# ESTIMATING THE RELATIONSHIP BETWEEN A TRANSIENT EFFECT AND THE ONSET OF AN ACUTE EVENT: A COMPARISON OF THE CASE-CROSSOVER DESIGN AND COHORT DESIGN

**BY CARLIN PATRICK BRICKNER**

**A dissertation submitted to the**

**School of Public Health**

**Rutgers, The State University of New Jersey**

**In partial fulfillment of the requirements**

**for the degree of**

**Doctor of Public Health**

**Written under the direction of**

**Dirk Moore, PhD & Shou-En Lu, PhD**

**and approved by**

_____

_____

_____

_____

**New Brunswick, New Jersey**

**January, 2015**

## ABSTRACT OF THE DISSERTATION

# Estimating the relationship between a transient effect and the onset of an acute event: a comparison of the case-crossover design and cohort design

**by CARLIN PATRICK BRICKNER**

**Dissertation Directors: Dirk Moore, PhD & Shou-En Lu, PhD**

The case-crossover design was first published in 1991 as an epidemiological method to estimate the transient effect of an exposure on an acute event in research where primary data collection is conducted. Since the inception of the case-crossover design, the quality and availability of data warehouses has become standard. Health care providers and insurers have migrated from recording routinely collected patient information on paper to using electronic health records which are stored in data warehouses. This development has enabled researchers to observe the same acute events and exposures of interest in the traditional case-crossover paradigm at any time the patient is in care without expending the resources associated with primary data collection. Recent epidemiological studies have implemented the case-crossover design in situations where the data necessary for a retrospective cohort design are readily available. The case-crossover design's main appeal is that it implicitly controls for time-invariant characteristics of each patient in the study, measured or unobserved, by utilizing conditional logistic regression. In a retrospective cohort, an investigator typically would choose between using a Cox Proportional Hazard Model or a longitudinal logistic regression model. Since researchers also are interested in studying the transient effect of an exposure on subsequent acute events in an observational

setting, and since developments in health information technology have provided researchers with more plentiful and detailed data than were available when the case-crossover design originally was proposed, researchers can now select from variety of methods. This thesis shows how the case-crossover design compares to a time-dependent covariate analysis in a cohort setting, and provides recommendations when one design preferable over the other. This thesis makes an important connection between the two designs, and proposes that the principle of lagged covariates can be applied in the case-crossover design. Furthermore, this thesis also proposes a two parameter, geometric lag estimation method which can describe a non-linear, deteriorating effect within the case-crossover design setting.

# Acknowledgments

I wish to thank, first and foremost, Dr. Moore and Dr. Lu for serving as the directors of this dissertation. I could not have completed this dissertation without their guidance and expertise. I would also like to thank Dr. Demissie and Dr. Rosati for sitting on my committee and providing thoughtful and insightful recommendations and discussion. I would like to thank my colleagues Bob, Tim, and Rocco who have provided support, advice and mentorship over the years. The accumulation of their work provided a foundation for the VNSNY Medication and Fall Study. I would also like thank my mother and father, who have always provided encouragement and steered me in the right direction throughout my academic journey. Their support has opened unlimited opportunities for which I am eternally grateful. Last but not least, I would most especially like to thank my wife Mihriban, who has provided unwavering support over the years and who has made as many sacrifices as I during the pursuit of this accomplishment. I share this feat with her.

# Dedication

I would like to dedicate this dissertation to my lovely wife Mihriban and my son Deren. Deren, you have come crashing into this world as I complete this chapter in my life and I look forward to starting the next chapter with you and your mom.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

In this thesis, a specialized epidemiological study design called the "case-crossover" design will be considered, in addition to a comparison of its properties to those of the much more common "prospective" design. To understand the case-crossover design, it is helpful to consider two other motivating designs. The randomized crossover deign, which is a repeated response design, where all subjects receive all treatments and the sequence of treatments is randomized. The effect of the treatment is estimated by comparing within subject responses, and as a result the self controlled nature of the design allows the estimates to be more efficient than a randomized treatment design. In a study where the acute event is rare, randomized studies can be infeasible due to the cost to conduct such a study. In addition, many times researchers desire to study the effect of an exposure that would be unethical to randomize to patients.

The other motivating design is the case-control design, which is an epidemiological design used in retrospective studies. Case-control studies observe subjects who experience an event and are referred to as a case; patients who did not experience the event are sampled from the population giving rise to the cases and are referred to as controls. In this analysis, the outcome is considered fixed and the covariate is assumed to be random, contrary to the assumptions of a cohort study. The case-control study is a retrospective, follow-up study which samples the cases and controls from a full census of the target population. The framework, properties, and methodology of both matched and non-matched case-control studies have been established primarily due to the work by Breslow and Day [11, 12].

According to the case paradigm, cohort studies should have a case-control counterpart, and the case-crossover design was proposed as one such counterpart to a cohort study, where patients cross between periods of exposure and nonexposure [59]. Maclure developed

the case-crossover design to study the effects of time-dependent exposures on the risk of an imminent event [59]. The case-crossover design only samples individuals who have experienced the event retrospectively and is classified as a case only design. It is applied in settings where the exposure is intermittent, the effect is immediate and transient, and the outcome is rare and acute. The case-crossover design combines features from both randomized crossover design and case-control designs; cases serve as their own controls and outcome-based sampling.

The case-crossover design samples cases from the cohort and deploys a within subject reference strategy, for which the distribution of exposures are compared between a case-time to that in a control-time. The case-time is selected by the investigator to capture an exposure which is hypothesized to elevate the hazard during the time when the event was observed. The investigator must also match this case-time to one or more reference periods in which the individual is assumed to be event free. Much of the theoretical work for the case-crossover design borrows from the work established for matched case-control studies [59, 62, 65]. Conditional logistic regression is often utilized to obtain an estimate of the hazard ratio for a one unit change in the exposure, for which any confounders that do not change over time are implicitly controlled for by the design. McNemar's odds ratio may be used when the exposure is a binary indicator. The within subject reference of individuals, who only experience the event, requires several assumptions, in order for estimates and inference to be valid. The assumptions include that the baseline hazard is small or constant, the exposure distribution in any of the time intervals is globally exchangeable within matched sets, censoring is non-informative, no carryover effects exist and the within subject correlation structure must be independent in applications where multiple events occur.

The case-crossover design is most prominently employed in the study of ambient air pollution and acute health effects[14, 15, 27, 47, 52, 53, 54, 78, 96, 93, 95]. It has also been used in a wide variety of other applications. Some examples include the association between anger, anxiety, curiosity and myocardial infarction [65]; cell phone use and automobile accidents [63, 80] ; emotional stress as a trigger of falls leading to hip or pelvic fractures [66]; alcohol and gout [110]; the effectiveness of condoms and sexually transmitted disease [105];

Medicare costs attributable to a fall [89]; folic acid antagonist use with birth defects [39]; the effects of vaccination [76]; and nurse staffing and nosocomial infections [43]. The case-crossover design is also prominently applied in pharmacoepidemiology, where it is common to explore how changes in medications may trigger an acute event [18, 64, 71, 87, 88, 104].

Several studies have utilized the case-crossover design to study how changes in patient medication regimen are associated with the risk of an imminent fall [71, 87, 88]. Patient exposure to medications for these studies were obtained from primary data collection methods in institutional settings. These studies also provided motivation to conduct a similar study to a population of home health care patients. Home health care consists of a mix of skilled nursing and therapy visits in the home, as well as home health aide services, to a mostly frail, elderly population, requiring post-acute services. The Visiting Nurse Service of New York (VNSNY) is the largest not-for-profit home health care agency in the nation. Between 2010 to 2011 there were 192,438 admissions into VNSNY's certified home health care program (CHHA). Of these admissions, 4.0% (7840) were documented in the electronic health record to have fallen at least once while receiving home health care services. In addition to recording falls, a wide range of patient characteristics, both the detailed timing of patient medication regimen and baseline patient characteristics, are recorded in the patient electronic health records. The VNSNY Fall and Medication Data invoke the research question, "Do changes in patient medication regimen increase the risk of an imminent fall in a home health care population?"

The VNSNY Fall and Medication Data contains baseline patient characteristics on both patients who experienced a fall and for those who did not. The case-crossover design, however, was developed in a primary data collection setting. Applying the case-crossover design to the VNSNY Medication and Fall Data discards all patients in the cohort who did not fall, accounting for 96% of the original data. The case-crossover design is a highly efficient design due to its within patient comparison. However, a design that utilizes the remaining 96% of the cohort, which the case-crossover design ignores, is likely to obtain estimates with much better efficiency. Given the data available in the VNSNY Fall and Medication Data, another approach to study the association of a medication change and the risk of falling, is to conduct a retrospective study, deploying a Cox hazards with

time-dependent covariates. The Cox hazard model can obtain estimates for the same hazard function assumed by the case-crossover design. The case-crossover design and the Cox hazard model have likelihood functions that share a similar form, but differ in that the case-crossover design makes with subject comparison and the Cox hazard model uses information from other patient at risk in the cohort. The Cox hazard model uses a cohort-based sampling, while the case-crossover design deploys an outcome-based sampling and reverses the cohort design's time scale. The purpose of this study is to explore the performance of both designs and compare performance under different circumstances.

Previous studies, which have explored the association between a medication change and falling, have only summarized the relationship assuming a simple, on and off relationship by using just a single parameter describing an elevated risk resulting from the exposure change amidst a common background risk [71, 87, 88]. This study, however, hypothesizes that if an association exists it may have a more complicated form than addressed by these studies. Furthermore, Maclure proposed the case-crossover design with the assumption that the risk is elevated for a period of time after the exposure change and than dissipates over a period of time, before returning to a constant, background risk. The second objective of this study is to propose an approach for estimating an effect in the case-crossover design, when a non-linear association is present.

This dissertation contains an additional five chapters. In Chapter 2, a literature review is presented, which covers the methodology of the case-crossover design, similarities to other designs, examples of application and the relevance of studying medication changes and the risk of falling. A methodological framework, for both the case-crossover design and the cohort design using a Cox hazard model, is presented in Chapter 3, along with drawing comparisons of the similarities between the designs. In addition, Chapter 3 provides an approach to exploring nonlinear transient effects and proposes a method to estimate a deteriorating effect, using the case-crossover design. Chapter 4 contains an evaluation of multiple simulation studies deploying both designs concurrently for various scenarios. Chapter 4 also contains simulations evaluating the proposed estimation method for nonlinear associations. A thorough analysis of the VNSNY Fall and Medication Data is presented in Chapter 5, which applies the methodology covered in Chapter 3. Finally, Chapter 6

provides a summary and discussion of this paper and makes recommendations to investigators.

# Chapter 2

# Literature Review

This chapter provides a thorough literature review on the case-crossover design history, methodology and similarities with other designs. The content and setting of the original case-crossover design is important to first understand since the original study occurred in a primary data collection setting. The next section presents the accumulation of knowledge about the methodology and assumptions of the case-crossover design, followed by a section reviewing the literature which has covered its similarities to other methodology. Another section is dedicated to reviewing the settings in which the case-crossover design has been applied. The second half of the literature review provides a background on fall and medication research; this includes the relevance, factors known to identify who is likely to fall, and also the factors which may trigger a fall.

## 2.1 Case-Crossover Design

At the time the case-cross over design was proposed, it was common belief that myocardial infarctions occur purely at random over time. Mittleman, Maclure and colleagues sought to study subject activities in the time leading up to a mycoardial infarction, thus the name, The Onset Study [59]. As a result, Maclure developed the case-crossover design to study the transient effect of a time-dependent exposure on an acute event in an observational setting [59].

Maclure assumed that there is usually a common, constant, background risk of an acute event. He also assumed that the temporal effect of an exposure, $Z_{it}$ for patient $i$ at time $t$, is initiated by a possible induction period followed by an elevated risk that dissipates over a period of times before returning to the background risk. Maclure proposed the epidemic curve as motivation for the case-crossover design and is reconstructed in Figure 2.1. By

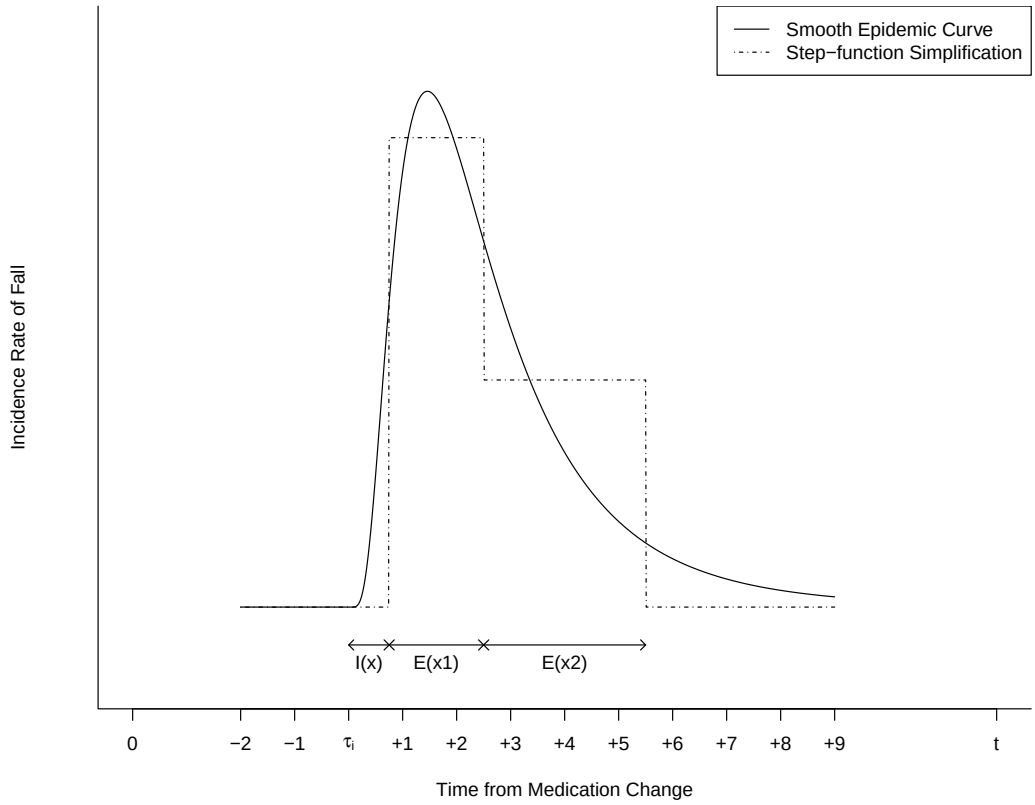


Figure 2.1: Illustration of Maclure's assumed epidemic curve for the case-crossover design [59]. The curve demonstrates incidence of an acute-onset event (i.e. fall) after observing a point exposure (i.e. change in medication). The population induction time is indicated by $I(x)$, and a step function is suggested to estimate the curve through $E(x_1)$ and $E(x_2)$. $E(x_1)$ indicates the time of highest risk effect period, and $E(x_2)$ a more moderate risk effect period.

fixing the event, and comparing the distribution of exposures at the case and control times, Maclure used features of both the cross-over and case-control designs to describe the transient effect of an exposure on an acute event. The case-crossover design estimates the average incidence rate ratio after observing a one unit increase in the exposure [59].

