©Copyright 2014

Sooraj Kuttykrishnan

A Marginal Structural Cox Model Based Analysis Of The Comparative Effectiveness Of Two Dialysis Therapies

Sooraj Kuttykrishnan

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science

University of Washington

2014

Committee:

Ken Rice, Chair

Adam Szpiro

Program Authorized to Offer Degree:

Biostatistics

University of Washington

Abstract

A Marginal Structural Cox Model Based Analysis Of The Comparative Effectiveness Of Two Dialysis Therapies

Sooraj Kuttykrishnan

Chair of the Supervisory Committee:

Ken Rice

Associate Professor

Department of Biostatistics

The Institute of Medicine identified comparing the effectiveness of renal replacement therapies as the only kidney-disease related topic among the top 100 initial national priorities for comparative effectiveness research. The median life expectancy of patients starting renal replacement therapy is only a little over three years and they spend on average 12 days in the hospital annually. The overwhelming majority of patients are treated with thrice-weekly in-center hemodialysis(TWICHD). However, an increasingly larger number of patients are being treated alternate dialysis modalities, including nocturnal in-center hemodialysis (NICHD), which is a modified hemodialysis regimen that include significantly longer treatment times. As such, it is critically important to perform a rigorous assessment of the benefit, if any, conferred by NICHD.

In this thesis, we present a careful analysis of a large cohort from a retrospective study of dialysis modalities in the US, using marginal structural Cox models. The data set

consists of 208,820 patients who initiated dialysis between January 1, 2007- December 31, 2011 at facilities owned by Davita corporation.

Prior to presenting the data analysis, we provide a overview of the relevant concepts from the theory of the counterfactual model and causal inference, and discuss marginal structural models for binary as well as time to event outcomes. For the primary analysis, we estimated the causal hazard ratio comparing the risk of death if patients received NICHD as opposed to TWICHD through their follow-up, contrary to their observed treatment history using a pre-specified marginal structural Cox model. We found that those who were treated with NICHD had a substantial reduction in their risk of death. We also performed a series of sensitivity analyses to assess the effect of potential sources of bias which were not accounted for in the primary analysis. We found that our main result was robust to these potential sources of bias. In addition, using a bootstrap procedure, we also investigated the effect of the parametric assumptions we made on the precision of our estimates. We found that the model we used provided a substantial gain in precision of our estimate of the causal effect of NICHD on survival.

Acknowledgements

First and foremost, I would like to thank my advisor Ken Rice for being a patient listener and his guidance with both the writing of this thesis and for understanding the relevant statistical concepts. I would also like to thank Adam Szpiro for his valuable suggestions to improve the thesis, and Scott Emerson for teaching me about all things critical about biostatistics.

Gitana Garofalo was very helpful with my questions about departmental requirements and rules. I thank the biostatistics department for providing me with partial funding.

I would also like to thank Rajnish Mehrotra from the Division of Nephrology, both for giving access to the data set and for his insights and suggestions about analyses presented.

Finally I have to thank my family for keeping me focused on the task, and being there for me in general. In particular Rebecca was always supportive throughout the trials and tribulations of the masters program. Without her infinite patience, this thesis would not have been possible. My kids were a joy, despite the sacrifices they had to make to give their daddy his "work time".

Contents

0 Introduction **10**

1 Causal Inference **13**

1.1 Causal effects 13

1.2 Endogeneity and time varying confounding. 17

1.2.1 Inference in the presence of endogenous covariates 18

2 Marginal Structural Models **21**

2.1 Marginal structural logistic model 21

2.1.1 IPTW and consistent estimation of parameters 23

2.1.2 Estimating stabilized IPT weights 25

2.1.3 Modifying IPTW to account for censoring 26

3 Time to event data and Marginal Structural Cox models **29**

3.1 Inverse probability of treatment weights for Cox models 31

3.1.1 Censoring 31

3.1.2 Estimating stabilized IPT weights 32

3.2 Fitting the model - multiple options 33

3.2.1 Simulation results from Xiao et al. (2010) 33

4 Comparative Effectiveness of Two Hemodialysis Therapies **36**

4.1 Background 36
4.2 Descriptive Analyses 38
4.2.1 Missing Values 45
4.3 Primary Analysis 47
4.3.1 Results 49

5 Sensitvity Analyses & Discussion of the results 52
5.1 Potential survivor bias and attribution of deaths 52
5.1.1 Hazard of death among those who survived the first 90 days of dialysis 53
5.1.2 Hazard of death estimated from a nested case-control cohort . . . 54
5.2 Facility level effects 56
5.3 Precision of the point estimate of the primary analysis 58
5.4 Discussion of the results 58

List of Figures

1.1 An example of a DAG 16
1.2 Inducing backdoor paths 17
1.3 An endogenous covariate - L 18
1.4 Induced unmeasured confounding in the presence of endogeneity 19

2.1 Time varying exposure with a binary outcome 22
2.2 Time varying exposure with a binary outcome (loss to follow up included) . 27

4.1 Distribution of stabilized weights sw_i^*. 51

5.1 Normality of bootstrap estimates 59

List of Tables

3.1 Comparison of different methods to fit an MSCM 35

4.1 Descriptive statistics of covariates comparing patients who were ever treated with NICHD to those who were only treated with TWICHD 39

4.2 Comparison of patients who transferred to NICHD from TWICHD to those remained on TWICHD 42

4.3 Characteristics of patients with missing values in select time varying covariates 46

4.4 Causal association between being treated with NICHD and mortality . . . 49

5.1 Sensitivity due to left truncation and lagged effect of therapy on the causal effect of being treated with NICHD on mortality 54

5.2 Sensitivity of the causal effect of NICHD on mortality to treatment history . 56

5.3 Sensitivity of the causal effect of NICHD on mortality to being treated in a dialysis facility offering NICHD 57

Chapter 0

Introduction

In this thesis, we present a careful analysis of a large data set, using a sophisticated and nonstandard statistical method. The data set is from a retrospective study of dialysis modalities in the US. It consists of 208,820 patients who initiated dialysis between January 1, 2007- December 31, 2011 at facilities owned by Davita corporation. It has information about demographics, lab variables, intravenous medications and access types. The goal of the analysis was to investigate the causal effect of choice of a dialysis modality on mortality.

Some of the lab variables as well as intravenous medications are potentially time varying confounders of the association between the choice of dialysis modality and mortality. In addition, the choice of a specific dialysis modality at any particular time points maybe both predictive of the choice of the modality at the next time point, and these choices are putatively causally associated with mortality. In other words the treatment status is an endogenous covariate. As such, traditional regression methods would likely give a biased estimate of this causal effect. Hence we analysed this data set using marginal structural Cox models to perform statistical inference on a parameter, interpretable as a hazard ratio that quantifies the causal effect of interest. Specifically, we investigated if

nocturnal in-center hemodialysis (NICHD) offers any potential survival benefit compared to the conventional three times a week in-center hemodialysis (TWICHD).

The thesis is organized in five chapters.

Chapter 1: Causal Inference. In this chapter we discuss concepts related to causal inference, time varying confounders and endogeneity.

Chapter 2: Marginal Structural Models. This Chapter provides motivation for use of *Marginal Structural Logistic Models*, a statistical approach by which we can estimate the causal effect of an endogenous covariate, in the presence of potential time varying confounders, on a binary outcome. In this chapter we define a marginal structural logistic model, and explain how to estimate the causal parameter of interest using stabilized inverse probability of treatment and censoring weights.

Chapter 3: Marginal Structural Cox Models. We present a variant of Marginal Structural Models, *Marginal Structural Cox Models* in this chapter. Marginal Structural Cox Models allows one to estimate the causal effect of a time varying and endogenous exposure on a time to event outcome, in the presence of potential time varying confounders. In Chapter 3, we explain how to estimate the causal parameter of interest using a weighted Cox proportional hazards model. In addition we also discuss results from a simulation analysis to discuss the benefit of using weighted Cox proportional hazards regression, as opposed to a discrete logistic regression, to estimate the causal effect of interest.

Chapter 4: Comparative effectiveness of two hemodialysis therapies. In this chapter, we compare the survival benefit of NICHD, if any, over TWICHD using marginal structural cox models. In addition, we also present descriptive statistics, for relevant covariates that may affect choice of the dialysis modality as well as mortality.

Chapter 5: Sensitvity Analyses & Discussion of the results We discuss a series of sensitivity analysis to assess potential biases that may arise from the assumptions

we make about the marginal structural model used in the primary analysis. Specifically, we investigate two sources of bias. First, we explore the effect of potential survival bias that may arise from the substantial interval from the initiation of dialysis and initiation of treatment with NICHD. Second we study the effects of being treated in a dialysis facilities that offer NICHD as opposed to facilities that do not. We also investigate the precision of our estimate of the causal effect using, a bootstrap procedure.

In addition, we also summarize the main findings of the thesis, and include a discussion of how our results fit into the existing literature.

Chapter 1

Causal Inference

1.1 Causal effects

We will use notation used in Hernan et al. (2000) to explain the relevant concepts from causal inference.

Suppose we are interested in the relationships between an exposure A and an outcome Y, and that A could take any value in a set $\mathscr{A}$. We call the outcome for an individual i *had their exposure status been* a a potential or *counterfactual* outcome. We will denote these counterfactual outcomes for each individual as $Y_i^{A=a}, a \in \mathscr{A}$, and note that each individual has a counterfactual outcome for each $a \in \mathscr{A}$. Note that, except for the factual outcome Y_i corresponding to the exposure status that actually occured, the counterfactual outcome(s) for individual i are not observed.

For example, if the exposure, A, is dichotomous, there are two possible exposure; $a = 1$ (exposed) and $a = 0$ (unexposed). We denote $Y_i^{a=1}$ as the outcome variable for individual i if they were exposed and $Y_i^{a=0}$ if they were unexposed. One of these variables $Y_i^{a=1}$ and $Y_i^{a=0}$, corresponding to the value of a actually seen will be factual, and the other is counterfactual. We can only observe the factual outcome for any given individual, not

their counterfactual outcome, or outcomes when more than one is considered.

This terminology is useful when defining the *average causal effect*; in the case of dichotomous A the standard definition of this effect is just the difference between the means of the counterfactual outcomes, i.e. $\mathbb{E}[Y^{a=1}] - \mathbb{E}[Y^{a=0}]$. But the term 'causal effect' could more generally be defined as a contrast of any functional of the distributions of counterfactual outcomes under different exposure histories. In other words, any comparison of $Y^{a=1}$ and $Y^{a=0}$ could define a causal effect; difference in means, medians, hazards, odds, or rates could all be considered.

Causal effect vs association

Formally, we say that there is an *association* between an exposure (A) and an outcome (Y) if the distribution of the outcome variable is different in categories defined by the exposure. Exposure A is therefore associated with Y if, for example, $\mathbb{E}[\,Y|A = 1\,] \neq \mathbb{E}[\,Y|A = 0\,]$. In other words, there is association if the mean of Y among the group of people who were exposed, is different from the mean of Y among those who were not. Association is implied by unequal *conditional* expectations. In contrast, A has a causal effect on Y if the means of counterfactual outcomes are unequal; if $E(Y^{a=1}) \neq E(Y^{a=0})$. That is, the mean of the outcome if everyone in the population were exposed is different from the mean of the outcome if none of them were exposed, keeping everything else the same in both situations.

Directed acyclic graphs (DAGs)

Following the pioneering work of (Pearl 2000, Robins 1986) con causal diagrams, directed acyclic graphs (DAGS)have become a standard and useful way to describe concepts and results used in causal inference. Here, we summarize the discussion in Greenland et al. (1999) about how to use such diagrams to identify potential confounders, the

biases that may arise in inference due to confounding, and how to appropriately control for the presence of confounders.

Causal diagrams summarize the causal assumptions about the hypothesized relationships between covariates of interest. For example, suppose that in addition to the causal relationship between A and Y in the previous section, there are other covariates of interest, say L, K and M. L directly influences A and Y, K directly influences A, but its influence of Y is through its relation with L. M directly influences Y, and influences A through its relations with L. These relationships are summarized in the graph in figure 1.1. Arrows from one variable to another indicate the direct influences described above; lack of an arrow between two variables means there is no direct influence.

