Survival Analysis Methods for Recurrent Medical Cost Data

Laura M. Yee

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2014

Reading Committee:

Kwun Chuen Gary Chan, Chair

Susanne May

Yingye Zheng

Program Authorized to Offer Degree:
Biostatistics

University of Washington

Abstract

Survival Analysis Methods for Recurrent Medical Cost Data

Laura M. Yee

Chair of the Supervisory Committee:
Professor Kwun Chuen Gary Chan
Department of Biostatistics

Cost data are being collected more frequently in randomized clinical trials in order to assess the cost-effectiveness of experimental treatments. As such, the goal of this dissertation is to study three separate topics which advance the analysis of medical cost data.

First, time between recurrent medical events may be correlated with the cost incurred at each event. As a result, it may be of interest to describe the relationship between recurrent events and recurrent medical costs by estimating a joint distribution. In this paper, we therefore formulate a nonparametric estimator for the joint distribution of recurrent events and recurrent medical costs in right-censored data. We also derive the asymptotic variance of our estimator, and present simulation studies to demonstrate the performance of our point and variance estimators. Our estimator is shown to perform well for a range of levels of correlation, demonstrating that our estimators can be employed in a variety of situations when the correlation structure may be unknown in advance. We apply our methods to hospitalization events and their corresponding costs in the second Multicenter Automatic Defibrillator Implantation Trial (MADIT-II), which was a randomized clinical trial studying the effect of implantable cardioverter-defibrillators in preventing ventricular arrhythmia.

Next, as the costs of medical care increase, more studies are evaluating cost in addition to effectiveness of treatments. Cost-effectiveness in randomized clinical trials has typically been evaluated only at the end of follow-up. However, cost-effectiveness may change over time. We therefore propose a nonparametric estimator to assess the incremental cost-

effectiveness ratio over time. We also derive the asymptotic variance of our estimator and present implementation of simultaneous confidence bands. Simulation studies demonstrate the performance of our proposed methods. We also illustrate our methods using data from a randomized clinical trial, the second Multicenter Automatic Defibrillator Implantation Trial (MADIT-II). This trial studied the effects of implantable cardioverter-defibrillators on patients at high risk for cardiac arrhythmia. Results show that our estimator performs well in large samples, indicating promising future directions in the field of cost-effectiveness.

Finally, in randomized clinical trials that study cost as well as effectiveness, a common complication is often noncompliance to assigned treatment. In situations where compliance in the trial may differ from compliance rates in the population, it may be of interest to study complier average cost-effectiveness. In this paper, we relate the standard intention-to-treat parameters to the complier average causal effects of two well-known measures of cost-effectiveness, the incremental net benefit (INB) and the incremental cost-effectiveness ratio (ICER). In particular, we show that the intention-to-treat effects are proportional to the complier average effect in the case of the INB, but that the intention-to-treat effect can be interpreted as the complier average effect for the ICER. We outline the assumptions required for these relationships to hold and we also present simulation studies confirming these properties. This work provides some incentive for employing the ICER over the INB when researchers are interested in complier average cost-effectiveness.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

DEDICATION

For Betsy, who supported me throughout this journey.

Chapter 1

BACKGROUND

1.1 Introduction

In the past thirty years, medical costs have been growing disproportionately quickly, compared to overall gross domestic product. In the US, health care costs were estimated to be about 9.2% of the GDP in 1980, while they made up 17.6% in 2009 (Bodenheimer and Grumbach, 2009). For this reason, policy makers and researchers have intensified their efforts to understand the determinants of medical cost.

There are numerous questions which can be answered by medical cost data. For example, how can we contain medical costs without sacrificing care? What is the societal cost burden of a specific disease? In chronic diseases, what is the distribution of costs for one episode of disease? How do we determine whether a new treatment is too costly? How can we interpret cost-effectiveness results that come from a randomized clinical trial?

This dissertation will focus on three statistical problems in the field of medical cost data. First, the time between recurring hospitalizations may be correlated with the cost incurred at each hospitalization. For this reason, this disseration will extend existing methods to develop a nonparametric estimator for the joint distribution of recurrent event times and medical cost.

Next, cost-effectiveness of new treatments may change over time. For example, a costly surgery may have front-loaded costs, resulting in costs that may not outweigh the effects in the short run. However, the effectiveness of the surgery might well outweigh the costs over a longer period of time. This dissertation will explore a method to examine a measure of cost-effectiveness over time.

Finally, it is sometimes of interest to know the cost-effectiveness of treatment amongst

participants who comply with treatment assignment in a randomized clinical trial with noncompliance. This dissertation will examine interpretations of two measures of cost-effectiveness using the theory of instrumental variables.

In this chapter, we will review the background literature which is essential to understanding the problems we plan to address in this dissertation.

1.2 Medical Cost Data

1.2.1 Describing Medical Cost

Throughout your life, you can expect to accumulate cost from medical treatments, exams, and hospitalizations. But how should we describe the pattern of medical cost? In this dissertation, we will describe medical cost data in two different ways. We will refer to these scales as cumulative cost and recurrent cost.

Suppose we are interested in an individual's hospitalization costs after receiving an implantable cardioverter-defibrillator (ICD). Over a period of several years, the patient visits the hospital four times before she dies. Her cost history over time is illustrated in Figure 1.1 on page 25.

Two examples of cumulative medical cost are lifetime medical cost or accumulated medical cost up to some point in time. The patient's cumulative lifetime medical cost (that is, from ICD implantation to death) is therefore the sum of her four hospitalization costs, denoted by Y_i (Figure 1.1A). Note that we only observe Y_i if we observe the patient's death (here, at time T_i), since costs may accrue further after a patient is censored. A variable which can only be observed if we reach the endpoint is known as a marked variable (Huang and Louis, 1998). Thus, lifetime medical cost is a marked variable.

We can also study the patient's cumulative medical cost at any given time after the start of the study if we have her medical cost history. For example, say that we wish to know the patient's cumulative cost at time t_3 (Figure 1.1A). Then her cumulative cost, $Y_i(t_3)$, is the sum of the hospitalization costs at times t_1, t_2, and t_3. Here, we observe cumulative cost at

all times up to the patient's death ($Y(T_i)$). Had she been censored before death, we would have observed her cumulative cost up to her censoring time ($Y(C_i)$).

An entirely different way to look at this patient's medical cost is to examine her costs at each hospitalization separately (Figure 1.1B). The cost incurred at her first hospitalization is called m_{i1}, at her second hospitalization m_{i2}, and so forth. These are called the patient's recurrent medical costs. The gap times between the previous hospitalization (or baseline) and the current hospitalization are called the recurrent event times, and are denoted by t_{ij}, j=1, . . . , 4. In this case, we only observe recurrent medical costs if a recurrent event occurs. Thus, recurrent medical cost is a marked variable for the recurrent event, hospitalization.

In the next section, we will introduce one of the main problems in medical cost data with right censoring, induced informative censoring.

1.2.2 Induced Informative Censoring

Since medical costs may be censored before our endpoint is reached, we need to consider methods to account for censoring. Nonparametric methods of describing the distribution of survival times–like the Kaplan-Meier survival curve or the Nelson-Aalen cumulative hazard–account for independent censoring on the time scale (Kaplan and Meier, 1958; Nelson, 1972; Aalen, 1978). However, when we wish to talk about medical cost, we encounter a problem known as induced informative censoring.

Let Y_i denote lifetime medical cost, T_i denote survival time, and C_i denote censoring time for individual i, i=1, . . . , n. We only observe lifetime medical cost if $T_i \leq C_i$. We also observe X_i= min(T_i, C_i) and Δ_i=I($T_i \leq C_i$).

Survival methods like the Kaplan-Meier estimator rely on the assumption of non-informative censoring on the time scale. That is, they assume that participants' survival times (T_i) and censoring times (C_i) are independent. This assumption justifies conditioning on risk sets (H(t) = $\sum_{i=1}^{n} I(X_i \geq t)$) when we are truly interested in the set of people who are still alive.

Induced informative censoring is a problem that has been documented on the cost scale (Huang and Louis, 1998; Lin et al., 1997; Bang and Tsiatis, 2000; Strawderman, 2000), and has also been studied in other contexts, such as quality-adjusted life and mean number of recurrent events (Glasziou et al., 1990; Zhao and Tsiatis, 1997; Cook and Lawless, 1997; Strawderman, 2000). Induced informative censoring refers to the situation where even if the non-informative censoring assumption holds on the time scale, censoring will be informative on the cost scale. For example, a patient who is quite sick might accumulate medical cost quickly, and will consequently have larger costs at both her survival time and her censoring time. This is illustrated in Figure 1.2 on page 26. This figure depicts a simplified situation where patient 1 and patient 2 accumulate cost at constant rates over time. In this figure, we see that if patient 1 is ill while patient 2 is relatively healthy, patient 1 accumulates more cost by both C and T, while patient 2 has lower medical costs at both C and T. Thus, we see that there is within individual correlation between Y(C) and Y(T). In general, as in this example, it is reasonable to believe that Y(C) and Y(T) are correlated within individuals, even if we assume that C and T are independent.

Initially, investigators studying medical cost did not explicitly account for censoring. When studying mean lifetime cumulative cost, researchers often chose to estimate the mean using all observed costs or costs from fully observed cases (Lin et al., 1997). Using all observed costs resulted in a systematic downward bias because costs accumulated after censoring times were not incorporated into the estimate. Employing costs only from fully observed cases was an attempt to correct this bias, but this estimate was also biased because it only used the costs of people with shorter survival times.

Subsequently, researchers began to use the Kaplan-Meier method to account for censoring (Lin et al., 1997). This estimator used medical cost instead of survival time, and could be expressed as follows:

$$\hat{S}_Y(t) = \prod_{u_j \leq t} \left[1 - \frac{dY_j}{H_j} \right]$$

where $dY_j = \sum_{i=1}^{n} I(Y(T_i) = u_j)$ is the sum of lifetime costs for individuals who die with cost u_j, and $H_j = \sum_{i=1}^{n} I(Y(X_i) \geq u_j)$ is the sum of costs for all individuals whose observed total cost ($\mathrm{Y}(X_i) = \min(\ \mathrm{Y}(T_i),\ \mathrm{Y}(C_i))$ is greater than or equal to u_j.

This modified Kaplan-Meier estimator ignored induced informative censoring. That is, since $\mathrm{Y}(T_i)$ is correlated with $\mathrm{Y}(C_i)$, the risk set (H_j) composed of $\mathrm{Y}(X_i)$s is generally not representative of all those individuals with lifetime cost ($\mathrm{Y}(T_i)$) greater than u_j. The next section will describe one way to estimate mean lifetime cost without having to assume that $\mathrm{Y}(T_i)$ is independent of $\mathrm{Y}(C_i)$.

1.2.3 Simple Weighted Estimator

In this section, we describe what we call the simple weighted estimator for estimating mean medical cost in the presence of censoring. This estimator was initially proposed by Bang and Tsiatis (2000), and is similar to the estimator proposed for the problem of quality-adjusted survival time (Zhao and Tsiatis, 1997). This estimator draws upon work from the field of missing data (Robins et al., 1994; Robins and Rotnitzky, 1992) to estimate mean cost without relying upon an assumption of non-informative censoring on the cost scale. The premise of the estimator is to use inverse probability weighting in order to up-weight observed lifetime costs at times when censoring was high.

Using the notation from section 1.2.2, we add the constraint that $\mathrm{T} \leq \tau$, and that $\mathrm{P}(C_i \geq \tau) > 0$. Requiring $\mathrm{T} \leq \tau$ means that we will evaluate mean medical cost for the length of time τ, and imposing $\mathrm{P}(C_i \geq \tau) > 0$ necessitates that we will observe some medical costs at time τ.

It should be noted that τ is generally chosen to be the length of study observation time. In the survival time setting, if some people are expected to survive beyond the period of study observation–that is, if $\mathrm{P}(\mathrm{T} \geq \tau) > 0$–we can only nonparametrically identify the survival time distribution up to the end of the study, τ. However, since we may not observe much lifetime cost information for people for whom $\mathrm{P}(\mathrm{T} \geq \tau) > 0$, we are not able to

nonparametrically identify the distribution of lifetime medical cost anywhere (Huang, 2009). For this reason, Bang and Tsiatis impose the time restriction of τ on medical cost, which allows us to evaluate the mean of the τ-restricted cost, which is truly $Y(\tau)$. $Y_i(\tau)$ is the cumulative cost up to τ for patient i, or the medical cost up to death or time τ. Huang (2009) noted that this restriction is a limitation of this method, and it can lead to interpretation difficulties when we are interested in talking about lifetime medical cost (Huang, 2009).

For the remainder of this section, let Y_i=$Y_i(\tau)$ for notational convenience. Now, the simple weighted estimator for μ_Y= E[Y] can be written as follows:

$$\hat{\mu}_{WT} = \frac{1}{n}\sum_{i=1}^{n}\frac{\Delta_i Y_i}{\hat{K}(T_i)} \tag{1.1}$$

In this formula, $\hat{K}$(t) denotes the estimator for the probability of not being censored by time t, P(C>t). This can be estimated via the method of Kaplan-Meier, using X as the observed time variable, and 1-Δ as the event indicator. Note that we have used the subscript WT for this estimator to denote the simple weighted (WT) estimator of Bang and Tsiatis.

This estimator sums the lifetime medical cost of each fully observed case (that is, people whose death is observed), and up-weights this cost by the inverse probability of not being censored at this point in the study. For example, consider Figure 1.3 on page 27, which displays K(t) for participants from the MADIT-II study who were randomized to receive ICD implantation (Moss et al., 2002). At the start of the study, the probability of not being censored is 1. Since everyone was still available for follow-up when the study began, we would not need to upweight any lifetime costs at this time. As time progresses, more participants experience censoring, so we begin to upweight costs more and more. For example, at 1.1 years, the proportion not censored is 0.75. Since we only have 3/4 of the participants not censored at 1.1 years, we need to count the lifetime medical cost for whoever dies at this time 4/3 times so that, on average, we have complete medical costs. At 1.9 years, only 50% of participants have not been censored. Therefore, we would need to further upweight medical cost at this time (upweight 1/.5 = 2 times).

Bang and Tsiatis (2000) showed that this estimator is unbiased for mean cost by using a conditioning argument:

$$\begin{aligned}
\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n}\frac{\Delta_i Y_i}{K(T_i)}\right] &= \mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\left(\frac{\Delta_i Y_i}{K(T_i)}|T_i, Y_i\right)\right] \\
&= \mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n}\frac{Y_i}{K(T_i)}E(I(C_i \geq T_i)|T_i, Y_i)\right] \\
&= \mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n}Y_i\right]
\end{aligned}$$

The simple weighted estimator of Bang and Tsiatis presented a solution to the problem of induced informative censoring when estimating mean cost. While this estimator has the advantage of being an unbiased estimator of mean cost, it is inefficient because it only uses cost data from fully observed participants. In the next two sections, we will present some notable extensions of the simple weighted estimator.

1.2.4 Estimator of Huang and Louis

Huang and Louis proposed estimating a nonparametric joint distribution of survival time and marked variables (Huang and Louis, 1998). As we recall from section 1.2.1, a marked variable is one that we observe only if the survival endpoint is reached. Since we only obtain lifetime medical cost if a patient dies during the observation period, lifetime medical cost is a marked variable for survival time. The estimator of Huang and Louis is a generalization of the simple weighted estimator because it sets up a framework for estimation involving any marked variable of survival time, and because we can obtain an estimate of the entire joint distribution of survival time and the marked variable instead of only the mean of the marked variable. Furthermore, as we will see, the estimator of Huang and Louis uses a survival-type rationale for its construction, while the simple weighted estimator used a missing data rationale.

First, let us lay out the notation for the Huang and Louis estimator. Huang and Louis were interested in estimating the joint distribution of survival time, T, and marked variable, Y^0. Since we are interested in the marked variable medical cost, for simplicity, we will henceforth refer to Y^0 as lifetime medical cost. In this setting, we observe X= min(T,C), where C is the censoring time, Y= Y^0I(T≤C), and Δ=I(T≤C). Then the joint distribution, F_{TY^0}(t,y) proposed by Huang and Louis can be written as follows:

$$\begin{aligned} F_{TY^0}(t,y) &= \int_0^t S_T(s-)\Lambda_{TY^0}(ds,y) \\ &= \int_0^t \prod_{[0,s)} [1-\Lambda_T(dv)]\Lambda_{TY^0}(ds,y) \end{aligned}$$

where $s-$ denotes the time right before time s, S_T(x) = P(T>x) is the survival function, Λ_{TY^0}(t,y) is a cumulative mark-specific hazard function, and Λ_T(t) = Λ_{TY^0}(t,∞) is the cumulative hazard function for T. The cumulative mark-specific hazard function is the novel parameter:

$$\Lambda_{TY^0}(t,y) = P(T \le t, Y^0 \le y | T \ge t-) = \int_0^t \frac{F_{TY^0}(ds,y)}{S_T(s-)}$$

We note that this hazard function can be estimated by assuming only that (T, Y^0) is independent of C. In other words, we once again do not require an assumption of independent censoring on the cost scale. Using this assumption, we have a parameter that we can estimate since

$$\Lambda_{TY^0}(t,y) = \int_0^t \frac{F_{TY^0}(ds,y)}{S_T(s-)} = \int_0^t \frac{F^u_{TY^0}(ds,y)}{S_X(s-)}$$

That is, the assumption that (T, Y^0) is independent of C implies that the uncensored joint distribution, $F^u_{TY^0}$(t,y) = P(T≤t, Y^0≤y, Δ=1) = $F^u_{TY^0}$(t,y)S_C(t), and that S_X(t) = S_T(t)S_C(t).

Then for iid random data points (X_i, Y_i, Δ_i), i=1, …, n, Huang and Louis proposed

the estimator for the joint distribution to be

$$\hat{F}_{TY^0}(t,y) = \int_0^t \prod_{[0,s)} [1 - \hat{\Lambda}_T(dv)] \hat{\Lambda}_{TY^0}(ds,y)$$

where

$$\hat{\Lambda}_{TY^0}(t,y) = \int_0^t \frac{F^u_{TY^0}(ds,y)}{\hat{H}(s)} = \int \frac{\sum_{i=1}^n I(X_i = s, Y_i \leq y, \Delta_i = 1)}{\sum_{i=1}^n I(X_i \geq s)}$$

and

$$\hat{\Lambda}_T(t) = \hat{\Lambda}_{TY^0}(t,\infty)$$

The estimator of Huang and Louis has taken a survival-based approach to constructing an estimator of F_{TY^0}. That is, Huang and Louis used survival functions and cumulative hazard functions to define their estimator. As described by Bang and Tsiatis (2000), we note that the joint nonparametric formulation is quite simlar to the missing data approach taken by Bang and Tsiatis:

$$\begin{aligned}
\hat{F}_{TY^0}(\tau,y) &= \int_0^\tau \prod_{[0,s)} \left[1 - \hat{\Lambda}_T(dv)\right] \hat{\Lambda}_{TY^0}(ds,y) \\
&= \int_0^\tau \hat{S}_T(s-) \frac{\sum_{i=1}^n I(X_i = s, Y_i \leq y, \Delta_i = 1)}{\sum_{i=1}^n I(X_i \geq s)} \\
&= \sum_{i=1}^n \frac{\hat{S}_T(T_i-)}{\hat{H}(T_i)} \Delta_i I(Y_i \leq y) I(X_i \leq \tau) \\
&= \sum_{i=1}^n \frac{\Delta_i I(Y_i \leq y) I(X_i \leq \tau)}{\hat{K}(T_i-)} \qquad (1.2)
\end{aligned}$$

since $\hat{H}(t) = \hat{S}(t-)\hat{K}(t-)$. If we compare equation (1.2) to equation (1.1), we see that the two have essentially the same formulation.

The joint distribution of Huang and Louis is identifiable in the region [0,τ] x [0,∞), which

is potentially preferable to the identifiability solution in the simple weighted estimator of Bang and Tsiatis. In section 1.2.3, we noted that the marginal distribution of lifetime medical cost is not identifiable nonparametrically, so Bang and Tsiatis proposed estimating the mean τ-restricted lifetime medical cost. The method of Huang and Louis therefore allows joint estimation of lifetime medical cost, instead of a hard to interpret marginal distribution of τ-restricted lifetime medical cost.

We will employ the framework of Huang and Louis in Chapter 2 to construct a joint estimator for recurrent survival time and recurrent medical cost. In the next section, we briefly review some further extensions to estimation of mean lifetime medical cost.

1.2.5 Other Extensions

Both the simple weighted estimator and the estimator of Huang and Louis use cost data only from participants observed to die. While these estimators are unbiased for mean lifetime cost, they are generally inefficient. Although we will only draw upon the work of some of the authors in the following extensions, we will briefly discuss them here. This will help indicate possible directions for future research.

Lin, Feuer, Etzioni, and Wax (1997) proposed an estimator for mean medical cost, which differed from the simple weighted estimator in that it partitioned the time scale into K subintervals (Lin et al., 1997). This estimator summed cost within each subinterval, weighted according to whether the individual was at risk in that interval. This estimator was derived using a survival-based argument similar to the one we discussed in 1.2.4. A strength of this estimator was that it took costs of censored individuals into account in the subintervals before they were censored. On the other hand, a weakness of this estimator was that it required the assumption that censoring times occurred either at the start or end of each subinterval, which is not generally true.

The partitioned estimator of Bang and Tsiatis (2000) was similar to the estimator of Lin et al (1997), but this new partitioned estimator did not require the assumption that

individuals were censored at the beginning or end of subintervals. This estimator only used cost information in subintervals which occurred entirely before an individual's censoring time. This estimator was formulated using a missing data approach similar to the one we discussed in 1.2.3.