The case-crossover design is applied to research questions where the exposure is intermittent, the effect of the point exposure's risk is immediate and transient, and the outcome is rare and acute. Maclure showed that only cases are required to estimate the transient effect of an exposure, and because the design is like a matched case-control study, subject level effects are not estimable. As a result, both measured and unmeasured confounders and are implicitly controlled for by the design.

Maclure also presented several threats to the validity of the case-crossover design which

should be considered in practice. Similar to the randomized crossover design, carryover and period effects must be considered. The case-crossover design is an exploration of a transient effect, and so any doubt about the duration of the effect should conservatively be considered during the deployment of the selection of case- and control-times. Failure to allow ample time for an effect to return to the background hazard in the reference strategy for the case-crossover design, results in the control time containing the elevated risk and estimates that are biased towards the null.

The second and third threats to validity pertain to within-individual confounding; which include the consideration of treatment sequencing and patient assignment, in addition to the definition of cross-over rules and timing [59]. The case-crossover design implicitly controls for any confounders that are constant over time. However, since the case-crossover deploys reference strategies within an individual it is susceptible to time-dependent confounders. Any additional time-dependent exposures that are associated with the event and the exposure of interest to the study must also be controlled for in the case-crossover design in order to obtain unbiased estimates. Careful consideration in understanding the sequence of how multiple time-dependent confounders are associated should be considered when deciding on the timing and duration of case- and control-times.

Selection bias is another of Maclure's threats to validity and should be considered when determining the reference strategy [59]. Some time-dependent exposures may influence patient participation in the study. Another type of selection bias may exist if the probability of observing the exposure is different in the case- versus the control-time. Recall bias results in pushing estimates away from the null in a case-crossover design. Systematic bias would exists in a case-crossover design analyzing medication changes and falls if selected control times stratify the date in which a new drug was released; resulting in the probability of exposure during the control would be zero. Similar to the assumption in survival analysis, in which censoring is assumed to be noninformative, the case-crossover design assumes that drop-outs are independent of the event and exposure. Finally, applying the methods used for matched studies, such as McNemar's test or conditional logistic regression, may provide inappropriate estimates of the effect in the presence of repeated outcomes [59].

### 2.1.1 Methodology

Maclure proposed using the Mantel-Haenszel method to estimate relative risk of a binary exposure change [59]. Marshall and Jackson presented a general maximum likelihood approach to analyze any type of exposure in the case-crossover design [62]. They derived a conditional likelihood assuming a proportional hazards model to describe the dependence between the time-dependent exposure and the risk of the event. Furthermore, they showed several special cases of applying this conditional likelihood including how the Mantel-Haenszel estimator approximates the conditional likelihood estimate, in scenarios when the exposure is continuous, arises from a mixture distribution, or the association is described by joint effects. Vines and Farrington demonstrated that in order for the conditional likelihood to be valid, a strong condition assuming the exposures are globally exchangeable within matched sets [102]. Greenland proposed a general framework and likelihood for the analysis of case only designs such as case-crossover, case-genotype, and case-specular studies while also providing an extension to leverage information from traditional controls into the analysis [31].

Mittleman, Maclure and Robins assessed five different reference strategies in their analysis of myocardial infarctions in the Onset Study [65]. They assessed the relative efficiency of estimates obtained from case-crossover designs selecting $M = 1$, 3, 4, 6, 8, 12, and 25 control times, and found that the efficiency of relative risk estimates varied greatly based on the selection of $M$ . They presented a calculation for the relative efficiency of estimates obtained by the case-crossover design compared those from a cohort study. This was based on previously work developed by Breslow and Day for case-control studies [12]. The relative risk of the case-crossover design or the case-control design depends on the number of cases, number of controls, the probability of exposure for the case and control, and the relative risk. When the association between the exposure and the event is null, the relative efficiency of a 1:M case-crossover design achieves $100 \cdot \frac{M}{M+1}$ % of that achieved by the cohort design [65].

Much of the methodological work surrounding the case-crossover design has focused on the issues that arise in the presence of unmeasured time-dependent confounders or

underlying trends in the exposure. This may become apparent in pharmacoepidemiology where the prescription of a therapeutic may indicate a decline in a subjects health status. Therefore, applying a case-crossover design to a health related event when this situation exists may result in biased estimates of the transient effect. Another source of bias in pharmacoepidemiology studies may arise when the health of the population is generally declining, and as a result, the probability of the exposure is increasing over the follow-up period. Applying the case-crossover design in this situation would likely mix together potential causal effects with the natural increase in the exposure.

Suissa was the first to propose an adjustment to the case-crossover design to account for trends in the exposures [85]. Suissa proposed a method, the case-time-control design, which adjusts the case-crossover estimates with estimates of bias due to exposure trends. These estimates of bias due to trends in exposure are obtained by applying the case-crossover design to comparable set of individuals who did not experience the event [85, 91]. Simulation studies demonstrated that it out performs the case-crossover design in scenarios where the treatment is correlated with time [1]. Further comparison of the two designs in applied settings have supported the expectations on performance in practice [39, 43].

Careful consideration has to then be taking into account when selecting control patients, because a potential imbalance in baseline risk may exists for patients with and without the event. Failure to do can reintroduce selection bias into the case-time-control design estimates either away or towards the null [30, 69]. The case-case-time-control design was proposed as a remedy to the case-time-control design. The case-case-control-design also sought to select another sample to control for time trends, but proposed to select future cases instead of patients without the event [103]. This design must assume that there is no within-subject correlation for patients with multiple events.

Several reference design strategies were proposed in the field of environmental epidemiology to account for trends in the exposure. Exposure trends are of particular importance when applying the case-crossover design to air pollution studies. Navidi showed that trends could be controlled for in certain settings by implementing a bidirectional sampling of the control times in the case-crossover design [69]. That is, a control time is selected both prior

to and after the event. This is only applicable when the exposure after the event is independent of the event, which environmental air pollution is one of the few fields which can make this assumption. Navidi's design is a full-stratum design where all days are selected as control days except for the day of the event. Several other bidirectional designs applied to environmental studies which built upon Navidi's full stratum design were proposed. These include the symmetric bidirectional design, the semisymmetric bidirectional design, and the time stratified design [7, 57, 70]. Autocorrelated exposures and overlap bias are additional forms of bias, which must be considered in air pollution studies deploying the bidirectional reference strategies for the case-crossover design. The presence of autocorrelated exposures and overlap bias result in the score function to the conditional maximum likelihood estimate are not valid since they does not have mean zero [44, 57].

The original case-crossover design was conceptualized and designed for univariate event data [60]. However, many acute events such as falls, gout attacks, automobile accidents, and sexually transmitted infections are recurrent in practice [5, 71, 105, 110]. When studying events that are recurrent in nature, it is necessary to account for the within-subject correlation among the recurrent events. For example, it may be naive to assume that later falls are independent of earlier medication changes and falls. Hoffman et al. showed that the Within-Cluster Resampling (WCR) method can be used so that univariate analysis can be applied to data which randomly selects, with replacement, a single observation within each cluster and averages the estimates from repeating the sampling and univariate analysis over large number of iterations [40]. By repeating the sampling a large number of times, parameters are robust to non-ignorable cluster size. Rieger and Weinberg applied WCR using conditional logistic regression on clustered binary outcome data in a case-control design [82]. Luo and Sorock extended the application to the case-crossover design and used simulations to demonstrate that a single event could be sampled for each subject along with the corresponding matched sets of case- and control-periods. This allows the researcher to leave the correlation structure among multiple matched sets within each subject unspecified [58]. Estimates remain valid when the number of matched sets or events per subject is related to the outcome. Luo named this application of the WCR method to case-crossover design as the Within-Subject Pairise Resampling (WSPR). In addition,

the WCPR is applied to the case-crossover design is the same as weighting estimating equations to solve for relative risk estimates. This connection was based on Williamson's work that showed if the WCR sampling algorithm followed a uniform distribution with a probability mass of $\frac{1}{n_{\tilde{t}_i}}$, then weighted estimating equations could be used to estimate the weighted score function from a pooled analysis approach [109]. Where $n_{\tilde{t}_i}$ is the number of events for subject $i$.

By definition the acute event in the case-crossover design is rare. When multiple exposures are of interest to the investigator, variable selection poses challenges due to small sample size in multivariate analysis. Avalos and colleagues were interested in exploring whether an exposure to any of 89 pharmacotherapeutic classes were related to 422 automobile accidents among an elderly population [5]. They tackled the issues associated with an exploration of many potential associations using the lasso and elastic net penalties when obtaining parameter estimates from the conditional likelihood as a variable selection method for case-crossover designs. They recommended the use of the lasso and elastic net over other related methods due to the superior performance in simulation studies and data analysis, with the exception that they yielded high false positive results in simulations where the correlation between parameters was negative.

### 2.1.2 Comparison to Other Designs and Methods

What was thought to be an alternative method to analyze daily environmental exposures and case-only data, log-linear time series regression models were traditionally used to estimate the effect of a one unit change in the exposure on the total number of events [23] . A property of environmental epidemiological studies of air pollutants is that the exposure, $Z_t = Z_{it}$, is the same for all individuals $i$ at time $t$. Lu and Zeger showed that when the exposure is the same for all subjects the estimate obtained from the conditional logistic regression in the case-crossover design is a special case of the time series log-linear model [56]. The standard errors obtained in each analysis may be different since the log-linear model may take overdispersion into account while the case-crossover design assumes the Poisson variance [56]. Lu and collaborators then used this relationship to recommend model checking for influential points and overdispersion [55]. The presence of overdisperion

often indicates the presence of an unmeasured time-dependent confounder.

The self-controlled case series method was proposed as an alternative design to the case-crossover design to alleviate some strong assumptions made by the case-crossover [20, 21, 22]. These assumptions include the requirement that the exposure distribution is stationary and globally exchangeable. The self-controlled case series design is similar to the case-crossover design in that it only samples cases, and controls for all fixed confounders. However, it differs from the case-crossover design in that the likelihood is built upon assuming a Poison cohort model and conditions on the exposure having occurred during the observation period. Limitations of the design includes the requirement that the exposure is independent of the occurrence of the event and for non-recurrent events it only works when the event risk is small over the observation time [107].

Vibound et al. compared the performance of the case-crossover design to a case-control design in simulations and found that the case-crossover design consistently has lower type I error for varying exposure incidence rates [101]. For both designs the type I error decreased with the prevalence of the exposure. In addition, they found that the case-control design had higher power for relative risk less than eight, but lower power when the relative risk exceeded eight.

The study by Whang and collaborators used the case-crossover analysis and the Cox hazards model to distinguish between the short and long term effects of exercise [106]. Other studies have compared results of both methods to assess short term effects of a transient exposure only. Lepeule and colleagues proposed that a Cox hazards model with time dependent covariaties could be applied to ecologic studies exploring associations between short-term air pollution and mortality in lieu of the case-crossover design or time series analysis [51]. Furthermore, they compared their results applying a Cox hazards model to an earlier study which utilized a case-crossover design on the same population [24]. Peters et al. compared a Poisson time series, case-crossover, and Cox hazards model analysis and observed similar results when exploring the relationship between ambient air pollution and exacerbation of cardiovascular disease [77].

### 2.1.3 Application

The first study to apply the case-crossover design was The Onset Study [65]. The study identified patients admitted to one of 22 hospitals with a myocardial infarction. The patients were interviewed and the self reported time of myocardial infarction was recorded as well as measures of anger, anxiety and curiosity during each of the 26 hours prior to the onset of the myocardial infarction. The patients also reported there average or usual annual frequency displays of anger. The case-crossover design was applied to compare the level of anger two hours prior to the myocardial infarction compared to same two hour period the day prior. For the anger measurement, another case-crossover design was used with the usual frequency measurement of anger as the control. The study found that an exposure above the 75th percentile for anger and curiosity increased the risk of a myocardial infarction in the following two hours by 1.9 (95%CI=[1.3, 2.7]) and 1.6 (95% CI=[1.1, 2.2]) times respectively. There was not enough evidence (p-value=0.70) to conclude that an exposure to curiosity in the two hours prior to myocardial infarction was associated with the event.

One of the most prominent and early applications of the case-crossover design is the Redelmeier and Tibshirani study of cell phone usage and motor vehicle collisions [80]. Similar to The Onset Study, it is unrealistic to randomize cellular phones to individuals and wait to observe motor vehicle collisions. Persons included in the study were identified, along with the date of the collision, if they reported to the Collision Reporting Centre in Toronto. Individuals only reported to this center if they were in a collision that did not involve injury, criminal activity, or the transport of dangerous goods. The study concluded that cellular phone activity increased the risk of a motor vehicle collision by 4.3 (95% CI = [3.0, 6.5]) times. A similar study published eight years later found that cell phone usage increased the risk of a crash resulting in hospitalization by 4.1 (95% CI=[2.2, 7.7]) times [63].

The study of the association of air pollution with various health-related outcomes is a discipline that commonly deploys the case-crossover design [15]. The ambient exposures

often studied in this field includes black smoke, carbon monoxide, particular matter, nitrogen dioxide, nitrogen oxide, ozone, sulfur dioxide with several health related outcomes such as mortality, respiratory events, hospitalizations, heart failure, COPD related events [15]. Particulate matter has been found to increase the risk of heart failure [78], asthma and other respiratory diseases among children [53, 54, 96], mortality [27, 47, 93], medical emergency calls [14]. There are several studies, however, which did not find associations between particulate matter and health related outcomes [52, 95]. Carracedo-Martinez and colleagues also found that some pollens increased the risk of medical emergency calls [14]. Measures of sulfur dioxide and ozone were found to increase the risk of mortality [47].