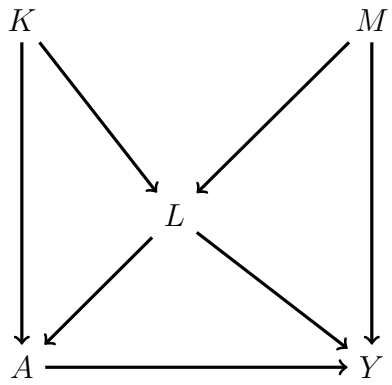
In graph-theoretic 'language', the points on the graphs representing the variables are called *nodes*, a *path* is any unbroken route traced through the arrows of the graph, a *directed path* is from one node to another is a path that is always entering an edge through the tail and exiting through the arrow head, and a *cycle* is a directed path beginning and ending at the same node. A *directed acyclic graph (DAG)* is a graph that has no cycles. In figure 1.1, A, Y, K and M are nodes, and the arrows are edges; an example of a directed path will be KAY. Also, this graph is an example of a DAG.

A *backdoor path* from A to Y is a path with an arrow pointing to A. All paths from A to Y in figure 1.1 are backdoor paths, except the direct path AY. A path *collides* at a node if the path enters and exits the node through arrow heads. The path ALY collides at L, and L is called a collider. A path is blocked if it has one or more colliders.

The graph-theoretic language makes it straightforward to define when a relationship is confounded. Greenland et al. (1999) gives the following algorithm to detect confounders in a situation represented by a DAG: Remove all single headed arrows in the graph emanating from the exposure, and if there are any unblocked paths from the exposure to the outcome in the new graph then there is confounding of the net exposure effect. Checking

if there is such an unblocked path in a DAG reduces to checking if the exposure and the outcome share a common cause. Confounding is precisely the bias that arises while estimating the causal effect of an exposure on an outcome by the associaiton of those two variables. When the exposure and the outcome have a common cause, the common cause is the confounding variable (Hernan & Robins 2013, Chapter 7, page 83).

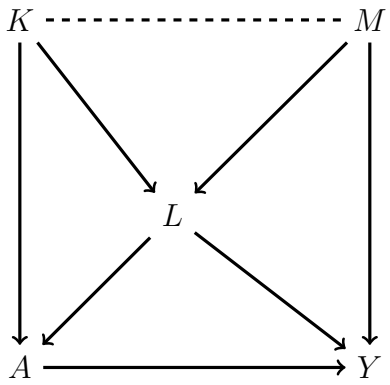
Figure 1.1: An example of a DAG



However, simply removing nodes on the raw graph until all backdoor paths are blocked does not rule out confounding. Greenland et al. (1999) points out a situation when adjustment for a confounder could *induce* backdoor paths that were not present prior to adjustment. Consider the graph in Figure 1.1. Neither K nor M are potential common causes of the exposure A and the outcome Y. However, adjusting for only L while estimating the association between A and Y can induce an association between K and M. That is, even if K and M are independent (denoted $K \coprod M$) in the general population, within the strata of L, K and M might be dependent. In other words, adjusting for L while estimating the association between A and Y could introduce confounding of the effect estimate by K or M. In particular, Berkson (1946) provided an example where K indicated presence of lung cancer, M indicated presence of tuberculosis, and L indicated admittance to hospital. Berkson showed that even if $K \coprod M$ in the general population

and both influenced admittance to hospital, then K and M would be associated among the hospital admittees.

Figure 1.2: Inducing backdoor paths



1.2 Endogeneity and time varying confounding.

We shall define a *time varying confounder* to be a time varying covariate that at any given time t is causally related to the outcome of interest as well as exposure at subsequent times, and an *intermediate variable* to be a variable that is affected by the exposure and also affects the outcome. To illustrate these, we give a hypothetical example where a variable is both a time varying confounder and an intermediate variable.

Example 1.2.1. Suppose data is available at follow-up times $t = 0, 1$. Let A_1 and A_2 be binary variables indicating whether a particular treatment is given at times $t = 0$ and 1 respectively. The outcome Y, measured after the second time point is also binary, with $Y = 1/0$ indicating presence/absence of disease. Treatment status of each subject at time $t = 0$ is decided based on their risk factor profile at baseline, denoted L_0. Based on the risk factor profile measured at time point one, denoted L_1, treatment A_1 is decided. U is an unmeasured covariate that causally affects L_1 and Y. The directed acyclic graph

in Figure 1.3 represents relevant covariates and the conditional causal relationships between them at the two time points.

The path A_1L_1Y and $A_0L_0L_1Y$ are unblocked backdoor paths from A_1 to Y and A_0 to Y. So L is a confounder for the association between A and Y at both time points. As such, L is a time varying confounder. In addition L_1 is also an intermediate variable, influenced by the exposure A_0 and influencing the outcome Y. L_1 is an example of an *endogenous covariate* – a covariate that is simultaneously a time varying confounder, and an intermediate variable.

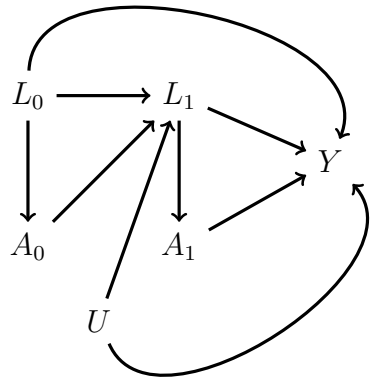
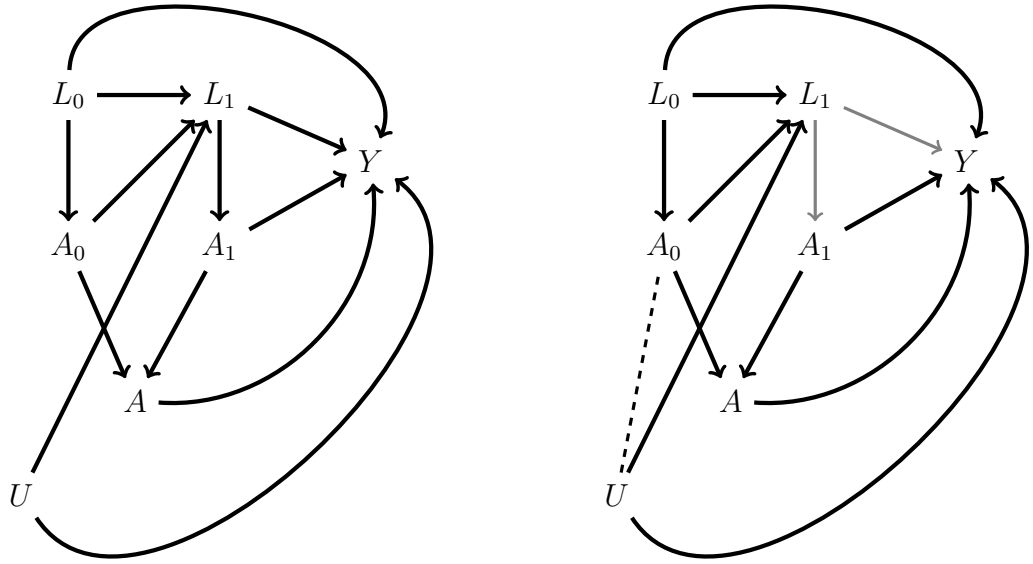


Figure 1.3: An endogenous covariate - L

1.2.1 Inference in the presence of endogenous covariates

Conventional methods could lead to a biased estimate of causal effect between an exposure and an outcome, if there is a confounder of the relationship of interest that is also an endogenous covariate. To see this, consider example 1.2.1, and suppose we are interested in the causal effect of the treatment at both time points on the outcome. In particular we would like to estimate the difference in means between two counterfactual outcomes; the outcome when all subjects are treated at both time points (so $A_0 = A_1 = 1$) and the



(a) (b)

Figure 1.4: Induced unmeasured confounding in the presence of endogeneity

outcome when none of the subjects are ever treated (so $A_0 = A_1 = 0$), regardless of other covariate values for the subjects. The causal assumptions about the variables involved are represented in Figure 1.4a. Formally, we are interested in the *average causal effect (ACE)*

$$ACE_{AY} = \mathbb{E}[\, Y^{\overline{A}=\{1,1\}}\,] - \mathbb{E}[\, Y^{\overline{A}=\{0,0\}}\,],$$

where following Hernan et al. (2000) notation, $\overline{A} = (A_0, A_1)$ is the treatment history of the subjects.

To obtain an unbiased estimate of this causal effect, we will have to estimate the causal effect of A_0 and A_1 on Y simultaneously and without systematic bias. $A_1 L_1 Y$ is an unblocked backdoor path from A_1 to Y, and as such L_1 is a confounder for the association between A_1 and Y. The conventional approach to control for the confounder

L_1 would be to adjust for it in a regression model, or equivalently to stratify on the levels of it. However adjusting for L_1 while estimating the association between A_1 and Y induces a path from U to A_0 in the DAG in figure 1.4b (Greenland et al. 1999, page 42). A_0UY is an unblocked backdoor path from A_0 to Y in figure 1.4b, and U is now an unmeasured confounder for the association between A_0 and Y. Note that prior to adjustment for L_1, U was an unmeasured covariate, but not a confounder for either association of interest.

We see that we cannot obtain an unbiased estimate of the association between A_0 and Y using this conventional approach. What is needed is a method that removes confounding by L on the effect estimates of the associations between A_0 and A_1 on Y, without inducing confounding by U. We will discuss one such method, marginal structural modeling, in the next chapter.

Chapter 2

Marginal Structural Models

Marginal structural models (MSMs) are a class of causal models developed by James Robins, Miguel Hernan and Babette Brumback in a series of articles published in 1997-2000 (Robins 1997, 1998, 1999*a*,*b*, Hernan et al. 2000, Robins et al. 2000). MSMs provide a way to obtain unbiased estimates of parameters representing the causal effect of a time varying exposure on an outcome, in the presence of time varying endogenous covariates. Marginal structural Cox models (MSCMs) are an extension of MSMs when the outcome of interest is time to an event in the presence of censoring. In this chapter we will give a brief overview of MSMs, explaining how to use inverse probability of treatment weights (IPTW) to consistently estimate parameters of an MSMs and MSCMs.

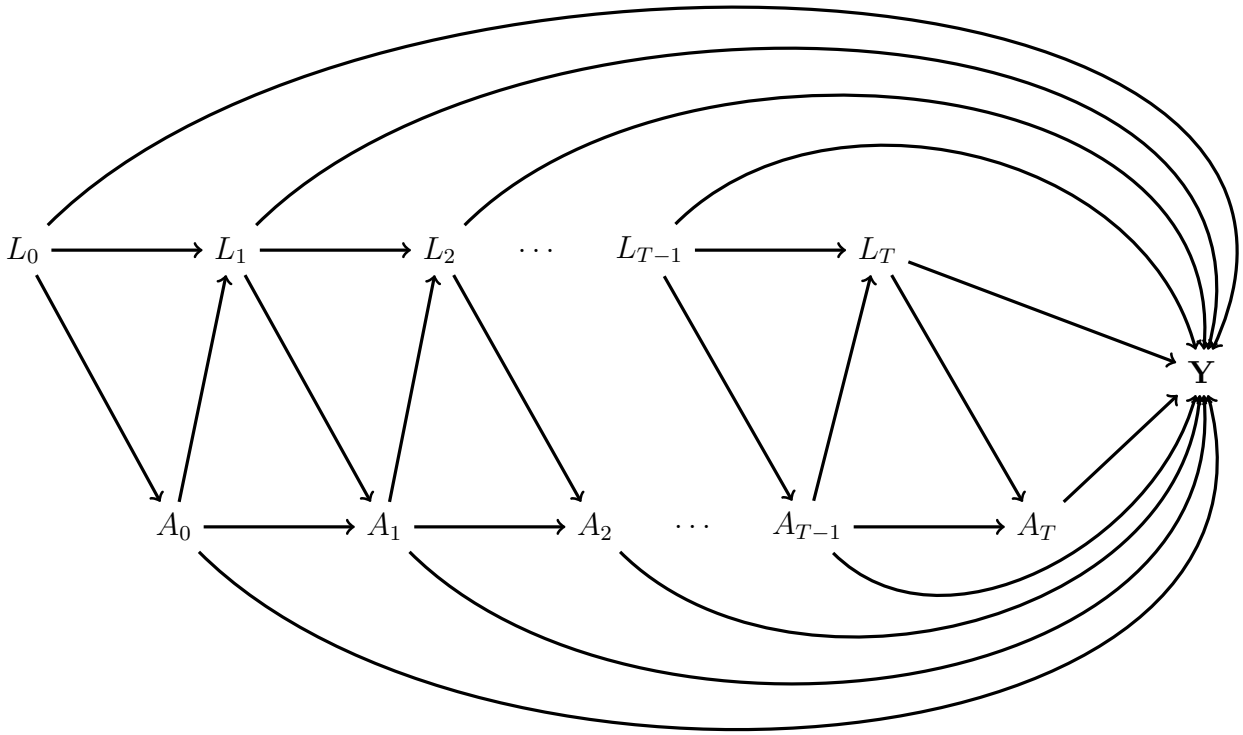
2.1 Marginal structural logistic model

We will use the example given by Robins et al. (2000) to explain the concepts related to marginal structural models when the outcome of interest in binary.