The estimator of Cook and Lawless (1997)– derived for the marked variable number of recurrent events– was a continuous version of the partitioned estimators of Lin et al. and Bang and Tsiatis. The estimator by Cook and Lawless can be written as follows:

$$\hat{\mu}_{CL} = \sum_{i=1}^{n} \int_0^{\infty} \frac{\hat{S}_T(s-)}{\hat{H}(s)} dY_i(s) \tag{1.3}$$

$$= \sum_{i=1}^{n} \int_0^{\infty} \frac{dY_i(s)}{\hat{K}(s-)} \tag{1.4}$$

since S(u-)K(u-)=H(u), dY_i(t) is the jump in the marked variable, K(t) once again denotes the P(C>t), S_T(t) = P(T>t), and H(t)=E[I(X>t)].

Thus, we see that while the estimator in (1.3) was formulated using a survival approach (as in section 1.2.4), it can be rewritten to take the missing data format of 1.2.3. That is, (1.4) is similar to the form of the simple weighted estimator (1.1), but that it now incorporates costs (Y_i(t)) for censored individuals before their censoring time. Therefore, we see that this estimator is the limit of the partitioned estimator by Bang and Tsiatis. That is, if the number of subintervals in the partitioned estimator go to infinity, then we obtain the estimator of Cook and Lawless. We will employ the estimator of Cook and Lawless in Chapter 3.

Finally, Bang and Tsiatis (2000), Strawderman (2000), and Zhao and Tian (2001) applied work from the field of missing data in order to obtain more efficient estimators. All three papers employed work by Robins and Rotnitzky (1992) to attempt to find the semiparametric efficiency bound. These estimators are more complex, but tend to be more efficient in realistic medical cost situations.

1.3 Recurrent Event Data

1.3.1 Recurrent Events: Gap times

In section 1.2.1, we briefly discussed recurrent event times and recurrent medical costs. In that section, we highlighted that we will focus on describing the gap times and gap costs between recurrent events (recall, for example, Figure 1.1 B).

Prentice, Williams, and Peterson presented semiparametric models for analyzing recurrent event data for both the entire course of study and for gap times (Prentice et al., 1981). The gap time data structure they presented is the same structure of data we will discuss in Chapter 2. In that chapter, we will focus on a nonparametric approach to dealing with gap time recurrent event times and gap time recurrent medical events. In the next section, we will discuss nonparametric estimation of gap time recurrent events.

1.3.2 Estimator of Wang and Chang

Wang and Chang (1999) proposed a nonparametric estimator for the recurrent survival distribution. Since we will employ the framework of Wang and Chang in Chapter 2, we summarize their estimator in this section.

First, let us lay out the notation of Wang and Chang. For individuals i, i=1, ..., n, who had an initial event at the time of admission to the study, we record their recurrent events j, j=1, ..., $k_i - 1$. We let T_{ij} be the recurrent event times. T_{ij} therefore denotes the time from the (j-1)th event to the jth event for individual i. Let C_i be the censoring time for individual i. Note that a censoring time in this setting may be the time of death, since death is not the recurrent event of interest. For the method of Wang and Chang, it is possible to avoid problems that may arise from calling death a censored event by considering a composite endpoint which includes both the event of interest and death. Let k_i be the

index indicating the censoring time. That is, our last observed recurrent event is $T_{i(k_i-1)}$:

$$\sum_{j=1}^{k_i-1} T_{ij} \leq C_i$$

and

$$\sum_{j=1}^{k_i} T_{ij} > C_i$$

We let k_i^*=max(k_i-1, 1). Finally, we observe the following:

$$y_{ij} = \begin{cases} t_{ij} & \text{if } j = 1, ..., k_i - 1 \\ t^+_{ik_i} & \text{if } j = k_i \end{cases}$$

where $t^+_{ik_i}$ is the time between $t_{i(k_i-1)}$ and C_i.

To help visualize this notation, we refer to Figure 1.1 B on page 25. Table 1.1 below corresponds to the data we saw in this figure. We observed four events, and the corresponding gap times are t_{i1}, ..., t_{i4}. The y_{ij} correspond to t_{ij} in these cases, and to the time between the fourth event and the censoring time for y_{i5}. The censoring time was denoted by T_i in Figure 1.1 B because the patient died at that time. In this situation, k_i=5 since censoring (here, death) occurred before the fifth recurrent event.

Table 1.1: Recurrent event data corresponding to the hypothetical participant in Figure 1.1B.

Gap Time (y_{ij})	**Recurrent Event Status**	**Event Number (j= 1, ..., k_i)**
t_{i1}	1	1
t_{i2}	1	2
t_{i3}	1	3
t_{i4}	1	4
t^+_{i5}	0	5

Now, the goal of the Wang and Chang estimator is to define a risk set and a mass at time t in order to create an estimator for the cumulative hazard of recurrent events. This

estimate of the cumulative hazard can then be used to estimate survival time.

The averaged risk set proposed by Wang and Chang was

$$\hat{H}_{WC}(t) = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{k_i^*}\sum_{j=1}^{k_i^*} I(y_{ij} \geq t)\right] \tag{1.5}$$

while the averaged mass was

$$\hat{F}_{WC}(t) = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{I(k_i \geq 2)}{k_i^*}\sum_{j=1}^{k_i^*} I(y_{ij} \leq t)\right] \tag{1.6}$$

Therefore, the estimate for the recurrent cumulative hazard function could be written as follows:

$$\hat{\Lambda}_{WC}(t) = \int_{[0,t]} \frac{\hat{F}_{WC}(ds)}{\hat{H}_{WC}(s)} \tag{1.7}$$

Both the risk set and the mass average within individuals before averaging across individuals. By doing so, Wang and Chang allow for different rates of recurrences within individuals.

Although it is inefficient, we could estimate the recurrent cumulative hazard function using only the y_{i1}, i= 1, . . . , n. In this simpler setting, we would have

$$\hat{\Lambda}_{simple}(t) = \int_0^t \frac{\hat{F}_{simple}(ds)}{\hat{H}_{simple}(s)}$$

where

$$\hat{F}_{simple}(t) = \frac{1}{n}\sum_{i=1}^{n} I(y_{i1} \leq t) I(k_i \geq 2)$$

and

$$\hat{H}_{simple}(t) = \frac{1}{n}\sum_{i=1}^{n} I(y_{i1} \geq t) = \frac{1}{n}\sum_{i=1}^{n} I(y_{i1} \geq t)I(k_i \geq 2) + \frac{1}{n}\sum_{i=1}^{n} I(y_{i1} \geq t)I(k_i = 1) \quad (1.8)$$

The second equality in equation (1.8) shows that the risk set in this simple situation can be split into those people for whom we observed a recurrence (I($k_i \geq$2)) and those for whom we didn't (I($k_i = 1$)). It is valid in this reduced setting to incorporate all y_{i1}– that is, both the t_{i1} and the t_{i1}^{+}–because from the origin, we assume that T_{i1} and C_i are independent.

Wang and Chang improve upon $\hat{\Lambda}_{simple}(t)$ by incorporating all observed recurrences in both the mass (equation (1.6)) and the risk set (equation (1.5)). We note that we can rewrite the risk set for the estimator of Wang and Chang in equation (1.5) as follows:

$$\hat{H}_{WC}(t) = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{k_i^*}\sum_{j=1}^{k_i^*} I(y_{ij} \geq t)\right] \quad (1.9)$$

$$= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{k_i^*}\sum_{j=1}^{k_i^*} I(y_{ij} \geq t)I(k_i \geq 2) + \frac{1}{n}\sum_{i=1}^{n} I(y_{i1} \geq t)I(k_i = 1) \quad (1.10)$$

Here, we see that the risk set (equations (1.5) and (1.10)) does not incorporate any $t_{ik_i}^{+}$ if $k_i \geq 2$ (another way to say this is the risk set only sums up to k_i^*). This means that we only include censored recurrence times if censoring occurred before the first recurrence.

Why do Wang and Chang only include censored recurrence times before the first recurrence? In the simple case (equation (1.8)), we still have that T_{i1} and C_i are independent. After the first recurrrence time, however, censoring is informative. For example, in Figure 1.4 on page 28, we see that person 1 has a higher frequency of recurrence, and therefore will tend to have shorter $t_{ik_i}^{+}$. On the other hand, person 2 has a lower frequency of recurrence, and will tend to have longer $t_{ik_i}^{+}$. Since censoring is informative in this setting, it is not valid to use I($y_{ij} \geq$ t) in the place of I($t_{ij} \geq$ t) for the censoring times. Thus, Wang and Chang chose to omit those censoring times from the risk set.

Now that we have described the estimator for $\hat{\Lambda}_{WC}(t)$, we can estimate the recurrent

survival function, S_{WC}(t), using

$$\hat{S}_{WC}(t) = \prod_{u \leq t} (1 - d\hat{\Lambda}_{WC}(t))$$

We have described the nonparametric estimator of recurrent survival time as proposed by Wang and Chang. We will employ this estimator in Chapter 2 when we present a joint distribution of recurrent survival time and medical cost.

1.4 Cost-Effectiveness Analysis

1.4.1 Introduction to Cost-Effectiveness Analysis

Cost-effectiveness analysis describes methods which weigh the costs of a new treatment against its efficacy, which is commonly referred to as its effectiveness. In cost-effectiveness studies, researchers are generally interested in comparing a new treatment against a standard treatment. This comparison is often completed in a randomized clinical trial in which both effectiveness and medical cost data are collected.

At the end of the study, there are four possible situations involving the cost-effectiveness of a new treatment (Willan and O'Brien, 1996; Stinnett and Mullahy, 1998). These situations are summarized in the following table. Compared to a standard treatment, a new treatment may be more effective or less effective, and may have higher cost or lower cost. From this table, we see that we may be interested in describing cost-effectiveness whenever there is a trade-off between cost and effectiveness– that is, when we have a more effective treatment that is also more costly, or when we have a less effective treatment that is less costly.

	Lower Cost	**Higher Cost**
Less Effective	Cost-effectiveness analysis	Reject new treatment
More Effective	Use new treatment	Cost-effectiveness analysis

Table 1.2: Four possible outcomes of a study measuring both cost and efficacy.

In the following two sections, we will present two different ways of describing cost-effectiveness.

1.4.2 Incremental Cost-Effectiveness Ratio

The incremental cost-effectiveness ratio (ICER) is one measure of cost-effectiveness that has been widely used by biomedical researchers (Mushlin et al., 1998; Zwanziger et al., 2006). In the setting of randomized clinical trials, the ICER evaluates the difference in cost between treatment and standard, divided by their difference in effectiveness. Two descriptors of effectiveness that are commonly employed are mean survival time and probability of survival. The ICER is typically evaluated at the end of follow-up in a randomized clinical trial.

For individual i, i= 1, ..., n, define T_i to be the survival time, C_i to be the censoring time, and X_i=min(T_i, C_i). We observe lifetime medical cost, Y_i, when δ_i= I($T_i \leq C_i$)=1. We also know whether patients are randomized to treatment (R_i=1) or standard (R_i=0). The duration of interest, typically defined by the end of follow-up in the randomized clinical trial, is denoted by τ. As in sections 1.2.3 and 1.2.4, we note that E(T) is typically not identifiable nonparametrically if P(T$\geq \tau$) > 0. Instead, we must use an estimator of E(T|T$\leq \tau$). This has the disadvantage that τ is defined by study design, instead of by a scientific goal.

We wish to estimate the parameters

$$\mathbf{ICER_S} = \frac{\mu_{Y_1} - \mu_{Y_0}}{S_1(\tau) - S_0(\tau)}$$
$$\mathbf{ICER}_{\mu_\mathbf{T}} = \frac{\mu_{Y_1} - \mu_{Y_0}}{\mu_{T_1} - \mu_{T_0}}$$

where, for randomization group k (k=1 for treatment, and k=0 for standard), μ_{Y_k}= E[Y | R=k] is the mean lifetime cost, $S_k(\tau)$ = P(T $> \tau$ | R=k) is the probability of survival at the end of follow-up, and μ_{T_k}= E[T | R=k] is the mean survival time.

Since there is induced informative censoring on the cost scale (see section 1.2.2), we can estimate μ_{Y_k} using one of the estimators in sections 1.2.3- 1.2.5. Furthermore, $S_k(\tau)$ can be

estimated via Kaplan-Meier, and μ_{T_k} can be estimated by the area under the Kaplan-Meier curve.

Several investigators have discussed evaluating the uncertainty around the ICER (O'Brien et al., 1994; Van Hout et al., 1994; Willan and O'Brien, 1996; Briggs et al., 1997; Chaudhary and Stearns, 1996; Polsky et al., 1997; Briggs et al., 1999; Heitjan et al., 1999; Zhao and Tian, 2001). Initially, a simple Taylor series expansion was proposed (O'Brien et al., 1994), but it was noted that a normal approximation may not be appropriate since the ICER is a ratio. For this reason, several authors advocate using some form of the bootstrap (O'Brien et al., 1994; Briggs et al., 1997; Heitjan et al., 1999). Finally, an application of Fieller's theorem was proposed. This method is computationally easier than the bootstrap, and often performs similarly well (Chaudhary and Stearns, 1996; Polsky et al., 1997; Briggs et al., 1999).

1.4.3 Incremental Net Benefit

The incremental net benefit (INB) is another descriptor of cost-effectiveness that has been proposed for the biomedical setting (Phelps, 1991; Ament and Baltussen, 1997; Stinnett and Mullahy, 1998; Willan and Lin, 2001). The INB was initially proposed to help mitigate some interpretation problems of the ICER (Stinnett and Mullahy, 1998; Willan and Lin, 2001). First, as a ratio, the same ICER can describe two completely different situations. For example, the ICER of a treatment that is both more effective (say, $\hat{\mu}_{T_1} - \hat{\mu}_{T_0} = 0.25$), but more costly ($\hat{\mu}_{Y_1} - \hat{\mu}_{Y_0} = \$10{,}000$) could have the same ICER as a new treatment that is less effective ($\hat{\mu}_{T_1} - \hat{\mu}_{T_0} =$ -0.25) but less costly ($\hat{\mu}_{Y_1} - \hat{\mu}_{Y_0} =$ -\$10,000). Second, it is possible for the confidence interval of the ICER to be undefined when using Fieller's method (Willan and O'Brien, 1996; Willan and Lin, 2001).

The incremental net benefit is a parameter that characterizes the net benefit, in monetary units, of giving a patient a new treatment over the standard treatment. That is, if we define

λ to be the dollar value of a unit of effectiveness, then the net benefit, b(λ), will be:

$$b_{u_T}(\lambda) = (\mu_{T_1} - \mu_{T_0})\lambda - (\mu_{Y_1} - \mu_{Y_0})$$
$$b_S(\lambda) = (S_1(\tau) - S_0(\tau))\lambda - (\mu_{Y_1} - \mu_{Y_0})$$

As in section 1.4.2, we can estimate $S_k(\tau)$, where k= 1 for treatment or 0 for standard, with the Kaplan-Meier estimator, μ_{T_k} with area under the Kaplan-Meier curve, and μ_{Y_k} using one of the estimators in sections 1.2.3- 1.2.5.

One noted drawback of the INB is that it relies heavily on the choice of λ (Phelps, 1991). However, in practice, researchers may evaluate b(λ) for a set of λ, or plot b(λ) versus λ in order to look at the net benefit for an acceptable range of λ.

Willan and Lin (2001) noted that the ICER is the horizontal intercept in a plot of λ versus b(λ) (Willan and Lin, 2001). This is because

$$\begin{aligned} b\left(\frac{\mu_{Y_1} - \mu_{Y_0}}{\mu_{T_1} - \mu_{T_0}}\right) &= (\mu_{T_1} - \mu_{T_0})\frac{\mu_{Y_1} - \mu_{Y_0}}{\mu_{T_1} - \mu_{T_0}} - (\mu_{Y_1} - \mu_{Y_0}) \\ &= 0 \end{aligned}$$

To illustrate the INB with respect to the four situations of cost-effectiveness, we examine Figure 1.5 on page 29. Figures 1.5 A-D respectively depict situations with lower cost and less effectiveness, higher cost and less effectiveness, lower cost and more effectiveness, and higher cost and more effectiveness. We see that a positive slope corresponds to a treatment which is more effective (Figures 1.5C and D). A positive vertical intercept indicates that the new treatment is less costly (Figures 1.5A and C), while a negative vertical intercept indicates that the treatment is more costly (Figures 1.5B and D). The ICER, the horizontal intercept, is negative in both the best (Figure 1.5C) and worst (Figures 1.5B) situations, while it is positive in both trade-off situations (Figures 1.5A and D).

We will employ both the ICER and the INB in Chapter 4, when we discuss complier average cost-effectiveness.

1.4.4 Time Dependent Cost-Effectiveness

Biomedical researchers have noted that time-dependent cost-effectiveness may be of interest (Zwanziger et al., 2006; Salukhe et al., 2004; van de Wetering et al., 2012). A follow-up study on implantable cardioverter-defibrillator (ICD) clinical trials noted that the initial cost of ICD intervention is high, but that the benefits accumulate long after implantation. They presented a nonparametric time-dependent measure of effectiveness: life-years gained per ICD implant (Salukhe et al., 2004). They estimated life-years gained per implant by calculating the cumulative area between the Kaplan-Meier curves of the intervention and standard treatment groups at each point in time. This measure did not explicitly take cost into account.

Researchers with the MADIT-II study also noted that costs of ICD implants are front-loaded (Zwanziger et al., 2006). To account for the discrepancy in cost-effectiveness over time, they projected survival and cost data for 12 years. The reasoning behind this was that 1) ICER decreases over time for ICD implants, and 2) the 12 year ICER may be more similar to lifetime cost-effectiveness than the few years of follow-up available. Their projections for survival were based on quadratic hazard rate models from external, matched data, while their projections for costs were based on linear mixed effects models. This method chose to evaluate cost-effectiveness at one time point (12 years) after randomization, so it did not explicitly account for time-dependence.

A recent editorial highlighted the need for time-dependent cost-effectiveness analyses in Europe, where cost-effectivness information is employed regularly to make policy decisions (van de Wetering et al., 2012). The authors described a motivating example: In the Netherlands, a study was completed to assess the cost-effectiveness of a new, digital mammography system versus the old, analog mammography. Based on the results, the government decided to make the switch from the old system to the new system. Since the original study was completed via a time-constant cost-effectiveness analysis, it was assumed that the savings would begin immediately, although there was actually a lag between imple-

mentation of the new system and cost savings. By reporting results that assumed constant effects over time, health providers may have doubted the cost savings of the new system, therefore causing slower implementation and even more initial losses.

1.5 Example: MADIT-II

1.5.1 Implantable Cardioverter-Defibrillators

Implantable cardioverter-defibrillators (ICDs) are small medical devices that are placed in a patient's chest in order to monitor and correct the heartbeat of patients who have experienced or may experience cardiac arrhythmia. Cardiac arrhythmia, or an irregular heartbeat, can sometimes cause the heart to stop beating. Serious ventricular arrhythmias– such as ventricular tachycardia (a very fast heartbeat) or ventricular fibrillation (a fast and irregular heartbeat)–can be especially dangerous since the ventricles pump blood to the rest of the body (Staff, 2010). Doctors recommend ICDs to patients who are at high risk for dangerous arrhythmias– such as patients who have had a heart attack, patients who have experienced cardiac arrest, or patients who have a history of serious arrhythmia (Staff, 2011b).

ICDs work by tracking the patient's heartbeat. If an abnormal heartbeat is detected, the device sends electrical shocks through wires connected to the ventricle or ventricles with the arrhythmia (Staff, 2011a). These shocks bring the heartbeat back to a regular rhythm.

Currently, ICDs are usually attached to the heart via a transvenous wire (Staff, 2011a). That is, the ICD box is placed under the skin on the chest, and a connecting wire is attached to the ventricle through a vein. ICDs can also be inserted surgically by attaching the wire to the outside of the heart. This method, which requires more serious surgery, was primarily used prior to FDA approval of transvenous leads (wires) in the early 1990s (Moss et al., 1996).

1.5.2 The Second Multicenter Automatic Defibrillator Implantation Trial (MADIT-II)

The second Multicenter Automatic Defibrillator Implantation Trial (MADIT-II) was a clinical trial in which patients at risk for cardiac arrhythmia were randomized to receive implantable cardioverter-defibrillator (ICD) or standard medical treatment in an attempt to prevent sudden cardiac death (Moss et al., 2002).

MADIT-II was conducted from July 11, 1997 to January 16, 2002. 1,232 participants were enrolled from 76 hospitals in the US and Europe. Participants were randomized at baseline to receive standard therapy (n=490) or transvenous ICD (n=742). Participants were considered for inclusion in the study if they had experienced a previous heart attack and had a left ventricular ejection fraction ≤ 0.30. Additional exclusion criteria are documented in the original study report (Moss et al., 2002). Study follow-up was 20 months, on average, and ranged from 6 days to 4.5 years.

In their paper on ICD efficacy, MADIT-II investigators reported the hazard ratio for all cause mortality comparing ICD to standard treatment to be 0.69, 95% CI (0.51, 0.93), p=0.02. Given that ICDs are effective but costly, several papers have discussed cost-effectiveness in the context of the MADIT and MADIT-II trials (Moss et al., 1996; Mushlin et al., 1997, 1998; Al-Khatib et al., 2005; Sanders et al., 2005; Zwanziger et al., 2006).

The paper by Zwanziger et. al (2006) employed medical utilization and cost data which was collected prospectively throughout MADIT-II follow-up. This paper employed data from 1,095 of the original 1,232 patients in the MADIT-II study, including 664 patients randomized to ICD and 431 patients randomized to standard therapy. This cost study excluded 109 patients at European hospitals, and 28 patients from US hospitals who had a great deal of missing cost data.