The case-crossover design has also recently grown rapidly in its application in pharmacoepidemiology. A recent systematic review found that since 2008 there have been at least nine publications on the topic of case-crossover designs in the area of pharmacoepidemiology [18]. The most common event used as an outcome in pharmacopidemiology case crossover design studies is hospitalizations followed by falls and motor vehicle collisions[18]. Other primary outcomes are disease related events such cardiovascular or cerebrovascular events, which was the most published disease, followed by gastrointestinal, heaptic or renal, and respiratory events[18].

In a study resembling the motivating example for this proposal, Meuleners et al. applied the case-crossover design to a database of 616 individuals who were hospitalized after a motor vehicle crash [64]. The individuals were identified from the database using ICD10 codes, and were linked to a government subsidized prescription medication database. They found that benzodiazepines, antidepressants and opioid analgesics increased the risk of a motor vehicle crashes which required hospitalization.

Möller and colleagues conducted a study exploring emotional stress as a trigger of falls leading to hip or pelvic fractures [66]. They interviewed 137 patients who fell in one of two hospitals in Stockholm and obtained emotional stress measures which were used as exposures in the case-crossover design. The authors concluded that anger, sadness and stress increased the risk of falling by (95% CI=[2.7,54.7]), 5.7 (95% CI=[1.1, 28.7]), and 20.6 (95% CI=[4.5, 93.5]) respectively.

A study deploying the case-crossover design, among two other designs, to study the

association of therapeutics with the risk of death in a population of Pennsylvancia Medicare beneficiaries [104]. Estimates from the case-crossover design suggested that exposure to Lipid lowering, blood pressure regulating, glaucoma, glucose regulating, or osteoporosis therapeutics decreased the risk of death in the following 30 days. Wang and collaborators provided additional estimates of risk, which took into consideration the authors concern of exposure trends and time-dependent confounders. These adjustments resulted in null results for each medication with the exception of Osteoporosis, which resulted in an increase risk of death in the following 30 days [104].

The application of the case-crossover design has been applied in many other disciplines. Some examples include alcohol and gout [110]; the effectiveness of condoms and sexually transmitted disease [105]; Medicare costs attributable to a fall [89]; and folic acid antagonist use with birth defects [39]. Another notable study found that nurse to patient staffing levels at several hospitals is associated with increased risk of infection. Chew and collaborators found that various measures of inflammatory activation is associated with acute coronary syndromes [17]. Hambdige and collaborators utilized the case-crossover design to explore the safety of influenza vaccinations in children of age 6-23 months old. They concluded that influenza vaccines did not increase the risk for any of the medical related events they studied, and in several outcomes were preventative [35].

The Nurses' Health Study followed 121,701 female, registered nurses between the ages of 30 to 55 years. Baseline information was obtained about medical history, coronary heart disease risk factors, and life style. The woman were followed up on a biannual basis via mailed questionnaires that obtained new diagnoses and information on physical activity from 1986-2000. The level of physical exertion that the individual experienced right before death was extracted from the medical record or from next of kin. 288 of 84888 who responded to 1980 questionnaire were observed to have experienced a cardiac death. Whang et. al wanted to understand the risk associated with sudden cardiac death during moderate to vigorous exertion, and to assess the long-term effects of exercise on sudden cardiac death. A case-crossover analysis found that the relative risk of cardiac death during moderate to vigorous exertion was increased by 2.38 [1.23-4.60]. To the contrary, a proportional hazards model with a time-dependent covariate found that the long term

effect of moderate to vigorous exercise was inversely related to mortality [106].

## 2.2 Falls

Injuries are the leading cause of death for individuals over the age of 65, and the leading type of fatal injury is due to a fall [25]. Falls are also the leading cause of non-fatal injuries for individuals over the age of 65 [25]. In any given year, it is estimated that about one third of all community dwelling adults, age 65 or greater, will experience a fall [37, 100]. Between 9-14% of these falls will result in a serious injury[98, 99]. Furthermore, it is estimated that 65% of the elderly who experience a serious injury when falling will not be able to get up, and 47% of those who do not experience a serious injury [99]. Particular consideration should be given to females, who are at 2.48 (2.16, 2.85) times the risk of experiencing a fracture compared to their male counterparts [111].

Stevens and colleagues estimated that there were 10,300 fatal falls, and 2.6 million non-fatal falls that required medical intervention in 2000 [89]. The same study estimated that fatal falls cost 200 million U.S dollars, and that the cost associated with a non-fatal fall in the twelve months following was 19 billion U.S. dollars in the year 2000 [89]. Bishop et. al estimated that injuries in the elderly cost Medicare 9 billion U.S. dollars in 1999, where 67% of these costs were attributable to a fracture [9].

## 2.3 Risk Factors for Falling

A large body of work has accumulated over the years that has identified factors that are associated with individuals experiencing a fall in elderly populations. The factor which research has consistently shown has the strongest relationship with falling, is a previous fall [41, 72, 99, 100, 108]. Measures of mobility and balance have also been shown to increase the risk of falling [37, 75, 100, 108]. Hausdorff et al. studied the variation of a subject's stride as a predictor of falling in the following twelve months, and found that an increase of a one standard deviation change in variability increased the odds of falling by 5.3 (1.01, 27.2) [37]. Other factors that have been linked to the risk of falling include age [75, 108]; dependency in activities of daily living [75]; fear of falling [75]; hearing and vision

impairment [75, 99]; poor health related quality of life scores [75]; arthritis [72]; parkinson [72]; and urinary incontinence [100].

There is a large body of research that has found associations that link medication classifications to patients who experience a fall. Leipzig and collaborators conducted an extensive literature review of forty non-randomized controlled trials published between 1966 and 1996 and conducted a meta analysis of the results [48]. They found that the use of any psychotropic medication increases the risk of falling by 1.73 (95% CI= 1.52-1.97), neuroleptics by 1.50 (95% CI =1.25-1.79), sedatives by 1.54 (95% CI = 1.40-1.70), any antidepressants by 1.66 (95% CI = 1.40-1.95), tricyclic antidepressant by 1.51 (95% CI = 1.14, 2.00), and both short and long acting benzodiazepines by 1.48 (95% CI = 1.23, 1.77). More recent studies have been able to reproduce the findings of psychotropic medications and risks of falling in patients in long term care or psychiatric institutions [46, 74, 108]. Leipzig et. al also conducted a similar meta analysis of twenty nine studies to evaluate the evidence that has linked cardiac and analgesic drugs to risk of falling [49]. They found that the use of diuretic increases the risk of falling by 1.08 (95% CI = 1.02-1.16), type Ia antiarrhythimics, and digoxin by 1.22 (95% CI = 1.05-1.42) times. Leipzig et. al did not find enough evidence to conclude an association between the following medications classes and falling: thiazide diuretics, loop diuretics, beta blockers, centrally acting antihypertensives, ACE inhibitors, calcium channel blockers, nitrates, narcotic and nonnarcotic analgesics, NSAID, aspirin [49]. Other research has focused on the complexity of medication regimen, where a popular measure for medication complexity is whether or not a patient is on five or more medications, which several studies have found associations with five or more medications and the risk of falling [72, 108].

## 2.4 Events Triggering a Fall

Recent research has shifted the focus from studying which factors may identify those who are most likely to fall, to a focus on understanding the events that may trigger a fall. Neutel et al. found that a new benzodiazepine or antipyschotic drug increases the risk of falling in the next two days by 11.4 (95% CI = 1.5-89.0) fold, and a change in any drug increases the risk of an imminent fall by 1.8 (95% CI = 1.0-3.2) times [71]. The study

applied a 1:1 case-crossover design, deploying a reference strategy that looked for a new medication during the two days prior to the fall, and compared it to the control-time at 8-9 days prior to the fall. Neutal and colleagues considered two analysis: one, defined the unit of analysis as falls and assumed that multiple falls per patient were independent; two, they repeated the analysis only using the first fall [71]. There was not enough evidence to support that a new central nervous system, cardiovascular, antibiotic or gastrointestinal drug increased the risk of an imminent fall.

Sorock et al. applied a 1:1 case-crossover design to 158 patients in a long term care setting and found that changes in central nervous drugs increased the risk of falling within the next 1-3 days by 3.38 (95% CI =1.20, 9.49) times [88]. This study defined a change in medication as any new medication, dosage change, as-needed, or discontinued medication. This association was found by testing several different reference strategies; 1 vs 9, 1-2 vs 8-9, 1-3 vs 7-9, 1-4 vs 6-9 days prior to the fall. An odds ratio of 3.51 (95% CI = 1.05-11.68) was found for a change in central nervous system medications in the 1-2 day vs 8-9 day comparison, and no associations at the other case-crossover time comparisons. Sorock justified these findings at 1-3 days prior to the fall based on the 20 hour elimination half-life for the central nervous system drugs haloperidol and risperidone. Sorock's study did not find any relationships when exploring changes in gastrointestinal, hypoglycemic, antibiotics, cardiovascular, analgesics, and any non-CNS medications [88]. The study allowed patients to have multiple falls included in the analysis by applying a method to weight the fall's contribution by the number of falls the patient experienced [58].

Suto and colleagues found that an initial use of hypnotic, anti-anxiety, antihypertensive, or antiparkinson drugs increases the risk of falling by 8.42 (95% CI =3.12, 22.72), 4.18 (95% CI = 1.75, 10.02), 3.25 (95% CI =1.62, 6.5) and 2.44 (95% CI =1.22, 4.51) times, respectively, in a sample of 349 patients in an acute care hospital in Japan [87]. They applied a 1:3 case-crossover design, fixing the case-time at 0-2 days before the fall and the control-times at 6-8, 9-11, 12-14 days prior to the fall. Suto et. al justified the selection of the case-time based on the elimination half-life of the prescribed medication. Suto et. al did not find any relationship with an initial use of antipsychotic agents, antihistamines, antidiabetic agents, diuretics, and anti-ulcer agents [87].

Gribbin et. al applied a self-controlled case series analysis and found that the first prescription of thiazide increases the incident rate of falling in the subsequent 21 days by a magnitude of 2.8 (95% CI = 1.7, 4.57) [34]. In addition, they found that risk of falling 22-60 days after the new prescription increased by 1.16 (95% CI = 1.04, 1.30) fold for beta-blockers and by 1.15 (95% CI = 1.04, 1.28) times for angiotensin-converting enzyme inhibitors. There was not enough evidence to support an association with angiotensin-II receptor antagonist or calcium channel blockers.

A study by Butt et. al used a self-controlled case series design to find that a first prescription for any type of antihypertensive drugs increases the risk of falling in the following 45 days is 1.43 (95% CI = 1.19, 1.72) times [13]. Further exploration found incident rate ratios of 1.33 (0.94, 1.88) for thiazide diuretics, 1.53 (95% CI = 1.12-2.10) for ACE inhibitors, 1.41 (95% CI = 0.65, 3.05) angiotensin II receptor antagonist/blockers, 1.30 (95% CI = 0.83-2.02) for calcium channel blockers, and 1.58 (95% CI = 1.01, 2.48) for beta-adrenergic blockers [13].

## 2.5 Clinical Plausibility for Medications Triggering Falls

Controlling for demographics, medications, physical conditions, mental disorders, and hypertension, Bolton and colleagues found that several psychotropic medications are associated with the risk of osteoporotic fractures [10]. This study found that the use of selective serotonin reuptake inhibitors increase the odds of osteoporotic fracture by 1.45 (95% CI = 1.32- 1.59) times. These findings are reinforced by studies that have found that selective serotonin reuptake inhibitors decrease bone density [19, 36]. Bolton et. al also found that benzodiazepines and other monoamine antidepressants increase the risk of osteoportic fractures, but Lithium medications may actually be protective against fractures with an odds ratio of 0.63 (95% CI =0.43-0.93) [10].

Diuretics, antidepressants, antiepileptics, antipsychotics, chemotherapeutics, and recreational party drugs have been shown to have an association with hyponatremia [6]. Hyponatremia is a condition in which there is not enough sodium in the body, and has been shown to increase the risk of falls and fractures in the elderly [81]. This is likely due to

its effect giving rise to a mild cognitive impairment which results in unsteady gait[6]. In addition, hyponatremia is likely to cause or magnify osteoporosis, leaving elderly patients who fall at elevated risks of experiencing a fracture [6].

Berlie and Gardwood reviewed the literature of diabetic medications and falls and concluded that metformin is associated with vitamin B12 deficiency, which is linked to neuropathy, and neuropathy is associated with increasing the risk of falls [8]. They also point out that insulin secretagogues and insulin have been identified as characteristics of patients that fall, but can't provide a clinical explanation except that in observational studies these medications maybe identifying patients with hypoglycemia which results in reduced balance, strength and gait abnormalities. Furthermore, they reviewed the literature that has shown that thiazolidinediones are associated with bone density loss and increased risk of fractures.

## 2.6 Relevance

Several of the studies reviewed here deployed the case-crossover design in settings utilizing electronic administrative data [35, 64, 104]. It is presumed that a cohort design could be constructed from each of these studies respective databases. Similarly, the VNSNY Medication and Fall Data enables the investigators to observe medication exposures and potential baseline confounders on all patients. Applying the case-crossover design in this setting, results in the exclusion of medication information from the patients who did not fall and is likely to result in substantial loss of efficiency. This literature review has identified which factors are known to be associated with falling, and can be observed in the VNSNY Medication and Fall data. An alternative approach to using the case-crossover design would be to conduct a well known, cohort design utilizing a Cox hazards model with time-dependent covariates. This model could control for factors that may be assumed to confound the relationship between falling and medication changes.

The case-time-control design was proposed as a bias adjustment for the case-crossover design. If information is readily available on individuals in the cohort who did not fall, a retrospective cohort design using a method such as the Cox hazards model may be more

appropriate. The comparison of the Cox hazards model to the case-crossover design has been limited to applied data analysis studies. There is an absence in the research that connects the likelihood and design methodology for each approach. Simulation studies under different assumptions could provide researchers with valuable information for choosing an approach to estimating the effect of an exposure on an event.