Consider a cohort study of HIV infected patients. Suppose that patients are followed for $T+1$ days. Let A_t be the dose of treatment with AZT on day t after start of follow up,

Y be the binary outcome that HIV RNA is detectable at the end of follow up on day $T+1$, and L_t be the vector of all measured risk factors for the outcome on day t (age, CD4 count etc). The causal relationships involved are summarized in the DAG in figure 2.1. The dose of treatment on day t is determined by the measured levels of risk factors on that day, and the dose of AZT on the previous day. The dose of treatment,and the levels of risk factors on the previous day affects levels of risk factors on the next day. Finally the dose of AZT and the levels of risk factors on each day affects the outcome Y.

Figure 2.1: Time varying exposure with a binary outcome



Let $\overline{A_t} = (A_0, A_1, \ldots, A_t)$ be the observed treatment history through day t. Also, let

$\overline{A} = \overline{A_T}, \overline{L_t} = (L_0, L_1, \ldots, L_t)$ and $\overline{L} = \overline{L_T}$ be the observed treatment history through the entire follow up time, the observed history of confounders through day t, and the observed history of confounders through the entire follow up respectively. We are interested in the causal effect of the cumulative dose through end of follow up, $cum(\overline{A}) = \sum_{t=0}^{t=T} A_t$ on the outcome Y.

Let $Y^{\overline{a}}$ be the value of Y had all subjects had a dose history of $\overline{a} = (a_0, a_1, \ldots, a_T)$ instead of the observed history $\overline{A}$. One measure of the causal effect of interest could be the causal odds ratio. It can be expressed in terms of the parameters in the following logistic model.

$$\text{logit } \mathbb{P}[Y^{\overline{a}} = 1] = \beta_0 + \beta_1 \text{cum}(\overline{a}) \tag{2.1}$$

The model in (2.1) is a *marginal structural model*. It is a *marginal* model, since it models the marginal distribution of the counterfactual random variables $Y^{\overline{a}}$. Counterfactual variables are referred to as *structural variables* in econometric and social science literatures. In that sense the above model is a *structural* model as well.

2.1.1 IPTW and consistent estimation of parameters

One way to obtain unbiased estimates of the parameters in (2.1) is to fit an appropriately weighted conventional logistic regression model. Consider the following model.

$$\text{logit } \mathbb{P}[Y = 1|\overline{A} = \overline{a}] = \beta_0' + \beta_1' \text{cum}(\overline{a}) \tag{2.2}$$

If the missingness in the outcome is completely at random, and if there is no selection bias or measurement error, then we can obtain unbiased estimates of β_0' and β_1' by fitting (2.2) using standard methods. If there were no confounding of the association between A_t and Y then the parameters in models (2.2) and (2.1) are the same. In this scenario, we can obtain consistent estimates of β_0 and β_1 by fitting (2.2). However, in the presence

of confounding due to L_t, $\beta_i \neq \beta_i'$.

Even when the association between the treatment variables A_t and the outcome Y are confounded, as shown in (Robins 1997), we can obtain consistent estimates of β_0 and β_1 by fitting the conventional logistic regression model (2.2) with subject i weighted with weights

$$w_i = \frac{1}{\prod_{t=0}^{T} \mathbb{P}(A_t = a_{ti} | \overline{A}_{(t-1)} = \overline{a}_{(t-1)i}, \overline{L}_t = \overline{l}_{ti})} \tag{2.3}$$

where $\overline{A}_{-1} = 0$. This method will work if conditional on the confounder L, the risk of the outcome for an individual treated with a dose a is the same as the risk for the individual treated with a different dose a', had they received the dose a'. That is $Y^a \coprod A|L$. This is condition is known as conditional exchangeability.

The denominator of the weights in (2.3) is informally the conditional probability that the subject i had their own observed treatment history through time T. The effect of weighting with w_i can be thought of as creating a pseudo population with w_i copies of each subject i.

The pseudo population above will have two properties. First, in the pseudo population, the association between A_t and Y is not confounded by L_t. Second, the causal odds ratio (risk difference and risk ratio as well) in the pseudo population is the same as the study population.

To see why these properties hold, for simplicity, we will further assume that L is a categorical variable and that A is dichotomous. The pseudo population will have twice the size of the original population because each individual will appear twice, treated and as untreated. Consequently, within the stratums of L_t, the proportion of treated and untreated are equal; hence $L_t \coprod A_t$ in the pseudo population. Hence, the treated and the untreated are exchangeable in the pseudo population. Thus the associational risk ratio is equal to the causal risk ratio in the pseudo population (Hernan & Robins 2013, Chapter 2).

The pseudo population discussed above, each individual has a probability of receiving the treatment equal to one, and hence it had twice the size of the original population. However, we can also construct similar populations where the probability of being treated is, say 0.5. In this case, the pseudo population will have the same size as the original population, and the corresponding weights are given by $0.5/w_i$. The effect estimates obtained from any of the pseudo populations with weights $p_i/w_i, 0 < p_i \leq w_i$ will be the same (Hernan & Robins 2013, Chapter 12). With the weights w_i given in Equation 2.3, the variance of the estimator could be large, with a markedly non-Normal distribution (Robins et al. 2000). As such, it is better to use the weights

$$sw_i = \frac{\prod_{t=0}^{T} \mathbb{P}(A_t = a_{ti} | \overline{A}_{t-1} = \overline{a}_{(t-1)i})}{\prod_{t=0}^{T} \mathbb{P}(A_t = a_{ti} | \overline{A}_{(t-1)} = \overline{a}_{(t-1)i}, \overline{L}_t = \overline{l}_{ti})} \tag{2.4}$$

where $\overline{A}_{-1} = 0$. These weights have an expected mean of 1, and result in narrower 95% confidence intervals. Furthermore, the estimate obtained by fitting the conventional logistic regression model (2.2) with subject i weighted with weights sw_i given in 2.4 will be consistent estimates of β_0 and β_1.

2.1.2 Estimating stabilized IPT weights

In this section, for simplicity, we will assume that A_t are dichotomous. In practice, each term in the denominator and numerator of the stabilized weights in (2.4) is estimated by modeling each of these conditional probabilities using a generalized linear model. In our case, for each $t \in 0, 1, \ldots T$, we model the conditional probability, $pr[A_t = 1 | \overline{A}_{t-1} = \overline{a}_{t-1}, \overline{L}_{(t-1)} = \overline{l}_{(t-1)}]$ using

$$\begin{aligned} \text{logit } \mathbb{P}[A_t = 1 | \overline{A}_{(t-1)} = \overline{a}_{(t-1)}, \overline{L}_{(t-1)} = \overline{l}_{(t-1)}] = \alpha_0 + \alpha_1 t + \alpha_2 a_{t-1} + \alpha_3 a_{t-2} \\ + \alpha_4 l_t + \alpha_5 l_{t-1} + \alpha_6 a_{t-1} l_k + \alpha_7 l_0. \end{aligned} \tag{2.5}$$

Similarly we will fit another logistic model to obtain the probabilities in the numerator,

$$\text{logit } \mathbb{P}[A_t = 1|\overline{A}_{(t-1)} = \overline{a}_{(t-1)}] = \alpha'_0 + \alpha'_1 t + \alpha'_2 a_{t-1} + \alpha'_3 a_{t-2}. \tag{2.6}$$

For each subject i, using the estimates of the parameters in the model in (2.5), we can obtain the maximum likelihood estimates for $pr[A_t = 1|\overline{A}_{t-1} = \overline{a}_{(t-1)i}, \overline{L}_{(t-1)i} = \overline{l}_{(t-1)i}]$, $\hat{p}_{0i}, \hat{p}_{1i}, \hat{p}_{2i}, \ldots \hat{p}_{Ti}$. Similarly we can also obtain the corresponding maximum likelihood estimates $\hat{p}'_{0i}, \hat{p}'_{1i}, \hat{p}'_{2i}, \ldots \hat{p}'_{Ti}$ for the probabilities in the numerator of the stabilized weights by fitting the model in (2.6). Using these estimates, we can obtain an estimate of the stabilized weights,

$$\widehat{sw}_i = \frac{\prod_{t=0}^{T} (\hat{p}'_{ti})^{a_{ti}} (1 - \hat{p}'_{ti})^{1-a_{ti}}}{\prod_{t=0}^{T} (\hat{p}_{ti}^{a_{ti}} (1 - \hat{p}_{ti})^{1-a_{ti}}} \tag{2.7}$$

If there are no unmeasured confounders, and if the model specified in (2.5) is correct, then the estimates we obtain by fitting the logistic model (2.2) with the weights given by (2.7) will provide consistent estimates of the causal parameter of interest, β_1 in (2.1). Further, the coverage probability of the 95% confidence intervals obtained in this fashion will be at least 0.95 (Robins et al. 2000).

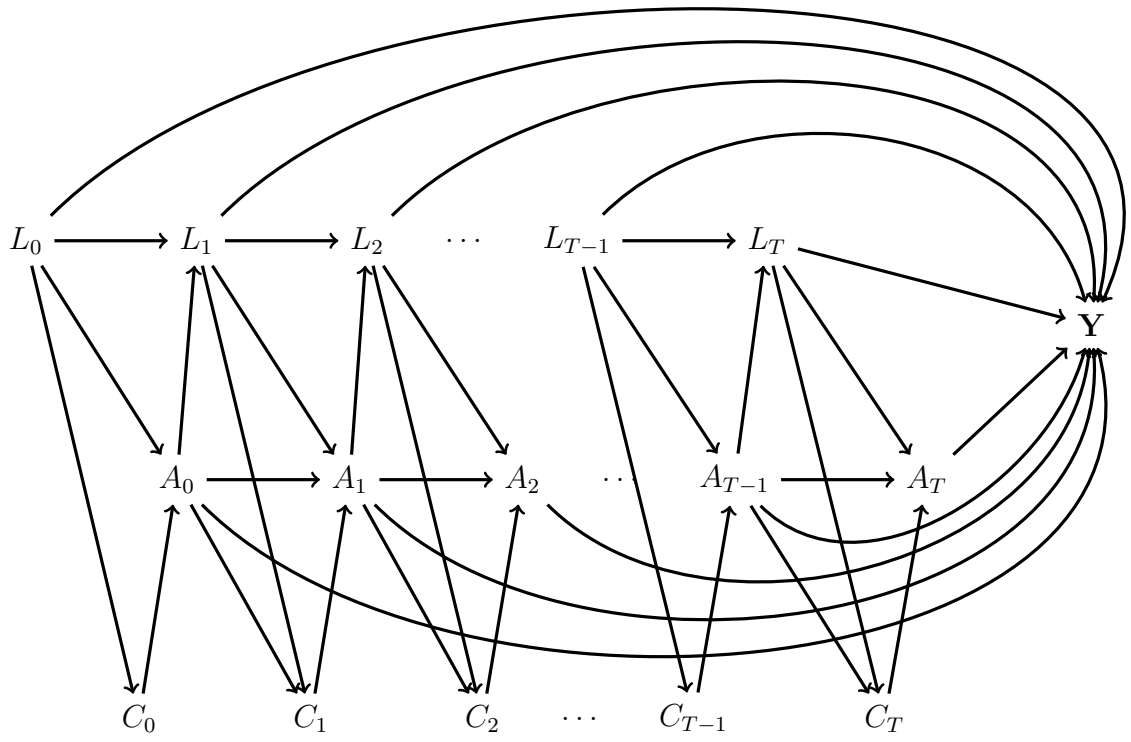
2.1.3 Modifying IPTW to account for censoring

Suppose that in the example represented by the DAG in the figure 2.1 some of the patients were censored due to being lost to follow up during the course of the study. Let $C_{ti} = 1$ if the subject i was lost to follow up at time t, and $C_{ti} = 0$ otherwise. We will need to account for C_t in our models. One way to do this is to think of C_t as another time varying treatment which affects A_t. We would like to do inference in the absence of this time varying treatment; we are interested in estimating the causal effect of $\overline{A}$ on Y if all the subjects had been uncensored, rather than each of them having their observed censoring history. In other words, we would like to estimate the parameter β_1 in the marginal

structural model (2.1), except that now $Y^{\overline{a}}$ refers to the counter factual outcome if the subject had the treatment history $\overline{a}$ and was not censored during the follow up period.