The cost data was collected as follows: Scheduled follow-up visits or telephone interviews occurred each month after randomization. At those times, all information regarding inpatient, outpatient, home care, other medical visits, and medication usage were collected. Hospital bills were then obtained for each reported medical utilization. A small percentage

of missing cost data was imputed using the utilization data. For this imputation, investigators assigned the average cost per day for the type of utilization reported.

In the following chapters, we will employ the MADIT-II survival data and corresponding cost data for the 1,095 patients in the paper reported by Zwanziger et al.

1.6 Outline of Research

In this chapter, we have described statistical problems associated with medical cost, and some ways to deal with these problems. In particular, we have introduced cumulative and recurrent medical cost, induced informative censoring, and cost-effectiveness analyses.

In chapter 2, we will develop the theory for a nonparametric estimator of the joint distribution of recurrent events and medical cost in right censored data. To do so, we will synthesize the work of Huang and Louis (1998) and Wang and Chang (1999), which we discussed in sections 1.2.4 and 1.3.2. We will present the estimator, its variance, numerical simulations, and an example using the MADIT-II data.

Chapter 3 will present a nonparametric, time-dependent estimator of cost-effectiveness, the time-dependent incremental cost effectiveness ratio (ICER(t)). This problem defines ICER(t) in the context of a randomized clinical trial. To solve this problem, we will utilize the estimator of Cook and Lawless (1997) as well as previous literature concerning the ICER. We have discussed background work for this problem in sections 1.2.5 and 1.3.2, while the motivation for this problem was outlined in section 1.4.4. We will present the estimator, pointwise variance, confidence bands, numerical simulations, and an application to the MADIT-II study.

Chapter 4 will demonstrate that, under some assumptions, the intention-to-treat ICER in a randomized clinical trial has an interpretation as a complier average ICER. We will also show that the incremental net benefit (INB), another measure of cost-effectiveness in randomized clinical trials, does not have this same property. The solultion to this problem will draw upon work by (Angrist et al., 1996) as well as literature involving the ICER and the INB. We have discussed background for the ICER and INB in sections 1.4.2 and 1.4.3.

We will present the assumptions required for these relationships to hold, and we conduct numerical simulations to confirm these properties.

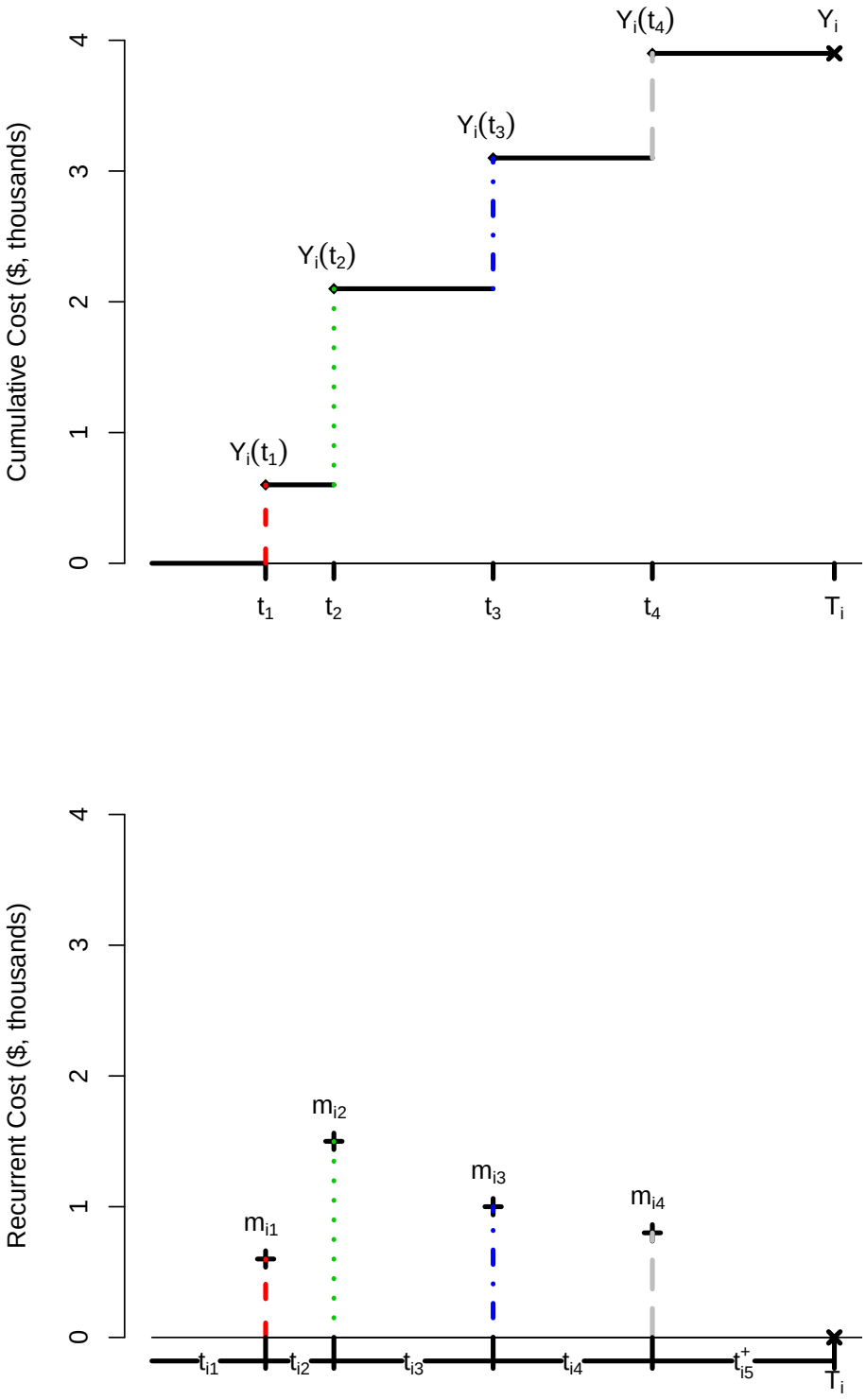

Figure 1.1: Cost measured cumulatively (Figure 1.1A) and recurrently (Figure 1.1B) for hypothetical patient i.

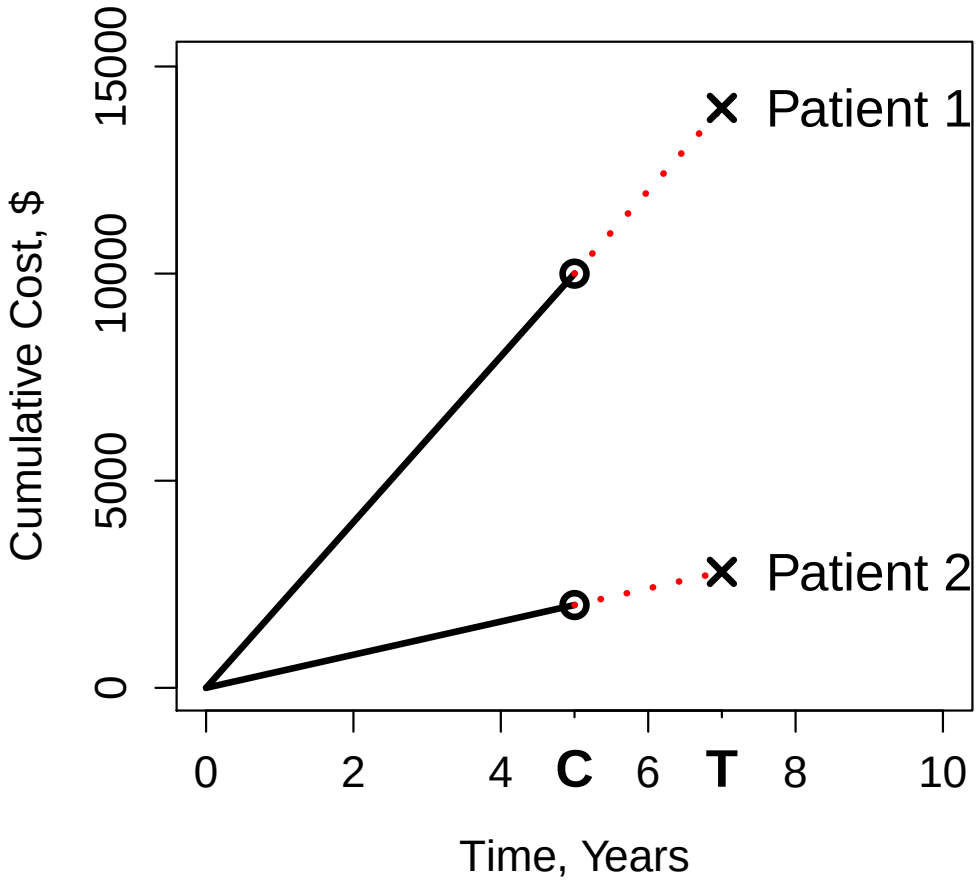

Figure 1.2: Illustration of induced informative censoring for two simplified scenarios.

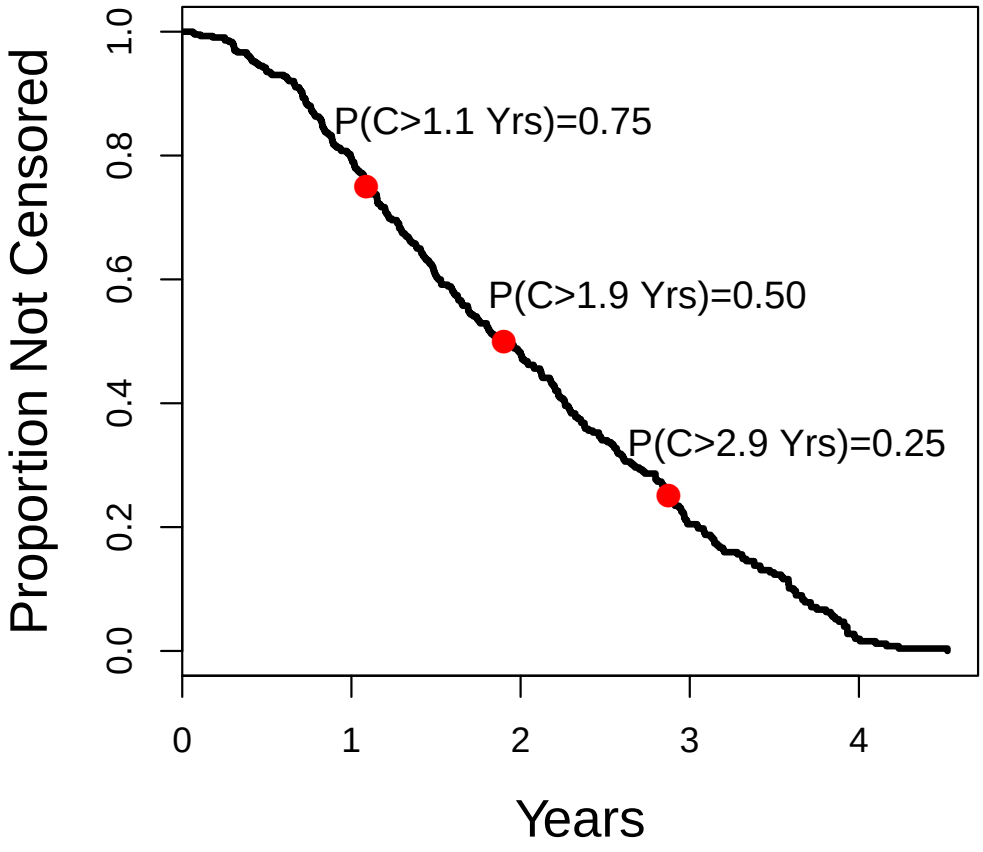

Figure 1.3: $\hat{K}$(t) for ICD group participants in the MADIT-II study.

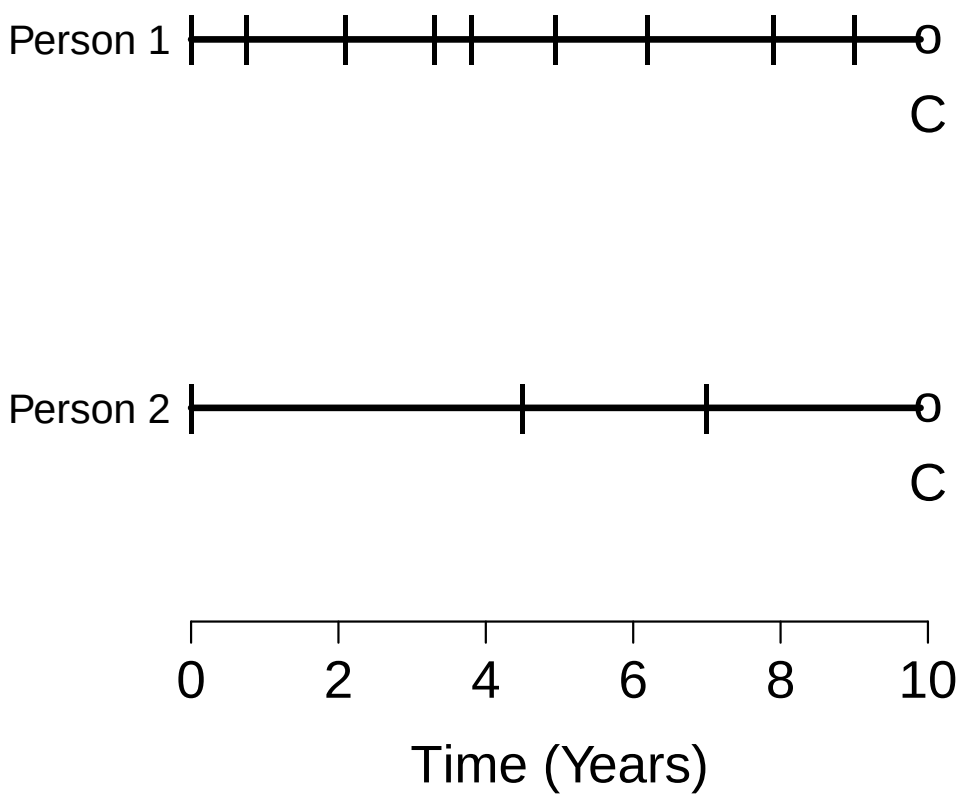

Figure 1.4: Informative censoring for recurrent events.

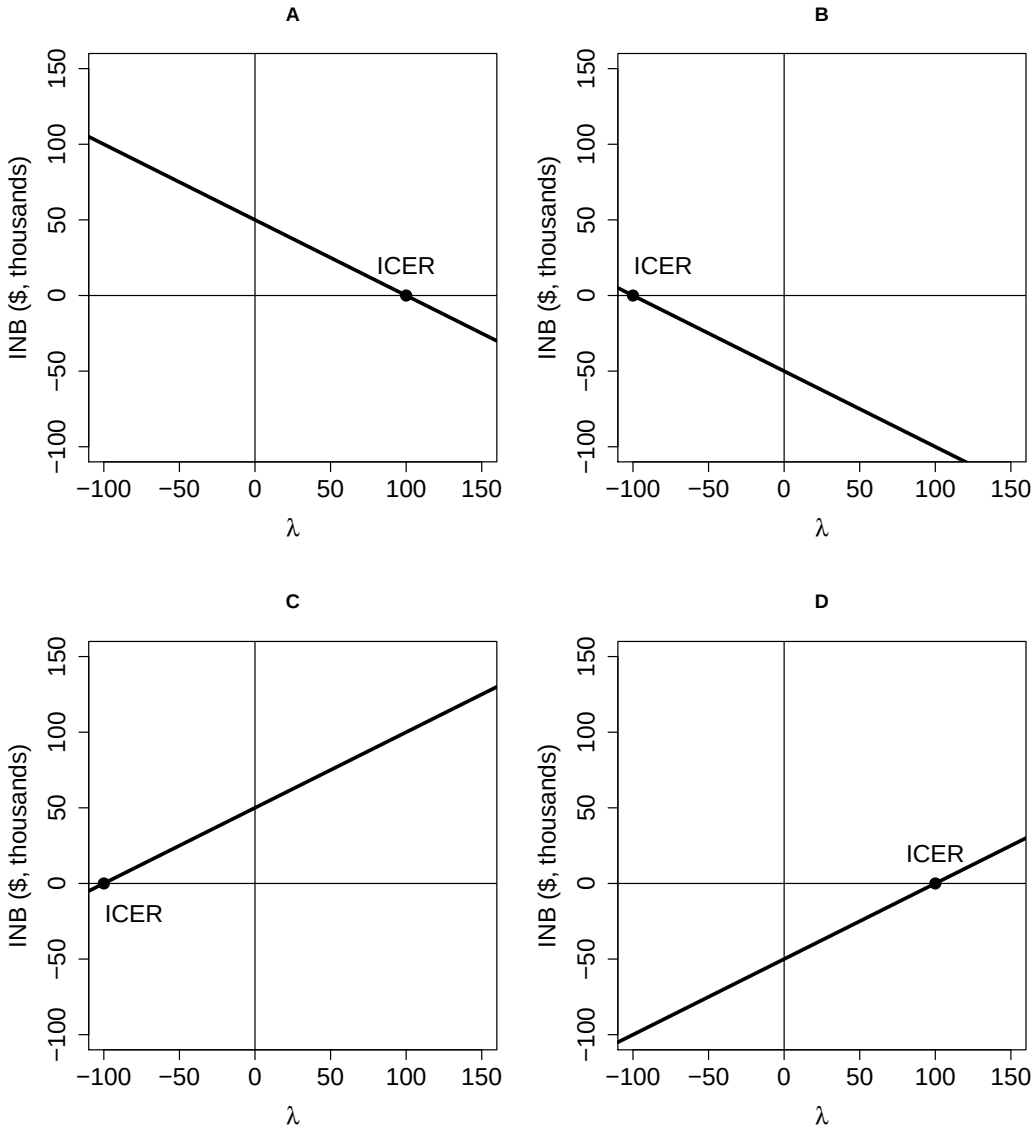

Figure 1.5: Incremental net benefit ($, in thousands) versus dollar amount (in thousands) per unit of effectiveness (λ) for the four possible situations of cost-effectiveness.

Chapter 2

NONPARAMETRIC INFERENCE FOR THE JOINT DISTRIBUTION OF RECURRENT MARKED VARIABLES AND RECURRENT SURVIVAL TIME

2.1 Introduction

Cost data are being collected more frequently in order to assess cost-effectiveness of new treatments. In particular, more studies are collecting longitudinal cost data, where cost information is collected at random medical event times. As such data becomes more common, it may be of interest to characterize the time between recurring events, and the medical cost associated with these events. For example, the time between recurring hospitalizations might be correlated with the cost incurred at each hospitalization. In such a situation, it may be of interest to inspect the joint distribution of recurrent hospitalization time and recurrent hospitalization cost.

Medical cost is known as a marked variable. When lifetime cost is of interest, we say that lifetime cost is a mark of survival. This means that lifetime cost is only observed at an event time when a patient has not been censored. In the presence of right censoring, a complication called induced informative censoring occurs on the cost scale (Glasziou et al., 1990; Lin et al., 1997; Huang and Louis, 1998; Strawderman, 2000; Bang and Tsiatis, 2000). For example, a patient who is quite sick might accumulate medical cost more quickly, and therefore his or her cumulative cost would tend to be elevated at both the censoring time and the survival time, even when the two times are independent.

Time between events that generate medical cost may be observed in a recurrent fashion. After enrollment due to an initial event (such as hospitalization due to myocardial infarction), events (such as subsequent hospitalization) incurring medical cost will recur periodically throughout the course of study. Recurrent time data also has the complication of informative censoring (Wang and Chang, 1999; Pena et al., 2001). In this situation, a patient who is quite sick will have more hospitalizations, and therefore the time between his

or her last recurrence time and censoring time is likely to be shorter. In addition, recurrent cost will be a marked variable for recurrent survival.

In previous work, Huang and Louis (1998) presented a nonparametric estimator for the joint distribution of non-recurrent survival time and marked variables. Huang and Louis used a survival analysis formulation to define a cumulative mark specific hazard which did not rely upon an assumption of non-informative censoring on the mark scale. In addition, they derived the asymptotic variance of their estimator using a martingale approach. We plan to extend the method of Huang and Louis to account for recurrent survival time and recurrent marked variables.

Focusing on recurrent event times, several authors have discussed nonparametric estimation for the recurrent survival distribution, which have been discussed in a recent review paper by Zhu (2013). In particular, Wang and Chang (1999) used a survival analysis approach to define a recurrent cumulative hazard function, which incorporated mass and risk sets which were weighted within an individual's recurrent events before weighting across individuals. Their method allowed for within-person correlation amongst recurrent survival times, which handles informative censoring of recurrent events. The approach proposed by Wang and Chang has been extended to estimate the distribution of alternating recurrent events (Huang and Wang, 2005). In addition, Luo and Huang (2011) proposed a resampling method in order to analyze recurrent events, which was shown to be equivalent to the method of Wang and Chang when the number of resamples tends to infinity. We will extend the method of Wang and Chang, along with the work of Huang and Louis, to define the estimator for our joint distribution of recurrent survival time and marked variables.

In this paper, we present a nonparametric estimator for the joint distribution of recurrent survival time and marked variables. We will also present its asymptotic variance, simulation results, and an example based on implantation of cardioverter-defribrillators in the second Multicenter Automatic Defibrillator Implantation Trial (MADIT-II).