Most applications of the case-crossover design covered in this literature review assume that the transient effect of an exposure on an acute event can be summarized by a simple, linear parameter. This assumption results in a naive estimation of the increased risk if Maclure's Epidemic Curve is the underlying true relationship [59]. There is a need to fill the gap in the case-crossover design literature for how to explore a nonlinear transient effect in the case-crossover design setting. Furthermore, methods to compare more complicated associations must be easily available to the investigator.

# Chapter 3

# Methods

This thesis proposes to use the principal of distributed lagged variables to model a non-linear, deteriorating effect within the case-crossover design setting. This thesis also demonstrates that under certain assumptions, a cohort design using a Cox hazards model with time-dependent covariates will yield the same log hazard estimate as the case-crossover design. Furthermore, the case-crossover design is shown to be preferred over the cohort design in the presence of unobserved confounders. This chapter discusses the methodology, and assumptions, for obtaining a log-hazard ratio in the case-crossover design and in the cohort setting.

The first section introduces notation, in the context of the VNSNY Medication and Fall data, assuming a cohort sample has been observed, and reviews the construction of the partial likelihood for a Cox hazards model with a single, time-dependent covariate. The second section adapts the notation to the conditional probability used to develop the likelihood function for the case-crossover design. The last section, which is broken into two subsections, proposes a new approach to estimating a non-linear, deteriorating association in the case-crossover design setting. First, the concept of distributed lags within the case-crossover design setting is introduced by extending the simple one parameter relationship, which is defined in the preceding sections, to include two parameters. This model describes a relationship that is slightly more complicated than an "on-off" relationship. The distributed lag concept is then extended to include a large, finite, number of lagged parameters. Finally, a functional form of the lagged parameters is proposed to account for a non-linear, deteriorating effect of the exposure within the case-crossover design setting.

## 3.1 Notation

Consider a cohort of $N$ patients who are admitted into the acute care program at Visiting Nurse Service of New York (VNSNY) during the calendar years 2010 and 2011. Time from entry ($t = 0$) begins with admission into home health care. Each patient is assessed by a clinician using the Outcome and Assessment Information Set (OASIS). This assessment tool is required by the Center for Medicare and Medicaid Services. This baseline assessment enables investigators to observe an array of patient characteristics (e.g. demographical, clinical, etc.) at the time of entry into home health care. These characteristics are assumed to be constant over the duration of care and are represented by the vector of covariates $\boldsymbol{X}$ for patient $i$. In addition, the medication regimen for each patient is observed on a daily basis throughout the episode of home health care. The status of a change in medication regimen for patient $i$ at time $t$ is denoted by the vector of the time-dependent covariates $\boldsymbol{Z_t}$.

Let $\tilde{T}_i$ and $C_i$ denote the event time (i.e. time to the first fall) and censoring time of patient $i$, respectively. The observed time for patient $i$ is defined by $T_i = min(\tilde{T}_i, C_i)$ with a fall indicator

$$\delta_i = \begin{cases} 1 & T_i = \tilde{T}_i \\ 0 & T_i = C_i \end{cases} \tag{3.1}$$

where censoring is assumed to be non-informative.

Under the cohort design, assume the relationship between the hazard of falling and the constant patient characteristics $\boldsymbol{X}$, and time-varying characteristics $\boldsymbol{Z_t}$, is given by

$$\lambda(t) = \lambda_0(t) e^{\boldsymbol{X\gamma} + \boldsymbol{Z_t\beta}} \tag{3.2}$$

where $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ are vectors of coefficients that describe the respective multiplicative effects of $\boldsymbol{X}$ and $\boldsymbol{Z_t}$. $\boldsymbol{Z_t}$ may describe one or more time-varying factors which specify the status of exposures at time $t$ . For instance, in the VNSNY Medication and Fall data, $\boldsymbol{Z_t}$ could include a continuous measure of medication regimen complexity, indicators of changes in one of many medication therapeutic classification, or a set of lagged-covariates that account for carry-over effects, etc.

To describe a simple association between a medication change occurring for patient $i$

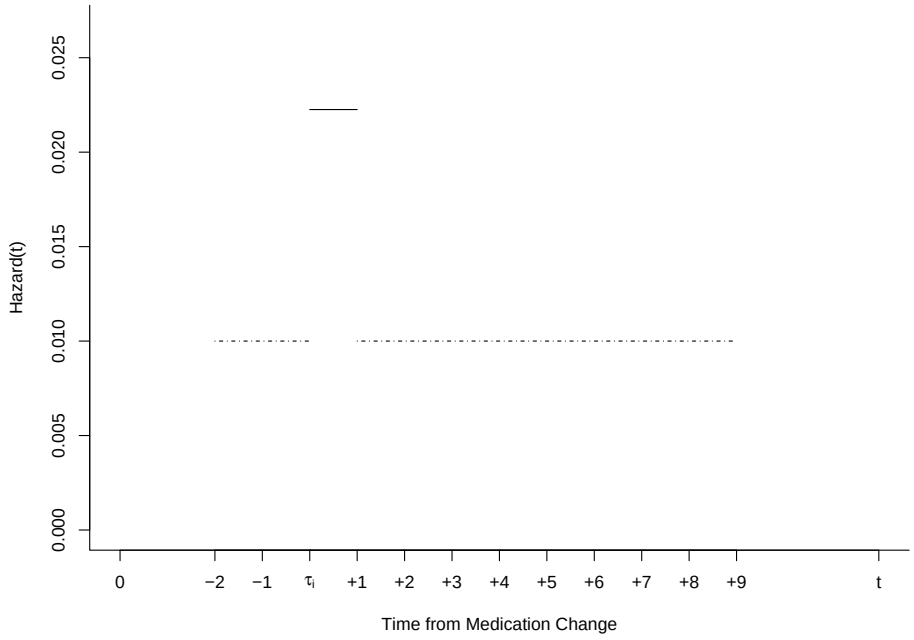


Figure 3.1: Illustration of simple, "on-off" transient effect

at $\tau_i$, a single time dependent variable, $Z_{it}$, may be specified to indicate the change and its effect, such as

$$Z_{it} = \begin{cases} 1 & t \in (\tau_i, \tau_i + \Delta] \\ 0 & t \notin (\tau_i, \tau_i + \Delta] \end{cases} \tag{3.3}$$

where $\Delta$ is the assumed duration of the transient effect of the medication change. This may be described as an "on-off" relationship, since the background risk is quickly elevated, due to the change in exposure, before quickly returning to the background risk. This simple effect can be written with the following hazard function

$$\lambda_i(t) = \lambda_0(t) e^{\boldsymbol{X}\boldsymbol{\gamma} + \beta Z_{it}} \tag{3.4}$$

with $Z_{it}$ defined by Equation (3.3). Figure 3.1 illustrates this simple, "on-off" transient effect, described by $\beta$, on the baseline hazard in Equation (3.4). This figure assumes that patient characteristics have no effect ($\gamma = 0$), the baseline hazard $\lambda_0(t) = 0.01$, the effect of the medication change increase the baseline hazard by 2.25 times, or $\beta = 0.8$, for a duration of $\Delta = 1$, and then returning to the baseline hazard of 0.01.

## 3.2 Cohort Design

The VNSNY Medication and Fall data includes both constant, patient characteristics and time-dependent covariates. The time-dependent covariates describe medication regimen changes on both patients who experienced a fall and those who did not. One method to estimate the immediate effect that a medication change has on falling is to fully utilize the data of the entire cohort and use a conventional Cox proportional hazards model where medication changes are treated as time-dependent covariates.

A partial likelihood to estimate $\beta$ and $\gamma$ from Equation (3.4) is constructed as

$$\begin{aligned} L(\beta,\gamma) &= \prod_{i=1}^{N}\left[\frac{\lambda_0(\tilde{t_i})e^{\boldsymbol{X_i}\boldsymbol{\gamma}+\beta Z_{i\tilde{t_i}}}}{\Sigma_{k\in R(\tilde{t_i})}\lambda_0(\tilde{t_i})e^{\boldsymbol{X_k}\boldsymbol{\gamma}+\beta Z_{k\tilde{t_i}}}}\right]^{\delta_i} \\ &= \prod_{i\in\{i;\delta_i=1\}}\frac{e^{\boldsymbol{X_i}\boldsymbol{\gamma}+\beta Z_{i\tilde{t_i}}}}{\Sigma_{k\in R(\tilde{t_i})}e^{\boldsymbol{X_k}\boldsymbol{\gamma}+\beta Z_{k\tilde{t_i}}}} \end{aligned} \tag{3.5}$$

where $R(\tilde{t_i})$ denotes the risk set of the full cohort at time $\tilde{t_i}$, and $\tilde{t_i}$ denotes the event time for patient $i$ [42]. The partial log-likelihood is given by

$$l(\beta,\gamma) = \sum_{i\in\{i;\delta_i=1\}}\left\{\boldsymbol{X_i}\boldsymbol{\gamma}+\beta Z_{i\tilde{t_i}} - log\left[\Sigma_{k\in R(\tilde{t_i})}e^{\boldsymbol{X_k}\boldsymbol{\gamma}+\beta Z_{k\tilde{t_i}}}\right]\right\} \tag{3.6}$$

Then, $\beta$ and $\gamma$ are estimated by taking the partial derivative of Equation (3.6) with respect to $\beta$ and $\gamma$, respectively, and solving the resulting partial score equation $U(\beta,\gamma)=0$, where

$$U(\beta,\gamma) = \begin{pmatrix} \frac{\partial l(\beta,\gamma)}{\partial\beta} \\ \frac{\partial l(\beta,\gamma)}{\partial\gamma} \end{pmatrix} = \begin{pmatrix} \sum_{i\in\{i;\delta_i=1\}}\left[Z_{i\tilde{t_i}} - \frac{\Sigma_{k\in R(\tilde{t_i})}Z_{k\tilde{t_i}}e^{\boldsymbol{X_k}\boldsymbol{\gamma}+\beta Z_{k\tilde{t_i}}}}{\Sigma_{k\in R(\tilde{t_i})}e^{\boldsymbol{X_k}\boldsymbol{\gamma}+\beta Z_{k\tilde{t_i}}}}\right] \\ \sum_{i\in\{i;\delta_i=1\}}\left[X_i - \frac{\Sigma_{k\in R(\tilde{t_i})}X_k e^{\boldsymbol{X_k}\boldsymbol{\gamma}+\beta Z_{k\tilde{t_i}}}}{\Sigma_{k\in R(\tilde{t_i})}e^{\boldsymbol{X_k}\boldsymbol{\gamma}+\beta Z_{k\tilde{t_i}}}}\right] \end{pmatrix} \tag{3.7}$$

The information matrix is obtained by taking the derivative of $U(\beta,\gamma)$ with respect to $\beta$ and $\gamma$. Variance estimates of $\beta$ and $\gamma$ can then be obtained by inverting the information matrix. In practice, estimates and standard errors for $\beta$ and $\gamma$ can be obtained by using `coxph` function in **R** [97].

## 3.3 Case-Crossover Design

Maclure described a sampling method for observational data that matches periods of time within a subject in order to estimate the association between a time dependent exposure

and an acute event [59]. The case-crossover design is fundamentally a matched case-control design, but where times are matched within patient rather than patient-to-patient matching in the traditional setting. In this sampling strategy, depicted in Figure 3.2, a time is selected for each patient who has experienced the event, during a period when the exposure is assumed to be attributable to the event, and is referred to as the case-time or hazard-time. One or more control-times are then matched to the case-time for the same subject. The control-times are assumed to occur at time when the patient is assumed to be event free. Therefore, the event is assumed to be fixed and the exposure is random.

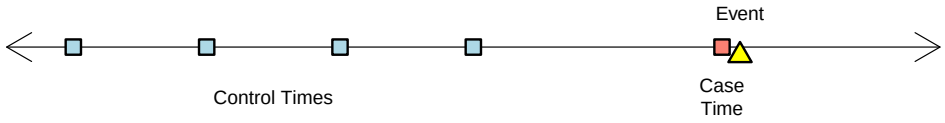


Figure 3.2: Illustration of a case-time matched to control-times for the same subject in a case crossover design

The association is then estimated by comparing the distribution of exposures during case- and control-times for all matched sets. The case-crossover design is most useful when only patients who have experienced the event are available to the investigator.

The simplest form of the case-crossover design is a 1:1 case-crossover design. This design consists of one case-time matched to one control-time. The number of control-times is often denoted by $M$, and when $M > 1$ control-times are used in the design, the study is called a 1:$M$ case-crossover design. In principle it is possible to select more than one case-time for each subject, but is rarely applied in practice.

The time scale in the case-crossover design is referenced to the time of the event; this is a reversal of the time scale used in the retrospective cohort design. A reference strategy is then carefully designed around this time scale for $M + 1$ intervals. The reference strategy is commonly fixed by the investigator based on prior knowledge of the exposure and the event. In this dissertation, the case- and control-times are always selected in the days prior to the fall, and $m = 0$ indexes the case-time which is matched with $m = 1, 2, \ldots, M$ control times.

Let the reference times be denoted by the within-patient temporal risk set $W_i(\tilde{t}_i)$. The times selected in this risk set are selected apriori by the investigator and reference the fall time $\tilde{t}_i$ for patient $i$. If case- and control-times are assumed to occur sequentially over the days immediately preceding the fall, the case-time is then said to occur at $\tilde{t}_i$ and $W_i(\tilde{t}_i) = \{\tilde{t}_i - (M+1), \ldots, \tilde{t}_i - 2, \tilde{t}_i - 1, \tilde{t}_i\}$. This reference strategy is illustrated in Figure 3.3. The dual x-axis in this plot illustrates how the time scale in the case-crossover design references the time of the fall, which is a reversal of that used in a retrospective cohort design (i.e. time from entry).