If a subject is censored at time point t then they are not treated in the subsequent time points. In addition the levels of the risk factors, L_t likely affects the censoring status at the same time, C_t. The DAG that represents the relevant causal relationships,including that of censoring status is in figure 2.1.3.

Figure 2.2: Time varying exposure with a binary outcome (loss to follow up included)



As in the absence of censoring, we can obtained consistent estimates of β_1 in the MSM logistic model (2.1) in the presence of censoring as described above by fitting the conventional logistic regression model in (2.2) with appropriate weights. Informally the weights, w_i^c are the inverse of the probability that a subject had their own treatment history

and were uncensored, and more formally we define

$$w_i^c = \frac{1}{\prod_{t=0}^{T} \mathbb{P}(A_t = a_{ti}, C_t = 0 | \overline{A}_{(t-1)} = \overline{a}_{(t-1)i}, \overline{L}_t = \overline{l}_{ti})}$$
$$= \frac{1}{\prod_{t=0}^{T} \mathbb{P}(A_t = a_{ti} | C_t = 0, \overline{A}_{(t-1)} = \overline{a}_{(t-1)i}, \overline{L}_t = \overline{l}_{ti}) \times \prod_{t=0}^{T} \mathbb{P}(C_t = 0 | \overline{A}_{(t-1)} = \overline{a}_{(t-1)i}, \overline{L}_t = \overline{l}_{ti})}$$

where $\overline{A}_{-1} = 0$.

Again, we can stabilize these weights appropriately to obtain weights that will provide efficient estimates. These weights are defined as,

$$sw_i^c = \frac{\prod_{t=0}^{T} \mathbb{P}(A_t = a_{ti} | C_t = 0, \overline{A}_{t-1} = \overline{a}_{(t-1)i}) \times \prod_{t=0}^{T} \mathbb{P}(C_t = 0 | \overline{C}_{t-1} = 0, \overline{A}_{t-1} = \overline{a}_{(t-1)i})}{\prod_{t=0}^{T} \mathbb{P}(A_t = a_{ti} | C_t = 0, \overline{A}_{(t-1)} = \overline{a}_{(t-1)i}, \overline{L}_t = \overline{l}_{ti}) \times \prod_{t=0}^{T} \mathbb{P}(C_t = 0 | \overline{A}_{(t-1)} = \overline{a}_{(t-1)i}, \overline{L}_t = \overline{l}_{ti})}, , \tag{2.8}$$

where $\overline{A}_{-1} = 0$. These weights can be interpreted as the number of copies of each patient in the pseudo-population in which there is no censoring, and that the time varying confounders, L, are not causally associated with the initiation of treatment.

There are a few assumptions necessary for the inferences from an MSM to be valid. First, the measured covariates, L are sufficient to adjust for both confounding and selection bias due to censoring. That is, there are no unmeasured confounders for the association between the exposure and the outcome, given L. Second, we assumed that the logistic models for estimating the weights are correctly specified. In other words, we assumed that the models for initiation of treatment and censoring are correctly specified. Finally, we also assumed that the marginal structural model, for the causal effect of treatment on the outcome, given in equation 2.1 is correctly specified.

Chapter 3

Time to event data and Marginal Structural Cox models

In the previous chapter we introduced marginal structural models, a class of models that can be used to estimate the causal effect of an exposure on a binary outcome, where the exposure is an endogenous covariate and where there are time varying confounders of the association between the exposure and outcome. Hernan et al. (2000) similarly introduced marginal structural Cox proportional hazards models (MSCMs) to estimate the effects of an exposure in similar settings, when the outcome of interest is time to a particular event, in the presence of censoring.

We will use the example used in Hernan et al. (2000) to explain the concepts underpinning MSCMs. Consider a cohort study of HIV infected patients. Let A_t be the dose of treatment with AZT on time t after start of follow up, L_t be a vector of all measured risk factors for the outcome at time t (age, CD4 count, etc), and V be a vector of risk factors for the outcome measured at baseline. We define T to be a patient's time of death, measured in months after the start of follow up.

Suppose we are interested in comparing the hazard of death (at any time t) had all

patients been treated through their follow up, to the hazard of death had all subjects been not treated through the follow-up, contrary to their actual treatment course. We can estimate the ratio of these two hazards using an MSCM as follows. Let $T_{\overline{a}}$ be the time of death for a subject if they had the treatment history $\overline{a} = \{a(t)|\ 0 \leq t < \infty\}$, possibly contrary to their observed treatment history. Then we can model the hazard of death at time t as

$$\lambda_{T_{\overline{a}}}(t|V) = \lambda_0(t)e^{\beta_1 a_t + \beta_2 V} \tag{3.1}$$

In (3.1) λ_0 is the unspecified baseline hazard. This model is an MSCM, because within the levels of the baseline covariates V, it is a causal model for the marginal distribution of the counterfactual random variable $T_{\overline{a}}$ (Hernan et al. 2000). Furthermore, the parameter β_1 can be interpreted as the causal effect of the treatment, as measured via a log rate ratio; e_1^{β} can be interpreted as ratio of hazard of death at time t had all patients been treated to the hazard had they all been not treated.

We can obtain consistent estimates of β_1 in (3.1) by fitting the following model, which models the conditional hazard of death given treatment history and baseline covariates;

$$\lambda_T(t|\overline{A}_t, V) = \lambda_0(t)e^{\gamma_1 A_t + \gamma_2 V}. \tag{3.2}$$

In fitting this model, we use appropriate weights to account for the missing counterfactual treatment/outcome combinations.

3.1 Inverse probability of treatment weights for Cox models

Asymptotically unbiased estimates of the association parameter γ_1 in (3.2) can be obtained by maximizing the Cox partial likelihood. However, without appropriate weighting, this estimate will not be a consistent estimate of the causal effect of the treatment on mortality. We therefore fit the model in (3.1) using stabilized inverse probability of treatment weights,

$$sw_i(t) = \frac{\prod_{k=0}^{t-1} \mathbb{P}(A_k = a_{ki} | \overline{A}_{k-1} = \overline{a}_{i(k-1)})}{\prod_{k=0}^{t-1} \mathbb{P}(A_k = a_{ik} | \overline{A}_{(k-1)} = \overline{a}_{i(k-1)}, \overline{L}_k = \overline{l}_{ik})} \tag{3.3}$$

where $\overline{A}_{-1} = 0$.

Heuristically, weighting by $sw_i(t)$ effectively creates , for risk sets at time t, a pseudopopulation in which $L(t)$ is not associated with $A(t)$, and the causal association between $A(t)$ and mortality is the same as in the study population Hernan et al. (2000). Then as argued in Robins et al. (2000) – and discussed in Chapter 2 – $\hat{\gamma_1}$ obtained by fitting (3.2) weighted by $sw_i(t)$ will consistently estimate the causal parameter of interest, β_1 in (3.1).

3.1.1 Censoring

Now suppose that some patients are lost to follow up during the study, prior to death. That is, these patients were right censored. Let $C_i(t) = 1$ if the patient i was censored at time t, and zero otherwise. Similar to the procedure described in Chapter 2, we can consistently estimate β_1 in (3.1) by fitting the model in (3.2), weighted by $sw_i^*(t) = sw_i(t) \times sw_i^c(t)$, where

$$sw_i^c(t) = \frac{\prod_{k=0}^{t} \mathbb{P}(C_k = 0 | \overline{C}_{k-1} = 0, \overline{A}_{k-1} = \overline{a}_{i(k-1)})}{\prod_{k=0}^{t} \mathbb{P}(C_k = 0 | \overline{A}_{(k-1)} = \overline{a}_{i(k-1)}, \overline{L}_k = \overline{l}_{ik})} \tag{3.4}$$

where $\overline{A}_{-1} = 0$.

3.1.2 Estimating stabilized IPT weights

Similar to the discussion in Chapter 2, when the outcome of interest is binary, each term in the denominator and numerator of the stabilized weights in (3.3) is estimated by modeling each of these conditional probabilities using a logistic regression. For each $k \in 0, 1, \ldots t$, we model the conditional probability, $pr[A_k = 1|\overline{A}_{k-1} = \overline{a}_{k-1}, \overline{L}_{(k-1)} = \overline{l}_{(k-1)}]$ using

$$\text{logit } pr[A_k = 1|\overline{A}_{(k-1)} = 0, \overline{L}_{(k-1)}] = \alpha_0 + \alpha_1 t + \alpha_2 a_{t-1} + \alpha_3 l_(t-1) + \alpha_4 V \tag{3.5}$$

We will fit another logistic model to obtain the probabilities in the numerator,

$$\text{logit } pr[A_k = 1|\overline{A}_{(k-1)} = \overline{a}_{(k-1)}] = \alpha'_0 + \alpha'_1 t + \alpha'_2 a_{t-1} + \alpha_3 V. \tag{3.6}$$

For each subject i, using the estimates of the parameters in the model in (3.5), we can obtain the maximum likelihood estimates for $pr[A_k = 1|\overline{A}_{k-1} = \overline{a}_{i(k-1)}, \overline{L}_{i(k-1)} = \overline{l}_{i(k-1)}]$, $\hat{p}_{i0}, \hat{p}_{i1}, \hat{p}_{i2}, \ldots \hat{p}_{ik}$. Similarly we can also obtain the corresponding maximum likelihood estimates $\hat{p}'_{i0}, \hat{p}'_{i1}, \hat{p}'_{i2}, \ldots \hat{p}'_{ik}$ for the probabilities in the numerator of the stabilized weights by fitting the model in (3.6). Using these estimates, we can obtain an estimate of the stabilized weights,

$$\widehat{sw_i(t)} = \frac{\prod_{k=0}^{t} (\hat{p}'_{ik})^{a_{ik}} (1 - \hat{p}'_{ik})^{1-a_{ik}}}{\prod_{k=0}^{t} (\hat{p}_{ik}^{a_{ik}} (1 - \hat{p}_{ik})^{1-a_{ik}}}. \tag{3.7}$$

We can estimate the censoring weights $sw_i^c(t)$ in a similar manner to the treatment weights, $sw_i(t)$, with $A(t)$ replaced by $C(t)$, conditioned on $C(t-1) = 0$, instead of $A(t-1) = 0$ and adding $A(t-1)$ as a covariate.

3.2 Fitting the model - multiple options

The subject specific weights, $sw_i^*(t)$ vary over time. Until recently, most standard software (e.g. Stata, SAS) did not allow the weights used to fit weighted Cox proportional hazards models to be time-varying. As such, Hernan et al. (2000) suggests fitting a pooled logistic regression model to estimate β_1 in (3.1). Specifically, one can obtain consistent estimates of β_1 by fitting a weighted logistic regression model,

$$\text{logit } \mathbb{P}[D(t) = 1 | D(t-1) = 0, \overline{A(t-1)}, V] = \beta_0' + \beta_1 A(t-1) + \beta_2 V, \qquad (3.8)$$

weighted by the product of the stable weights estimated in (3.7) and the stable censoring weights.

Another approach is to fit the Cox proptional hazards model in (3.2) weighted by the estimates of the stable weights $sw_i^*(t)$. Fitting such a weighted model is possible using the `coxph()` function in the package `survival` in R. In a simulation study Xiao et al. (2010) have demonstrated that using weighted time-dependent Cox proportional hazards models to estimate the causal parameter of interest in an MSCM, rather than pooled logistic regression models, leads to estimates with decreased bias and increased efficiency.

