_0)$ should be "close" under the null hypothesis. This procedure is essentially the same idea behind the partial score test, and although the Wald and score tests are asymptotically equivalent, it would be interesting to see the benefits of using such a test in a dynamic fashion.

6.2.2 On Summarizing the Internal Covariate Process

The current modelling within Chapter 4 assumes that the quantity ν_i in (4.2) serves as an adequate summary of the OAT dispensation process. Although ν_i is interpretable and serves as a starting point, we can increase the model complexity in (4.2) to capture important

features within the data. For example, as touched upon in Chapter 4.6, we could include covariates $\boldsymbol{\mathcal{X}}(\cdot)$ into the model, or specify ν_i as a polynomial or non-parametric function of time (e.g. splines).

Another extension would be to incorporate an additional layer to our modelling by introducing latent classes, where the group membership distribution depends on the internal covariate process. This approach was recently considered by Wong *et al.* (2022), where they used the EM-algorithm to estimate model parameters. To improve the computational complexity of their approach, one can again adopt the idea behind the conditional score approach, where this time, the correction for "measurement error" is made in both the hazard and latent group membership models. Depending on the number and nature of the strata levels, be it ordinal or nominal, this can itself be an extension of the conditional score approach of Stefanski and Carroll (1987) for generalized linear models.

6.2.3 Latent Treatment Usage States

Implementing classes into the modelling of OAT dispensation records can be valuable when the focus is on estimating the effects of *titrating*, *maintenance*, or *tapering* dosage on mortality risk. In principle, we can classify individuals into states based in conjunction with clinical guidelines (British Columbia Centre on Substance Use 2017), and on the sign and magnitude of the dose's rate of change. In practice however, classifying individuals is a challenge, as individuals may have an opioid tolerance level that is not adequately reflected by clinical guidelines, and many dose trajectories are highly non-smooth. In other words, there is potential for individuals to have their treatment usage state to be misclassified.

The approach adopted by Wong *et al.* (2022) referenced in Chapter 6.2.2 can also serve as a potential solution, where a model is proposed to reflect *titrating*, *maintenance*, and *tapering* dosage states. We can again view the states as "latent" and adopt the conditional score approach as an alternative to the (computationally expensive) EM-algorithm.

6.2.4 Informative Censoring & Truncation Times

We briefly discussed in Chapter 2 that the *time since first recorded OAT dispensation* is not necessarily the *time since first OAT dispensation*. For individuals with their first OAT dispensation near the date of data collection (01/01/1996), it is very likely that their first OAT dispensation occurred before the data collection period. In other words, their starting date of treatment usage is *left censored*. Moreover, only individuals with an OAT dispensation during the data extraction period were included in our analysis; therefore, OAT users with a death record prior to 1996 are omitted. It would be worthwhile to take into account both the potentially informative left censoring and left truncation.

6.2.5 Causal Inference

There has been much attention recently towards estimating causal effects from observational data (Igelström *et al.* 2022). The basic idea is to account for all potential *confounders* in the analysis so that individuals in both treatment arms are comparable.

The goal remains the same within our setting: to estimate the causal effect of the OAT usage process on mortality risk. However, there are additional challenges present within our setting. The "treatment assignment" is not truly random, as the prescribed treatment was selected to best serve a patient's individualized need. Adjusting for all possible confounders is not only challenging to ensure the "no unmeasured confounding" assumption is satisfied, but difficult to implement when the covariate / confounder space is high dimensional. There is also the issue to account for the OAT status changing over time, which can be viewed as a form of *treatment non-compliance*. This adds an additional level of complexity to the problem.

Bibliography

Andersen, P.K. and R.D. Gill (1982). Cox's regression model for counting processes: A large sample study. *The Annals of Statistics* **10**(4), 1100–1120.

Andrinopoulou, E.R., P.H.C. Eiler, J.M. Takkenberg and D. Rizopoulos (2018). Improved dynamic predictions from joint models of longitudinal and survival data with time-varying effects using P-splines. *Biometrics* **74**(2), 685–693.

Barbieri, S., S. Mehat, B. Wu, C. Bharat, C. Poppe, L. Jorm and R. Jackson (2022). Predicting cardiovascular risk from national administrative databases using a combined survival analysis and deep learning approach. *International Journal of Epidemiology* **51**(3), 931–944.