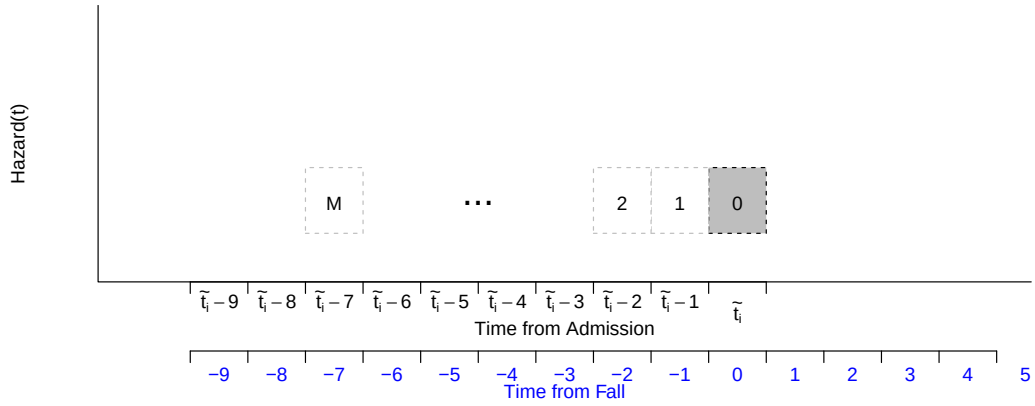


Figure 3.3: Illustration of a 1:M case-crossover design with time scales from both case-crossover and cohort designs

The notation (0) and $(m)$ is introduced to map the case- and $m^{th}$ control-time strata to the time scale used in the cohort design, or the time from entry. Let $Z_{i(0)}$ denote the exposure during the case-time for patient $i$, $Z_{i(1)}, \ldots Z_{i(M)}$ denote the exposures during the $M$ control times for patient $i$, and $\Psi_i = \{Z_{i(0)}, Z_{i(1)}, \ldots Z_{i(M)}\}$ denote the unordered set of exposures for patient $i$, where $\Psi$ is defined by the reference strategy design. Assume that falls are rare and $\lambda_0(t)$ is constant or small [102]. It then follows, the conditional probability that the covariate value during the case-time is precisely, $Z_{i(0)} = z_{i(0)}$, given that $Z_{i(0)}$ lies in $\Psi_i$

$$P(Z_{i(0)} = z_{i(0)} | \tilde{T} = \tilde{t}_i, \Psi_i) = \frac{P(\tilde{T} = \tilde{t}_i | Z_{i(0)} = z_{i(0)}, \Psi_i) P(Z_{i(0)} = z_{i(0)} | \Psi_i)}{\sum_{j \in W_i(\tilde{t}_i)} P(\tilde{T} = j | Z_{ij} = z_{ij}, \Psi_i) P(Z_{ij} = z_{ij} | \Psi_i)} \tag{3.8}$$

The exposure distribution in any $M+1$ successive time intervals must be globally exchangeable within the matched set $\Psi$, i.e.

$$P\left(Z_{i(0)} = z_{i(0)}, \ldots, Z_{i(M)} = z_{i(M)}\right) = P\left(Z_{i\kappa(0)} = z_{i\kappa(0)}, \ldots, Z_{i\kappa(M)} = z_{i\kappa(M)}\right)$$

for all permutations $\kappa$ of $\{0, 1, \ldots, M\}$. Based on the model in Equation (3.4), the conditional probability can then be written as [102]

$$\begin{aligned} \frac{P(\tilde{T} = \tilde{t}_i | Z_{i(0)} = z_{i(0)})}{\sum_{j \in W_i(\tilde{t}_i)} P(\tilde{T} = j | Z_{ij} = z_{ij})} &= \frac{\lambda_0(t) e^{\boldsymbol{X_i}\boldsymbol{\gamma} + \beta Z_{i(0)}}}{\sum_{j \in W_i(\tilde{t}_i)} \lambda_0(t) e^{\boldsymbol{X_i}\boldsymbol{\gamma} + \beta Z_{ij}}} \\ &= \frac{e^{\beta Z_{i(0)}}}{\sum_{j \in W_i(\tilde{t}_i)} e^{\beta Z_{ij}}} \end{aligned} \tag{3.9}$$

Expression (3.9) is of the same form as the conditional probability used in a conditional logistic regression. From this conditional probability, the likelihood function of $\beta$ can be constructed as

$$L_{cco}(\beta) = \prod_{i \in \{i; \delta_i = 1\}} \frac{e^{\beta Z_{i(0)}}}{\sum_{j \in W_i(\tilde{t}_i)} e^{\beta Z_{ij}}} \tag{3.10}$$

The corresponding log conditional likelihood and score function for $\beta$ is

$$l_{cco}(\beta) = \sum_{i \in \{i; \delta_i = 1\}} \left[ \beta Z_{i(0)} - log \left( \sum_{j \in W_i(\tilde{t}_i)} e^{\beta Z_{ij}} \right) \right] \tag{3.11}$$

$$U_{cco}(\beta) = \sum_{i \in \{i; \delta_i = 1\}} \left[ Z_{i(0)} - \frac{\sum_{j \in W_i(\tilde{t}_i)} Z_{ij} e^{\beta Z_{ij}}}{\sum_{j \in W_i(\tilde{t}_i)} e^{\beta Z_{ij}}} \right] \tag{3.12}$$

respectively, $\beta$ can be estimated by solving $U_{cco}(\beta) = 0$. Because patient specific characteristics cancel out in the conditional probability in Equation (3.9) , and subsequently in the conditional likelihood and score function, estimation of $\beta$ implicitly controls for time invarying characteristics. However, time-varying confounders must still be controlled for in any analysis. In practice, the parameter estimates and standard errors can be obtained by using software that can specify a conditional logistic regression model, such as the `clogit` function in the survival package in **R** [97].

The case-crossover design's conditional likelihood function in Expression (6.2) has the same form as the Cox partial likelihood function in Equation (3.6). The product for each likelihood function is taken over the number of patients who fell. The numerator is constructed from taking the hazard function in Equation (3.4) at time $\tilde{t}_i$ for the Cox

partial likelihood, and at a time where the exposures is attributable (assumed) to the event $\tilde{t_i}$ for the case-crossover design for the $i^{th}$ patient who fell. The two likelihood functions are distinguishable by the risk set used in each functions respective denominator. The risk set used in the Cox partial likelihood is comprised across other patients in the cohort who are at risk of falling at time $\tilde{t_i}$, including both those who fell and those who did not. Whereas, the risk set used in the case-crossover design constitutes of times selected within the $i^{th}$ patient who fell.

A theoretical cohort of patients displayed in Figure 3.4 is used to illustrate the distinction between the two designs. Eight patients are followed from time of entry to represent a sample that could be observed in the VNSNY Medication and Fall data. Patients are followed until each experiences an acute event or is lost to follow-up. In this example, falls are observed at times $\tilde{t_i} = 1, 2, 3, 4, 5, 6, 7$ for patients $1, 2, 3, 4, 5, 6, 7$ respectively and patient 8 is lost to follow up and censored at time $c_i = 8$. In addition, a binary exposure is observed for patients $1, 3, 4, 7, 8$. The across patient risk set for the Cox hazards model is indicated by the red, dashed lines. The temporal, within subject risk set for the case-crossover design is drawn in the blue, dotted lines. Patients 1-7 all have a risk set for both designs. Patient eight contributes to the risk sets in the Cox partial likelihood, but is not included at all in the case-crossover design since no event was observed.

## 3.4 Specifiction of a Non-Linear Relationship

Maclure hypothesized that the transient effect resembled the smooth curve plotted in Figure 3.5. He assumed that all patients have a constant low risk of experiencing the event until the point exposure occurs; which may or may not be followed by an induction time. After the induction period, the patient enters a period where he or she is at highest risk. This risk is then assumed to deteriorate the further in time after the point exposure. A naive approach to this deteriorating, non-linear relationship would be to force an on-off relationship by assuming the hazard in Equation (3.4) and time-dependent covariate in Equation (3.3).

A nonlinear relationship in some cases may be justified by aprori information about the

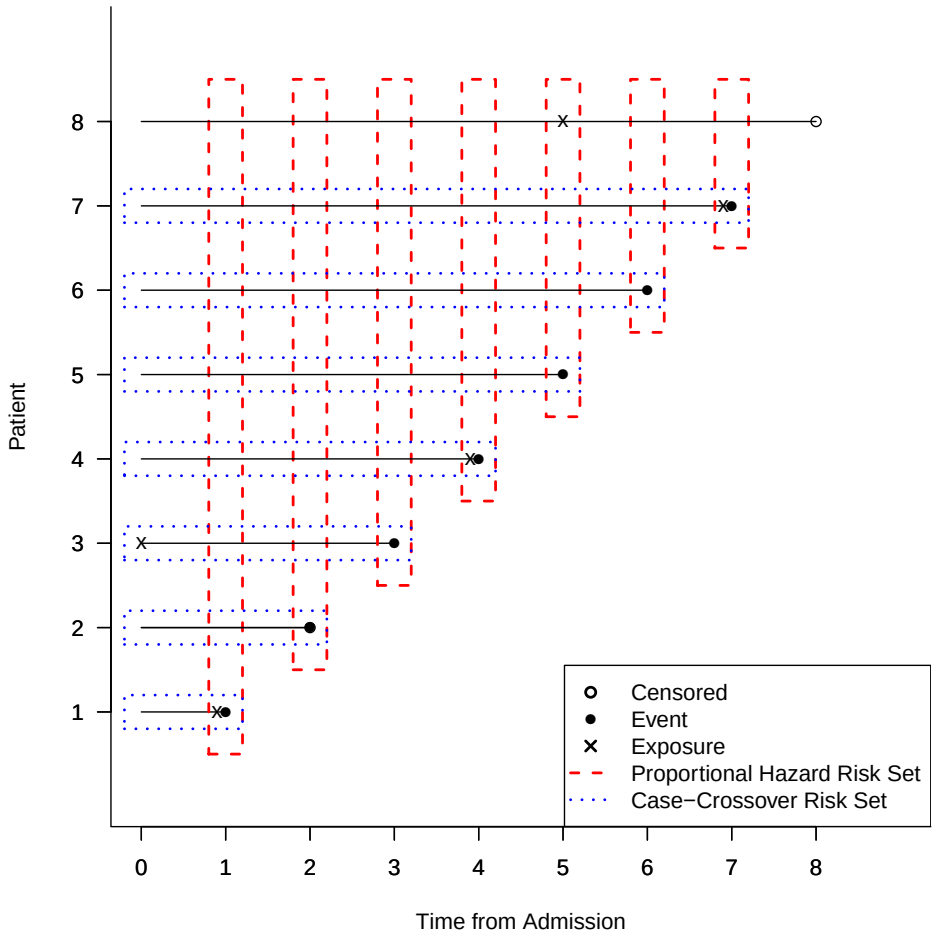


Figure 3.4: Comparison of case-crossover design to cox proportional hazards model for a theoretical cohort

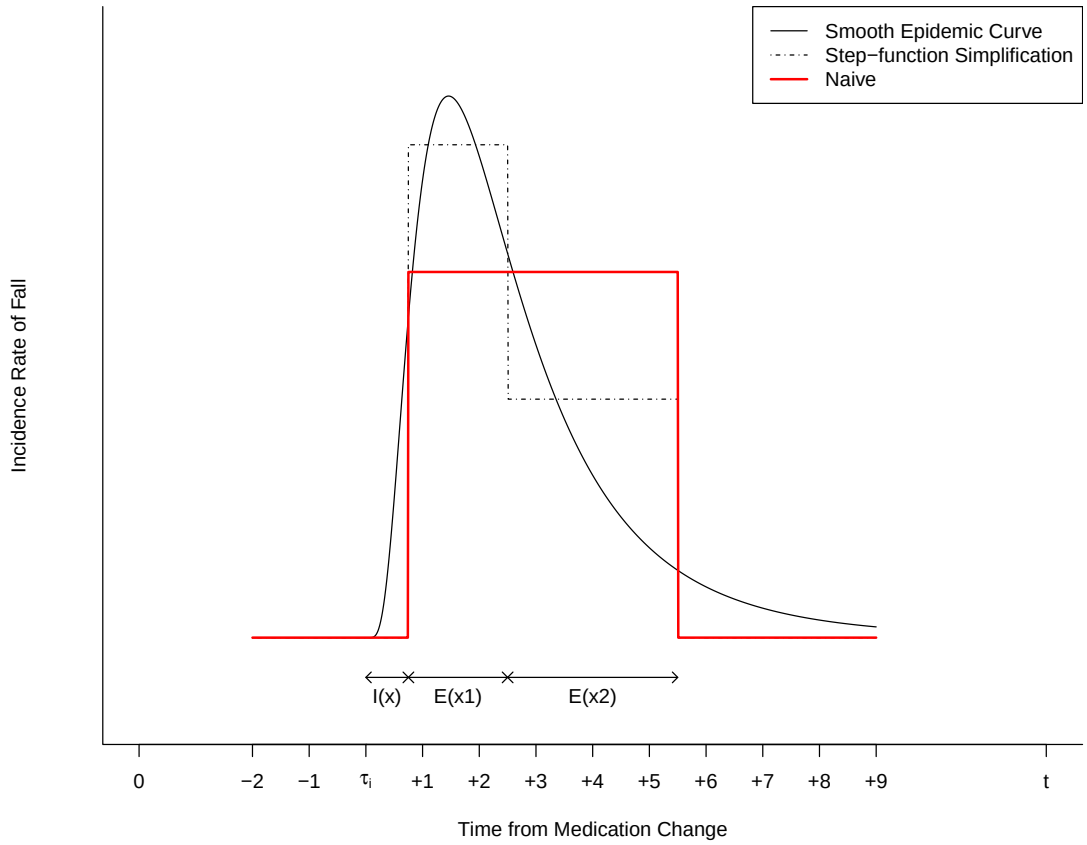


Figure 3.5: Illustration of Maclure's assumed epidemic curve for the case-crossover design [59]. The curve demonstrates incidence of an acute-onset event (i.e. fall) after observing a point exposure (i.e. change in medication). The population induction time is indicated by $I(x)$, and a step function is suggested to estimate the curve through $E(x_1)$ and $E(x_2)$. $E(x_1)$ indicates the time of highest risk effect period, and $E(x_2)$ a more moderate risk effect period. The naive or "on-off" relationship is overlayed with the thicker red line

relationship between the exposure change and the event. For example, the pharmacokinetical properties of a medication (i.e. absorption, distribution, metabolism, and excretion) may lead the investigator to assume the risk of falling after an increase in a specific therapeutic medication classification of follows the non-linear deteriorating effect, possibly resembling Maclure's epidemic curve illustrated in Figure 3.5 [59]. Forcing the "on-off" relationship can underestimate the risk during the $E(x_1)$ period, and overestimates the risk during the $E(x_2)$ period. This dissertation proposes a new approach to estimate the curve hypothesized by Maclure within the case-crossover design framework.