3.2.1 Simulation results from Xiao et al. (2010)

In their article, Xiao et al. (2010) performed various simulation studies to evaluate the performance of different approaches to fitting an MSCM using IPTW. We replicated the subset of their results that are relevant to our discussion, and will summarize them below.

Following Xiao et al, we simulated a data set of N HIV-positive patients, to assess the effect of highly active anti-retroviral therapy. In the data set, we also simulated a time

varying confounder, CD4 cell count, which acted both as a time varying confounder and an intermediate variable.

Data generation

Patients were assumed to have a study visit every 6 months from the start of follow up. Let T_i denote the time from start of follow up to an AIDS defining event, and $A_i(t)$ the treatment status (1- treated, 0- not treated) and $L_i(t)$ the CD4 levels at time point t for patient i. For patients $i = 1 \dots N$, covariates were generated using the following procedure.

1. $L_i(1) \sim$ *lognormal*$(6, 1)$
2. For $t = 1, \dots m$, logit $\mathbb{P}[A_i(t)|A_i(t-1), L_i(j)] = 3.623 - 2.605I[L_i(t) > 500] - 0.002[L_i(t-200) + 0.009(L_i(t)-200)I[L_i(t) > 500] + 0.405A_i(t-1)$, with $A_i(0) = 0$.
3. For $t = 2, \dots m$, $L_i t(t) = L_i(t-1) + 70A_i(t-1) + \Delta_i + \epsilon_i(t)$, where $\Delta_i \sim$ *uniform*$(-80, -5)$ and $\epsilon_i(t) \sim N(0; 3^2)$.

The interval survival time t^*, time from the beginning of an interval between two consecutive visits, was generated by inverting the hazard,

$$\lambda_{i,t}[t^*|A_i(t), L_i(t)] = 0.12e^{-0.6931L_i(t) - 0.0016A_i(t)}$$

For each patient, all t^* values greater than 0.5 were interpreted as the patient being alive at the end of the 6 month period. As such, $T_i = 0.5(k-1) + t_i^*(k)$ where k is the earliest period in which $t_i^* \leq 0.5$. Also, censoring covariate, C_i was assumed to follow a $U(0, 40)$ distribution, and the observed follow-up time for patient i were defined as $t_i = \min(T_i, C_i, 5)$. Data was generated for 2000 independent samples, with $N = 2000$, with a maximum follow-up time of $m = 10$.

Each simulated sample was analyzed using four different methods. We fitted, an unweighted Cox model, adjusting for $A_i(t), A_i(t-1)$ and $L_i(t)$. In addition, we also fitted an MSCM to this data with weights calculated to account for confounding by $L_i(t)$ as discussed in the previous section. We used an MSCM in which the hazard for death was modeled as $\lambda_i(t) = \lambda_0(t)e^{\beta_1 L_i(1)+\beta_2 A_i(2)+\beta_3 A_i t(t-1)}$. We fitted this model in three different ways: as a weighted time-dependent Cox model with unstabilized weights as well as with stabilized weights, and as a pooled logistic regression model with stabilized weights.

The biases of the estimates of β_2 and β_3 were estimated as the mean difference between the 2000 estimates from the 2000 samples, and the true values of these parameters. In addition we also calculated the empirical standard deviations of the 2000 point estimates. The results are summarized in Table 3.1.

Table 3.1: Comparison of different methods to fit an MSCM

	β_2 (true:=-0.693)		β_3 (true:=-0.112)	
	Bias	SD	Bias	SD
Unweighted Cox model	0.134	0.091	0.047	0.101
Weighted Cox model (non stable weights)	-0.003	0.755	0.126	0.785
Weighted Cox model (stable weights)	-0.009	0.504	-0.013	0.543
Pooled logistic model (stable weights)	-0.023	0.711	0.042	0.755

All MSCMs provide estimates with small bias compared to that of the unweighted Cox model. In addition, the weighted Cox model with stable weights appear to be notably more efficient than the other methods investigated.

Chapter 4

Comparative Effectiveness of Two Hemodialysis Therapies

4.1 Background

Over the last 30 years, the risk of death or hospitalizations for patients undergoing maintenance dialysis in the United States has declined significantly; yet, challenges remain. The median life expectancy of patients starting renal replacement therapy is only a little over three years and they spend on average 12 days in the hospital annually. The overwhelming majority of patients are treated with thrice-weekly in-center hemodialysis (TWICHD) and most of the rest perform peritoneal dialysis (PD) treatments at home. However, an increasingly larger number of patients are being treated with modified hemodialysis (HD) regimens that include significantly longer treatment times (nocturnal), or different frequency (2-6 times/week), or alternative platforms (viz., NxStage System One).

Treatment with these alternative dialysis regimens significantly alters and in many instances, increases the burden of treatment on patients. It is critically important to perform a rigorous assessment of the true nature of benefit, if any, of these alternative HD regi-

mens on patient-centered outcomes. Illustrating the importance of this issue, the Institute of Medicine identified comparing the effectiveness of renal replacement therapies as the only kidney-disease related topic among the top 100 initial national priorities for comparative effectiveness research. Even though several studies have recently compared the outcomes of patients treated with these alternative HD regimens, most of these considered only differences in patient characteristics at the time of start of maintenance dialysis. We undertook this study to assess the association between being treated with nocturnal in-center hemodialysis, as opposed to TWICHD, and survival, accounting for differences in baseline and time-varying patient- and facility-level characteristics. We base our evaluation on a contemporary and large multi-ethnic incident cohort of dialysis patients in the United States.

Study Population and Data Source

This cohort comprised all patients that started maintenance dialysis in calendar years 2007-2011 and received treatment in one of the facilities operated by DaVita Inc. Patients < 18 years of age, or who did not receive treatment for at least 60 days were excluded and our study population is comprised of 162,671 individuals. All data were obtained from electronic records at DaVita.

The entire follow-up period for each patient was divided into quarters comprising successive 91 day periods from the date of first dialysis of that patient; follow-up was available for up to 20 quarters. Each patient was assigned one of six different dialysis modalities to each quarter - TWICHD, PD, Less frequent in-center hemodialysis (less Frequent HD; less than or equal to two times/week), Home hemodialysis (Home HD), Frequent in-center hemodialysis (Frequent HD: greater than three times/week), and nocturnal in-center hemodialysis (NICHD). Each patient was assigned to be treated with a

given modality if s/he was treated with that particular modality for at least 60 consecutive days. The modality assigned for any given quarter was the therapy with which the patient was treated for ≥ 45 days of the quarter.

Hemodynamic, other dialysis-related, and laboratory parameters were summarized for each quarter as arithmetic means. Similarly, summary values of each parenteral medication were computed for each quarter. The dialysis access with which the patient was treated for more than 45 days was assigned as the access for the quarter. Each patient was also assigned a dialysis facility with attribution to the facility where the patient received care for ≥ 45 days in a given patient quarter.

Patients who were treated at least one quarter with PD, Less Frequent HD, Home HD, Frequent HD, or NICHD were categorized as *ever PD*, *ever Less Frequent HD*, *ever Home HD*, *ever Frequent HD* and *ever NICHD* respectively. Patients who were only treated with TWICHD during follow up were grouped as *only TWICHD*. From this cohort we selected patients who were either in the *ever NICHD* or *only TWICHD* categories.

4.2 Descriptive Analyses

We calculated descriptive statistics for patients who were ever treated with NICHD and only treated with TWICHD (see Table 4.1). Of the 1452 patients ever treated with NICHD, 1340 entered the cohort within the first 91 days of start of dialysis; NICHD was the initial dialysis modality for 29%. Compared to individuals who treated only with TWICHD, NICHD patients were younger, more likely to be male, Hispanic, or Asian, and had insurance other than Medicare or Medicaid. They were more likely to have a history of kidney transplant, diabetes, hypertension, congestive heart failure, atherosclerotic heart disease, or dyslipidemia, and had higher body weight. In the first 91-day period from the initiation of dialysis, they were more likely to have been dialyzed with an arteriovenous

fistula, and had lower serum ferritin and phosphorous levels.

Table 4.1: Descriptive statistics of covariates comparing patients who were ever treated with NICHD to those who were only treated with TWICHD

	Ever Nocturnal HD (N=1134)	**Only Thrice-Weekly In-Center HD (N= 113,141)**
Interval from start of dialysis to start of modality, months	9 [0, 20]	0 [0, 0]
Demographics		
Age, years	51±13	62±15
Race, %		
White	3	47
Asian	37	3
Black	47	31
Hispanic	10	15
Other	3	4
% Male	70	57
Primary Health Insurance,%		
Medicare	38	55
Medicaid	6	7
Other insurance	55	39
Cause of ESRD, %		
Diabetes	44	46
Hypertension	27	29
Glomerular Disease	13	10
Other	16	15
h/o previous transplant, %	4	1
Comorbidities, %		
Diabetes	66	56
Hypertension	58	51
Congestive Heart Failure	54	35

	Ever Nocturnal HD (N=1134)	Only Thrice-Weekly In-Center HD (N= 113,141)
Atherosclerotic Heart Disease	20	14
Other Cardiovascular	19	15
Dyslipidemia	31	24
Hospitalized	27	31
Access Type		
AV Fistula	66	15
Central Venous Catheter	20	74
AV Graft	3	4
PD Catheter	3	0
Unknown	8	7
Treatment Variables		
Length of treatment with In-center hemodialysis, minutes (per 30 minutes)	276±84	211±24
Pre-dialysis systolic BP, mm Hg (per 10 mm Hg)	152±19	147±19
Pre-dialysis diastolic BP, mm Hg (per 10 mm Hg)	83±12	77±12
Max BP Change per treatment (per 5 mm Hg)	31±12	30±12
Body Weight greater than 100kg, %	53	32
Weight Change per treatment	2.6±1.1	2.0±0.9
Weekday IDWG , % (per 1 %)	2.2 [1.5, 3]	2.3 [1.5, 3.2]
Weekend IDWG, %	3.0 [2, 4]	3.1 [2.1, 4.2]
Lab Variables		
Hemoglobin, g/dL (per 1g/dL)	11.1±1.2	11.1±1.2
Iron Saturation, %	21 [17, 25]	22 [17, 27]
Serum Ferritin, ng/mL	235 [139, 373]	283 [164, 486]
Serum Albumin, g/dL (per 1g/dL)	3.6±0.5	3.5±0.5
spKt/V (per 0.1 units)	1.4 [1.2, 1.8]	1.4 [1.3, 1.6]

	Ever Nocturnal HD (N=1134)	Only Thrice-Weekly In-Center HD (N= 113,141)
Serum Calcium, mg/dL (per 1mg/dL)	9.1±0.6	9.1±0.6
Serum Phosphorous, mg/dL (per 1mg/dL)	5.3±1.2	4.9±1.2
Parathyroid hormone, pg/mL (per 100pg/mL)	377 [242, 568]	314 [197, 486]
Alkaline phosphatase, u/L	87 [69, 111]	87 [69, 115]
Hemoglobin A1C, %	6.6±1.3	6.4±1.2
IV medication		
Cumulative Iron, mg (per 100mg)	900 [0, 1400]	800 [0, 1300]
ESA Median Week Dose (per 1000 units)	4401 [1524, 11917]	4694 [1496, 11943]
Geographic Location		
Northeast, %	9	13
Midwest, %	30	18
West, %	21	25
South, %	40	44
Year of Incidence		
2007	27	21
2008	25	21
2009	23	21
2010	17	20
2011	8	16
Median First Time on Modality, months*	9	25
Median Total Time on Modality, months*	11	25
Expired prior to transferring to any other modality, %	6	24
Expired, %	11	24
Received transplant prior to transferring to any other modality, %	3	2
Transplant, %	4	2

In addition, to compare characteristics that might be related to transferring to an alternative modality from TWICHD, we used a nested case-control design that matched on treatment history and the year of incidence. Specifically, we 1:1 matched patients that transferred from TWICHD into NICHD (cases) with patients who continued treatment with TWICHD up to the quarter of transfer (controls), and the year of incidence. We then caalculated descriptive statistics for relevant covariates for patients who transferred, and their associated matched controls (see Table 4.2). After a median treatment of 14 months with TWICHD, 827 patients transferred to NICHD; 30% of these individuals transferred to another dialysis facility at the same time as the change in dialysis modality. In the 91-day period immediately preceding the transfer, patients who transferred to NICHD were more likely to be hospitalized, compared to matched controls.