Barrett, J. and L. Su (2017). Dynamic predictions using flexible joint models of longitudinal and time-to-event data. *Statistics in Medicine* **36**, 1447–1460.

Beyersmann, J. and M. Schumacher (2008). Time-dependent covariates in the proportional subdistribution hazards model for competing risks. *Biostatistics* **9**(4), 765–776.

Bharat, C., L. Degenhardt, T. Dobbins, S. Larney, M. Farrella and S. Barbieri (2021). Using administrative data to predict cessation risk and identify novel predictors among new entrants to opioid agonist treatment. *Drug and Alcohol Dependence* **228**, 109091.

Bilias, Y., M. Gu and Z. Ying (1997). Towards a general asymptotic theory for Cox model with staggered entry. *Annals of Statistics* **25**, 662–682.

British Columbia Centre on Substance Use (2017). *A guideline for the clinical management of opioid use disorder*.

British Columbia Ministry of Health (2018*a*). *Discharge Abstract Database (Hospital Separations)*. British Columbia Ministry of Health.

British Columbia Ministry of Health (2018*b*). *Medical Services Plan Payment Information File*. British Columbia Ministry of Health.

British Columbia Ministry of Health (2018*c*). *PharmaNet*. British Columbia Ministry of Health.

British Columbia Vital Statistics Agency (2018). *Vital Statistics Deaths*. British Columbia Ministry of Health.

Brookmeyer, R. and N. Abdalla (2019). Multistate models and lifetime risk estimation: Application to Alzheimer's disease. *Statistics in Medicine* **38**(9), 1558–1565.

Carroll, R.J., D. Ruppert, L.A. Stefanski and C.M. Crainiceanu (2006). *Measurement Error in Nonlinear Models: A Modern Perspective - Second Edition*. Chapman & Hall/CRC.

Clair, L., H. Anderson, C. Anderson, O. Ekuma and H.J. Prior (2022). Cardiovascular disease and the risk of dementia: a survival analysis using administrative data from Manitoba. *Canadian Journal of Public Health* **113**, 455–464.

Cook, R.J., J.F. Lawless and B. Xie (2022). Marker-dependent observation and carry-forward of internal covariates in Cox regression. *Lifetime Data Analysis* **28**, 560–584.

Cortese, G. and P.K. Andersen (2009). Competing risks and time-dependent covariates. *Biometrical Journal* **51**(6), 138–158.

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

Ding, J. and J.L. Wang (2008). Modeling longitudinal data with nonparametric multiplicative random effects jointly with survival data. *Biometrics* **64**(2), 546–556.

Eddelbuettel, D. (2013). *Seamless R and C++ Integration with Rcpp*. Springer. New York. ISBN 978-1-4614-6867-7.

Eddelbuettel, D. and J.J. Balamuta (2018). Extending R with C++: A Brief Introduction to Rcpp. *The American Statistician* **72**(1), 28–36.

Eddelbuettel, D. and R. François (2011). Rcpp: Seamless R and C++ integration. *Journal of Statistical Software* **40**(8), 1–18.

Fan, J. and R. Li (2002). Variable selection for Cox's proportional hazards model and frailty model. *The Annals of Statistics* **30**(1), 74–99.

Godambe, V.P. (1991). *Estimating Functions*. Oxford University Press, USA.

Gong, Q. and D.E. Schaubel (2013). Partly conditional estimation of the effect of a time-dependent factor in the presence of dependent censoring. *Biometrics* **69**, 338–347.

Government of British Columbia (2018). *Opioid use disorder: Diagnosis and management in primary care*.

Green, P.J. and B.W. Silverman (1994). *Nonparametric Regression and Generalized Linear Models*. Chapman and Hall. London.

Health Canada (2016). *Health Canada to propose regulatory change to enable consideration of applications under the special access programme to facilitate treatment of chronic relapsing opioid dependence*.

Health Canada (2019). *Government of Canada approves new treatment options for opioid use disorder and supports research, treatment and harm reduction projects in Ontario*.