### 3.4.1 Transient Effect Described by Two Lagged Parameters

Maclure recommended simplifying the epidemic curve in Figure 3.5 by employing a step function, illustrated by the dashed line, to approximate the hazard over two intervals of high risk $E(x_1)$ and moderate risk $E(x_2)$. It would be naive to force a simple "on-off" relationship for the transient effect resembling Maclure's epidemic curve by estimating the effect assuming the hazard function in Equation (3.4). A better model may assume that the hazard function is described by two covariates, such that

$$\lambda_i(t) = \lambda_0(t)e^{\beta_1 Z_{it}^{(1)} + \beta_2 Z_{it}^{(2)}} \tag{3.13}$$

where

$$Z_{it}^{(1)} = \begin{cases} 0 & t \leq \tau_i \\ 1 & t \in (\tau_i, \tau_i + \Delta_1] \\ 0 & t > \tau_i + \Delta_1 \end{cases} \quad Z_{it}^{(2)} = \begin{cases} 0 & t \leq \tau_i + \Delta_1 \\ 1 & t \in (\tau_i + \Delta_1, \tau_i + \Delta_1 + \Delta_2] \\ 0 & t > \tau_i + \Delta_1 + \Delta_2 \end{cases} \tag{3.14}$$

where $\Delta_1$ is similar to the duration of $E(x_1)$, and $\Delta_2$ to $E(x_2)$, in Maclure's epidemic curve in Figure 3.5, and the induction time, $I(x)$, is zero. Notation is then eased by assuming $\Delta_1 = \Delta_2 = 1$, and $E(x_1)$ and $E(x_2)$ occur sequentially after the time of medication change $\tau_i$. $Z_{it}^{(1)}$ and $Z_{it}^{(2)}$ can then be reexpressed using the lagged covariates $Z_{i,t}$ and $Z_{i,t-1}$. An example of this relationship 3.6 and specified by the hazard function

$$\lambda_i(t) = \lambda_0(t)e^{\beta_1 Z_{i,t} + \beta_2 Z_{i,t-1}} \tag{3.15}$$

where $Z_{i,t}$ and $Z_{i,t-1}$ represent the status of an exposure change at time $t$ and at one lagged period prior to $t$, respectively. Furthermore, these two lagged covariates are defined as

$$Z_{i,t} = \begin{cases} 0 & t \leq \tau_i \\ 1 & t \in (\tau_i, \tau_i + 1] \\ 0 & t > \tau_i + 1 \end{cases} \quad Z_{i,t-1} = \begin{cases} 0 & t - 1 \leq \tau_i \\ 1 & t - 1 \in (\tau_i, \tau_i + 1] \\ 0 & t - 1 > \tau_i + 1 \end{cases} \tag{3.16}$$

This two-parameter lagged effect is illustrated in Figure 3.6. The baseline hazard is assumed to be 0.01, where the risk is elevated by 2.2 fold at the time of medication change (i.e. $\beta_1 = 0.8$), followed by a period where the baseline risk is elevated by 1.5 times (i.e. $\beta_2 = 0.4$), and then returning to the baseline hazard.

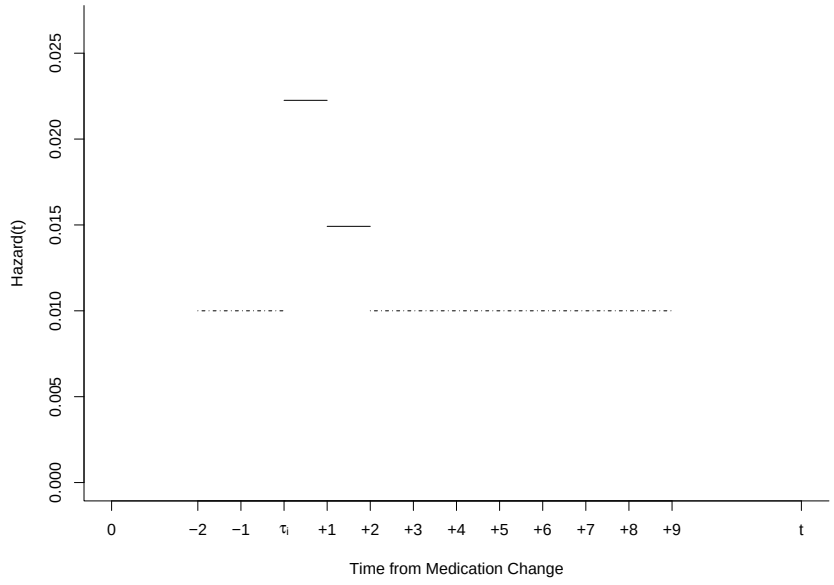


Figure 3.6: Illustration of a transient effect described by two lagged covariates in the cohort setting where $\beta_1 = .8, \beta_2 = .4, \lambda_0(t) = 0.01$ in Equation (3.15
)

### 3.4.2 Distributed Lagged Parameters

Building on section 3.4.1, increasing the number of intervals that construct the step function from two to $L$ is likely to provide better fit to Maclure's epidemic curve in Figure 3.5. In this paper $L$ is assumed to be fixed and based on some apriori information. Notation is again eased by assuming the step intervals are of length one and occur immediately after the medication change. The hazard function to describe this step function of $L$ intervals may then be written as

$$\lambda_i(t) = \lambda_0(t)e^{\beta_1 Z_{i,t} + \beta_2 Z_{i,t-1} + \ldots + \beta_L Z_{i,t-(L-1)}} \tag{3.17}$$

where the $L$ lagged covariates are used to specify each step can be written as

$$\begin{aligned}
Z_{i,t} &= \begin{cases} 0 & t \leq \tau_i \\ 1 & t \in (\tau_i, \tau_i + 1] \\ 0 & t > \tau_i + 1 \end{cases} \\
Z_{i,t-1} &= \begin{cases} 0 & t - 1 \leq \tau_i \\ 1 & t - 1 \in (\tau_i, \tau_i + 1] \\ 0 & t - 1 > \tau_i + 1 \end{cases} \\
&\vdots \qquad \vdots \\
Z_{i,t-(L-1)} &= \begin{cases} 0 & t - (L-1) \leq \tau_i \\ 1 & t - (L-1) \in (\tau_i, \tau_i + 1] \\ 0 & t - (L-1) > \tau_i + 1 \end{cases}
\end{aligned} \tag{3.18}$$

Describing the relationship between an outcome and an explanatory variable using its current and lagged values, as shown in Equation (3.17), is often referred to as a distributed lag model [2]. Since the parameters of the lags are unrestricted, it is referred to as an unconstrained distributed lag model [2]. Figure 3.7 shows an example of a such a hazard function with $l = 5$ covariates. This example assumes a baseline hazard of 0.01, which the risk elevated by hazard ratios $e^{.8}, e^{.4}, e^{.2}, e^{.1}$, and $e^{.05}$ the five days after the medication change, and then finally returning to the baseline hazard.

Figure 3.8 illustrates how a 1:$M$ case-crossover design with reference windows of length five could be applied to the transient effect assumed in Figure 3.7. The presentation of the transient effect in this plot is the mirror image of that presented for the cohort setting in Figure 3.7. The reverse image is because the time scale in case-crossover design is in reference to the time of the fall. The dual axis again shows the difference in time scales between the two designs under consideration. The solid bordered, gray box immediately preceding the time of event $\tilde{t_i} = 0$ highlights the case-time in the case-crossover design with $l = 1, 2, 3, 4, 5$ lagged covariates. The other white boxes denote the 1:$M$ control-times chosen by the reference strategy design with the $l = 1, 2, 3, 4, 5$ lagged covariates. For both reference windows the lagged covariates are enumerated from right-to-left due to the change in time-scale in the case-crossover design.

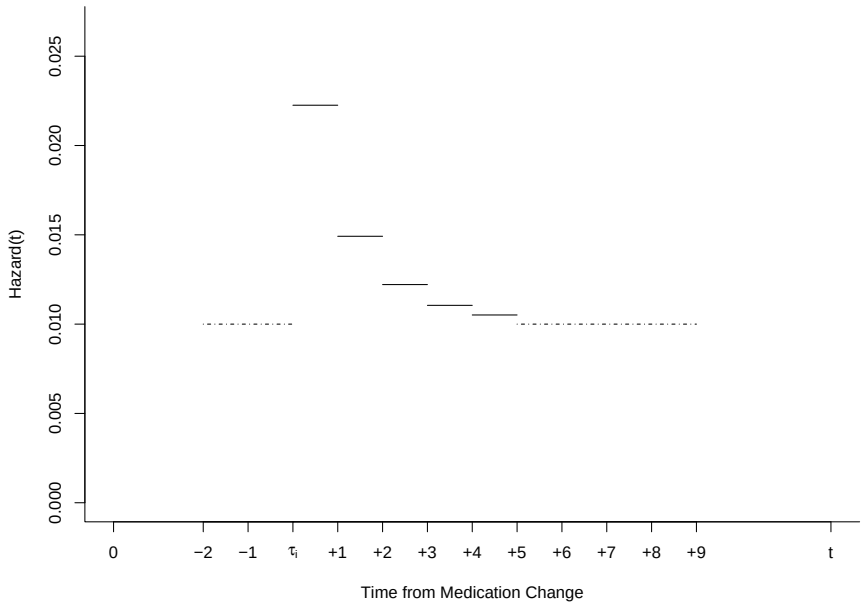


Figure 3.7: Illustration of a transient effect described by five lagged covariates in cohort setting where $\beta_1 = .8, \beta_2 = .4, \beta_3 = .2, \beta_4 = .1, \beta_5 = .05$, and $\lambda_0(t) = 0.01$ in Equation (3.17)

The previous example motivates the introduction of additional notation. Let $\omega(m, l)$ denote a function that maps the case- or control strata, $m$, at the lagged effect, $l$, to the time from admission $t$. For example, if a medication change occurs three days prior to the fall, the lagged variables are parameterized during the case-time as

$$\left\{Z_{i\omega(0,1)} = 0, Z_{i\omega(0,2)} = 0, Z_{i\omega(0,3)} = 1, Z_{i\omega(0,4)} = 0, Z_{i\omega(0,5)} = 0\right\}$$

Similarly, if a medication change happens at $l = 4$ in the $m = 2$ control-time the lagged variables are parameterized as

$$\left\{Z_{i\omega(2,1)} = 0, Z_{i\omega(2,2)} = 0, Z_{i\omega(2,3)} = 0, Z_{i\omega(2,4)} = 1, Z_{i\omega(2,5)} = 0\right\}$$

The conditional likelihood for one-to-$M$ case-crossover design for the hazard in Equation ( 3.17) can then be constructed as

$$L_{cco}(\beta_1, \beta_2, \ldots \beta_L) = \prod_{i \in \{i;\delta_i=1\}} \frac{e^{\beta_1 Z_{i\omega(0,1)} + \beta_2 Z_{i\omega(0,2)} + \ldots + \beta_L Z_{i\omega(0,l)}}}{\sum_{m=0}^{M} e^{\beta_1 Z_{i\omega(m,1)} + \beta_2 Z_{i\omega(m,2)} + \cdots + \beta_L Z_{i\omega(m,l)}}} \tag{3.19}$$

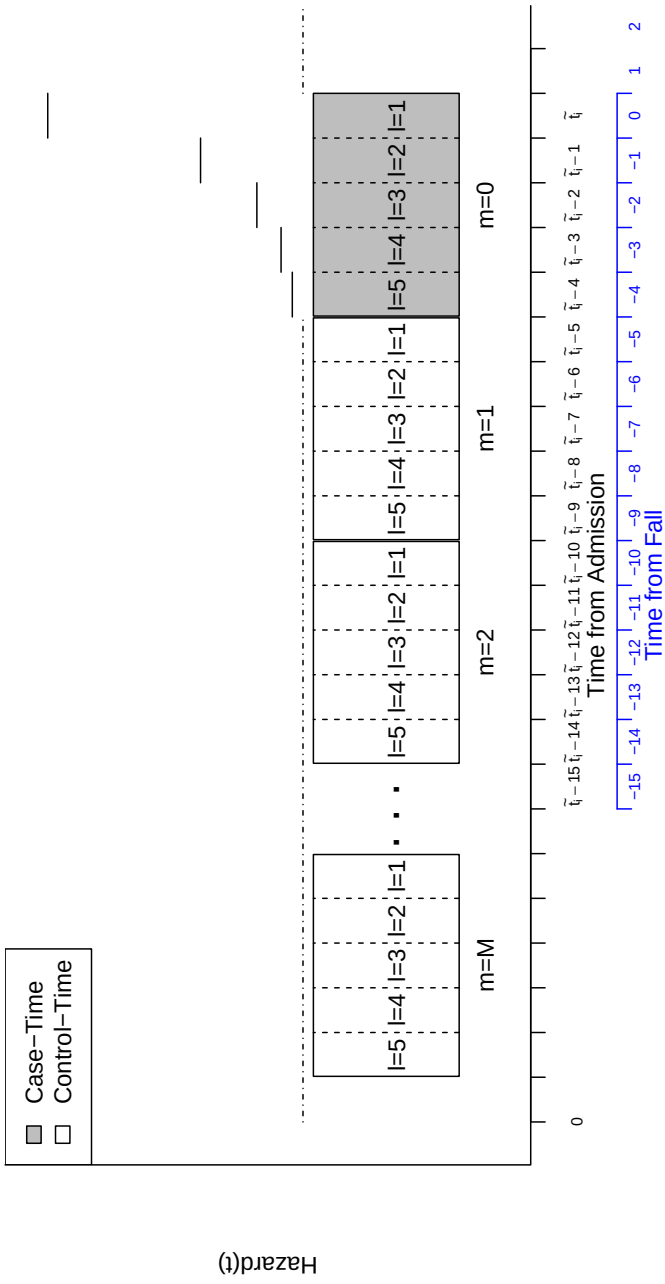


Figure 3.8: Illustration of a 1:$M$ case-crossover design being applied to a transient effect deteriorating over five days and assumes a baseline hazard $\lambda_0(t) = 0.01$ with log hazard ratios of $\beta_1 = .8, \beta_2 = .4, \beta_3 = .2, \beta_4 = .1, \beta_5 = .05$

with log conditional likelihood

$$\begin{aligned} l_{cco}(\beta_1, \beta_2, \ldots \beta_L) &= \sum_{i \in \{i;\delta_i=1\}} \Big[\beta_1 Z_{i\omega(0,1)} + \beta_2 Z_{i\omega(0,2)} + \ldots + \beta_L Z_{i\omega(0,L)} \\ &- log\left(\sum_{m=0}^{M} e^{\beta_1 Z_{i\omega(m,1)} + \beta_2 Z_{i\omega(m,2)} + \ldots + \beta_L Z_{i\omega(m,L)}}\right)\Big] \end{aligned} \tag{3.20}$$

Taking the derivative for each of the $L$ lagged covariates yields the score function

$$U_{cco}(\beta_l) = \sum_{i \in \{i;\delta_i=1\}} \left[Z_{i\omega(0,1)} - \frac{\sum_{m=0}^{M} Z_{i\omega(m,l)} e^{\beta_1 Z_{i\omega(m,1)} + \beta_2 Z_{i\omega(m,2)} + \ldots + \beta_l Z_{i\omega(m,L)}}}{\sum_{m=0}^{M} e^{\beta_1 Z_{i\omega(m,1)} + \beta_2 Z_{i\omega(m,2)} + \ldots + \beta_L Z_{i\omega(m,L)}}}\right] \tag{3.21}$$

Estimates for $\beta_l$ can be obtained by solving $U_{cco}(\beta_l)$. Estimates and standard errors of $\beta_l$ can be obtained in practice using `clogit` function in **R** [97].