2.2 Methods

2.2.1 Notation

For individual $i, i = 1, \ldots, n$, we record recurrent events j, with the time of the initial event set to 0. Define T_{ij} to be recurrent event times. That is, T_{ij} denotes the gap time between the $(j-1)$th event to the jth event for the ith individual. Let C_i be the censoring time for individual i, and let M_{ij} be the observed marked variable, which is observed if

$$\sum_{l=1}^{j} T_{il} \leq C_i.$$

Adopting notation similar to that of Wang and Chang (1999), let k_i be the index indicating the censoring time. That is, person i's last observed recurrent event is indexed by $j = (k_i - 1)$, where

$$\sum_{j=1}^{k_i-1} T_{ij} \leq C_i, \qquad \text{and} \qquad \sum_{j=1}^{k_i} T_{ij} > C_i.$$

Since we only observe marked variables and recurrent events before C_i, we define Y_{ij} to be the observed recurrent event times as follows:

$$Y_{ij} = \begin{cases} T_{ij} & \text{if } j = 1, ..., k_i - 1 \\ T^{+}_{ik_i} & \text{if } j = k_i \end{cases}$$

where $T^{+}_{ik_i}$ is the time from the $(j-1)$th event to C_i. We additionally make two assumptions:

1. There is a subject-level random variable X such that for all individuals i, the recurrent event times and marked random variables (T_{i1}, M_{i1}), (T_{i2}, M_{i2}), ... are independent and identically distributed given X_i.

2. The censoring time, C_i, is independent of $(X_i, \{T_{i1}, M_{i1}\}, \{T_{i2}, M_{i2}\}, \ldots)$.

The first assumption therefore allows for dependence of the recurrent data within a subject, while the second assumption is the independent censoring condition. Like the standard

independent censoring assumption, this second assumption would be violated if participants drop out of the study due to illness. In this case, participants would be exiting the study right before or during events that would also generate high medical cost.

2.2.2 Estimation of the joint marked recurrent distribution

Let us first define the parameters that we will employ in the definition of our joint distribution. Let $S(t)$ be the recurrent survival function, $S(t) = P(T_{ij} > t)$. Once again following the definitions of Wang and Chang (1999), define $H(t) = E[I(T_{i1 \geq t})I(C_i \geq t)]$ and $F(t) = E[I(T_{i1} \leq t)I(T_{i1} \leq C_i)]$. Now, let $\Lambda(t, m)$ be the joint cumulative hazard for recurrent events and the marked variable. As noted by Huang and Louis (1998), $\Lambda(t, m)$ can be written as follows:

$$\Lambda(t, m) = \int_{[0,t]} \frac{F(ds, u)}{H(s)}, \tag{2.1}$$

where $F(t, m) = E[I(T_{i1} \leq t, M_{i1} \leq m)I(T_{i1} \leq C_i)]$ and $F(t, \infty) = F(t) = 1 - S(t)$. $F(t, m)$ is therefore the cumulative joint distribution function of recurrent events and recurrent marked variables amongst uncensored participants.

Then the goal is to estimate the joint distribution of the recurrent event times and the marked variable, $G(t, m)$:

$$G(t, m) = P(T_{ij} \leq t, M_{ij} \leq m) = \int_0^t S(s)\Lambda(ds, m). \tag{2.2}$$

To estimate $G(t, m) = \int_0^t S(s)\Lambda(ds, m)$, we will first focus on estimating $\Lambda(t, m)$ of equation (2.1) under a recurrent event setting. As defined by Wang and Chang (1999), an estimator for $H(t)$, the weighted mean risk set, is defined as follows:

$$\hat{H}(t) = \frac{1}{n} \sum_{i=1}^{n} \left[\frac{1}{k_i^*} \sum_{j=1}^{k_i^*} I(y_{ij} \geq t) \right],$$

where $k_i^* = \max(k_i\text{-}1,\ 1)$. Now, extending the work of Wang and Chang (1999) to include

marked variables, we let

$$\hat{F}(t,m) = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{I(k_i \geq 2)}{k_i^*}\sum_{j=1}^{k_i^*} I(y_{ij} \leq t, m_{ij} \leq m)\right]$$

be the mass at recurrent event time t and marked variable value m, first averaging over a person's recurrent events. Similarly to Wang and Chang (1999), this estimate excludes the last recurrent event time since we only sum up to k_i* for participants for whom I(k_i $\geq$2). Then, an estimator for $\hat{\Lambda}(t,m)$ can be defined as follows:

$$\hat{\Lambda}(t,m) = \int_{[0,t]} \frac{\hat{F}(ds,m)}{\hat{H}(s)}.$$

To estimate $G(t,m)$, we also need to estimate $S(t)$. Note that $\hat{\Lambda}(t,\infty) = \hat{\Lambda}(t)$, where $\hat{\Lambda}(t)$ is the cumulative hazard estimator of recurrent event times proposed by Wang and Chang (1999):

$$\hat{\Lambda}(t) = \int_{[0,t]} \frac{\hat{F}(ds)}{\hat{H}(s)},$$

where $\hat{F}(t) = \hat{F}(t,\infty)$. Now, we let $\hat{S}(t)$ be the estimated recurrent survival function by Wang and Chang (1999), where $\hat{S}(t) = \prod_{[0,t]}[1 - \hat{\Lambda}(ds)]$. Finally, we propose to estimate $G(t,m)$ by plugging in $\hat{\Lambda}(t,m)$ for $\Lambda(t,m)$, and $\hat{\Lambda}(t)$ for $\Lambda(t)$ in the product-limit estimator of $S(t)$. Then our estimator for the joint distribution of recurrent survival time and recurrent marked variables, $\hat{G}(t,m)$, is written as follows:

$$\hat{G}(t,m) = \int_{[0,t]} \prod_{[0,s)} [1 - \hat{\Lambda}(dv)]\hat{\Lambda}(ds,m).$$

When there is only a single survival event, $\hat{S}(t)$ reduces to the Kaplan-Meier estimator. Furthermore, in this case, $\hat{\Lambda}(t,m)$ and $\hat{G}(t,m)$ reduce to the corresponding estimators proposed by Huang and Louis (1998).

The estimated recurrent survival distribution, $\hat{S(t)}$, is identifiable on [0,τ], where τ = sup(t : $S(t) > 0$) can be thought of as the end of study follow-up. As was discussed by

Huang and Louis (1998), although we are not able to identify the full survival distribution, the joint distribution, $\hat{G}(t,m)$, is identifiable for the region [0,τ] x $\mathbb{R}$.

2.2.3 Variance estimation

We shall now discuss the asymptotic iid representation of $\hat{G}(t,m)-G(t,m)$ and the resulting estimator for the variance of $\hat{G}(t,m)$. In the Appendix, we use the functional delta method (Van der Vaart, 1998) to show that there exists a function $\eta_i(t,m)$ such that

$$(\hat{G}(t,m) - G(t,m)) = \frac{1}{n}\sum_{i=1}^{n}\eta_i(t,m) + o_p(1/\sqrt{n}) \text{ , and}$$
$$\eta_i(t,m) = \frac{1}{n}\sum_{i=1}^{n}\left[\int_0^t S(s-)\phi_i(ds,m) + \int_0^t \phi_i^{WC}(s)\Lambda(ds,m)\right] + o_p\left(\frac{1}{\sqrt{n}}\right).$$

In the above equations, $\phi(t,m)$ and $\phi_i^{WC}(t)$ are iid asymptotic representations of $(\hat{\Lambda}(t,m) - \Lambda(t,m))$ and $(\hat{S}(t) - S(t))$, respectively. The formulations of $\phi_i(t,m)$ and $\phi_i^{WC}(t)$ are also included in the Appendix. In addition, a consistent estimator for the variance of $\hat{G}(t,m)$ can be written as follows:

$$\begin{aligned}\widehat{Var}[\hat{G}(t,m)] &= \frac{1}{n}\sum_{i=1}^{n}\left[\int_0^t \hat{S}(s-)\hat{\phi}_i(ds,m) + \int_0^t \hat{\phi}_i^{WC}(s)\hat{\Lambda}(ds,m)\right]^2 \\ &= \frac{1}{n}\sum_{i=1}^{n}[\hat{\eta}_i(t,m)]^2,\end{aligned}$$

where $\hat{\eta}_i(t,m)$, $\hat{\phi}_i(t,m)$, and $\hat{\phi}_i^{WC}(t)$ are also given in the Appendix.

2.3 Simulation

We used numerical simulation to illustrate the performance of the proposed methods in Section 2.2. We employed simulation parameters based on the simulation experiments conducted by Huang and Louis (1998) and Wang and Chang (1999). Simulation results were based on 1,000 replications.

We generated n=200 or n=400 uniform(0,1) distributed frailty values, x_i, $i = 1,\ldots,n$. Next, correlated sets of recurrent survival times and marked variables were generated using

Frank's bivariate family, a copula model (Genest, 1987). To generate data according to this model, we first sampled independent (u_{ij}, v_{ij}), where $i = 1, \ldots, n$ and $j = 1, \ldots$ were standard uniform distributed. Next, we set $R_{ij} = \alpha^{u_{ij}} + (\alpha - \alpha^{u_{ij}})\, v_{ij}$, where $\alpha = \alpha(x_i) = exp(c \times x_i)$. The constant c allowed for different levels of correlation between t_{ij} and m_{ij}. Finally, we defined $t_{ij} = 2\left(-\frac{1}{x_i} log(1 - u_{ij})\right)^{1/1.5}$ and $m_{ij} = log_\alpha\left[R_{ij}/(R_{ij} + (1-\alpha)v_{ij})\right]$. This resulted in recurrence times, t_{ij}, that were Weibull distributed with the survival function $exp(-x_i(t/2)^{1.5})$, which was the survival setup used by Wang and Chang (1999). This also resulted in recurrent marked variables, m_{ij}, that were standard uniform distributed, as studied by Huang and Louis (1998). In our simulations, we let c vary between -10, 0, and 10, which corresponded to Kendall's rank correlations between t_{ij} and m_{ij} of 0.32, 0, and -0.32.

A summary of the simulation results can be found in Table 2.1. We summarize at values of t and m, where t is equal to 1.0 or 2.0, and m takes values in 0.25, 0.5, 0.75, and 1.0. For each pair (t, m), we present three rows of information: 1) the true value ($\times 10^3$) followed by the empirical variance of the estimators ($\times 10^5$); 2) our estimated value ($\times 10^3$) followed by our estimated variance ($\times 10^5$); and 3) the empirical coverage of the 95% confidence intervals. From our results, we see that the bias of the estimator is small in all situations. The estimated variance is also close to the empirical variance for all t and m, confirming the validity of our derived asymptotic variance formula. The empirical coverage is close to nominal levels for nearly all m and t. However, empirical coverage is quite poor for small values of both m and t in the case of c=10. This is due to the fact that there was nearly no mass in the joint distribution for these values.

2.4 Second Multicenter Automatic Defibrillator Implantation Trial (MADIT-II)

We illustrate our methods using data from the second Multicenter Automatic Defibrillator Implantation Trial (MADIT-II) (Moss et al., 2002). MADIT-II was a clinical trial that aimed to prevent sudden cardiac death in patients at high risk for cardiac arrhythmia. Participants were randomized to receive an implantable cardioverter-defibrillator (ICD) or standard treatment. Cost and healthcare utilization data were also collected, and were

reported by Zwanziger et al. (2006).

Conducted from July 11, 1997 to January 16, 2002, participants were considered for inclusion in the study if they had a previous myocardial infarction and had a left ventricular ejection fraction less than 0.30. Investigators enrolled 1,232 participants from 76 hospitals in the US and Europe. Zwanziger and others reported cost-effectiveness results for 664 ICD participants and 431 standard care participants (Zwanziger et al., 2006). This cost study further excluded participants from European hospitals and participants with large amounts of missing medical cost data. We demonstrate our methods using the 1,095 participants and data presented by Zwanziger and others. Further information about the study has been documented by the original investigators (Moss et al., 2002; Zwanziger et al., 2006).

In order to mimic hospitalization costs, equal consecutive daily costs were summed and any resulting cost events less than \$500 were omitted. In addition, the start time in the ICD group was set to the day after each participant's ICD surgery was completed, since we were interested in studying the joint distribution of recurrent medical cost and survival time after ICD implantation. For this analysis, gap times were defined to be the time between the previous admission time and the current admission time. We assigned medical costs for an entire hospitalization to the date of hospital admission.

Participants were followed for up to 4.5 years. Observed follow-up time was 1.8 years (SD=1.08 years) and 1.7 years (SD=1.05 years), on average, for the ICD group and the standard treatment group. On the person level, we observed an average of 4.2 (SD=6.4) and 2.9 (SD=5.5) number recurrent hospitalizations amongst ICD participants and standard treatment participants, respectively.

Table 2.2 summarizes the results of our analysis of the MADIT-II data. The estimate of the joint distribution and the associated standard deviation (in parentheses) are reported for a variety of recurrent survival times and recurrent costs (t=1, 2, 3 years and m=\$1000, \$2000, \$7000, \$100000). In addition, we compare our proposed joint distribution (left) to the method of Huang and Louis (1998) applied only to the first recurrent event (right).

Referring to the results of our proposed method (labeled "All Recurrences"), we see that standard care participants have relatively few low recurrent cost events for any length of recurrent survival time. We see that after ICD implantation, participants in the ICD group

tended to have a higher proportion of low recurrent costs and more frequent recurrent times. In other words, for the ICD patients, we observe a joint distribution that is more concentrated in the region of low recurrent costs and shorter recurrence times. Now we turn to the results employing the estimator of Huang and Louis (1998) (labeled "First Recurrence"), which only uses the first recurrent hospitalization time and first recurrent medical cost. The point estimates in this case are similar to the point estimates of our proposed method, but as expected, the estimated standard deviations are larger since information on subsequent recurrences are not utilized. Compared to the method of Huang and Louis (1998), our estimator reduces up to 33% of the estimated standard deviation in this example.

2.5 Discussion

This paper discussed a method to estimate the joint distribution between recurrent survival time and a recurrent marked variable. We presented a nonparametric estimator and variance estimator, and illustrated our methods via simulation and a data example.

A potential limitation of this method is that since we are dealing with a joint distribution including a marked variable (here, medical cost), we are unable to easily consider death as part of a composite event if the participant did not have a corresponding medical cost associated with the death. In addition, our data application does not account for length of hospital stay. This could affect our analyses by requiring longer gap times for participants who have extended stays in the hospital since they are not at risk for a readmission at that time. We also assigned medical costs for the entire hospitalization to the participant's admission date. This could affect our analyses by attributing larger costs to participants with the aforementioned longer gap times.

Our proposed method assumes the same joint distribution across recurrences. In our analysis of the MADIT-II data, we saw that estimates across all recurrences were similar to those using only the first recurrence, indicating that this assumption was not violated. However, this assumption may not be tenable in some situations, especially when recurrences may be viewed as progression of disease (Wang and Chen, 2000). As such, future work may extend this framework to a more general setting such as multivariate recurrences.

2.6 Appendix

We wish to find the influence function of $\hat{G}$(t,m). To do so, we will first find the influence function of $\hat{\Lambda}$(t,m), and then use those results to get the final influence function of interest. We use the same notation and setup of Section 2.2 of the paper.

2.6.1 Influence Function: $\hat{\Lambda}(t,m)$

First, note that $\hat{\Lambda}(t,m) = \int_{[0,t]} \hat{F}(ds,m)/\hat{H}(s)$ depends on the pair ($\hat{F}$(t,m), $\hat{H}$(t)) through two maps:

$$(A,B) \to (A, \frac{1}{B}) \to \int_0^t \frac{1}{B} dA.$$

Then, the functional derivative of the maps at (F,H) is:

$$(\alpha,\beta) \to (\alpha, -\frac{\beta}{H^2}) \to \int_0^t \frac{d\alpha}{H} - \int_0^t \frac{\beta dF}{H^2},$$

where $\alpha = \hat{F}$-F and $\beta = \hat{H}$-H. Now, by the functional delta method (Van der Vaart, 1998) and simplification we have:

$$\begin{aligned}
\hat{\Lambda}(t,m) - \Lambda(t,m) &= \int_0^t \frac{1}{H(s)}[\hat{F}(ds,m) - F(ds,m)] - \int_0^t \frac{\hat{H}(s) - H(s)}{(H(s))^2} F(ds,m) + o_p\left(\frac{1}{\sqrt{n}}\right) \\
&= \int_0^t \frac{\hat{F}(ds,m)}{H(s)} - \Lambda(t,m) - \int_0^t \frac{\hat{H}(s)F(ds,m)}{(H(s))^2} + \Lambda(t,m) + o_p\left(\frac{1}{\sqrt{n}}\right) \\
&= \frac{1}{n}\sum_{i=1}^n \left[\frac{I(k_i \geq 2)}{k_i^*} \sum_{j=1}^{k_i^*} \left(I(y_{ij} \leq t, m_{ij} \leq m) \frac{1}{H(y_{ij})} \right) - \int_0^t \frac{F(ds,m)}{(H(s))^2} \frac{1}{k_i^*} \sum_{j=1}^{k_i^*} I(y_{ij} \geq s) \right] \\
&\qquad + o_p\left(\frac{1}{\sqrt{n}}\right) \\
&= \frac{1}{n}\sum_{i=1}^n \phi_i(t,m) + o_p\left(\frac{1}{\sqrt{n}}\right)
\end{aligned}$$

where

$$\phi_i(t,m) = \frac{I(k_i \geq 2)}{k_i^*} \sum_{j=1}^{k_i^*} \left(I(y_{ij} \leq t, m_{ij} \leq m) \frac{1}{H(y_{ij})} \right) - \int_0^t \frac{F(ds,m)}{(H(s))^2} \frac{1}{k_i^*} \sum_{j=1}^{k_i^*} I(y_{ij} \geq s)$$

is the influence function of $\hat{\Lambda}$(t,m) - Λ(t,m). Then a consistent estimator for ϕ_i(t,m), $\hat{\phi}_i$(t,m), can be calculated by plugging in $\hat{H}(t)$ and $\hat{F}(t,m)$ for $H(t)$ and $F(t,m)$.

2.6.2 Influence Function: $\hat{G}$(t,m)

Now, we wish to find the influence function for $\hat{G}$(t,m), the estimated joint distribution of recurrent survival time and marked variable, where $\hat{G}(t,m) = \int_0^t \hat{S}(s-)\hat{\Lambda}(ds,m)$. The influence function for the recurrent survival function, $\hat{S}(t)$, as presented in Wang and Chang, is written as follows:

$$\phi_i^{WC}(t) = S(t) \left[\int_0^t \frac{1}{H^2(u)} \left(\frac{1}{k_i^*} \sum_{j=1}^{k_i^*} I(y_{ij} \geq u) \right) F(du) - \frac{I(k_i \geq 2)}{k_i^*} \sum_{j=1}^{k_i^*} \frac{I(y_{ij} < t)}{H(y_{ij})} \right]$$

Now, completing steps similar to those in the previous section of the Appendix, we use both $\phi_i(t,m)$ and $\phi_i^{WC}(t)$ to obtain the influence function for $G(t,m)$:

$$\begin{aligned}
\hat{G}(t,m) - G(t,m) &= \int_0^t S(s-) \left[\hat{\Lambda}(ds,m) - \Lambda(ds,m) \right] + \int_0^t \left[\hat{S}(s-) - S(s-) \right] \Lambda(ds,m) + o_p\left(\frac{1}{\sqrt{n}}\right) \\
&= \frac{1}{n} \sum_{i=1}^{n} \left[\int_0^t S(s-)\phi_i(ds,m) + \int_0^t \phi_i^{WC}(s)\Lambda(ds,m) \right] + o_p\left(\frac{1}{\sqrt{n}}\right) \\
&= \frac{1}{n} \sum_{i=1}^{n} \eta_i(t,m) + o_p\left(\frac{1}{\sqrt{n}}\right)
\end{aligned}$$

Then an estimator for $\eta_i(t,m)$, $\hat{\eta}_i(t,m)$, can be estimated by plugging in $\hat{S}(t)$, $\hat{\phi}_i(t,m)$, $\hat{\phi}_i^{WC}(t)$, and $\hat{\Lambda}(t,m)$ for $S(t)$, $\phi_i(t,m)$, $\phi_i^{WC}(t)$, and $\Lambda(t,m)$.

Table 2.1: Simulation results for the joint distribution of recurrent events and recurrent marked variables for each time (t=1 or 2) and marked category (m=0.25, 0.5, 0.75, or 1): row 1– true values $\times 10^3$ (empirical variance $\times 10^5$), row 2– empirical means of the estimator $\times 10^3$ (empirical means of the variance estimator $\times 10^5$), and row 3– empirical coverage at the 95% nominal level.

			m=0.25	m=0.5	m=0.75	m=1
C=-10						
	N=200	t=1	101 (43)	142 (57)	154 (62)	158 (63)
			101 (45)	143 (61)	154 (66)	158 (67)
			0.95	0.96	0.96	0.96
		t=2	164 (66)	287 (100)	348 (113)	368 (115)
			164 (68)	286 (101)	348 (112)	367 (115)
			0.95	0.95	0.94	0.95
	N=400	t=1	101 (21)	142 (31)	154 (33)	158 (33)
			101 (23)	142 (31)	154 (33)	158 (34)
			0.95	0.95	0.95	0.95
		t=2	164 (35)	287 (54)	348 (61)	368 (61)
			164 (34)	287 (51)	349 (56)	368 (58)
			0.94	0.94	0.95	0.95
C=0						
	N=200	t=1	39 (15)	79 (33)	118 (50)	158 (63)
			39 (16)	78 (31)	118 (46)	158 (61)
			0.93	0.93	0.94	0.95
		t=2	92 (39)	184 (71)	276 (97)	368 (115)
			92 (38)	183 (69)	275 (95)	367 (114)
			0.93	0.95	0.95	0.95
	N=400	t=1	39 (8)	79 (18)	118 (26)	158 (33)
			39 (8)	79 (16)	118 (23)	158 (31)
			0.93	0.93	0.94	0.94
		t=2	92 (20)	184 (40)	276 (54)	368 (61)
			92 (19)	184 (35)	276 (48)	368 (58)
			0.95	0.94	0.93	0.95
C=10						
	N=200	t=1	4 (2)	16 (7)	57 (25)	158 (63)
			4 (2)	15 (7)	57 (23)	158 (61)
			0.59	0.86	0.93	0.95
		t=2	20 (10)	81 (36)	204 (76)	368 (115)
			20 (9)	81 (35)	203 (75)	367 (114)
			0.89	0.93	0.95	0.95
	N=400	t=1	4 (1)	16 (3)	57 (13)	158 (33)
			4 (1)	15 (3)	57 (12)	158 (31)
			0.76	0.91	0.93	0.94
		t=2	20 (4)	81 (18)	204 (42)	368 (61)
			20 (5)	81 (18)	205 (38)	368 (58)
			0.93	0.94	0.94	0.95

Table 2.2: Results from the second Multicenter Automatic Defibrillator Implantation Trial (MADIT-II). Proposed estimation of the joint distribution of recurrent hospitalization time (at t=1, 2, or 3 years) and recurrent medical cost (at \$1,000, \$2,000, \$7,000, or \$100,000) for the proposed method (left) or the method of Huang and Louis applied only to the first recurrence (right).