Table 4.2: Comparison of patients who transferred to NICHD from TWICHD to those remained on TWICHD

	Transferred to NICHD (N=827)	Controls (N=827)
Interval from initiation of dialysis to switch to alternate modality, months	14 [6, 24]	Not applicable
Transferred to another facility when modality changed, %	30	2
Demographics		
Age, years	51$\pm$13	61$\pm$15
Race, %		
White	3	4
Asian	38	31
Black	45	45
Hispanic	11	16
Other	2	4
% Male	71	58
Primary Health Insurance,%		
Medicare	41	55
Medicaid	8	7

	Transferred to NICHD (N=827)	Controls (N=827)
Other insurance	51	38
Cause of ESRD, %		
Diabetes	45	48
Hypertension	28	29
Glomerular Disease	13	10
Other	14	13
h/o previous transplant, %	4	1
Comorbidities, %		
Diabetes	68	66
Hypertension	57	53
Congestive Heart Failure	60	41
Atherosclerotic Heart Disease	22	18
Other Cardiovascular		19
Dyslipidemia		30
Hospitalized		23
Access Type		
AV Fistula	52	46
Central Venous Catheter	34	39
AV Graft	10	11
PD Catheter	Not applicable	Not applicable
Unknown	4	4
Treatment Variables		
Length of treatment with In-center hemodialysis, minutes (per 30 minutes)	269±61	212±25
Pre-dialysis systolic BP, mm Hg (per 10 mm Hg)	152±19	149±19
Pre-dialysis diastolic BP, mm Hg (per 10 mm Hg)	82±12	78±12
Max BP Change per treatment (per 5 mm Hg)	34±13	33±14
Body Weight greater than 100kg, %	44	18
Weight Change per treatment	3.1±1.3	2.4±1

	Transferred to NICHD (N=827)	Controls (N=827)
Weekday IDWG , % (per 1 %)	2.8[2.1,3.8]	2.8[1.9,3.7]
Weekend IDWG, %	3.9[2.7,5]	3.7[2.6,4.9]
Lab Variables		
Hemoglobin, g/dL (per 1g/dL)	11.4±1	11.5±1.1
Iron Saturation, %	25.8[20.3,32.7]	26.3[20.5,33.6]
Serum Ferritin, ng/mL	447.3[257.1,685.5]	422[223,691]
Serum Albumin, g/dL (per 1g/dL)	3.9±0.4	3.7±0.4
spKt/V (per 0.1 units)	1.4±1.6	1.4±1.5
Serum Calcium, mg/dL (per 1mg/dL)	9±0.5	9.1±0.6
Serum Phosphorous, mg/dL (per 1mg/dL)	5.5±1.3	5.2±1.3
Parathyroid hormone, pg/mL (per 100pg/mL)	317[220,501]	267.1[191.3,399.7]
Alkaline phosphatase, u/L	85.2[67,111.7]	89[68.3,117]
Hemoglobin A1C, %	7±1.8	6.6±1.4
IV medication		
Cumulative Iron, mg (per 100mg)	650[250,1100]	400[0,1000]
ESA Median Week Dose (per 1000 units)	10060.5 [4125,19272]	10000 [4293,20167]
Geographic Location		
Northeast, %	8	14
Midwest, %	38	44
West, %	29	18
South, %	25	24
Year of Incidence		
2007	33	33
2008	27	27
2009	22	22
2010	12	12
2011	5	5

	Transferred to NICHD (N=827)	Controls (N=827)
Transfer to thrice-weekly in-center HD, %	27	Not applicable
Transfer to another modality, %	3	Not applicable
Expired prior to transferring to any other modality, %	6	Not applicable
Expired, %	12	23
Received transplant prior to transferring to any other modality, %	2	Not applicable
Transplant, %	3	2

4.2.1 Missing Values

At baseline, data for spKt/V were missing 8%, pre dialysis body weight were missing 6%, hemoglobin, hematocrit, iron saturation ratio, ferritin, albumin, calcium, phosphorous, parathyroid hormone and alkaline phosphatase were missing 1-2%. We imputed these covariates using multiple imputation. Specifically, we used the chained equations method, and used a linear regression model with variables representing each of these covariates for the twenty patient quarters against age, gender, race and select comorbidities. To characterize the patients with missing values in potential time varying confounders, we compared descriptive statistics for baseline covariates among those with complete data for the time varying covariates through follow-up, and among those who have a missing value in any one of the covariates at any point during the follow-up (see Table 4.3).

Compared to those have complete data, patients who have missing values appear to be younger and more likely to have Medicare as their primary insurance. They are less likely to be diabetic, have CHF, and are at less risk of death during follow-up. Moreover, the difference in risk of death between those who have complete data and missing values, is more pronounced among patients who are treated with only TWICHD, compared to

those who are ever treated with NICHD.

Table 4.3: Characteristics of patients with missing values in select time varying covariates

	Ever NICHD		Only HD	
	Complete Case (N=1040)	**Ever Missing (410)**	**Complete Case (N=104548)**	**Ever Missing (N=29892)**
Age	52+/-13	49 +/-14	63 +/- 15	60 +/- 15
Male, %	70	72	57	58
Race				
White	46	48	46	49
Black	37	36	32	31
Hispanic	11	10	15	14
Asian	3	3	3	3
Other	3	3	4	3
Insurance				
Medicare	36	45	53	59
Medicaid	7	5	7	5
Other	57	50	40	36
ESRD Reason				
Diabetes	47	36	46	45
Hypertension	26	30	30	28
GN	13	15	9	11
Other	14	19	15	16
Comorbidities				
Diabetes	69	59	58	48
Hypertension	56	61	51	50
CHF	56	49	37	29
ASHD	20	21	14	13
Dyslipidemia	30	34	25	21

	Ever NICHD		Only HD	
	Complete Case (N=1040)	**Ever Missing (410)**	**Complete Case (N=104548)**	**Ever Missing (N=29892)**
Other Cardio Vascular	19	17	15	14
Facility Regions				
North East	9	6	13	12
West	20	26	25	24
Midwest	31	26	18	17
South	40	42	44	47
Morality rate, per 1000 person years (using incident patients in 2007 or 2008)	11	9	37	21

4.3 Primary Analysis

Our primary analysis was to assess survival benefit, if any, that NICHD provides to patients compared to being treated with TWICHD alone. As such, we fitted marginal structural Cox models, with various levels of control for time varying and baseline confounding. We fitted five different models.

1. Model 1 was a conventional unweighted time dependent Cox model, with the outcome variable as time to death, regressing on current treatment status (1 - NICHD; 0 - TWICHD). Specifically, we fitted the model $\lambda_T(t|A(t)) = \lambda_0(t)e^{\beta_1 A(t)}$ where $\lambda_T(t|A(t))$ is the conditional hazard of death given treatment history $A(t)$.

2. Model 2 was also a conventional unweighted time dependent Cox model, with the same covariates as Model 1, but adjusted further for baseline covariates. We adjusted for baseline age, gender, race, primary insurance, diabetes, hypertension, dyslipidemia, other cardiovascular diseases, CHF, ASHD and year of incidence.

3. Model 3 was an MSCM, with the same covariates as Model 2, but accounting for

a limited set of time varying covariates (pre dialysis BMI, facility regions, cvc days and vascular accesstype) using stable weights.

We calculated the stable weights in the following manner. First, we created an augmented data set, in which each row corresponded for a patient quarter. Then we used the two logistic regressions models to calculate the denominator and the numerator of the stable weights, as described in Chapter 3 (3.7). For the denominator, we regressed treatment status for current patient quarter on the treatment status, pre dialysis BMI, facility regions, CVC days and vascular accesstype from the quarter preceding the current one, as well as baseline covariates - age, gender, race and cause of ESRD. For the numerator, we used a simpler logistic regression model, where we regressed the current treatment status on treatment status from the preceding quarter, and baseline covariates - age, gender, race and cause of ESRD. We used similar models to calculate the censoring weights, except that the outcome variable was censoring status of the current quarter (1- censored; 0 not censored) and had censoring status in the preceding quarter as an additional covariate in the models.

4. Model 4 was an MSCM as well. In addition to the covariates used in Model 3, we accounted for confounding by a more extensive list of covariates using stable weights. These covariates were SBP, BMI, Weekday IDWG, Weekend IDWG, hemoglobin, Iron Saturation, Ferritin, Albumin, spKt/V, Calcium, Phosphorous, PTH, Alkaline Phosphorous, Cumulative Iron Dose, Median ESA Dose, facility regions, CVC days and access type. We estimated stable IPTW and censoring weights using the same procedure in Model 3, except that we included all the time varying covariates above in the logistic regression model used to calculate the denominator of the weights.

5. The only difference between Models 4 and 5 was the data set used. For Model

4, we restricted to a complete case data set. That is, we excluded patients for whom data were missing for any of the covariates in any of the patient quarters of follow up. We fitted Model 5 to a larger data set in which all time varying covariates used in the model, except for vascular access type, were imputed using a multiple imputation model. We created 5 imputed copies of the data set, fitted the MSCM to each of the data set, and then combined the estimates using Rubin's rules. The distribution of stable weights used to fit Model 5, combining treatment weights and censoring weights, is shown in Figure 4.1. The distribution of stable weights has a mean around 1 at all time points, and its variance increases over time.