The lagged variables in Equation (3.17) are all describing the effect from the same exposure series. The use of a large number of lags is likely to yield unstable estimates for small samples. An exploratory analysis of $l$ covariates, assuming the hazard function in Equation (3.17), may suggest an underlying relationship between the covariates, or perhaps a scientific explanation exists that warrants constraining the $\beta_l$ coefficients by some function $\Theta_l(\Omega)$. Where $\Omega$ is a vector of parameters of size less than $l$.

Almon was the first to propose constraining the lagged coefficients as a function of the lag for economic time series [2]. The true relationship as shown in Figure 3.7 resembles the effect illustrated by the smooth curve proposed in Maclure's Epidemic Curve in Figure 3.5. This relationship spikes immediately after the medication change and deteriorates over time. One example of a constrained distributed lag model is the discrete geometric function, which is commonly used in time-series analysis. A geometric lag function for discrete lags, can be written as

$$\Theta_l(\beta, \theta) = \beta\theta^{(l-1)}$$

When $0 < \theta < 1$, the effect, $\beta$, deteriorates by $\theta$ over each time interval. If this constraint

is applied to Equation (3.17) so that

$$\begin{aligned} \beta_1 &= \beta \\ \beta_2 &= \beta\theta \\ \beta_3 &= \beta\theta^2 \\ \vdots &\quad \vdots \\ \beta_L &= \beta\theta^{L-1} \end{aligned} \tag{3.22}$$

the hazard function can then be written as

$$\lambda_i(t) = \lambda_0(t)e^{\beta Z_{i,t}+\beta\theta Z_{i,t-1}+\beta\theta^2 Z_{i,t-2}+\ldots+\beta\theta^{(L-1)}Z_{i,t-(L-1)}} \tag{3.23}$$

The likelihood function for a 1:$M$ case-crossover design for $\beta$ and $\theta$ from the hazard function in Equation (3.23) is

$$L(\beta,\theta)_{cco} = \prod_{i\in\{i;\delta_i=1\}} \frac{e^{\beta Z_{i\omega(0,1)}+\beta\theta Z_{i\omega(0,2)}+\beta\theta^2 Z_{i\omega(0,3)}+\ldots+\beta\theta^{(l-1)}Z_{i\omega(0,l)}}}{\sum_{m=0}^{M} e^{\beta Z_{i\omega(m,1)}+\beta\theta Z_{i\omega(m,2)}+\beta\theta^2 Z_{i\omega(m,2)}+\cdots+\beta\theta^{(l-1)}Z_{i\omega(m,l)}}} \tag{3.24}$$

with log-likelihood

$$\begin{aligned} l(\beta,\theta)_{cco} = \sum_{i\in\{i;\delta_i=1\}} \Bigg[ & \beta Z_{i\omega(0,1)} + \beta\theta Z_{i\omega(0,2)} + \beta\theta^2 Z_{i\omega(0,3)} + \ldots + \beta\theta^{(l-1)} Z_{i\omega(0,l)} \\ & - log\left(\sum_{m=0}^{M} e^{\beta Z_{i\omega(m,1)}+\beta\theta Z_{i\omega(m,2)}+\beta\theta^2 Z_{i\omega(m,2)}+\cdots+\beta\theta^{(l-1)}Z_{i\omega(m,l)}}\right)\Bigg] \end{aligned} \tag{3.25}$$

Taking the derivative of the log-likelihood with respect to both $\beta$ and $\theta$ yields the following score functions

$$
\begin{aligned}
U_{cco}(\beta,\theta) &= \begin{pmatrix} \frac{\partial l(\beta,\theta)}{\partial \beta} \\ \frac{\partial l(\beta,\theta)}{\partial \theta} \end{pmatrix} \\
&= \begin{pmatrix}
\sum_{i\in\{i;\delta_i=1\}} \left[ Z_{i\omega(0,1)} + \theta Z_{i\omega(0,2)} + \theta^2 Z_{i\omega(0,3)} + \ldots + \theta^{(l-1)} Z_{i\omega(0,l)} \right. \\
- \sum_{m=0}^{M} \left( Z_{i\omega(m,1)} + \theta Z_{i\omega(m,2)} + \theta^2 Z_{i\omega(m,3)} + \ldots + \theta^{(l-1)} Z_{i\omega(m,l)} \right) \\
\left. \cdot \frac{e^{\beta Z_{i\omega(m,1)} + \beta\theta Z_{i\omega(m,2)} + \theta^2 Z_{i\omega(m,3)} + \cdots + \beta\theta^{(l-1)} Z_{i\omega(m,l)}}}{\sum_{m=0}^{M} e^{\beta Z_{i\omega(m,1)} + \beta\theta Z_{i\omega(m,2)} + \beta\theta^2 Z_{i\omega(m,3)} + \theta^2 Z_{i\omega(m,3)} + \cdots + \beta\theta^{(l-1)} Z_{i\omega(m,l)}}} \right] \\
\\
\sum_{i\in\{i;\delta_i=1\}} \left[ \beta Z_{i\omega(0,2)} + 2\beta\theta Z_{i\omega(0,3)} + \ldots + (l-1)\beta\theta^{(l-2)} Z_{i\omega(0,l)} \right. \\
- \sum_{m=0}^{M} \left( \beta Z_{i\omega(m,2)} + 2\beta\theta Z_{i\omega(m,3)} + \ldots + (l-1)\beta\theta^{(l-2)} Z_{i\omega(m,l)} \right) \\
\left. \cdot \frac{e^{\beta Z_{i\omega(m,1)} + \beta\theta Z_{i\omega(m,2)} + \theta^2 Z_{i\omega(m,3)} + \cdots + \beta\theta^{(l-1)} Z_{i\omega(m,l)}}}{\sum_{m=0}^{M} e^{\beta Z_{i\omega(m,1)} + \beta\theta Z_{i\omega(m,2)} + \theta^2 Z_{i\omega(m,3)} + \cdots + \beta\theta^{(l-1)} Z_{i\omega(m,l)}}} \right]
\end{pmatrix}
\end{aligned}
\tag{3.26}
$$

Estimates of $\beta$ and $\theta$ can be estimated by solving the score equation, $U_{cco}(\beta,\theta) = 0$. In practice, estimates of $\beta$ and $\theta$ can be obtained by using the quasi-newton method with the R function `optim` [79]. This can be done by specifying the log likelihood function in Equation (3.25) as the function which `optim` should maximize with respect to $\beta$ and $\theta$. This requires the investigator to provide starting values for $\beta$ and $\theta$. To improve the convergence properties o teh optimization procedure, a profile likelihood approach is suggested in practice. The profile likelihood approach fixes one of the parameters in the log-likelihood as a constant and estimates the other. Denote the constant value of $\beta$ as $\beta_c$ and $\theta$ as $\theta_c$. Estimates for $\beta$ and $\theta$ can be obtained via profile likelihood by following this outline

1. Obtain estimates of $\beta_1$ and $\beta_2$ from Equation (3.23). Set initial values for $\beta^{(0)} = \beta_1$ and $\theta^{(0)} = \frac{\beta_1}{\beta_2}$, since initial estimates of $\beta_1$ and $\beta_2$ are readily available

2. Repeat $r$ times until convergence

(a) Estimate $\theta$ while assuming $\beta$ is constant, denoted $\beta_c$, by maximizing the function

$$\begin{aligned} l(\theta;\beta_c)_{cco} &= \sum_{i\in\{i;\delta_i=1\}} \Big[\beta_c\theta Z_{i\omega(0,2)} + \beta_c\theta^2 Z_{i\omega(0,3)} + \ldots + \beta_c\theta^{(l-1)} Z_{i\omega(0,l)} \\ &- log\left(\sum_{m=0}^{M} e^{\beta_c\theta Z_{i\omega(m,2)}+\beta_c\theta^2 Z_{i\omega(m,2)}+\cdots+\beta_c\theta^{(l-1)} Z_{i\omega(m,l)}}\right)\Big] \end{aligned} \tag{3.27}$$

using the `optim` function where the starting value for $\theta$ is $\theta^{(r-1)}$ and $\beta_c = \beta^{(r-1)}$. Set the resulting estimate from 3.27 to $\theta^{(r)} = \hat{\theta}$ .

(b) Estimate $\beta$ while assuming $\theta$ is constant, denoted $\theta_c$, by maximizing the function

$$\begin{aligned} l(\beta;\theta_c)_{cco} &= \sum_{i\in\{i;\delta_i=1\}} \Big[\beta Z_{i\omega(0,1)} + \beta\theta_c Z_{i\omega(0,2)} + \ldots + \beta\theta_c^{(l-1)} Z_{i\omega(0,l)} \\ &- log\left(\sum_{m=0}^{M} e^{\beta Z_{i\omega(m,1)}+\beta\theta_c Z_{i\omega(m,2)}+\cdots+\beta\theta_c^{(l-1)} Z_{i\omega(m,l)}}\right)\Big] \end{aligned} \tag{3.28}$$

using the `optim` function where the starting value for $\beta$ is $\beta^{(r-1)}$ and $\theta_c = \theta^{(r)}$. Set the resulting estimate from 3.28 to $\beta^{(r)} = \hat{\beta}$ .

where the convergence criteria is satisfied once $\|\beta^{(r-1)} - \beta^{(r)}\| < \epsilon_\beta$ and $\|\theta^{(r-1)} - \theta^{(r)}\| < \epsilon_\theta$ for some small $\epsilon_\beta$ and $\epsilon_\theta$, where $\beta^{(r)}$ and $\theta^{(r)}$ represent the estimate from the $r^{th}$ iteration. The respective Hessian matrix for both $\beta^{(r)}$ and $\theta^{(r)}$ is retained on the last iteration

3. The final estimates of $\hat{\beta}$ and $\hat{\theta}$ are $\beta^{(r)}$ and $\theta^{(r)}$, respectively. The standard errors are obtained by taking the square root of the inverse Hessian matrix from the $r^{th}$ iteration.

In practice , the logical first step is to determine whether the one parameter model is sufficient in explaining the effect of the exposure change. A likelihood ratio test can be constructed to determine if an unconstrained distributed lag model provides a significant gain over the one parameter model such that

$$\begin{aligned} H_o(Nested) &: \lambda(t) = \lambda_0(t)e^{\beta Z_{i,t}+\beta Z_{i,t-1}+\beta Z_{i,t-2}+\ldots+\beta Z_{i,t-(L-1)}} \\ H_A(Full) &: \lambda(t) = \lambda_0(t)e^{\beta_1 Z_{i,t}+\beta_2 Z_{i,t-1}+\ldots+\beta_L Z_{i,t-(L-1)}} \end{aligned} \tag{3.29}$$

where $2 * (loglikelihood(Full) - loglikelihood(Nested)) \sim \chi^2_{.05}(L-1)$

If the previous test rejects the null hypothesis, the next step is to determine whether the geometric model fits the data as well as the unconstrained distributed lag model. The expressions in Equation (3.22) details how the geomtric lag hazard function in Equation (3.23) is nested within the larger $L$ parameter model in 3.17. This allows a formal statistical test to verify whether the $L$ parameter model provides a significant gain over the two parameter, geometric distributed lag model. This test can serve to evaluate the goodness of fit for the two parameter model. Such that a likelihood ratio test can be constructed as $2 * (loglikelihood(Full) - loglikelihood(Nested)) \sim \chi^2_{.05}(L-2)$ for the hypothesis test

$$\begin{aligned} H_o(Nested) &: \lambda(t) = \lambda_0(t) e^{\beta Z_{i,t} + \beta\theta Z_{i,t-1} + \beta\theta^2 Z_{i,t-2} + \ldots + \beta\theta^{(L-1)} Z_{i,t-(L-1)}} \\ H_A(Full) &: \lambda(t) = \lambda_0(t) e^{\beta_1 Z_{i,t} + \beta_2 Z_{i,t-1} + \ldots + \beta_L Z_{i,t-(L-1)}} \end{aligned} \tag{3.30}$$

## 3.5 Conclusion

By showing the derivation for both likelihoods, this section illustrates that the likelihood for the Cox proportional hazards model with a time dependent covariate has the same form as the case-crossover design. They differ in that the Cox proportional hazards model uses a risk set of subjects, while the case-crossover design uses a within-subject risk set. The latter results in discarding information on those who did not experience the fall. Additionally, the idea of distributed lag covariates can be used describe a non-linear association. This chapter also showed how a two-parameter geometric constraint can be added to the distributed lag when the relationship deteriorates non-linearly after the medication change.

# Chapter 4

# Simulations

This chapter demonstrates how the geometric, two-parameter model, proposed in this thesis, is preferred over a fully parameterized, distributed lag model. In addition, the two-parameter model is shown to provide significant gain in goodness of fit over the naive one-parameter model. This chapter also confirms that both designs considered in this thesis obtain unbiased estimates of the hazard ratio associated with a change in exposure. In addition, it is shown that the case-crossover design is preferred over the Cohort design in the presence of an unmeasured, baseline confounder.