	Recurrent Cost (\$, thousands)							
	All Recurrences				First Recurrence (H&L)			
Yrs	1	2	7	100	1	2	7	100
Standard								
1	0.148	0.277	0.471	0.616	0.161	0.275	0.45	0.598
	(0.017)	(0.02)	(0.022)	(0.019)	(0.025)	(0.03)	(0.03)	(0.024)
2	0.177	0.329	0.579	0.756	0.194	0.334	0.568	0.76
	(0.025)	(0.029)	(0.034)	(0.028)	(0.034)	(0.04)	(0.044)	(0.035)
3	0.209	0.368	0.623	0.812	0.226	0.38	0.621	0.83
	(0.042)	(0.046)	(0.052)	(0.053)	(0.049)	(0.061)	(0.068)	(0.073)
ICD								
1	0.268	0.365	0.542	0.714	0.252	0.336	0.521	0.687
	(0.025)	(0.025)	(0.023)	(0.014)	(0.029)	(0.032)	(0.028)	(0.017)
2	0.321	0.438	0.639	0.824	0.311	0.427	0.639	0.819
	(0.041)	(0.048)	(0.054)	(0.056)	(0.047)	(0.058)	(0.063)	(0.062)
3	0.366	0.486	0.694	0.887	0.367	0.492	0.717	0.91
	(0.087)	(0.096)	(0.108)	(0.124)	(0.111)	(0.13)	(0.149)	(0.185)

Chapter 3

NONPARAMETRIC INFERENCE FOR TIME-DEPENDENT INCREMENTAL COST-EFFECTIVENESS RATIOS

3.1 Introduction

As medical costs rise, it becomes increasingly important to be able to assess the cost of effective treatments. The incremental cost-effectiveness ratio (ICER) is one measure of cost-effectiveness that has been widely used by biomedical researchers (Mushlin et al., 1998; Salukhe et al., 2004; Zwanziger et al., 2006). In the setting of randomized clinical trials, the ICER evaluates the difference in mean cost between a new treatment and standard practice, per unit of effectiveness gained by the new treatment. Two descriptors of effectiveness that are commonly employed are mean survival time and probability of survival. These ICERs can be interpreted as the additional cost per year of life saved or the additional cost per 100-percentage point increase in survival probability, respectively.

The ICER has typically been evaluated at the end of follow-up (Zhao and Tian, 2001; Willan et al., 2002). However, some biomedical researchers have noted that time-dependent cost-effectiveness may be of interest (van de Wetering et al., 2012). The problem of cost-effectiveness changing over time has been highlighted in the implantable cardioverter-defibrillator (ICD) literature (Mushlin et al., 1998; Salukhe et al., 2004; Zwanziger et al., 2006), but to our knowledge, the associated statistical problems have not yet been addressed explicitly. An example of a study with a time-varying cost-effectiveness is the second Multicenter Automatic Defibrillator Implantation Trial (MADIT-II), which is the study we will use to illustrate our proposed methods (Moss et al., 2002). In this case, the initial cost of intervention is high, but benefits accumulate long after implantation. In this research, we plan to address the problem of changing cost-effectiveness over time by evaluating a nonparametric estimator, $\widehat{ICER}(\mathrm{t})$.

When cost data are right-censored, we encounter a problem known as induced informative censoring. Induced informative censoring has been documented at length on the cost scale (Lin et al., 1997; Bang and Tsiatis, 2000; Strawderman, 2000), and has also been studied in other contexts, such as quality-adjusted life and mean number of recurrent events (Glasziou et al., 1990; Zhao and Tsiatis, 1997; Cook and Lawless, 1997; Strawderman, 2000). The nonparametric estimator proposed by Cook and Lawless estimated the mean number of recurrent events by weighting number of recurrent events by the probability of being censored (Cook and Lawless, 1997). This estimator can be applied to cost data, and is a continuous version of the well-known partitioned estimator of Lin et al. (1997). We employ the estimator of Cook and Lawless, and derive analytical variance and covariance for the corresponding ICER functions using empirical process theory.

Once we have estimators for the corresponding variance and covariance components, we turn to creating pointwise intervals for the ICER. Some authors have discussed normal-based approximations, which yield symmetric intervals (O'Brien et al., 1994; Chaudhary and Stearns, 1996; Polsky et al., 1997; Briggs et al., 1999). These intervals have drawbacks such as lower coverage probability when the ICER has a skewed distribution. For this reason, some authors have applied the Fieller interval (Fieller, 1954; Willan and O'Brien, 1996; Chaudhary and Stearns, 1996; Polsky et al., 1997; Briggs et al., 1999; Zhao and Tian, 2001; Willan et al., 2002). The Fieller interval is not symmetric, and has been shown to have good coverage probabilities in many situations (Willan and O'Brien, 1996; Polsky et al., 1997; Briggs et al., 1999). We will compare both types of intervals for our method.

Since we are interested in making inference over a period of time, we will construct simultaneous confidence bands. Lin et al. (1994) proposed a method for creating confidence bands for proportional hazards regression. They used a multiplier bootstrap algorithm to approximate the limiting distribution of the survival function. We extend the method of Lin, Fleming, and Wei to construct confidence bands for our estimator of ICER(t).

The goal of this paper is to introduce a nonparametric estimator for the ICER over

time, ICER(t). In section 3.2, we will discuss estimation, interval estimation, and how to evaluate confidence bands. Section 3.3 will demonstrate the performance of our estimators via simulation. In section 3.4, we will present a data example using the MADIT-II data.

3.2 Methods

3.2.1 Notation

For individual i, $i = 1, ..., n$, we define T_i to be the survival time, C_i to be the censoring time, and $X_i = \min(T_i, C_i)$. Let $Y_i(t)$ be the observed cumulative medical cost up to time t for individual i. That is, we define $Y_i(t) = Y_i^0(\min(t, C_i))$, where Y_i^0 is the ith individual's true cumulative medical cost up to time t if no censoring were to occur. Let $\mu_Y(t) = E[Y^0(t)]$ be the true mean cumulative cost up to time t. Furthermore, let $\hat{Y}(t) = \frac{1}{n}\sum_{i=1}^{n} Y_i(t)$ be the sample mean of individual observed cumulative medical costs up to time t, and $Y(t) = E[\hat{Y}(t)]$ be the mean observed cumulative cost up to time t. Assume that $(T_i, Y_i(\cdot))$ are independent of C_i. This requires that censoring is independent of survival time and observed medical cost, but allows T_i and $Y_i(\cdot)$ to be correlated.

We define $S(t) = P(T > t)$ to be the survival distribution, $K(t) = P(C > t)$ to be the survival distribution for the censoring times, $H(t) = E[I(T \geq t)I(C \geq t)]$ to be the average at risk population at time t, and $\Lambda_Y(t) = \int_0^t E\left[dY(s)|T \geq s\right] = \int_0^t E\left[dY(s)|X \geq s\right] = \int_0^t dY(s)/H(s-)$. We also define μ_T(t)=$\int_0^t S(u)du$ to be the restricted mean survival time at time t.

Since our focus is on randomized clinical trials, we define a randomization variable, R_i. R_i=1 if patient i was randomized to the new treatment, and R_i=0 for standard care. For notational ease, we will therefore employ the subscripts 1 and 0 to correspond to parameters and estimators which describe the treatment and standard groups, respectively.

In this paper, we evaluate cost-effectiveness only after it is possible for an effect of treatment to take place. Thus, we will restrict t to be larger than some clinically meaningful time after which treatment may start to take effect, τ^0. In the absence of a clinical recom-

mendation, τ^0 can be chosen such that the null hypothesis of equal effectiveness is rejected at a given level of significance. Furthermore, due to identifiability issues when estimating $\mu_Y(t)$ in the presence of censoring (Lin et al., 1997; Bang and Tsiatis, 2000), we also only evaluate cost-effectiveness up to some time τ, where τ is chosen to be at or near the end of study observation. Therefore, we restrict t to $0< \tau^0 \leq t \leq \tau$. In the following sections, we will provide estimators, variance estimators, and a method to estimate confidence bands for ICER(t).

3.2.2 Estimation

Our parameters of interest are $\text{ICER}_T(\text{t})$ and $\text{ICER}_S(\text{t})$, where

$$\mathbf{ICER_T(t)} = \frac{\mu_{Y_1}(t) - \mu_{Y_0}(t)}{\mu_{T_1}(t) - \mu_{T_0}(t)} \qquad \text{and} \qquad \mathbf{ICER_S(t)} = \frac{\mu_{Y_1}(t) - \mu_{Y_0}(t)}{S_1(t) - S_0(t)}.$$

We estimate $S(t)$ using the standard Kaplan-Meier estimator, $\hat{S}(t)$. Similarly, we estimate $\mu_T(t)$ by integrating the Kaplan-Meier estimator for survival times: $\hat{\mu}_T(t) = \int_0^t \hat{S}(u)du$. To estimate $\mu_Y(t)$, we employ the estimator of Cook and Lawless (1997), which was originally proposed for estimating number of recurrent events. A similar discrete-time estimator for mean medical cost was proposed by Lin et al. (1997). The weighted estimator of Cook and Lawless, adapted to estimate mean cost, can be written as follows:

$$\begin{aligned}
\hat{\mu}_Y(t) &= \int_0^t \frac{d\hat{Y}(u)}{\hat{K}(u-)} \\
&= \int_0^t \frac{\hat{S}(u-)}{\hat{H}(u)} d\hat{Y}(u) \qquad \text{since } \hat{S}(u-)\hat{K}(u-) = \hat{H}(u) \\
&= \int_0^t \hat{S}(u-) d\hat{\Lambda}_Y(u).
\end{aligned}$$

Here, $\hat{K}(t)$ is the Kaplan-Meier estimator with X_i as the observed time variable and $I(X_i = C_i)$ as the event indicator, where $I(\cdot)$ is the indicator function. Furthermore, $\hat{H}(t) = \frac{1}{n}\sum_{i=1}^n I(X_i \geq t)$ estimates the average observed risk set. Finally, $\hat{\Lambda}_Y(t) = \int_0^t \hat{Y}(ds)/\hat{H}(s-)$

can be estimated by plugging in the previously described estimators $\hat{H}(t)$ and $\hat{Y}(t)$. We can then use the above estimators to obtain $\widehat{ICER}_T(t) = (\hat{\mu}_{Y_1}(t) - \hat{\mu}_{Y_0}(t))/(\hat{\mu}_{T_1}(t) - \hat{\mu}_{T_1}(t))$ and $\widehat{ICER}_S(t) = (\hat{\mu}_{Y_1}(t) - \hat{\mu}_{Y_0}(t))/(\hat{S}_1(t) - \hat{S}_0(t))$.

3.2.3 Variance

Empirical process theory can be used to derive the asymptotic variances of $\widehat{ICER}_T$(t) and $\widehat{ICER}_S$(t). In this section, we present the consistent estimators for these asymptotic variances, which are derived in the Appendix.

For notational convenience, let $\hat{\Delta}_Y(t) = \hat{\mu}_{Y_1}(t) - \hat{\mu}_{Y_0}(t)$ be the difference in cost at time t, $\hat{\Delta}_S(t) = \hat{S}_1(t) - \hat{S}_0(t)$ be the difference in survival probability at time t, and $\hat{\Delta}_T(t) = \hat{\mu}_{T_1}(t) - \hat{\mu}_{T_0}(t)$ be the difference in mean survival time at time t. We will further simplify the difference in effectiveness measures to be $\hat{\Delta}_K(t)$ for measure K, K $\in \{S, T\}$, when appropriate. In the Appendix, we derive the influence functions of $\widehat{ICER}_K$(t). Based on the results, the variance estimators of $\widehat{ICER}_K$(t) take the form

$$\begin{aligned}\widehat{Var}[\widehat{ICER}_K(t)] &= \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\hat{\eta}_i^Y(t)}{\hat{\Delta}_K(t)} - \hat{\eta}_i^K(t)\left(\frac{\hat{\Delta}_Y(t)}{[\hat{\Delta}_K(t)]^2}\right)\right]^2 \\ &= \frac{1}{n}\sum_{i=1}^{n}\left[\hat{\varphi}_i^{ICER_K}(t)\right]^2,\end{aligned}$$

where $\hat{\eta}_i^Y$(t), $\hat{\eta}_i^T$(t), $\hat{\eta}_i^S$(t), and $\hat{\varphi}_i^{ICER_K}$(t) are the estimated influence functions for $\hat{\Delta}_Y(t)$, $\hat{\Delta}_T(t)$, $\hat{\Delta}_S(t)$, and $\widehat{ICER}_K$(t), respectively. The definitions and derivations of $\hat{\eta}_i^Y$(t), $\hat{\eta}_i^T$(t), $\hat{\eta}_i^S$(t), and $\hat{\varphi}_i^{ICER_K}$(t) are described in detail in the Appendix.

Let $z_{\alpha/2}$ be the desired standard normal critical value. Then the symmetric confidence interval for $\widehat{ICER}_K$(t) can therefore be formulated by normal approximation:

$$\widehat{ICER}_K(t) \pm z_{\alpha/2}\widehat{Var}^{\frac{1}{2}}[\widehat{ICER}_K(t)]/\sqrt{n}. \tag{3.1}$$

It has been noted that normal approximations tend to perform poorly when constructing

confidence intervals for the ICER since its distribution is typically skewed (Willan and O'Brien, 1996; Chaudhary and Stearns, 1996; Polsky et al., 1997; Briggs et al., 1999). Therefore, it is also of interest to estimate Fieller intervals, which have been shown to perform well in a variety of situations (Willan and O'Brien, 1996; Polsky et al., 1997; Briggs et al., 1999). To construct the Fieller intervals, we need to estimate the variances and covariances of $\hat{\Delta}_Y(t)$, $\hat{\Delta}_T(t)$, and $\hat{\Delta}_S(t)$. Once again using the empirical influence functions derived in the Appendix, for K $\in \{S, T\}$, we have that
$\hat{\sigma}^2_Y(t) \equiv \widehat{Var}[\hat{\Delta}_Y(t)] = \frac{1}{n}\sum_{i=1}^n [\hat{\eta}_i^Y(t)]^2, \quad \hat{\sigma}^2_K(t) \equiv \widehat{Var}[\hat{\Delta}_K(t)] = \frac{1}{n}\sum_{i=1}^n [\hat{\eta}_i^K(t)]^2$, and
$\hat{\sigma}_{YK}(t) \equiv \widehat{Cov}[\hat{\Delta}_Y(t), \hat{\Delta}_K(t)] = \frac{1}{n}\sum_{i=1}^n [\hat{\eta}_i^Y(t)\hat{\eta}_i^K(t)]$.

Then, as described by Zhao and Tian (2001), the pointwise Fieller intervals can be constructed by plugging the estimates into the following formula:

$$\frac{\hat{\Delta}_Y(t)\hat{\Delta}_K(t) - z_{\alpha/2}\hat{\sigma}_{YK}(t) \pm \left[\left(\hat{\Delta}_Y(t)\hat{\Delta}_K(t) - z^2_{\alpha/2}\hat{\sigma}_{YK}(t)\right)^2 - \left(\hat{\Delta}^2_Y(t) - z^2_{\alpha/2}\hat{\sigma}^2_Y(t)\right)\right]^{1/2}}{\hat{\Delta}^2_K(t) - z_{\alpha/2}\hat{\sigma}^2_K(t)}. \tag{3.2}$$

3.2.4 Confidence Bands

We can construct pointwise confidence intervals using either the symmetric approach of equation (3.1) or the Fieller approach of equation (3.2). Since we are describing the ICER over time, it is also of interest to assess simultaneous confidence bands so that we can make inference over the entire time span (τ^0, τ). We aim to replace the $z_{\alpha/2}$ of equations (3.1) and (3.2) with larger values, b_S and b_F, respectively. These larger critical values will yield the correct simultaneous coverage probabilities in each situation. We outline methods for computing b_S and b_F, which are based on a formulation of a confidence band for the Cox proportional hazards model (Lin et al., 1994). To compute b_S for the symmetric approach, we complete the following steps:

1. Generate random multipliers $\{G_i, i = 1, \ldots, n\}$, which are independent and standard

normally distributed. Define $W(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} G_i \varphi_i^{ICER_K}(t)/\hat{\sigma}(\varphi_i^{ICER_K}(t))$, where $\hat{\sigma}(\varphi_i^{ICER_K}(\mathrm{t}))$ is the estimated sample variance of $\varphi_i^{ICER_K}(\mathrm{t})$ at time t.

2. Repeat step 1 m times to obtain $W_l(t)$, $l = 1, \ldots, m$.

3. Obtain b_S, the (100-α) percentile of $\max_{(\tau^0,\tau)} \{|W_l(t)|\}$.

This approach is based on a normal assumption of the whole ratio, which may not hold since the ICER often has a skewed distribution, as noted above. Therefore, we also consider computing b_F, a critical value to employ with the Fieller confidence interval formula, which is based on separate normal approximations for the numerator and the denominator. We modify the above steps as follows:

1. Generate random multipliers $\{G_i, i = 1, \ldots, n\}$ and $\{J_i, i = 1, \ldots, n\}$, which are independent and standard normally distributed. Define $A(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} G_i \eta_i^Y(t)/\hat{\sigma}(\eta_i^Y(t))$ and
   $B^K(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} J_i \eta_i^K(t)/\hat{\sigma}(\eta_i^K(t))$, where $\hat{\sigma}$ once again denotes the estimated sample variance.

2. Repeat step 1 m times to obtain $A_l(\mathrm{t})$ and $B_l^K(\mathrm{t})$, $l = 1, \ldots, m$.

3. Obtain b_F, the (100-α) percentile of $\max_{(\tau^0,\tau)} \left\{|A_l(t)|, |B_l^K(t)|\right\}$.

3.3 Simulation

In this section, we present simulation studies which evaluate the proposed ICER(t), variance estimators, and the empirical coverage of the simultaneous confidence bands. We use a simulation setup similar to that of Zhao and Tian (2001), which has also been employed to some degree in other cost-related studies (Lin et al., 1997; Bang and Tsiatis, 2000; Strawderman, 2000; Chen and Zhao, 2013).

The total cost for each individual was composed of a diagnostic cost (incurred at the beginning of the study), a yearly cost (incurred randomly within each year), and an end

of life cost (incurred randomly in the year before death). For the treatment group, the diagnostic cost was uniform on [20000, 30000], the yearly cost was uniform on [1000, 2000], and the end of life cost was uniform on [5000, 10000]. For the standard group, the costs were uniform on [1000, 3000], uniform on [1000, 3000], and uniform on [5000, 10000], respectively.

We examined two different situations of survival time distributions. In the first situation, the survival time of the treatment group followed an exponential distribution with a mean of 8 years, while the standard group followed an exponential distribution with a mean of 5 years. In the second situation, the survival time of the treatment group followed an exponential distribution with a mean of 12.8 years, and the standard group followed a Weibull distribution with a mean of 7 years (scale=7.29, shape=1.11). The first situation is an example similar to that of Zhao and Tian (2001). The second situation is more realistic, and is based on the parameters obtained from a Weibull survival regression of the MADIT-II data.

We examined cost and survival up to τ=10 years. Censoring time was uniformly distributed on [0, 12.5], which resulted in censoring at 10 years of 49% for standard care and 65% for new treatment in the first survival time example, and 37% for standard care and 51% for new treatment in the second survival time example. A τ^0 of 1.5 years was chosen for both simulation examples. The results are based on 1000 simulation runs. Sample size was 1000 per treatment group. Confidence band critical values are based on 1000 resamples.

Table 3.1 presents $\text{ICER}_S(\text{t})$ and $\text{ICER}_T(\text{t})$ estimates, bias, pointwise estimated standard errror (ESE), pointwise sample standard error (SSE), and 95% coverage probability of the estimated confidence bands over time. In the first set of simulations–where $T_1 \sim \text{Exp}(8)$ and $T_0 \sim \text{Exp}(5)$–we see that our estimates are slightly biased, but the bias becomes less pronounced over time. The estimated standard error tends to be lower than the sample standard error early in the study, but our estimated standard error does quite well after the first few years of observation. We see similar patterns in the second set of simulations–where $T_1 \sim \text{Exp}(12.8)$ and $T_0 \sim$ Weib(scale=7.3, shape=1.1). Here, the estimates and standard

error estimates perform more poorly at 2 years, but perform well in subsequent years. The poorer performance in the earlier years is due to the small difference in effectiveness at those times, thus leading to unstable estimation of the ratios.