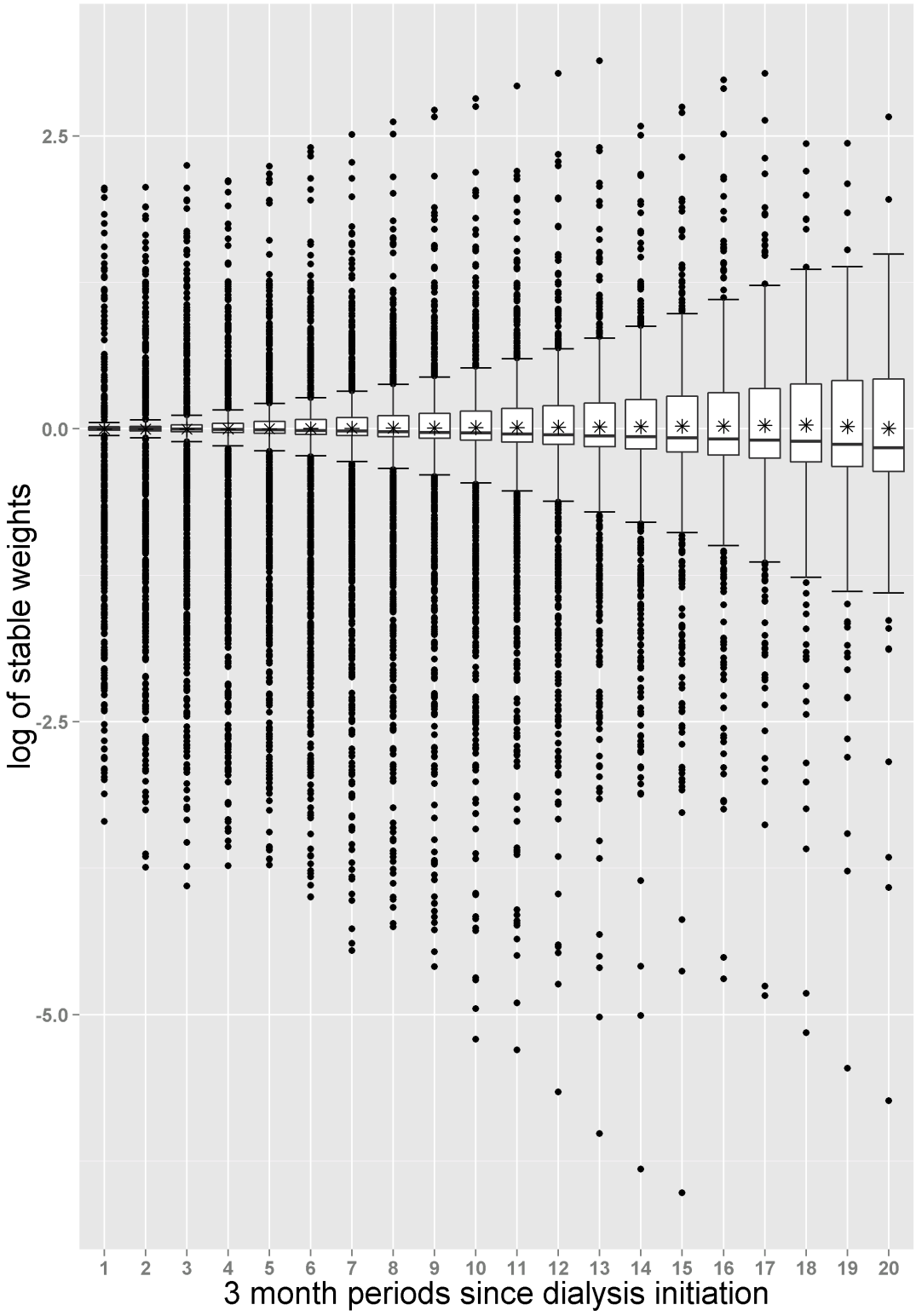
4.3.1 Results

Using a conventional unweighted Cox proprtional hazards model, we estimated the crude hazard ratio for all cause mortality, comparing NICHD to TWICHD as 0.39 (0.30,0.51). When adjusted for baseline covariates, the effect attenuated to a hazard ratio of 0.46 (0.33,0.71). We accounted for a limited set of time varying confounders, and baseline confounders using an MSCM (Model 3). The estimated causal hazard ratio, comparing patients who were treated with NICHD through their follow up, to those who were treated with TWICHD through their follow up (contrary to their observed treatment status) is 0.54 (0.34,0.77). When we account for even more potential time varying confounders in the MSCM (Model 4), and fit it to a complete case data set, this estimate of the causal hazard ratio attenuates to 0.58 (0.40,0.83). Fitting the same MSCM as in Model 4 to a data set in which the time varying coovariates were imputed using multiple imputation, our estimate of the causal hazard ratio of all cause mortality for NICHD is 0.57 (0.39, 0.82). Assuming that the relevant modeling assumptions are not grossly violated, these results suggest that patients treated with NICHD have better survival compared to those treated with TWICHD.

Table 4.4: Causal association between being treated with NICHD and mortality

	N (Ever NICHD)	Hazard Ratio	95% CI	P value
Model 1. Time dependent Cox Model, unadjusted	1448	0.39	(0.30, 0.51)	< 0.01
Model 2. Time dependent Cox Model, adjusted for baseline covariates	1448	0.46	(0.33, 0.71)	< 0.01
Model 3. MSCM, adjusted for limited covariates	1448	0.54	(0.34, 0.77)	< 0.01
Model 4. MSCM, fully adjusted	880	0.58	(0.40, 0.83)	< 0.01
Model 5. MSCM with multiply imputed data, fully adjusted	1418	0.61	(0.39, 0.82)	< 0.01

Figure 4.1: Distribution of stabilized weights sw_i^*. For each time point, the location of the mean is indicated with a *, median is the horizontal line at the middle of the box, quartiles are the upper and lower ends of the box, and the vertical lines extend to $1.5\times$ IQR.



Chapter 5

Sensitvity Analyses & Discussion of the results

In this chapter we will investigate some of the assumptions made in Chapter 4, and how sensitive the results are to those assumptions. Specifically, we will explore some variables that we did not consider in our main analysis, but that could potentially introduce biases int the point estimates from our primary result. We will also assess the validity of the precision of our estimates using a bootstrap procedure.

5.1 Potential survivor bias and attribution of deaths

In our cohort, the median interval from initiation of dialysis to initiation of NICHD is 9 months. So it is possible that a patient has to survive being treated with TWICHD for some time before they are treated with NICHD. As such our results could potentially be affected by survivor bias (Zhou et al. 2005).

Furthermore, it is possible that the effect of dialysis modality on death is lagged. That is, for a dialysis modality to have adverse effects on a patients health, they may have to

be treated with that modality for a certain period of time. In the literature, such effects are investigated by attributing the deaths of patients to a dialysis modality that they were treated with thirty and sixty days prior to death, rather than the modality they were on immediately prior to death (McDonald & Russ 2002, Afolalu et al. 2009, Johnson et al. 2009).

To investigate the effect of survivor bias and potential lagged effect of therapy on our results, we performed a series of analyses.

5.1.1 Hazard of death among those who survived the first 90 days of dialysis

Insurance coverage for the patients who are eligible but not receiving Medicare benefits prior to the first day of dialysis treatment begins on day 90. As such, the quality of the data available in the first 90 days after dialysis initiation tends to be less complete. Furthermore, it is well known that patients are at much higher risk of death in the first 90 days of dialysis as opposed to later periods of follow up. Our interest is to assess the differences in long term survival between NICHD and TWICHD. Hence, we performed an analysis, treating the dialysis modality 90 days after the first service date as be the initial modality. In other words, following the literature, for the survival analysis, we set time point 0 as the 90^{th} day after initiation of dialysis.

We fitted an MSCM to a data set where all patients were left truncated at 91 days after initiation of dialysis. As in Model 4 from previous chapter, we accounted for confounding by time varying covariates using stable weights. The covariates were SBP, BMI, Weekday IDWG, Weekend IDWG, hemoglobin, Iron Saturation, Ferritin, Albumin, spKt/V, Calcium, Phosphorous, PTH, Alkaline Phosphorous, Cumulative Iron Dose, Median ESA Dose, facility regions, CVC days and access type. Furthermore, in the marginal structural

Table 5.1: Sensitivity due to left truncation and lagged effect of therapy on the causal effect of being treated with NICHD on mortality

Deaths attributed to	Left truncated at	Hazard Ratio	95% CI	P value
Current modality	0 days	0.54	(0.39, 0.75)	<0.01
Current modality	90 days	0.58	(0.41, 0.84)	<0.01
Modality 30-days prior	0 days	0.53	(0.38, 0.74)	<0.01
Modality 30-days prior	90 days	0.58	(0.40, 0.83)	<0.01
Modality 60-days prior	0 days	0.51	(0.37, 0.72)	<0.01
Modality 60-days prior	90 days	0.58	(0.40, 0.82)	<0.01

The number of patients who were ever treated with NICHD in the data set used to fit the above models was 1450.

Cox model, we adjusted for age, sex, race, primary insurance, diabetes, hypertension, dyslipidemia, other cardiovascular diseases, CHF, ASHD and year of incidence.

The estimated putatitively-causal hazard ratio of death after 91 days of initiation of dialysis, comparing patients who were treated with NICHD through their follow up, to those who were treated with TWICHD through their follow up (contrary to their observed treatment status) is 0.58 (0.41,0.84) (see Table 5.1). This result is similar to the one we obtained when we did not left truncate patients at 90 days after dialysis initiation. Attributing the deaths to modalities 30 days or 60 days prior to deaths do not appear to have an important effect on the estimate of reduced risk of death among those treated with NICHD.

5.1.2 Hazard of death estimated from a nested case-control cohort

To assess the sensitivity of results to the length of time between initiation of dialysis and initiation of NICHD, we also fitted the same MSCM as in Chapter 4's Model 4 to a nested case control cohort, where each case was matched to a control based on treatment history and the year of incidence. Each patient in the cohort who transferred to NICHD

was matched to a patient who was only treated with TWICHD up to and including the patient quarter in which the transfer occured, and who had the same year of incidence. For the patients who started dialysis with NICHD, the matched control was a patient who was only treated with TWICHD, started dialysis with TWICHD in the first quarter, and had the same year of incidence. The total sample size for this cohort was 2078; 1039 cases and 1039 individually-matched controls.

As in Model 4 from Chapter 4, we accounted for confounding by time varying covariates using stable weights. These covariates were SBP, BMI, Weekday IDWG, Weekend IDWG, hemoglobin, Iron Saturation, Ferritin, Albumin, spKt/V, Calcium, Phosphorous, PTH, Alkaline Phosphorous, Cumulative Iron Dose, Median ESA Dose, facility regions, cvc days and access type. Furthermore, in the marginal structural Cox model, we adjusted for age, sex, race, primary insurance, diabetes, hypertension, dyslipidemia, other cardiovascular diseases, CHF, ASHD and year of incidence. The models fitted were weighted Cox proportional hazards models, stratified on the matched pairs. When calculating the robust standard errors, we clustered on patient IDs to account for the different periods at risk coming from the same patient.

The estimated hazard ratio for death, comparing patients who were treated with NICHD through their follow up to those who were treated with TWICHD through their follow up (contrary to their observed treatment status) is 0.78(0.50,1.21); see Table 5.2. The estimate of the causal effect is also similar when we left-truncated the patients at 90 days post dialysis initiation or changed the attribution of the deaths to modalities 30 or 60 days prior to death.

In this cohort the patients who were only treated with TWICHD were the ones who had survived just as long as their matches who transferred to NICHD. In other words the comparators chosen were potentially healthier, and the ones who lived longer, among those who were treated only with TWICHD. This might be the reason why the estimated

Table 5.2: Sensitivity of the causal effect of NICHD on mortality to treatment history

Modality deaths are attributed	Left truncated	Hazard Ratio	95% CI	P value
Current	0 days	0.78	(0.50, 1.21)	0.27
Current	90 days	0.77	(0.49, 1.19)	0.24
30-days prior	0 days	0.76	(0.48, 1.19)	0.23
30-days prior	90 days	0.77	(0.49, 1.19)	0.24
60-days prior	0 days	0.71	(0.45, 1.12)	0.14
60-days prior	90 days	0.77	(0.49, 1.19)	0.24

The number of patients who were ever treated with NICHD in the data set used to fit the above models was 1039.

lower risk of death is attenuated in this cohort compared to our primary analysis. The estimates of the effect obtained here are in the same direction as in the primary analysis – lower risk of death among those who treated with NICHD - suggesting that our primary results are robust to this concern. However, not accounting for survivor bias in an effective way is potentially a weakness of our primary analysis.

5.2 Facility level effects

Over the five-year period, the study cohort was treated in 2217 facilities in 45 states. Of these facilities, 183 offered NICHD. Among the patients who were ever treated with NICHD, 30% transferred to another dialysis facility at the same time as the change in dialysis modality, whereas among those who were only treated with TWICHD this number was 2%. We did two sets of analyses to assess the impact on our results, if any, of the effect of facilities offering NICHD. First we fitted the same MSCM as in Model 4 of Chapter 4, to a data set with only those patients who were ever treated in facilities that offered NICHD. The sample size for this cohort (Facility Cohort I) was 18614. Second, we fitted

the same MSCM to a data set of patients who were treated in the 183 facilities that offered NICHD as well as those patients who were treated in a facility from which the patients transferred to facilities that offered NICHD. The sample size for this cohort (Facility Cohort II) was 106,250. The second analysis was performed to assess the effects of including the patients who could have transferred to another facility offering NICHD but did not.

Table 5.3: Sensitivity of the causal effect of NICHD on mortality to being treated in a dialysis facility offering NICHD

		Facility Cohort I (N=18,614)			**Facility Cohort II (N=106,250)**		
Modality deaths are attributed	Left truncated	Hazard Ratio	95% CI	P value	Hazard Ratio	95% CI	P value
Current	0 days	0.59	(0.43, 0.81)	<0.01	0.50	(0.33, 0.75)	<0.01
Current	90 days	0.54	(0.36, 0.82)	<0.01	0.53	(0.35, 0.79)	<0.01
30-days prior	0 days	0.58	(0.42,0.80)	<0.01	0.56	(0.38, 0.83)	<0.01
30-days prior	90 days	0.58	(0.38, 0.87)	<0.01	0.56	(0.37, 0.84)	<0.01
60-days prior	0 days	0.56	(0.40, 0.78)	<0.01	0.54	(0.37, 0.82)	<0.01
60-days prior	90 days	0.58	(0.39, 0.88)	<0.01	0.56	(0.38, 0.83)	<0.01

The estimates of the putatively-causal hazard ratio of death comparing being treated with NICHD through follow-up vs being treated with TWICHD through follow-up in these restricted cohorts (see Table 5.3) are further away from the null than in our primary analysis. In our primary analysis we had estimated that those who were treated with NICHD had a 39% lower risk of death, whereas the reduction in risk of death is estimated to be 41% and 50% in facility cohorts I and II respectively. Furthermore, the confidence intervals for the hazard ratio for death in these restricted cohorts are slightly narrower than in the primary analysis. However, the causal effects obtained here are similar to the estimates we obtained in our primary analysis, suggesting that our main result is not

sensitive to the effects of dialysis facilities.