Henderson, R., P. Diggle and A. Dobson (2000). Joint modelling of longitudinal measurements and event time data. *Biostatistics* **1**(4), 465–480.

Hernan, M.A., S.R. Cole, J. Margolick, M. Cohen and J. Robins (2005). Structural accelerated failure time models for survival analysis in studies with time-varying treatments. *Pharmacoepidemiology & Drug Safety* **14**(7), 477–491.

Hinds, A., L.M. Lix, M. Smith, H. Quan and C. Sanmartin (2016). Quality of administrative health databases in Canada: A scoping review. *Canadian Journal of Public Health* **107**, e56–e61.

Hsieh, F., Y.K. Tseng and J.L. Wang (2006). Joint modelling of survival and longitudinal data likelihood approach revisited. *Biometrics* **62**, 1037–1043.

Hu, X.J., M. Lorenzi, J.J. Spinelli, S.C. Ying and M.L. McBride (2011). Analysis of recurrent events with non-negligible event duration, with application to assessing hospital utilization. *Lifetime Data Analysis* **17**(2), 215–233.

Igelström, E., P. Craig, J. Lewsey, J. Lynch, A. Pearce and S.V. Katikireddi (2022). Causal inference and effect estimation using observational data. *Journal of Epidemiology & Community Health* **76**(11), 960–966.

Iron, K. and D.G. Manuel (2007). *Quality assessment of administrative data (QuAAD): an opportunity for enhancing Ontario's health data.*. ICES Investigative Report. Toronto: Institute for Clinical Evaluative Sciences.

Jacobsen, M. (1989). Existence and unicity of MLEs in discrete exponential family distributions. *Scandinavian Journal of Statistics* **16**, 335–349.

Kalbfleisch, J. D. and R. L. Prentice (2002). *The Statistical Analysis of Failure Time Data*. New York: Wiley.

Karatzas, I. and S.E. Shreve (1988). *Brownian Motion and Stochastic Calculus*. New York: Springer.

Laing, M.K., K.W. Tupper and N. Fairbairn (2018). Drug checking as a potential strategic overdose response in the fentanyl era. *International Journal of Drug Policy* **62**, 59–66.

Laird, N.M. and J.H. Ware (1982). Random-effects models for longitudinal data. *Biometrics* **38**(4), 963–974.

Larney, S., N.R. Jones, M. Hickman, S. Nielsen, R. Ali and L. Degenhardt (2023). Does opioid agonist treatment reduce overdose mortality risk in people who are older or have physical comorbidities? Cohort study using linked administrative health data in New South Wales, Australia, 2002-17. *Addiction* pp. 1–13.

Lewis, J.W. (1985). Buprenorphine. *Drug and Alcohol Dependence* **14**(3), 363–372.

Lin, D.Y. and Z. Ying (1995). Semiparametric inference for the accelerated life model with time-dependent covariates. *Journal of Statistical Planning and Inference* **44**, 47–63.

Lin, D.Y., L.J. Wei, I. Yang and Z. Ying (2000). Semiparametric regression for the mean and rate functions of recurrent events. *Journal of the Royal Statistical Society, Series B* **62**(4), 711–730.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer.

Ministry of Public Safety and Solicitor General (2018). *British Columbia Corrections Dataset*. British Columbia Ministry of Health.

National Academies of Sciences, Engineering, Medicine (2019). *Medications for opioid use disorder save lives*.

Paulus, I. and R. Halliday (1967). Rehabilitation and the narcotic addict: Results of a comparative methadone withdrawal program. *Canadian Medical Association Journal* **96**, 655–659.

Pearce, L.A., J.E. Min, M. Piske, H.C. Zhou, F. Homayra, A. Slaunwhite, M. Irvine, G. McGowan and B. Nosyk (2020). Opioid agonist treatment and risk of mortality during an opioid overdose public health emergency: A population-based retrospective cohort study. *BMJ* **386**, m772.

Piske, M., F. Homayra, M.E. Jeong, H. Zhou, C. Marchand, A. Mead, J. Ng, M. Woolner and B. Nosyk (2021). Opioid use disorder and perinatal outcomes. *Pediatrics* **148**(4), e2021050279.