Coverage of the confidence bands varied depending on which type of interval (symmetric or Fieller) and critical value (b_S or b_F) were used. There was undercoverage in both simulation setups when we calculated the symmetric critical value, b_S, and plugged it into the symmetric interval to get confidence bands. Coverage improved when using the Fieller interval. We observed a slight undercoverage when we calculated bands using the symmetric-based critical value in the Fieller interval (Fb_S). The undercoverage was less pronounced in the first simulation setting (coverage = 0.944 for both $\widehat{ICER}_T(\text{t})$ and $\widehat{ICER}_S(\text{t})$), than in the second setting (coverage = 0.928 for $\widehat{ICER}_T(\text{t})$ and coverage = 0.922 for $\widehat{ICER}_S(\text{t})$). On the other hand, we observed a slight overcoverage in both simulation examples when we calculated bands using the non-symmetric critical value in the Fieller interval (Fb_F).

3.4 Second Multicenter Automatic Defibrillator Implantation Trial (MADIT-II)

The second Multicenter Automatic Defibrillator Implantation Trial (MADIT-II) was a clinical trial in which patients at risk for cardiac arrhythmia were randomized to receive implantable cardioverter-defibrillator (ICD) or standard medical treatment in an attempt to prevent sudden cardiac death (Moss et al., 2002). ICDs have a high initial cost, but their benefits accrue for years after implantation. For this reason, the ICD has often been the focus of cost-effectiveness analyses (Mushlin et al., 1998; Salukhe et al., 2004; Zwanziger et al., 2006).

MADIT-II was conducted from 1997 to 2002. Participants were enrolled from 76 hospitals in the US and Europe. Participants were considered for inclusion in the study if they had experienced a previous heart attack and had a left ventricular ejection fraction less than or equal to 0.30. Additional exclusion criteria and details about the study design are documented in the original report (Moss et al., 2002). We illustrate our methods with

the 664 ICD participants and 431 standard therapy participants who were analyzed in the cost-effectiveness paper by Zwanziger et al. (2006).

For our analyses, both costs and survival times were discounted at a 3% yearly rate using the formulas employed by Zhao and Tian (2001). We chose a τ^0 of 1.5 years after randomization because that was the time at which we might expect to start seeing a survival difference between treatment groups. Like Zwanziger et al. (2006), we chose τ to be 3.5.

Figure 3.1 shows (a) estimated mean cost, (b) estimated probability of survival , and (c) estimated mean survival time for participants in the ICD and standard treatment groups. In Figure 3.1 (a), we see that there is a high initial cost for the ICD group relative to the standard treatment group due to the high cost of intervention. However, cost accrual remains similar in ICD and standard groups after randomization. Figures 3.1 (b) and (c) show that the differences in probability of survival and mean survival time grow larger over time.

Now discussing ICER estimates, the point estimate and 95% Fieller confidence interval are comparable to what was presented by Zwanziger et al. (2006). If we chose to evaluate the estimates at only 3.5 years as they did, we would report an $\widehat{ICER}_T$ = \$213,000 per year of life saved (Fieller interval \$110,000 - \$1,035,000).

Figure 3.2 shows the estimated (a) $ICER_T$(t) and (b) $ICER_S$(t) along with the two most extreme confidence bands as determined by our simulation studies. We therefore illustrate both the symmetric band calculated with the symmetric-based critical value (Sb_S, thin line) and the Fieller confidence bands calculated with the Fieller critical value (Fb_F, dotted line). We note that the Fieller based band with the symmetric interval (Fb_S) was quite similar to the Fb_F band in this example, and was therefore omitted from the figures for simplicity of presentation.

The estimated $ICER_T$ (Figure 3.2(a), thick line) decreases over time, showing that ICDs become more cost-effective over time. The lower symmetric band (Sb_S) in this case is negative for almost the entire time, indicating that if we were to employ these bands, we

wouldn't be able to conclude that the additional cost per year of life saved is positive–even in such an extreme additional cost situation. We now inspect the corresponding Fieller based confidence band (Fb_F). The upper band in this case is infinite (not depicted), which means that when we account for making simultaneous probability statements over time, we are not able to rule out large values of $ICER_T$ anywhere on the curve. The lower band only goes as low as \$97,000, indicating that it would be unlikely for the cost of ICD per year of life saved to go below \$100,000.

The estimated $ICER_S$ (Figure 3.2(b), thick line) also decreases slightly over time, but tends to jump up and down. The jumps occur because the differences in Kaplan-Meier estimated survival probabilities take large jumps, mostly due to the sample size of the MADIT-II data. As with the $ICER_T$(t) example, we note that the lower symmetric band (Sb_S) is negative, while the upper Fieller confidence band is always infinite (not depicted). The lower band does not reach below \$123,000, indicating that it would be unlikely for the cost of the ICD per 100 percentage point increase in survival probability to go below \$125,000.

In our analysis of the MADIT-II data, we see that the confidence intervals and confidence bands often include very large values. This reflects our uncertainty about the ICER(t) estimates. It should be noted that the original study design was powered for detecting a difference in survival, and cost-effectiveness was a secondary analysis (Moss et al., 2002). As such, larger sample sizes or longer follow-up time might yield cost-effectiveness results that are more conclusive.

3.5 Discussion

In this paper, we have presented a nonparametric method to evaluate the incremental cost-effectiveness ratio over time. We presented an estimator, derived an analytical variance formula, and formulated confidence bands.

Several authors have used bootstrapping to estimate confidence intervals for the ICER (Efron and Tibshirani, 1993; Briggs et al., 1997; Chaudhary and Stearns, 1996; Polsky et al.,

1997). However, as was noted by Lin, Fleming, and Wei, it is not clear how to justify using the bootstrap for estimating a function over time (Lin et al., 1994). We therefore employed analytical interval formulas and applied a resampling method to calculate critical values for our confidence bands.

Our simulation results summarized the performance of our estimators using a rather large sample size. We believe that suggesting a larger sample size for our method is reasonable since we are interested in the whole time period (τ^0, τ) rather than just at τ. It should be noted that $\widehat{ICER}_S(\mathrm{t})$ and its corresponding confidence bands can be quite unstable, as we saw in the data example. This is due to the fact that the difference in Kaplan-Meier estimated survival probabilities may jump substantially when the sample size is small. In comparison, the $\widehat{ICER}_T(\mathrm{t})$ is more stable since a single event will not have such a large effect on the estimated mean survival time. Given its more desirable interpretation, many biomedical authors have employed the $ICER_T(\mathrm{t})$. We also recommend using this estimator due to the instability of the estimates and bands for $ICER_S(\mathrm{t})$.

Although it is clear that cost-effectiveness may change over time, statistical literature in the area has been lacking. We provided a starting point for this research by proposing a nonparametric estimator, variance, and confidence bands for the time-dependent incremental cost-effectiveness ratio. Future work in this area might develop regression methods so that researchers may account for covariates in the setting of observational studies.

3.6 Appendix

We wish to find the influence functions of $\widehat{ICER}_S(\mathrm{t})$ and $\widehat{ICER}_T(\mathrm{t})$. We first find the influence function of $\hat{\Lambda}_Y(\mathrm{t})$. We then use this influence function to find the influence functions of $\hat{\mu}_Y(\mathrm{t})$ and $\hat{\mu}_T(\mathrm{t})$. Finally, we find the influence functions of $\widehat{ICER}_S(\mathrm{t})$ and $\widehat{ICER}_T(\mathrm{t})$.

3.6.1 Influence Function for $\hat{\Lambda}_Y(t)$

In this section, we wish to find the influence function of $\hat{\Lambda}_Y(t)$, where $\Lambda_Y(t) = \int_0^t dY(s)/H(s-)$. $\hat{\Lambda}_Y(t)$ depends on the pair $(\hat{Y}(t), \hat{H}(t))$ through two maps:

$$(A, B) \to (A, \frac{1}{B}) \to \int_0^t \frac{1}{B} dA.$$

Then the derivative of these maps at (Y, H) is:

$$(\alpha, \beta) \to (\alpha, -\frac{\beta}{H^2}) \to \int_0^t \frac{d\alpha}{H} - \int_0^t \frac{\beta dY}{H^2}.$$

Let $\alpha = \hat{Y}(t) - Y(t)$ and $\beta = \hat{H}(t) - H(t)$. By the functional delta method (**?**) we have:

$$\begin{aligned}
\hat{\Lambda}_Y(t) - \Lambda_Y(t) &= \int_0^t \frac{1}{H(s)}[\hat{Y}(ds) - Y(ds)] - \int_0^t \frac{\hat{H}(s) - H(s)}{H^2(s)} Y(ds) + o_p\left(\frac{1}{\sqrt{n}}\right) \\
&= \int_0^t \frac{\hat{Y}(ds)}{H(s)} - \int_0^t \frac{\hat{H}(s)}{H^2(s)} Y(ds) + o_p\left(\frac{1}{\sqrt{n}}\right) \\
&= \frac{1}{n}\sum_{i=1}^n \left[\int_0^t \frac{\hat{Y}_i(ds)}{H(s)} - \int_0^t \frac{I(X_i \geq s)Y(ds)}{H^2(s)}\right] + o_p\left(\frac{1}{\sqrt{n}}\right) \\
&= \frac{1}{n}\sum_{I=1}^n \phi_i^{\Lambda_Y}(t) + o_p\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}$$

A consistent estimator for $\phi_i^{\Lambda_Y}(t)$ can be calculated by plugging in $\hat{H}(t)$ and $\hat{Y}(t)$ for $H(t)$ and $Y(t)$, respectively.

3.6.2 Influence Functions for $\hat{\mu}_Y(t)$ and $\hat{\mu}_T(t)$

Here, we wish to find the influence functions of $\hat{\mu}_Y(t)$ and $\hat{\mu}_T(t)$. First note that the influence function for $\hat{S}(t)$ is written as follows:

$$\phi_i^{KM}(t) = S(t)\left[\frac{\Delta_i I(X_i \leq t)}{H(X_i)} - \int_0^t \frac{I(X_i \geq s)}{H^2(s)} dF_1(s)\right],$$

where $X_i = \min(T_i, C_i)$, $\Delta_i = \mathrm{I}(T_i \leq C_i)$, and $F_1(x) = P(X \leq x, \Delta = 1)$. Employing steps similar to those in Section 3.6.1 of the Appendix, we use both $\phi_i^{\Lambda_Y}(\mathrm{t})$ and $\phi_i^{KM}(\mathrm{t})$ to obtain the following influence functions:

$$\begin{aligned}
&\hat{\mu}_Y(t) - \mu_Y(t) \\
&\qquad = \int_0^t S(s-)[\hat{\Lambda}_Y(ds) - \Lambda_Y(ds)] + \int_0^t [\hat{S}(s-) - S(s-)]\Lambda_Y(ds) + o_p\left(\frac{1}{\sqrt{n}}\right) \\
&\qquad = \frac{1}{n}\sum_{i=1}^n \left[\int_0^t S(s-)\phi_i^{\Lambda_Y}(ds) + \int_0^t \phi_i^{KM}(s-)\Lambda_Y(ds)\right] + o_p\left(\frac{1}{\sqrt{n}}\right) \\
&\qquad = \frac{1}{n}\sum_{i=1}^n \phi_i^{\mu_Y}(t) + o_p\left(\frac{1}{\sqrt{n}}\right), \text{ and} \\
\hat{\mu}_T(t) - \mu_T(t) &= \int_0^t [\hat{S}(u) - S(u)]du \\
&\qquad = \frac{1}{n}\sum_{i=1}^n \left[\int_0^t \phi_i^{KM}(u)du\right] + o_p\left(\frac{1}{\sqrt{n}}\right) \\
&\qquad = \frac{1}{n}\sum_{i=1}^n \phi_i^{\mu_T}(t) + o_p\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}$$

Define $\hat{F}_1(x) = \frac{1}{n}\sum_{i=1}^n I(X_i \leq x, \Delta_i = 1)$ to be the empirical estimator of $F_1(x)$. Then consistent estimators for $\phi_i^{KM}(t)$, $\phi_i^{\mu_Y}(t)$, and $\phi_i^{\mu_T}(t)$ can be calculated by replacing $\hat{F}_1(x)$ for $F_1(x)$, $\hat{H}(t)$ for $H(t)$, $\hat{S}(t)$ for $S(t)$, $\hat{\Lambda}_Y(t)$ for $\Lambda_Y(t)$, and $\hat{\phi}_i^{\Lambda_Y}(t)$ for $\phi_i^{\Lambda_Y}(t)$.

3.6.3 Influence Functions for $\widehat{ICER}(t)$

First, let n_1 and n_0 denote the number of participants on treatment and standard, respectively, such that $n = n_1 + n_0$. We find the influence function for $\hat{\Delta}_Y(\mathrm{t})$:

$$\begin{aligned}
\hat{\Delta}_Y(t) - \Delta_Y(t) &= [\hat{\mu}_{Y_1} - \hat{\mu}_{Y_0}] - [\mu_{Y_1} - \mu_{Y_0}] \\
&= \frac{1}{n}\sum_{i=1}^n \left[\frac{n}{n_1}R_i\phi_i^{\mu_Y}(t) - \frac{n}{n_0}(1 - R_i)\phi_i^{\mu_Y}(t)\right] + o_p\left(\frac{1}{\sqrt{n}}\right) \\
&= \frac{1}{n}\sum_{i=1}^n \eta_i^Y(t) + o_p\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}$$

Similarly, the influence functions for $\hat{\Delta}_S(\mathrm{t})$ and $\hat{\Delta}_T(\mathrm{t})$ take the form:

$$\eta_i^S(t) = \frac{n}{n_1} R_i \phi_i^{KM}(t) - \frac{n}{n_0}(1 - R_i)\phi_i^{KM}(t), \text{ and}$$
$$\eta_i^T(t) = \frac{n}{n_1} R_i \phi_i^{\mu_T}(t) - \frac{n}{n_0}(1 - R_i)\phi_i^{\mu_T}(t).$$

Now, by the functional delta method once again, we obtain the influence functions of $\widehat{ICER}_S(\mathrm{t})$ and $\widehat{ICER}_T(\mathrm{t})$. With $K \in \{S, T\}$, $\widehat{ICER}_K(\mathrm{t})$ takes the following format:

$$\begin{aligned}
\widehat{ICER}_K(t) - ICER_K(t) &= \frac{\hat{\Delta}_Y(t) - \Delta_Y(t)}{\hat{\Delta}_K(t)} - \left[\hat{\Delta}_K(t) - \Delta_K(t)\right] \frac{\hat{\Delta}_Y(t)}{\hat{\Delta}_K^2(t)} \\
&= \frac{1}{n} \sum_{i=1}^{n} \left[\frac{\eta_i^Y(t)}{\hat{\Delta}_K(t)} - \eta_i^K(t) \frac{\hat{\Delta}_Y(t)}{\hat{\Delta}_K^2(t)} \right] + o_p\left(\frac{1}{\sqrt{n}}\right) \\
&= \frac{1}{n} \sum_{i=1}^{n} \left[\varphi_i^{ICER_K}(t) \right] + o_p\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}$$

Finally, the consistent estimators for $\eta_i^Y(t)$, $\eta_i^S(t)$, $\eta_i^T(t)$, $\varphi_i^{ICER_S}(t)$, and $\varphi_i^{ICER_T}(t)$ are calculated by replacing $\hat{\phi}_i^{\mu_Y}(t)$, $\hat{\phi}_i^{KM}(t)$, and $\hat{\phi}_i^{\mu_T}(t)$ for $\phi_i^{\mu_Y}(t)$, $\phi_i^{KM}(t)$, and $\phi_i^{\mu_T}(t)$, respectively, in the above formulas.

Table 3.1: Summary of simulation results for $\widehat{ICER}_T$(t) and $\widehat{ICER}_S$(t) for two different survival time scenarios. The table describes estimation, bias, sample standard error (SSE), estimated standard error (ESE), and coverage probability for simultaneous confidence bands using the symmetric critical value b in the symmetric interval (Sb_S), symmetric b in the Fieller interval (Fb_S), and the Fieller b in the Fieller interval (Fb_F). Results reflect 1,000 replications, confidence bands are based on m=1,000 resamples, and there are n=1,000 people in each treatment group.

t (Yrs)	Ests ($1000s)	Bias (%)	SSE ($1000s)	ESE ($1000s)	Cov. Sb_S	Cov. Fb_S	Cov. Fb_F
$\widehat{ICER}_T$(t): $T_1 \sim$ Exp(8), $T_0 \sim$ Exp(5)							
2	189.5	5.5	52.6	48.0	0.904	0.944	0.982
4	55.6	3.0	10.0	9.6			
6	30.0	2.1	4.5	4.3			
8	20.6	1.6	2.8	2.7			
10	16.2	1.3	2.0	2.0			
$\widehat{ICER}_S$(t): $T_1 \sim$ Exp(8), $T_0 \sim$ Exp(5)							
2	210.4	4.9	60.6	53.7	0.838	0.944	0.963
4	139.0	2.6	24.5	23.0			
6	127.9	2.5	21.5	21.0			
8	133.1	2.0	23.9	23.6			
10	150.4	2.8	32.6	31.4			
$\widehat{ICER}_T$(t): $T_1 \sim$ Exp(12.8), $T_0 \sim$ Weib(scale=7.3, shape=1.1)							
2	405.0	11.1	2284.3	449.8	0.851	0.928	0.989
4	82.5	3.5	22.3	20.9			
6	36.3	1.6	6.6	6.5			
8	21.7	1.1	3.2	3.2			
10	15.3	0.8	2.0	2.0			
$\widehat{ICER}_S$(t): $T_1 \sim$ Exp(12.8), $T_0 \sim$ Weib(scale=7.3, shape=1.1)							
2	354.4	8.4	234.3	151.0	0.801	0.922	0.963
4	160.9	2.0	33.1	32.2			
6	117.5	1.5	19.2	18.6			
8	102.6	1.5	15.7	15.3			
10	99.0	1.7	16.1	15.7			

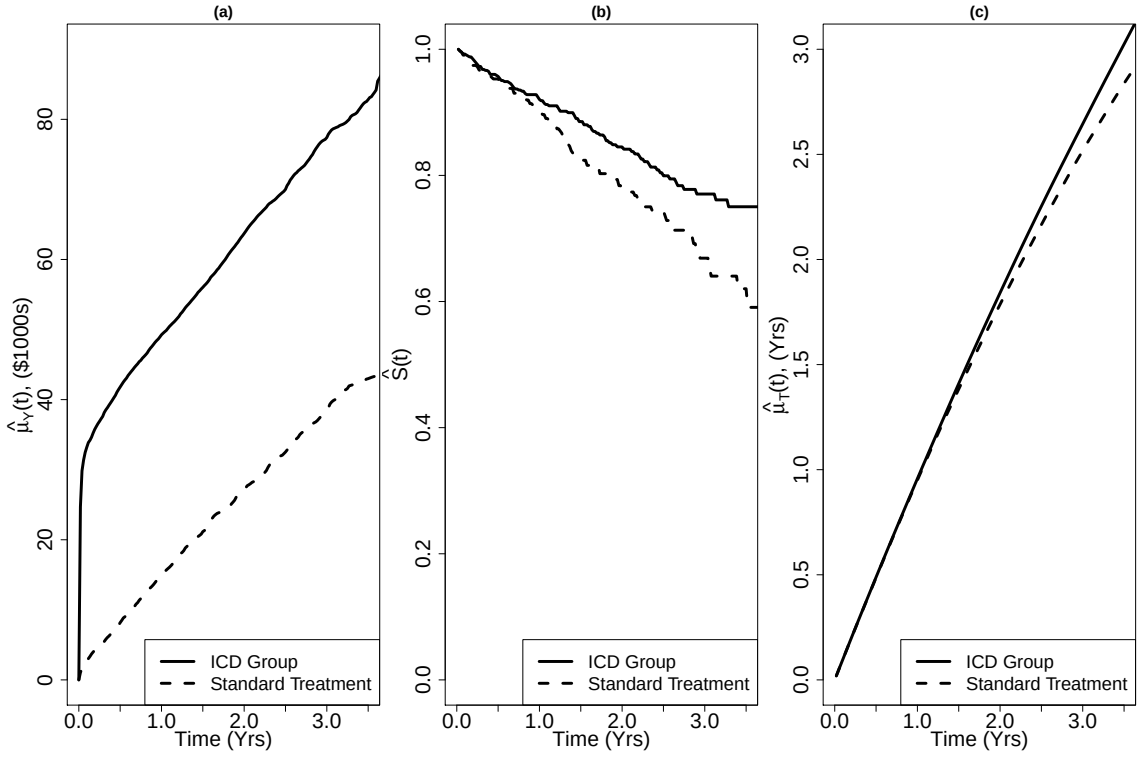

Figure 3.1: (a) Estimated mean cost, (b) probability of survival, and (c) mean survival time for the MADIT-II data.