5.3 Precision of the point estimate of the primary analysis

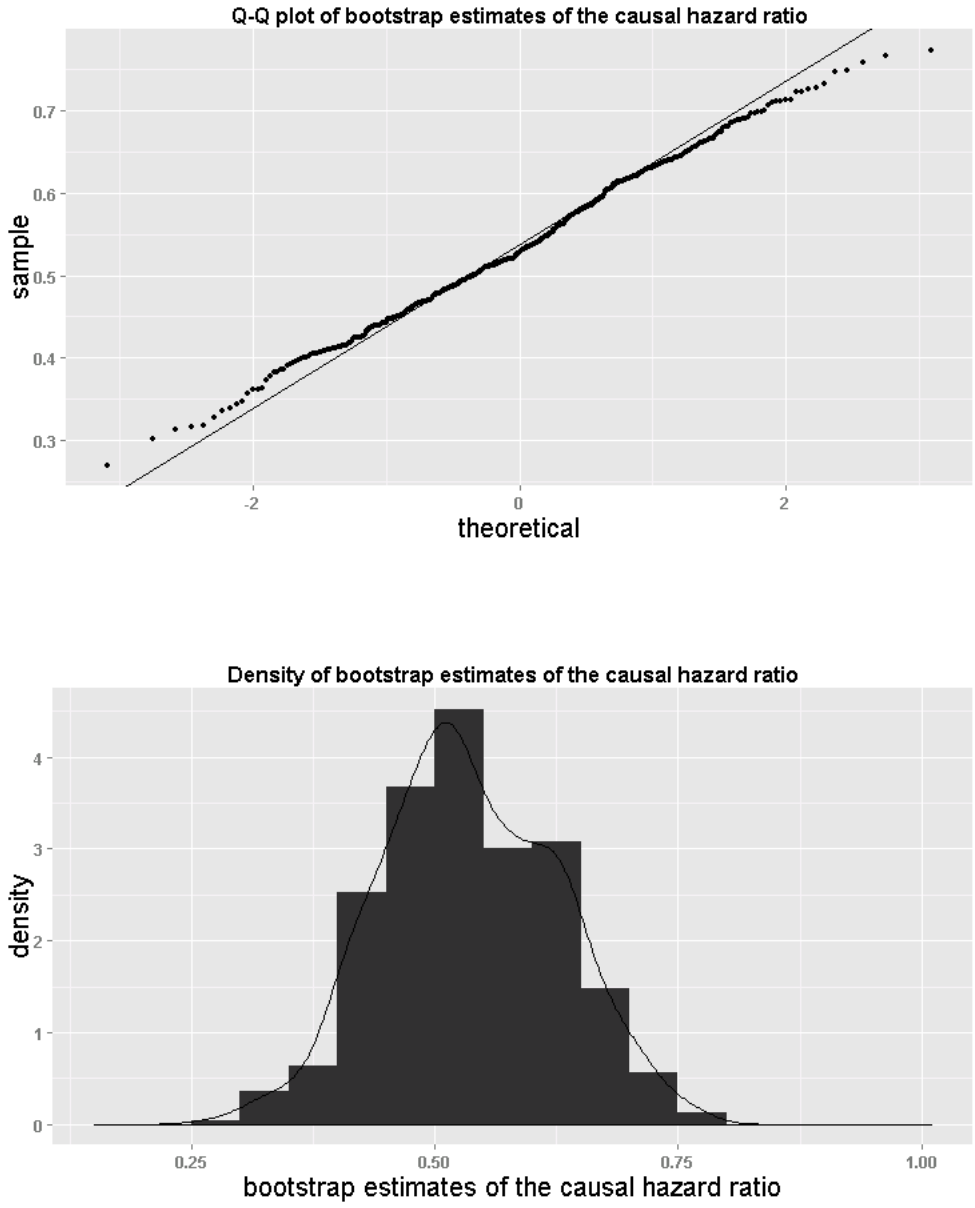
To assess the validity of the precision of our estimates of the causal hazard ratio, and in particular sensitivity to the parametric parts of the assumptions made therein, we performed a sensitivity analysis by calculating non-parametric bootstrap standard errors instead. Specifically we generated 1000 samples, by sampling with replacement from an imputed dataset of our cohort of 135, 890 patients. We then fitted the MSCM described in Chapter 4 (*Model 5* in the section *Primary Analysis*) to each of these data sets and calculated the mean and standard deviation of the 1000 estimates.

The bootstrap estimates of the log causal hazard ratio of death for being treated with NICHD appears to be normally distributed (see Figure 5.1). The bootstrap estimate of the standard error of our estimate of the log hazard ratio was 0.27 whereas the model based estimate was 0.21. This suggests an 65% gain in precision in terms of the standard error that is coming from the parametric model. However, in both estimates, the confidence intervals do not overlap zero. Hence our results from the primary analysis are not importantly sensitive to these parametric assumptions, although the large differences in the standard error estimates do indicate differences between the approaches.

5.4 Discussion of the results

Longer treatment times with TWICHD are associated with better survival (Saran et al. 2006, Miller et al. 2010, Troidle et al. 2007, Marshall et al. 2006). Treatments times with NICHD are considerably longer than the ones tested in these studies of conventional

Figure 5.1: Normality of bootstrap estimates



TWICHD. Whether these considerably longer treatment times with NICHD further lower death risk remains untested. All but one report of NICHD are small, single center studies with sample sizes from n=13 to 146.(Lacson et al. 2010, Powell et al. 2009, Bugeja et al. 2009, David et al. 2009).

The only observational study of "hard" outcomes with NICHD (n=655) showed lower hospitalizations but was did not show a significant lower death risk compared to TWICHD (Lacson et al. 2010). However, in that study, control for bias was limited. Specifically, the authors did not account for any potential confounders that were time varying. Furthermore, in that study, patients who were ever treated with NICHD were considered to remain on the therapy from initiation with NICHD to the end of follow-up. This does not appear to be the case in practice. For example, in our cohort, 30% of the patients who were treated with NICHD transfer to another dialysis modality after initiation of treatment with NICHD (see Chapter 4, Table 4.2).

Our analyses adds to the body of evidence comparing patients treated with NICHD to those treated with TWICHD within the United States. Our results are consistent with the literature in that our data does indicates survival benefit for NICHD over TWICHD.

Bibliography

Afolalu, B., Troidle, L., Osayimwen, O., Bhargava, J., Kitsen, J. & Finkelstein, F. O. (2009), 'Technique failure and center size in a large cohort of peritoneal dialysis patients in a defined geographic area', *Perit Dial Int* **29**(3), 292–296.

Berkson, J. (1946), 'Limitations of the application of fourfold table analysis to hospital data', *Biometrics* **2**(3), 47–53.

Bugeja, A., Dacouris, N., Thomas, A., Marticorena, R., McFarlane, P., Donnelly, S. & Goldstein, M. (2009), 'In-center nocturnal hemodialysis: another option in the management of chronic kidney disease', *Clin J Am Soc Nephrol* **4**(4), 778–783.

David, S., Kumpers, P., Eisenbach, G. M., Haller, H. & Kielstein, J. T. (2009), 'Prospective evaluation of an in-centre conversion from conventional haemodialysis to an intensified nocturnal strategy', *Nephrol. Dial. Transplant.* **24**(7), 2232–2240.

Greenland, S., Pearl, J. & Robins, J. M. (1999), 'Causal diagrams for epidemiologic research', *Epidemiology* **10**(1), 37–48.

Hernan, M. A., Brumback, B. & Robins, J. M. (2000), 'Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men', *Epidemiology* **11**(5), 561–570.

Hernan, M. A. & Robins, J. M. (2013), 'Causal Inference '.
URL: *http://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/*

Johnson, D. W., Dent, H., Hawley, C. M., McDonald, S. P., Rosman, J. B., Brown, F. G., Bannister, K. M. & Wiggins, K. J. (2009), 'Associations of dialysis modality and infectious mortality in incident dialysis patients in Australia and New Zealand', *Am. J. Kidney Dis.* **53**(2), 290–297.

Lacson, E., Wang, W., Lester, K., Ofsthun, N., Lazarus, J. M. & Hakim, R. M. (2010), 'Outcomes associated with in-center nocturnal hemodialysis from a large multicenter program', *Clin J Am Soc Nephrol* **5**(2), 220–226.

Marshall, M. R., Byrne, B. G., Kerr, P. G. & McDonald, S. P. (2006), 'Associations of hemodialysis dose and session length with mortality risk in Australian and New Zealand patients', *Kidney Int.* **69**(7), 1229–1236.

McDonald, S. P. & Russ, G. R. (2002), 'Survival of recipients of cadaveric kidney transplants compared with those receiving dialysis treatment in Australia and New Zealand, 1991-2001', *Nephrol. Dial. Transplant.* **17**(12), 2212–2219.

Miller, J. E., Kovesdy, C. P., Nissenson, A. R., Mehrotra, R., Streja, E., Van Wyck, D., Greenland, S. & Kalantar-Zadeh, K. (2010), 'Association of hemodialysis treatment time and dose with mortality and the role of race and sex', *Am. J. Kidney Dis.* **55**(1), 100–112.

Pearl, J. (2000), *Causality*, Cambridge University Press, Cambridge. Models, reasoning, and inference.

Powell, J. R., Oluwaseun, O., Woo, Y. M., Padmanabhan, N., Narasinghan, E., Latta, C., Tortolano, J., Jardine, A. G. & Geddes, C. C. (2009), 'Ten years experience of in-center thrice weekly long overnight hemodialysis', *Clin J Am Soc Nephrol* **4**(6), 1097–1101.

Robins, J. (1986), 'A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect', *Math. Modelling* **7**(9-12), 1393–1512. Mathematical models in medicine: diseases and epidemics, Part 2.
URL: *http://dx.doi.org/10.1016/0270-0255(86)90088-6*

Robins, J. M. (1997), Causal inference from complex longitudinal data, *in* 'Latent variable modeling and applications to causality (Los Angeles, CA, 1994)', Vol. 120 of *Lecture Notes in Statist.*, Springer, New York, pp. 69–117.

Robins, J. M. (1998), 'Correction for non-compliance in equivalence trials', *Stat Med* **17**(3), 269–302.

Robins, J. M. (1999*a*), 'Association, causation, and marginal structural models', *Synthese* **121**(1-2), 151–179. Statistics and causation.

Robins, J. M. (1999*b*), Marginal structural models.

Robins, J. M., Hernan, M. A. & Brumback, B. (2000), 'Marginal structural models and causal inference in epidemiology', *Epidemiology* **11**(5), 550–560.

Saran, R., Bragg-Gresham, J. L., Levin, N. W., Twardowski, Z. J., Wizemann, V., Saito, A., Kimata, N., Gillespie, B. W., Combe, C., Bommer, J., Akiba, T., Mapes, D. L., Young, E. W. & Port, F. K. (2006), 'Longer treatment time and slower ultrafiltration in hemodialysis: associations with reduced mortality in the DOPPS', *Kidney Int.* **69**(7), 1222–1228.

Troidle, L., Hotchkiss, M. & Finkelstein, F. (2007), 'A thrice weekly in-center nocturnal hemodialysis program', *Adv Chronic Kidney Dis* **14**(3), 244–248.

Xiao, Y., Abrahamowicz, M. & Moodie, E. E. M. (2010), 'Accuracy of conventional and marginal structural Cox model estimators: a simulation study', *Int. J. Biostat.* **6**(2), Art. 13, 30.

URL: *http://dx.doi.org/10.2202/1557-4679.1208*

Zhou, Z., Rahme, E., Abrahamowicz, M. & Pilote, L. (2005), 'Survival bias associated with time-to-treatment initiation in drug effectiveness evaluation: a comparison of methods', *Am. J. Epidemiol.* **162**(10), 1016–1023.