Piske, M., H. Zhou, J.E. Min, N. Hongdilokkul, L.A. Pearce, F. Homayra, M.E. Socias, G. McGowan and B. Nosyk (2020). The cascade of care for opioid use disorder: a retrospective study in British Columbia, Canada. *Addiction* **115**(8), 1482–1493.

Pollard, D. (1990). *Empirical Processes: Theory and Applications*. Hayward: Institute of Mathematical Statistics.

Prentice, R. (1982). Covariate measurement errors and parameter estimates in a failure time regression model. *Biometrika* **69**, 331–342.

Prentice, R.L., B.J. Williams and A.V. Peterson (1981). On the regression analysis of multivariate failure time data. *Biometrika* **68**(2), 373–379.

Province of British Columbia (2018). *Escalating BC's response to the overdose emergency*. British Columbia Ministry of Mental Health and Addictions.

Putter, H. and H.C. van Houwelingen (2022). Landmarking 2.0: Bridging the gap between joint models and landmarking. *Statistics in Medicine* **41**(11), 1901–1917.

R Core Team (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.

Rice, J.A. and C.O. Wu (2001). Nonparametric mixed-effects models for unequally sampled noisy curves. *Biometrics* **57**, 253–259.

Rizopoulos, D. (2010). JM: An R package for the joint modelling of longitudinal and time-to-event data. *Journal of Statistical Software* **35**(9), 1–33.

Rizopoulos, D. (2011). Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics* **67**, 819–829.

Rizopoulos, D., G. Molenberghs and E.M.E.H. Lesaffre (2017). Dynamic predictions with time-dependent covariates in survival analysis using joint modeling and landmarking. *Biometrical Journal* **59**(6), 1261–1276.

Rizopoulos, D., L.A. Hatfield, C.P. Bradley and J.M. Takkenberg (2014). Combining dynamic predictions from joint models for longitudinal and time-to-event data using bayesian model averaging. *Journal of the American Statistical Association* **109**(508), 1385–1397.

Robins, J. and A.A. Tsiatis (1992). Semiparametric estimation of an accelerated failure time model with time- dependent covariates. *Biometrika* **79**(2), 311–319.

Rockafeller, R.T. (1970). *Convex Analysis*. Princeton: Princeton University Press.

Santo Jr, T., M. Clark, B. Hickman, J. Grebely, G. Campbell, L. Sordo, A. Chen, L.T. Tran, C. Bharat, P. Padmanathan, G. Cousins, J. Dubouy, Kelty E., R. Muga, B. Nosyk, J.E. Min, R. Pavarin, M. Farrell and L. Degenhardt (2021). Association of opioid agonist treatment with all-cause mortality and specific causes of death among people with opioid dependence: A systematic review and meta-analysis. *JAMA Psychiatry* p. e210976.

SAS Institute Inc. (2013). *SAS/Stat Software, Version 9.4.*

Schaubel, D.E., A. Wolfe, C.S. Sima and R.M. Merion (2009). Estimating the effect of a time-dependent treatment by levels of an internal time-dependent covariate: Application to the contrast between liver wait-list and posttransplant mortality. *Journal of the American Statistical Association* **104**(485), 49–59.

Shorack, G.R. and J.A. Wellner (1986). *Empirical Processes with Applications to Statistics*. New York: Wiley.

Simoneau, G., E.E.M. Moodie, J.S. Nijjar, R.W. Platt and Scottish Early Rheumatoid Arthritis Inception Cohort Investigators (2020). Estimating optimal dynamic treatment regimes with survival outcomes. *Journal of the American Statistical Association* **115**(531), 1531–1539.

Song, S., M. Davidian and A.A. Tsiatis (2002). An estimator for the proportional hazards model with multiple longitudinal covariates measured with error. *Biostatistics* **3**(4), 511–528.

Sordo, L., G. Barrio, M.J. Bravo, B.I. Indave, L. Degenhardt, L. Wiessing, M. Ferri and R. Pastor-Barriuso (2017). Mortality risk during and after opioid substitution treatment: systematic review and meta-analysis of cohort studies. *BMJ* **357**, j1550.

Spreafico, M. and F. Ieva (2021). Functional modeling of recurrent events on time-to-event processes. *Biometrical Journal* **63**, 948–967.