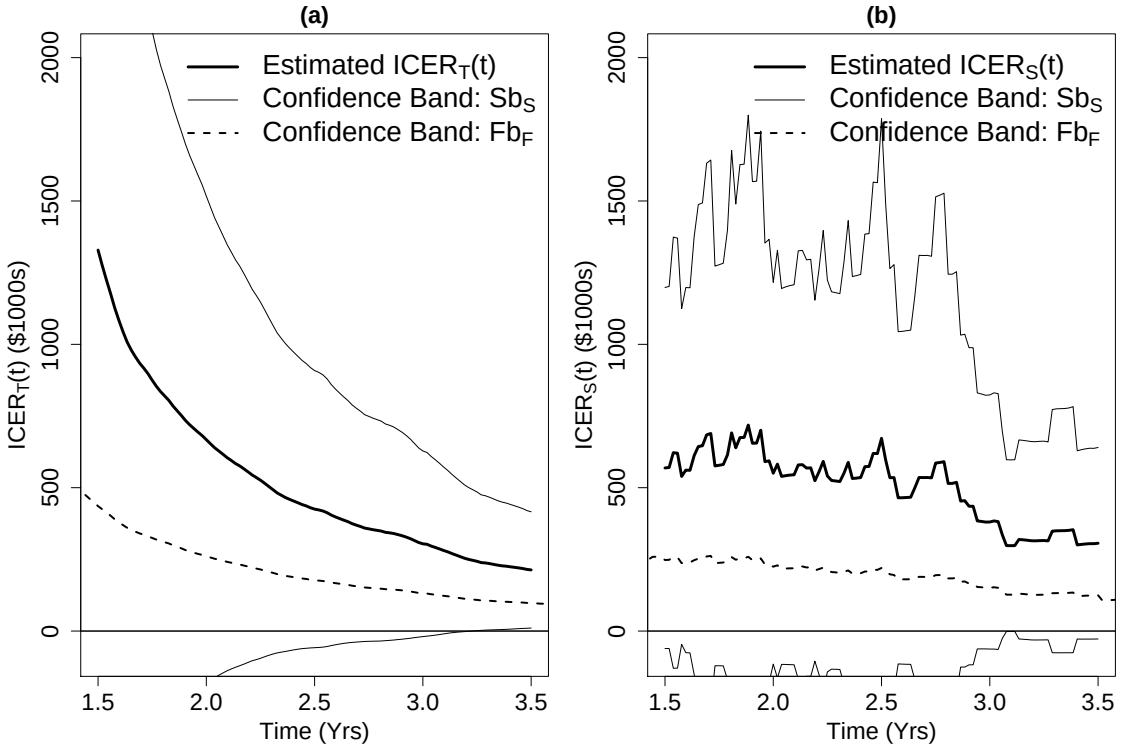

Figure 3.2: (a) Estimated $ICER_T$(t) and (b) $ICER_S$(t) along with the 95% confidence interval and simultaneous confidence bands. Fb_F denotes the Fieller confidence band calculated with the Fieller-based critical value.

Chapter 4

INTENTION-TO-TREAT EFFECTS AND COMPLIER AVERAGE CAUSAL EFFECTS FOR COST-EFFECTIVENESS ANALYSIS

4.1 Introduction

In randomized clinical trials studying the effect of a new treatment versus a standard treatment, some researchers also aim to examine the corresponding cost-effectiveness. Additionally, in trials where participants may not comply with their assigned treatment, it may also be of interest to study the cost-effectiveness amongst participants who comply with the treatment assignment, which we will call the complier average causal effect of cost-effectiveness.

While the intention-to-treat (ITT) effect is often the primary focus in a randomized trial, it has been acknowledged that the complier average causal effect may also be of interest in order to describe treatment efficacy or to help provide context when compliance levels may differ in another setting (Cheng et al., 2009; Nie et al., 2011). Angrist et al. (1996) presented a framework for examining complier average causal effects, which can be interpreted as the average effect of the new treatment over standard treatment amongst participants who comply with study treatment. They also provided a formal set of assumptions required in order to interpret their estimators as complier average causal effects. The framework of Angrist et al. (1996) has been extended to randomized trials that also have administrative censoring (Nie et al., 2011), which often occurs in trials where follow-up ends after a specified amount of time. In this paper, we plan to examine the complier average causal effects of measures of cost-effectiveness in settings both with and without censoring.

For cost-effectiveness analyses, we will focus on weighing lifetime cost against measures of effectiveness. Lifetime cost is defined to be all medical costs accrued by a participant from randomization to death or a prespecified end of study time, L. Baker (1998) reported

on the cost-effectiveness of breast cancer screening in a setting with noncompliance, but did not employ lifetime cost data. In addition, Baker (1998) did not discuss interpretation of complier average causal cost-effectiveness, which we plan to focus on here.

Two measures of cost-effectiveness that are commonly employed are the incremental cost-effectiveness ratio (ICER) and the incremental net benefit (INB) (Chaudhary and Stearns, 1996; Zhao and Tian, 2001; Willan and Lin, 2001; Willan et al., 2002). The ICER is often defined to be the difference in mean cost between the experimental treatment and standard treatment, divided by the difference in mean survival time. If we let Y denote lifetime cost, R denote randomization ($R = 1$ for experimental treatment, and $R = 0$ for standard treatment), and T be time to death, then the standard, intention-to-treat ICER is written as follows:

$$ICER = \frac{E[Y|R=1] - E[Y|R=0]}{E[T|R=1] - E[T|R=0]}. \tag{4.1}$$

The ICER in (4.1) can therefore be interpreted as the additional cost of being randomized to experimental treatment versus standard treatment, per year of life saved. The ICER has been used extensively in biomedical applications(Chaudhary and Stearns, 1996; Mushlin et al., 1998; Zwanziger et al., 2006), mainly due to this simple interpretation. In general, the measure of effectiveness in the denominator of the ICER can be chosen to be a quantity other than the difference in mean survival time such as the difference in probability of survival or the difference in quality-of-life adjusted mean survival time (Willan et al., 2002). However, for simplicity, we will focus on the difference in mean survival time as the measure of effectiveness.

The second measure of cost-effectiveness, the INB, is defined to be the difference in mean survival time, multiplied by the value of a year of life saved, minus the difference in mean cost. If we let λ be some constant that is chosen to be the monetary value of a year of life saved, on average, then the INB is written as follows:

$$INB = (E[T|R=1] - E[T|R=0])\,\lambda - (E[Y|R=1] - E[Y|R=0])\,. \tag{4.2}$$

The INB in (4.2) is said to be the monetary net benefit of the experimental treatment over the standard treatment.

As with the treatment effect, the complier average causal cost-effectiveness may be of interest to provide additional context on cost-effectiveness when compliance in the trial may differ from compliance levels in other settings of interest. In this paper, we therefore aim to define the complier average causal cost-effectiveness for both the ICER and the INB. We will focus on interpreting the relationships between these complier average causal effects and the corresponding ITT estimators. In particular, we will show that for the ICER, the ITT analyses can also be directly interpreted as the complier average causal effect, but the same is not true for the INB. Finally, we will complete simulation experiments to confirm these reported relationships.

4.2 Complier Average Causal Cost-Effectiveness

For the time being, let us discuss the situation where there is no censoring– that is, we have event times and lifetime medical costs for all individuals. We employ notation similar to that of Nie et al. (2011). For individual i, $i = 1, \ldots, n$, let R_i again denote treatment assignment, where $R_i = 1$ for participants assigned to the experimental treatment, and $R_i = 0$ for participants assigned to standard treatment. Furthurmore, let $A_i^{\mathbf{R}}$ denote the potential treatment subject i would receive given randomization assignment vector $\mathbf{R}$, where $A_i^{\mathbf{R}} = 1$ if the person receives treatment, and $A_i^{\mathbf{R}} = 0$ if the person receives placebo. Define $T_i^{\mathbf{R,A}}$ to be the ith person's potential failure time given randomization vector $\mathbf{R}$ and vector of treatment received $\mathbf{A}$. Similarly, let $Y_i^{\mathbf{R,A}}$ be the ith person's potential lifetime medical cost under vectors $\mathbf{R}$ and $\mathbf{A}$. We say that $A_i^{\mathbf{R}}$, $T_i^{\mathbf{R,A}}$ and $Y_i^{\mathbf{R,A}}$ are potential outcomes because although we may only observe A, T, and Y for individual i under one randomization and treatment scenario, we are interested in discussing the other unobserved scenarios.

At this point, we may define the different compliance classes. In our setting of all-or-none compliance with treatment, there are four compliance classes. The first class are called "compliers," who are the participants who would take either assigned treatment. Therefore,

in terms of potential treatment received on the individual level, if they were assigned to experimental treatment they would take experimental treatment and if they were assigned to standard treatment they would take standard ($A_i^1 - A_i^0 = 1 - 0 = 1$). The second class, dubbed "never-takers," would never take experimental treatment regardless of their potential assignment ($A_i^1 - A_i^0 = 0 - 0 = 0$). The third type, known as "always-takers," always take experimental treatment ($A_i^1 - A_i^0 = 1 - 1 = 0$). Finally, "defiers" always take the opposite treatment of what they are assigned ($A_i^1 - A_i^0 = 0 - 1 = -1$). For many analyses, it is common to make the assumption that there are no defiers.

In order to interpret our estimators as the complier average causal cost-effectiveness, we must make several assumptions, which were originally described by Angrist et al. (1996), and were also mentioned in a similar manner by Nie et al. (2011). The assumptions we make are as follows:

1. Letting $\mathbf{R}$ denote the vector of randomizations for all individuals i and l denote the column vector with each element equal to one, then for all $\mathbf{c}$ and $\mathbf{c'}$ such that $l^T\mathbf{c} = l^T\mathbf{c'}$, $P(\mathbf{R} = \mathbf{c}) = P(\mathbf{R} = \mathbf{c'})$.

2. For all individuals i, if $R_i = R_i'$, then $A_i^{\mathbf{R}} = A_i^{\mathbf{R'}}$. In addition, for all individuals i, if $R_i = R_i'$ and $A_i = A_i'$, then $T_i^{\mathbf{R},\mathbf{A}} = T_i^{\mathbf{R'},\mathbf{A'}}$ and $Y_i^{\mathbf{R},\mathbf{A}} = Y_i^{\mathbf{R'},\mathbf{A'}}$.

3. For all R_i, R_i', A_i, and subjects i, $T_i^{\mathbf{R},\mathbf{A}} = T_i^{\mathbf{R'},\mathbf{A}}$ and $Y_i^{\mathbf{R},\mathbf{A}} = Y_i^{\mathbf{R'},\mathbf{A}}$.

4. $E\left[A_i^1 - A_i^0\right] \neq 0$.

5. For all subjects i, $A_i^1 \geq A_i^0$.

The first assumption stipulates that R is random, which generally holds in randomized clinical trials. Assumption 2 was named the stable unit treatment value (Rubin, 1978; Angrist et al., 1996), and means that changes in other participants' randomization and treatment values should not change the outcomes of participant i. When we assume 2, we

are able to write $T_i^{R_i,A_i}, Y_i^{R_i,A_i}$, and $A_i^{R_i}$ instead of $T_i^{\mathbf{R},\mathbf{A}}, Y_i^{\mathbf{R},\mathbf{A}}$, and $A_i^{\mathbf{R}}$. Assumption 3 means that randomization affects T_i and Y_i only through its effect on treatment received. Given this assumption, we can simplify our notation to $T_i^{A_i}$ and $Y_i^{A_i}$ instead of $T_i^{R_i,A_i}$ and $Y_i^{R_i,A_i}$. The fourth assumption means that randomization has some effect on treatment received. The final assumption means that there are no defiers, or subjects who always take the opposite treatment that is assigned.

Using the notation outlined above, the ITT INB and ITT ICER can be rewritten as follows:

$$INB = E[T_i^{1,A_i} - T_i^{0,A_i}]\lambda - E[Y_i^{1,A_i} - Y_i^{0,A_i}]$$
$$ICER = \frac{E[Y_i^{1,A_i} - Y_i^{0,A_i}]}{E[T_i^{1,A_i} - T_i^{0,A_i}]}.$$

We now wish to show the relationship of these ITT measures of cost-effectiveness with the complier average causal effect for measures of cost-effectiveness.

Following the argument of Angrist et al. (1996), we will first describe the complier average causal effect of lifetime costs. First, we examine the subject-level potential effect of randomization on lifetime cost:

$$Y_i^{1,A_i} - Y_i^{0,A_i} = \left[Y_i^1 \times A_i^1 + Y_i^0 \times (1 - A_i^1)\right] - \left[Y_i^1 \times A_i^0 + Y_i^0 \times (1 - A_i^0)\right] \quad (4.3)$$
$$= (Y_i^1 - Y_i^0)(A_i^1 - A_i^0). \quad (4.4)$$

The left side of (3) is therefore just the subject-level potential effect of lifetime cost, which we can inspect on its own by assumptions 2 and 3. On the right side of (3), the first portion describes potential outcomes for participants who were randomized to experimental treatment and the second portion is for participants who were randomized to standard treatment. Via factoring, we see that the subject-level potential effect of treatment on lifetime cost can be written as the product of the causal effect of treatment received (A) on lifetime cost (Y) and the causal effect of treatment assigned (R) on treatment received

(A). To obtain the relationship between the ITT effect and complier average causal effect of treatment received on lifetime cost, we take expectations on both sides to obtain the following:

$$\begin{aligned}
E[Y_i^{1,A_i} - Y_i^{0,A_i}] &= E[(Y_i^1 - Y_i^0)|(A_i^1 - A_i^0) = 1]P[(A_i^1 - A_i^0) = 1] \\
&\quad - E[(Y_i^1 - Y_i^0)|(A_i^1 - A_i^0) = -1]P[(A_i^1 - A_i^0) = -1] \\
&\quad + E[(Y_i^1 - Y_i^0)|(A_i^1 - A_i^0) = 0]P[(A_i^1 - A_i^0) = 0] \\
&= E[(Y_i^1 - Y_i^0)|(A_i^1 - A_i^0) = 1]P[(A_i^1 - A_i^0) = 1].
\end{aligned}$$

The first equality rewrites the expected value to be weighted by the probability of being a complier (first term), a defier (second term), or an always-taker or a never-taker (third term). The second equality removes the defier term by assumption 5, and the always/never-taker term by assumption 4. Therefore, we see that the ITT effect of randomization on lifetime cost ($E[Y_i^{1,A_i} - Y_i^{0,A_i}]$) is equal to the complier average causal effect of treatment received on lifetime cost amongst compliers ($E[(Y_i^1 - Y_i^0)|(A_i^1 - A_i^0) = 1]$) multiplied by the probability of being a complier ($P[(A_i^1 - A_i^0) = 1]$). This same argument follows for relating the ITT effect of randomization of treatment on survival time to the complier average causal effect of treatment received on survival time amongst compliers. That is,

$$E[T_i^{1,A_i} - T_i^{0,A_i}] = E[(T_i^1 - T_i^0)|(A_i^1 - A_i^0) = 1]P[(A_i^1 - A_i^0) = 1].$$

Given these relationships between the ITT effects and the complier average causal effects, we can now write the corresponding relationships for our measures of cost-effectiveness, the INB and the ICER. For the INB we obtain:

$$\begin{aligned}
E[T_i^{1,A_i} - T_i^{0,A_i}]\lambda - E[Y_i^{1,a} - Y_i^{0,a}] &= E[(T_i^1 - T_i^0)|(A_i^1 - A_i^0) = 1]P[(A_i^1 - A_i^0) = 1]\lambda \\
&\quad - E[(Y_i^1 - Y_i^0)|(A_i^1 - A_i^0) = 1]P[(A_i^1 - A_i^0) = 1] \\
&= P[(A_i^1 - A_i^0) = 1]\left\{E[(T_i^1 - T_i^0)|(A_i^1 - A_i^0) = 1]\lambda - E[(Y_i^1 - Y_i^0)|(A_i^1 - A_i^0) = 1]\right\}.
\end{aligned}$$

Therefore, similarly to the relationship between the ITT and complier average causal mean effects, we have that the ITT INB is equal to the complier average causal INB amongst compliers times the probability of being a complier. However, for the ICER, we note that

$$\begin{aligned}\frac{E[Y_i^{1,A_i} - Y_i^{0,A_i}]}{E[T_i^{1,A_i} - T_i^{0,A_i}]} &= \frac{E[(Y_i^1 - Y_i^0)|(A_i^1 - A_i^0) = 1] \times P[(A_i^1 - A_i^0) = 1]}{E[(T_i^1 - T_i^0)|(A_i^1 - A_i^0) = 1] \times P[(A_i^1 - A_i^0) = 1]} \\ &= \frac{E[(Y_i^1 - Y_i^0)|(A_i^1 - A_i^0) = 1]}{E[(T_i^1 - T_i^0)|(A_i^1 - A_i^0) = 1]}.\end{aligned}$$

In this case, we see that the ITT ICER is equal to the complier average causal ICER, which is defined to be the complier average causal effect of treatment received on lifetime medical cost, divided by the complier average causal effect of treatment received on survival time.

4.3 Estimation

4.3.1 No Censoring

In the case of no censoring, we can estimate the INB and ICER using sample means. Let $\hat{\mu}_{Y^{1,A}}$, $\hat{\mu}_{Y^{0,A}}$, $\hat{\mu}_{T^{1,A}}$, $\hat{\mu}_{T^{1,A}}$, $\hat{\mu}_{A^1}$, and $\hat{\mu}_{A^0}$ be the sample means for estimating lifetime cost for participants randomized to experimental treatment ($R_i = 1$), lifetime cost when $R_i = 0$, survival time when $R_i = 1$, survival time when $R_i = 0$, treatment received when $R_i = 1$, and treatment received when $R_i = 0$. Then an estimator for the complier average INB, $\widehat{INB}^{comp}(\lambda)$, can be written as follows:

$$\begin{aligned}\widehat{INB}^{comp}(\lambda) &= \frac{(\hat{\mu}_{T^{1,A}} - \hat{\mu}_{T^{0,A}})\lambda - (\hat{\mu}_{Y^{1,A}} - \hat{\mu}_{Y^{0,A}})}{[\hat{\mu}_{A^1} - \hat{\mu}_{A^0}]} \\ &= \frac{\widehat{INB}^{ITT}(\lambda)}{[\hat{\mu}_{A^1} - \hat{\mu}_{A^0}]},\end{aligned}$$

where $\widehat{INB}^{ITT}(\lambda)$ is the corresponding estimate of the ITT INB. Similarly, we can estimate the complier average ICER, $\widehat{ICER}^{comp}$ as follows:

$$\begin{aligned}\widehat{ICER}^{comp} &= \frac{\hat{\mu}_{Y^{1,A}} - \hat{\mu}_{Y^{0,A}}}{\hat{\mu}_{T^{1,A}} - \hat{\mu}_{T^{0,A}}} \\ &= \widehat{ICER}^{ITT},\end{aligned}$$

where $\widehat{ICER}^{ITT}$ is the estimate of the ITT ICER.

4.3.2 With Censoring

When censoring occurs, participants may not be followed-up until time to death and lifetime medical cost can be recorded. In this case, standard methods for dealing with censoring can be applied to estimates describing survival time. For example, allowing for censoring, we can estimate mean survival time by calculating the area under the Kaplan-Meier estimator for the probability of survival. However, an additional problem known as induced informative censoring must be addressed when estimating mean lifetime cost (Lin et al., 1997; Bang and Tsiatis, 2000; Zhao and Tian, 2001; Willan and Lin, 2001).

When estimating mean lifetime medical cost, induced informative censoring occurs because even when the standard assumption of independence between survival times and censoring times is appropriate, the cost at survival times and costs at censoring times will not be independent. While many estimators have been proposed for estimating mean lifetime cost (Lin et al., 1997; Bang and Tsiatis, 2000; Zhao and Tian, 2001), we present an estimator originally proposed by Cook and Lawless (1997) for estimating average number of recurrent events. In our setting of medical costs which are recorded whenever there is a medical event generating cost, this estimator makes use of cost information up to an individual's censoring time. By contrast, other common nonparametric estimators used only cost information from participants who were observed to die (Lin et al., 1997; Bang and Tsiatis, 2000), which does not make use of information that is often readily available in

these types of studies.

In order to present these estimators, we first must introduce some additional notation, and add a standard non-informative censoring assumption on the time scale. Let us define C_i to be the censoring time for individual i, $\Delta_i^{\mathbf{R,A}} = I(T_i^{\mathbf{R,A}})$ to be the indicator of whether we observed a patient to die, $X_i^{\mathbf{R,A}} = min(T_i^{\mathbf{R,A}}, C_i)$ be the observed follow-up time, and $Y_i^{\mathbf{R,A}}$ be the observed medical cost up to time $X_i^{\mathbf{R,A}}$, as opposed to the true lifetime medical cost which we will denote in this section by $Y_i^{0\,\mathbf{R,A}}$. We also add the following assumption to the previous assumptions 1-5:

6. The censoring times C are independent of $(Y^{\mathbf{R,A}}, T^{\mathbf{R,A}})$.

This assumption is the non-informative censoring assumption, and allows for correlation of costs within an individual, but requires that C is not associated with the outcomes of interest.

We wish to compute $\hat{\mu}^c_{T^{1,A}}$, $\hat{\mu}^c_{T^{0,A}}$, $\hat{\mu}^c_{Y^{1,A}}$, and $\hat{\mu}^c_{Y^{0,A}}$, which are the estimates for mean survival time for participants randomized to experimental treatment ($R_i = 1$), mean survival time when $R_i = 0$, mean lifetime cost when $R_i = 1$, and mean lifetime cost when $R_i = 0$ when allowing for censoring. Recall that we have defined the pre-specified end of study time, the constant L, which will be the time at which we will evaluate our estimators. Let $\hat{S}^{\mathbf{R,A}}(L) = \hat{P}(T^{\mathbf{R,A}} > L)$ be the standard Kaplan-Meier estimator. Furthermore, define $\hat{K}^{\mathbf{R,A}}(t) = \hat{P}(C > t)$ to be the Kaplan-Meier estimate of the censoring times at time t. Finally, $\hat{Y}^{\mathbf{R,A}}(t) = \frac{1}{n}\sum_{i=1}^{n} Y_i^{\mathbf{R,A}}(t)$ denotes the sample mean of medical costs up to time t, and $d\hat{Y}^{\mathbf{R,A}}(t) = \hat{Y}^{\mathbf{R,A}}(t) - \hat{Y}^{\mathbf{R,A}}(t-)$ is the jump in estimated average medical cost at t. Then our estimators, allowing for censoring, can be written as follows:

$$\hat{\mu}^c_{T^{R,A}} = \int_0^L \hat{S}^{\mathbf{R,A}}(t)dt \qquad \text{and}$$
$$\hat{\mu}^c_{Y^{R,A}} = \int_0^L \frac{d\hat{Y}^{\mathbf{R,A}}(t)}{\hat{K}(t)}dt.$$

Then our estimators for complier average INB and complier average ICER, allowing for

censoring will be

$$\widehat{INB}^{c,\,comp}(\lambda) = \frac{(\hat{\mu}^c_{T^{1,A}} - \hat{\mu}^c_{T^{0,A}})\lambda - (\hat{\mu}^c_{Y^{1,A}} - \hat{\mu}^c_{Y^{0,A}})}{[\hat{\mu}^c_{A^1} - \hat{\mu}^c_{A^0}]} = \frac{\widehat{INB}^{c,\,ITT}(\lambda)}{[\hat{\mu}^c_{A^1} - \hat{\mu}^c_{A^0}]} \quad \text{and}$$

$$\widehat{ICER}^{c,\,comp} = \frac{\hat{\mu}^c_{Y^{1,A}} - \hat{\mu}^c_{Y^{0,A}}}{\hat{\mu}^c_{T^{1,A}} - \hat{\mu}^c_{T^{0,A}}} = \widehat{ICER}^{c,\,ITT}.$$

4.4 Simulation Studies

We completed numerical simulation studies in order to confirm relationships that we described in the previous section. We first generated compliance groups and treatment assignment. We assigned compliance status such that compliers, always takers, and never takers were generated with probabilities of 0.6, 0.2, and 0.2, respectively. Defiers were not allowed due to the fact that assumption 5 requires that there are no defiers. Treatment assignment was assigned with probabilities of 0.5 for both experimental treatment and standard treatment groups.