Stefanski, L.A. and R.J. Carroll (1987). Conditional scores and optimal scores for generalized linear measurement error models. *Biometrika* **74**(4), 703–716.

Sud, A., M. McGee, B. Mintzes and M. Herder (2022). Permissive regulation: A critical review of the regulatory history of buprenorphine formulations in canada. *International Journal of Drug Policy* **105**, 103749.

Suresh, K., J.M.G. Taylor, D.E. Spratt, S. Daignault and A. Tsodikov (2017). Comparison of joint modeling and landmarking for dynamic prediction under an illness-death model. *Biometrical Journal* **59**(6), 1277–1300.

Tetrault, J.M. and D.A. Fiellin (2012). Current and potential pharmacological treatment options for maintenance therapy in opioid-dependent individuals. *Drugs* **72**, 217–228.

Therneau, Terry M (2015). *A Package for Survival Analysis in S.* version 2.38.

Tsiatis, A.A. (1981). A large sample study of Cox's regression model. *The Annals of Statistics* **9**(1), 93–108.

Tsiatis, A.A. (2006). *Semiparametric Theory and Missing Data.* Springer Series in Statistics.

Tsiatis, A.A. and M. Davidian (2001). A semiparametric estimator for the proportional hazards model with longitudinal covariates measured with error. *Biometrika* **88**(2), 447–458.

Tsiatis, A.A. and M. Davidian (2004). Joint modelling of longitudinal and time-to-event data: An overview. *Statistica Sinica* **14**, 809–834.

Tsiatis, A.A., V. DeGruttola and M.S. Wulfsohn (1995). Modeling the Relationship of Survival to Longitudinal Data Measured with Error. Applications to Survival and CD4 Counts in Patients with AIDS. *Journal of the American Statistical Association* **90**(429), 27–37.

van Houwelingen, H.C. (2007). Dynamic prediction by landmarking in event history analysis. *Scandinavian Journal of Statistics* **34**, 70–85.

Vansteelandt, S. and M. Joffe (2014). Structural nested models and G-estimation: The partially realized promise. *Statistical Science* **29**(4), 707–731.

Wang, Y. and J.M.G. Taylor (1997). Jointly modeling longitudinal and event time data with application to acquired immunodeficiency syndrome. *Journal of the American Statistical Association* **96**(455), 895–905.

Wong, K.Y., D. Zeng and D.Y. Lin (2022). Semiparametric latent-class models for multivariate longitudinal and survival data. *Annals of Statistics* **50**(1), 487–510.

Wulfsohn, M.S. and A.A. Tsiatis (1997). A joint model for survival and longitudinal data measured with error. *Biometrics* **53**, 330–339.

Xu, J. and S.L. Zeger (2001). Joint analysis of longitudinal data comprising repeated measures and times to events. *Journal of the Royal Statistical Society: Applied Statistics, Series C* **50**(3), 375–387.

Yang, S., K. Pieper and F. Cools (2020). Semiparametric estimation of structural failure time models in continuous-time processes. *Biometrika* **107**(1), 123–136.

Yi, G.Y. (2017). *Statistical Analysis with Measurement Error or Misclassification: Strategy Method and Application.* Springer.

Zhao, J., X. Xin, G.X. Xie, P.P. Palmer and Y.G. Huang (2012). Molecular and cellular mechanisms of the age-dependency of opioid analgesia and tolerance. *Molecular Pain* **10**(38), 1–12.

Zheng, Y. and P.J. Heagerty (2005). Partly conditional survival models for longitudinal data. *Biometrics* **61**, 379–391.

Zhong, R. and J. Zhang (2022). *SLFPCA: Sparse Logistic Functional Principal Component Analysis*. R package version 2.0.

Zhong, R., S. Liu, H. Li and J. Zhang (2021). Sparse logistic functional principal component analysis for binary data. *ArXiv*.

Zhou, R. and J. Sun (2022). Estimation of the proportional mean residual life model with internal and longitudinal covariates. *Statistics in Biosciences* **14**, 550–563.

Zubieta, J.K., R.F. Dannals and J.J. Frost (1999). Gender and age influences on human brain mu-opioid receptor binding measured by PET. *American Journal of Psychiatry* **156**, 842–848.