We next generated cost and survival data in two different simulation setups. Simulation 1 illustrates a situation in which the survival difference and cost difference is larger than the situation in Simulation 2. For each simulation, survival time was distributed exponentially with varying mean parameters in the exponential distribution. Cost was simulated similarly to previous papers (Bang and Tsiatis 2000; Zhao and Tian 2001), and was made up of a diagnostic cost at time of randomization, yearly costs, and a death cost. The diagnostic cost can be thought of as the initial cost of administering the assigned treatment, yearly costs are any medical costs that occur within a given year after randomization, and the death cost accounts for additional medical costs that may occur shortly before a participant dies. Cost components were distributed according to the uniform distribution with varying minima and maxima.

In the first simulation setup, which we will call "Setup 1," the survival times and medical costs were distributed in the following manner. For compliers who received treatment, $T \sim$ Exp(mean=9), diagnostic costs $\sim$ Uniform(3000, 5000), and yearly costs $\sim$ Uniform(8000,

10000). For always-takers, $T \sim$ Exp(mean=9.5), diagnostic costs $\sim$ Uniform(3000, 5000), and yearly costs $\sim$ Uniform(6000, 8000). For never-takers, $T \sim$ Exp(mean=4.5), diagnostic costs $\sim$ Uniform(3000, 5000), and yearly costs $\sim$ Uniform(4000, 6000). Finally, for compliers who did not receive treatment, $T \sim$ Exp(mean=5), diagnostic costs $\sim$ Uniform(1000, 2000), and yearly costs $\sim$ Uniform(2000, 4000). All death costs were were distributed Uniform(5000, 10000). The rationale behind choosing these parameters is that the experimental treatment is more costly in terms of diagnostic cost and yearly cost, but leads to longer survival.

In the second simulation setup, "Setup 2," the corresponding survival times and yearly costs were changed slightly, while the diagnostic and death costs were left unchanged. For compliers who received treatment, $T \sim$ Exp(mean=8.5), and yearly costs $\sim$ Uniform(6000, 10000). For always-takers, $T \sim$ Exp(mean=9.5) and yearly costs $\sim$ Uniform(5000, 8000). For never-takers, $T \sim$ Exp(mean=4.5) and yearly costs $\sim$ Uniform(2000, 4000). Finally, for compliers who did not receive treatment, $T \sim$ Exp(mean=6.5) and yearly costs $\sim$ Uniform(2000, 4000). This simulation setup has the same rationale as Setup 1, but has smaller differences in cost and survival time.

For both simulation setups, $C_i \sim$ Uniform(0, 20), and we evaluated the estimators at 10 years. We set λ=100,000 when calculating the INB estimate. We varied the sample size for each simulation setup with sample sizes per randomization group equal to 500, 1000, or 2000, on average. For each simulation, we completed 10,000 replications.

Table 4.1 shows the average results across replications. Across sample sizes and simulation setups, we see that the average $\widehat{INB}^{c,\,ITT}$, the estimated INB for the intention-to-treat situation, is 0.6 times the estimated INB for the complier average situation, $\widehat{INB}^{c,\,comp}$. Since we defined the probability of being a complier to be 0.6, this confirms the relationship for the INB that we highlighted above. In addition, we see that the estimates for both INB situations are stable across sample sizes.

For the ICER, we confirm the relationships outlined in the previous sections, since we

see that the average ratio of the estimated ICER in the ITT situation, $\widehat{ICER}^{c,\,ITT}$, to the estimated ICER in the complier average situation, $\widehat{ICER}^{c,\,comp}$, is close to 1 in nearly all of our simulated examples. In the N=500 simulation of Setup 2, however, we see that the average ratio is 0.620. This speaks to the potential performance problems of the ICER in general in modest sample sizes when the difference in mean survival time may be small. In this situation, we saw a few very influential ICER values (both in the negative and positive direction) for some simulated datasets, which drove the empirical mean of the simulations to change dramatically. When we omitted the ten largest and smallest $\widehat{ICER}^{c,\,ITT}$ and $\widehat{ICER}^{c,\,comp}$ in this simulation, we obtained a truncated mean ratio of 0.998. For all of these influential replications, the difference in simulated mean survival time was extremely small, thus causing the estimate of the ICER to become very large. For the same reason, we also see somewhat less stable estimates for both ICER situations compared to the INB.

Table 4.1: Average estimated incremental net benefit and incremental cost-effectiveness ratio for the intention-to-treat effect and the complier average effect for 10,000 simulated datasets. N is the sample size per treatment assignment group.

	N	$\widehat{INB}^{c,\,ITT}$	$\widehat{INB}^{c,\,comp}$	Ratio	$\widehat{ICER}^{c,\,ITT}$	$\widehat{ICER}^{c,\,comp}$	Ratio
Setup 1	500	87,100	145,000	0.600	20,400	19,900	1.026
	1,000	87,500	145,800	0.600	19,700	19,500	1.010
	2,000	87,500	145,600	0.601	19,400	19,400	1.003
Setup 2	500	31,800	53,200	0.597	25,800	41,600	0.620
	1,000	31,800	53,000	0.600	40,800	39,000	1.046
	2,000	31,600	52,700	0.600	38,500	37,100	1.039

4.5 Discussion

In this paper, we saw that the complier average causal incremental net benefit is proportional to the intention-to-treat incremental net benefit in the standard way. More interesting was the fact that the intention-to-treat incremental cost-effectiveness ratio can be interpreted also as the complier average causal incremental cost-effectiveness ratio. This property may be of interest to investigators who wish to understand the cost-effectiveness for participants

who comply with treatment, but do not wish to delve too deeply into causal inference. Due to this property, some might advocate for employing the incremental cost-effectiveness ratio over the incremental net benefit when causal cost-effectiveness is of interest. In such a scenario, our results should be weighed against the advantages of the incremental net benefit, which were succinctly described by Willan et al. (2002).

Chapter 5

CONCLUSION

This dissertation discussed an array of statistical methods pertaining to the analysis of medical cost data. In the first chapter, we discussed the problem of induced informative censoring, and surveyed the background literature on estimating mean cost when data is right censored. We also discussed the notion of recurrent survival times, introduced two measures of cost-effectiveness analysis, and described the data from the second Multicenter Automatic Defibrillator Implantation Trial, which we employed as the data example in Chapters 2 and 3.

In chapter 2, we studied joint nonparametric estimation of recurrent survival time and a recurrent marked variable. In particular, we applied the proposed methods to the recurrent marked variable medical cost, although the work can also be employed for any other recurrent marked variable such as disease severity. This estimator required very few assumptions and allowed for correlation within individuals for both the recurrent survival times and the recurrent medical costs, which makes it an attractive estimator in many situations. The estimator can be employed to characterize the correlation between recurrent survival time and the recurrent marked variable of interest, as we saw in our example of the MADIT-II data. Future work in this area may include extending estimation to multivariate recurrent events.

Chapter 3 examined inference for a nonparametric estimator of the incremental cost-effectiveness ratio (ICER) over time. In this chapter, we discussed estimation, variance estimation, and creation of confidence bands, which allowed us to determine the uncertainty around the ICER estimates over a period of time. We believe that studying cost-effectiveness over time may be useful for several reasons. First, looking at cost-effectiveness over time

may aid in overall understanding of cost-effectiveness. For example, in the data illustration with the MADIT-II study, we saw that there was a sharp decline in cost per year of life saved in the first few years of study. Second, such an estimator can help researchers to characterize how quickly something may become reasonably cost-effective. Finally, these proposed methods may allow researchers to easily and directly compare cost-effectiveness across trials. As this is one of the first topics to discuss cost-effectiveness over time, there are many avenues for future research. For example, future studies may allow for covariate adjustment or develop cost-effectiveness over time for the incremental net benefit, another measure of cost-effectiveness.

Finally, chapter 4 discussed the relationships between intention-to-treat cost-effectiveness and complier average causal cost-effectiveness. For the incremental net benefit, the standard relationship–originally noted by Angrist et al. (1996)– applies. That is, the intention-to-treat effect is proportional to the complier average causal effect by a factor equal the probability of being a complier, where a complier is defined to be a participant who would take the assigned treatment regardless of which treatment was assigned. For the incremental cost-effectiveness ratio, the intention-to-treat estimate can be interpreted as the complier average effect. An understanding of these relationships can help provide more context to researchers who wish to understand about cost-effectiveness when study compliance may differ from other situations. Future research in this area could include completing more extensive simulation studies and applying the methods to a dataset where complier average effects would be of interest–such as in the setting of mental health treatment.

All three projects in this dissertation described nonparametric methods in cumulative medical cost data in the presence of right censoring. The overarching statistical themes in this work therefore pertained to survival analysis. One main difficulty that was addressed in each topic was the problem of induced informative censoring on the cost scale. In each project, we employed estimators with inverse probability weighting to avoid using Kaplan-Meier type estimators, which would require non-informative censoring on the cost scale.

Additionally, we used empirical process methods in order to justify our proposed inference procedures throughout the dissertation. While all three projects had much that was alike, this project was also quite varied. In this work, we were able to learn and touch upon topics such as recurrent survival (gap-time) data, measures of cost-effectiveness, methods for dealing with confidence bands, and statistical challenges when dealing with trials with noncompliance.

BIBLIOGRAPHY

O. Aalen. Nonparametric inference for a family of counting processes. *The Annals of Statistics*, 6(4):701–726, 1978.

S. M. Al-Khatib, K. J. Anstrom, E. L. Eisenstein, E. D. Peterson, J. G. Jollis, D. B. Mark, Y. Li, C. M. O'Connor, L. K. Shaw, and R. M. Califf. Clinical and economic implications of the Multicenter Automatic Defibrillator Implantation trial-II. *Annals of Internal Medicine*, 142(8):593, 2005.

A. Ament and R. Baltussen. The interpretation of results of economic evaluation: explicating the value of health. *Health Economics*, 6(6):625–635, 1997.

J. D. Angrist, G. W. Imbens, and D. B. Rubin. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434):444–455, 1996.

S. G. Baker. Analysis of survival data from a randomized trial with all-or-none compliance: estimating the cost-effectiveness of a cancer screening program. *Journal of the American Statistical Association*, 93(443):929–934, 1998.

H. Bang and A. A. Tsiatis. Estimating medical costs with censored data. *Biometrika*, 87 (2):329–343, 2000.

T. S. Bodenheimer and K. Grumbach. *Understanding Health Policy: A Clinical Approach.* McGraw Hill, New York, 5th edition, 2009.

A. H. Briggs, D. E. Wonderling, and C. Z. Mooney. Pulling cost-effectiveness analysis up by its bootstraps: a non-parametric approach to confidence interval estimation. *Health Economics*, 6(4):327–340, 1997.

A. H. Briggs, C. Z. Mooney, and D. E. Wonderling. Constructing confidence intervals for cost-effectiveness ratios: an evaluation of parametric and non-parametric techniques using monte carlo simulation. *Statistics in Medicine*, 18(23):3245–3262, 1999.

M. A. Chaudhary and S. C. Stearns. Estimating confidence intervals for cost-effectiveness ratios: an example from a randomized trial. *Statistics in Medicine*, 15(13):1447–1458, 1996.

S. I. Chen and H. Zhao. Estimating incremental cost-effectiveness ratios and their confidence intervals with different terminating events for survival time and costs. *Biostatistics*, 14 (3):422–432, 2013.

D. S. Cheng, J.and Small, Z. Tan, and T. R. Ten Have. Efficient nonparametric estimation of causal effects in randomized trials with noncompliance. *Biometrika*, 96(1):19–36, 2009.

R. J. Cook and J. F. Lawless. Marginal analysis of recurrent events and a terminating event. *Statistics in Medicine*, 16(8):911–924, 1997.

B. Efron and R. J. Tibshirani. *An Introduction to the Bootstrap*. Chapman & Hall, New York, NY, 1993.

E. C. Fieller. Some problems in interval estimation. *Journal of the Royal Statistical Society. Series B*, 16(2):175–185, 1954.

C. Genest. Frank's family of bivariate distributions. *Biometrika*, 74(3):549–555, 1987.

P. P. Glasziou, R. J. Simes, and R. D. Gelber. Quality adjusted survival analysis. *Statistics in Medicine*, 9(11):1259–1276, 1990.

D. F. Heitjan, A. J. Moskowitz, and W. Whang. Problems with interval estimates of the incremental cost-effectiveness ratio. *Medical Decision Making*, 19(1):9–15, 1999.

C. Y. Huang and M. C. Wang. Nonparametric estimation of the bivariate recurrence time distribution. *Biometrics*, 61(2):392–402, 2005.

Y. Huang. Cost analysis with censored data. *Medical Care*, 47(7S):115–119, 2009.

Y. Huang and T. A. Louis. Nonparametric estimation of the joint distribution of survival time and mark variables. *Biometrika*, 85(4):785–798, 1998.

E. L. Kaplan and P. Meier. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282):457–481, 1958.

D. Y. Lin, T. R. Fleming, and L. J. Wei. Confidence bands for survival curves under the proportional hazards model. *Biometrika*, 81(1):73–81, 1994.

D. Y. Lin, E. J. Feuer, R. Etzioni, and Y. Wax. Estimating medical costs from incomplete follow-up data. *Biometrics*, 53(2):419–434, 1997.

X. Luo and C. Y. Huang. Analysis of recurrent gap time data using the weighted risk-set method and the modified within-cluster resampling method. *Statistics in Medicine*, 30 (4):301–311, 2011.

A. J. Moss, W. Zareba, W. J. Hall, H. Klein, D. J. Wilber, D. S. Cannom, J. P. Daubert, S. L. Higgins, M. W. Brown, and M. L. Andrews. Prophylactic implantation of a defibrillator in patients with myocardial infarction and reduced ejection fraction. *New England Journal of Medicine*, 346(12):877–883, 2002.

W. J. Moss, A. J .and Hall, D. S. Cannom, J. P. Daubert, S. L. Higgins, H. Klein, J. H. Levine, S. Saksena, A. L Waldo, D. Wilber, M. W. Brown, and H. Moonseong. Improved survival with an implanted defibrillator in patients with coronary disease at high risk for ventricular arrhythmia. *New England Journal of Medicine*, 335(26):1933–1940, 1996.

A. Mushlin, J. Zwanziger, E. Gajary, M. Andrews, and R. Marron. Approach to cost-effectiveness assessment in the MADIT trial. *American Journal of Cardiology*, 80(5B): 33F–41F, 1997.

A. I. Mushlin, W. J. Hall, J. Zwanziger, E. Gajary, M. Andrews, R. Marron, K. H. Zou, and A. J. Moss. The cost-effectiveness of automatic implantable cardiac defibrillators:: Results from madit. *Circulation*, 97(21):2129–2135, 1998.

W. Nelson. Theory and applications of hazard plotting for censored failure data. *Technometrics*, 14(4):945–966, 1972.

H. Nie, J. Cheng, and D. S. Small. Inference for the effect of treatment on survival probability in randomized trials with noncompliance and administrative censoring. *Biometrics*, 67 (4):1397–1405, 2011.

B. J. O'Brien, M. F. Drummond, R. J. Labelle, and A. Willan. In search of power and significance: issues in the design and analysis of stochastic cost-effectiveness studies in health care. *Medical Care*, 32(2):150–163, 1994.

E. A. Pena, R. L. Strawderman, and M. Hollander. Nonparametric estimation with recurrent event data. *Journal of the American Statistical Association*, 96(456):1299–1315, 2001.

A. I. Phelps, C. E .and Mushlin. On the (near) equivalence of cost-effectiveness and cost-benefit analyses. *International Journal of Technology Assessment in Health Care*, 7(1): 12–21, 1991.

D. Polsky, H. A. Glick, R. Willke, and K. Schulman. Confidence intervals for cost-effectiveness ratios: a comparison of four methods. *Health Economics*, 6(3):243–252, 1997.

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

J. M. Robins and A. Rotnitzky. Recovery of information and adjustment for dependent censoring using surrogate markers. In N. P. Jewell, K. Dietz, and Farewell V. T., editors, *Aids Epidemiology, Methodological Issues*, pages 297–331. Birkhauser, Boston, 1992.

J. M. Robins, A. Rotnitzky, and L. P. Zhao. Estimation of regression coefficients when some regressors are not always observed. *Journal of the American Statistical Association*, 89 (427):846–866, 1994.

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

T. V. Salukhe, K. Dimopoulos, R. Sutton, A. J. Coats, M. Piepoli, and D. P. Francis. Life-years gained from defibrillator implantation. *Circulation*, 109(15):1848–1853, 2004.

G. D. Sanders, M. A. Hlatky, and D. K. Owens. Cost-effectiveness of implantable cardioverter–defibrillators. *New England Journal of Medicine*, 353(14):1471–1480, 2005.

Cleveland Clinic Staff. Implantable cardioverter-defibrillator? URL: www.myclevelandclinic.org/heart/services/tests/procedures/icd.aspx, November 2011a. Accessed: 03/12/2013.

Mayo Clinic Staff. Implantable cardioverter-defibrillators. URL: http://www.mayoclinic.com/health/implantable-cardioverter-defibrillator/MY00336, November 2010. Accessed: 03/12/2013.

NHLBI Staff. Who needs an implantable cardioverter-defibrillator? URL: www.nhlbi.nih.gov/health/health-topics/topics/icd/whoneeds.html, November 2011b. Accessed: 03/12/2013.

A. A. Stinnett and J. Mullahy. Net health benefits a new framework for the analysis of uncertainty in cost-effectiveness analysis. *Medical Decision Making*, 18(2):S68–S80, 1998.

R. L. Strawderman. Estimating the mean of an increasing stochastic process at a censored stopping time. *Journal of the American Statistical Association*, 95(452):1192–1208, 2000.

Gijs van de Wetering, Willem H Woertman, and Eddy M Adang. Time to incorporate time in cost-effectiveness analysis. *The European Journal of Health Economics*, 13(3):223–226, 2012.

A. W. Van der Vaart. *Asymptotic Statistics*. Cambridge University Press, New York, 1998.

B. A. Van Hout, M. J. Al, G. S. Gordon, and F. F. H. Rutten. Costs, effects and c/e-ratios alongside a clinical trial. *Health Economics*, 3(5):309–319, 1994.

M. C. Wang and S. H. Chang. Nonparametric estimation of a recurrent survival function. *Journal of the American Statistical Association*, 94(445):146–153, 1999.

M. C. Wang and Y. Q. Chen. Nonparametric and semiparametric trend analysis for stratified recurrence times. *Biometrics*, 56(3):789–794, 2000.

A. R. Willan and D. Y. Lin. Incremental net benefit in randomized clinical trials. *Statistics in Medicine*, 20(11):1563–1574, 2001.

A. R. Willan and B. J. O'Brien. Confidence intervals for cost-effectiveness ratios: An application of Fieller's theorem. *Health Economics*, 5(4):297–305, 1996.

A. R. Willan, D. Y. Lin, R. J. Cook, and E. B. Chen. Using inverse-weighting in cost-effectiveness analysis with censored data. *Statistical Methods in Medical Research*, 11(6): 539–551, 2002.

H. Zhao and L. Tian. On estimating medical cost and incremental cost-effectiveness ratios with censored data. *Biometrics*, 57(4):1002–1008, 2001.

H. Zhao and A. A. Tsiatis. A consistent estimator for the distribution of quality adjusted survival time. *Biometrika*, 84(2):339–348, 1997.

H. Zhu. Non-parametric analysis of gap times for multiple event data: An overview. *International Statistical Review*, 82(1):106–122, 2013.

J. Zwanziger, W. J. Hall, A. W. Dick, H. Zhao, A. I. Mushlin, R. M. Hahn, H. Wang, M. L. Andrews, C. Mooney, H. Wang, and Moss A. J. The cost effectiveness of implantable cardioverter-defibrillators: results from the multicenter automatic defibrillator implantation trial (MADIT)-II. *Journal of the American College of Cardiology*, 47(11):2310–2318, 2006.

VITA

Laura was born in New York City in 1986. She attended the George Washington University in Washington, DC for college, where she received her Bachelor of Science in statistics in 2008. For graduate school, she attended the University of Washington, obtaining a Master's of Science in biostatistics in 2011. After graduating, she plans to move back to Washington, DC, where she will start a position at the US Food and Drug Administration.