This chapter presents a series of simulation studies, which can be broken into two sections: an evaluation of the performance of the case-crossover design compared to the retrospective cohort design and an assessment of the proposed two-parameter, geometric lag model. The former assumes that the effect of the exposure on the event follows a linear relationship, which can be estimated by using a single parameter. The performance of each design is compared using the following criteria:

- Bias
- Relative efficiency of the case-crossover design compared to the cohort design
- Type I error
- Power
- Coverage probability

The second set of simulations evaluate the performance of the proposed two-parameter model in the presence of a non-linear deteriorating effect. The goodness of fit for the

two-parameter model is compared to two other models: a model assuming a naive "on-off" relationship and a model that is fully parameterized for each time interval over the assumed duration of the effect. The goodness of fit is assessed by

- Likelihood ratio tests
- AIC

## 4.1 Comparison of Case-Crossover and Cohort Design

The hazard function in Equation (3.4) is simplified to assume just one possible baseline patient characteristic, denoted $X_i$, such that the hazard is

$$\lambda_i(t) = \lambda_0(t)e^{\gamma X_i + \beta Z_{it}} \tag{4.1}$$

This first set of simulation studies assumes the exposure variable indicates a change in status, and the effect of the change has a duration of one time interval (i.e. $\Delta = 1$). Furthermore, these simulation studies are broken down into five scenarios which invovle different assumptions about the relationships between the time-dependent exposure, baseline covariate, and event. The first two scenarios assume that there is no baseline covariate, and therefore the $\gamma X_i$ term is dropped from Equation (4.1). Scenario one assumes that the time-dependent covariate is independent of the event, while scenario two assumes the exposure is positively associated with the event. These two scenarios are illustrated in Figure 4.1.a and Figure 4.1.b respectively. The third scenario, illustrated in Figure 4.1.c, assumes that both the baseline covariate and the time-dependent exposure are associated with the event but are uncorrelated with each other. Figure 4.1.d illustrates the fourth scenario; this scenario assumes the baseline covariate is associated with both the exposure and the event, but the exposure is independent of the event. Finally, the fifth scenario, illustrated in Figure 4.1.e, introduces a setting where the exposure's relationship is confounded by the baseline covariate; the exposure is associated with the event and the baseline covariate is associated with both the event and exposure.

The purpose of introducing the baseline covariate into the simulations for scenarios 3-5 is to provide further evaluation on the performance of both design's ability to estimate $\beta$

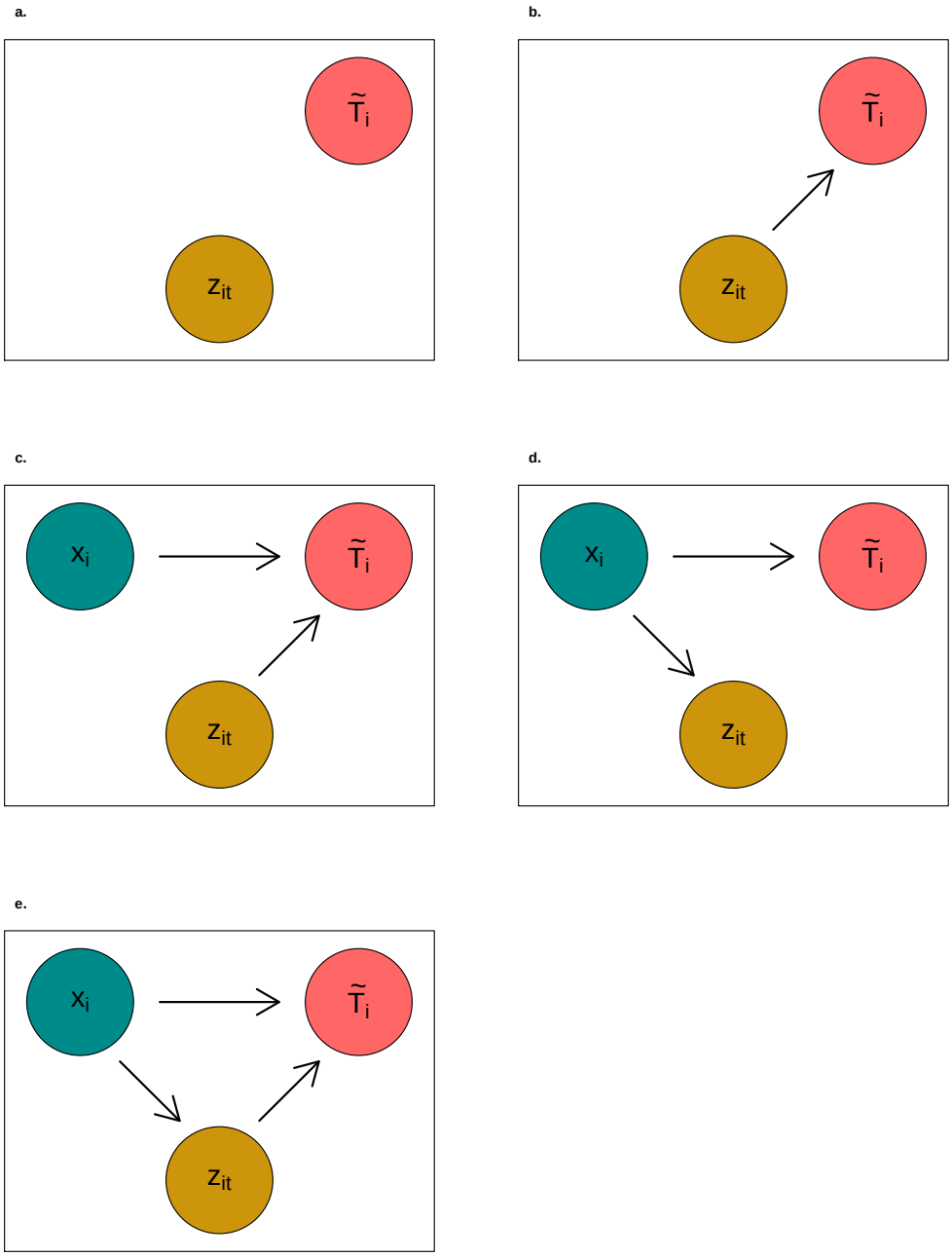


Figure 4.1: Causal diagram illustrating the five scenarios of interrelationships between three factors: baseline covariate ($x_i$), time-dependent exposure ($z_{it}$), and the time to event ($\tilde{T}_i$). The five scenarios are described as: *a.*) the time-dependent exposure is independent of the event, *b.*) the time-dependent exposure is associated with the event, *c.*) both the baseline covariate and the time-dependent exposure are independently associated with the event, *d.*) the baseline covariate is associated with the time-dependent exposure and the event, but the time-dependent exposure is independent of the event, *e.*) the baseline covariate confounds the relationship between the time-dependent exposure and the event

when the baseline covariate is not observed. The derivation of the Cox proportional hazards model in Equation (3.6) and the case-crossover design likelihood function in Equation (6.2) demonstrate that confounders measured at baseline must be controlled for in the Cox proportional hazards model but not the case-crossover design. As a result, the simulation studies for scenarios 3-5 demonstrate the implications associated with the investigators ability to observe and control for the confounder; simulation results labeled "model 1" assume that the covariate is observed, whereas those labeled "model 2" assume that it was not observed.

Austin outlined a data generating process for the Cox proportional hazards model with time-dependent covariates by assuming survival times follow an exponential, Weibull, or Gompertz distributions [3]. The exponential distribution is chosen for all simulations in this thesis to satisfy the assumptions made by the case-crossover design, this includes assuming the baseline hazard is constant, and therefore, $\lambda_0(t) = \lambda$ in Equation (4.1). Following the guidelines by Austin, the cohort is simulated from the algorithm outlined as follows

1. For scenarios three and four the baseline covariate is simulated from a binomial distribution. For scenarios four and five, the confounding relationship is setup so that the probability of observing a change in exposure at any time, denoted $Z_{i\cdot}$, is conditional on the status of the binary, baseline covariates. That is, $P(Z_{i\cdot} = 1)$ is not equal to $P(Z_{i\cdot} = 1 \mid X_i = 1)$. The baseline covariate and the chance to observe the exposure change at any time are generated from a correlated, multivariate, binary distribution using the `rmvbin` function in the bindata package in **R** [50].

2. Change of exposure times $\tau_i$ are simulated from a $Uniform(0, 60)$ distribution and are rounded to replicate the discrete, time scale observed in the VNSNY Medication and Fall data. For scenarios 4 and 5, no change in exposure time is generated if $Z_{i\cdot} = 0$.

3. The cumulative hazard, denoted $u$, is simulated from $Uniform(0, 1)$ distribution

4. Censoring times, denoted $C_i$, are simulated from a $Uniform(30, 60)$ distribution. These censoring times may also censor the time-dependent exposures from step 2 in addition to the event.

5. The time-to-event is then calculated from the inverse cumulative hazard function, which is a function of the random components: the baseline covariate, time-dependent exposure, and the cumulative hazard such that

$$\tilde{T} = \begin{cases} \frac{-log(u)}{\lambda e^{\gamma X_i}} & -log(u) < \lambda e^{\gamma X_i}\tau_i \\ \frac{-log(u)+\lambda e^{\gamma X_i}\left(e^{\beta}\tau_i-\tau_i\right)}{\lambda e^{\gamma X_i+\beta}} & -log(u) \in \left[\lambda e^{\gamma X_i}\tau_i, \lambda e^{\gamma X_i}(\tau_i+e^{\beta})\right) \\ \frac{-log(u)+\lambda e^{\gamma X_i}\left(1-e^{\beta}\right)}{\lambda e^{\gamma X_i}} & -log(u) \geq \lambda e^{\gamma X_i}(\tau_i+e^{\beta}) \end{cases} \tag{4.2}$$

6. Using the entire simulated cohort, log hazard estimates for the Cox hazard model is obtained using the `coxph` function from the survival package in **R** [97] .

7. To apply the case-crossover design, the cohort is filtered so that only patients with an event are included. The data is structured to meet the requirements of several different case-crossover reference strategies. The case-crossover designs assume that the case time is fixed immediately prior to the fall and $m = 1, \ldots, M$ control-times are fixed sequentially in the times preceding the case-time. Reference strategies of 1:1, 1:2, 1:3, 1:5, 1:10, and 1:25 are considered. When the correct number of controls are not available, because they occur prior to entry, the maximum number of available control times are selected, where a minimum of one control time is required. Log-hazard estimates for the conditional logistic regression are obtained using the `clogit` function in the survival package in **R** [97].

The case-crossover assumes that the event is rare and the incidence of falls in the VNSNY data is 4.4%. As a result, the parameters of the simulation studies are set to achieve incidence rates of approximately 5% and 10%. These incidence rates will be generated for cohorts of size 2500, 5000, 10000, 25000, 50000, 100000. All simulations are carried out for $Q = 1000$ iterations. The competing methods are evaluated based on measures of bias, relative efficiency, Type I error, Power, and coverage probability. Relative efficiency in these simulations is variance of the estimates for the case-crossover design compared to the Cox proportional hazards model.

### 4.1.1 Event is Independent of Exposure

The first set of simulations assume that no baseline covariate exists and the cohort arises from the hazard function

$$\lambda_i(t) = \lambda e^{\beta Z_{it}} \tag{4.3}$$

These simulations will generate data and assume that a change in the exposure has no effect on the risk of an event ($\beta = 0$). Simulations set $\lambda = 0.0025$ to obtain an incidence of approximately 10% and $\lambda = 0.00125$ for a 5% incidence. The algorithm estimating $\beta$ for the conditional logistic regression and Cox proportional hazards model failed to converge for a few iterations. This only occurred in the smaller cohorts, and is due to the sparseness of the event and exposure.

The results of simulations assuming no effect are displayed in Table 4.1 and 4.2 for cohorts with 10% and 5% incidence rates, respectively. Bias is present in the estimates of $\beta$ for cohorts of smaller size. The 1:1 case-crossover design slight over estimates $\beta$ for cohorts with 500 events or less. This is likely due to the slightly skewed right distribution, shown in Figure 4.3, as the median estimate is equal to zero at up to four decimal places. For the cohort with 5% incidence rate among 5,000 patients, the average estimate is substantially smaller than the true value of zero for the 1:25 case-crossover design and Cox proportional hazards model; resulting in average estimates of $\beta$ is -0.09 and -0.10 respectively.

The percentile intervals and standard deviations of the estimate decrease with larger cohorts size, incidence rates, and the number of matched controls, which can be observed in Figures 4.2 and 4.4. The type I error rate displayed in Tables 4.1 and 4.2 is calculated as the number of simulations where the null hypothesis assuming no effect was rejected assuming a type I error rate of 5%. The Type I error rate is lower than expected for smaller cohort sizes. This is likely due to the larger standard errors resulting from the small sample size.

Figure 4.5 displays the gains in relative efficiency for the case-crossover design compared to the Cox proportional hazards model. The 1:5 case-crossover design applied to cohorts with 10% incidence rate typically achieve almost 80% efficiency. The highest relative efficiency measures are observed in the larger cohorts with a 10% incidence rate and largest

number of matched controls. A strange pattern of relative efficiency is noticeable among the different cohort sizes. The standard deviations of the simulation estimates consistently get smaller with larger $N$ and $M$, and is shown in Figure 4.4. The relatively efficiency is calculated from the MSE which is defined as

$$MSE = (\overline{\widehat{\beta}} - \beta)^2 + (SE(\widehat{\beta}))^2$$

The MSE calculation contains the amount of bias summed with the standard error. Figure 4.6 displays a ratio of the bias squared compared to the standard error squared. This plot shows that while the amount of bias relative to the standard error decreases with larger N, it tends to be the highest for the Cox and 1:25 case-crossover design for a fixed cohort size, but a comparison of the ratio in smaller matched case-crossover design varies dramatically by incidence rate. For larger sample sizes, relative efficiency of the case-crossover design to the cohort design roughly follows the $\frac{M}{M+1}$ property derived by Breslow and Day for case-control studies [12].

| | Mean (SD*) | Median | Type I | Avg Std Error** | Relative Efficiency | Converged |
|---|---|---|---|---|---|---|
| **Cohort Size (N=5000 , # event=474)** | | | | | | |
| Case−Crossover (1:1) | 0.0012 (0.5562) | 0.0000 | 2.80% | 0.55 | 47.54% | 99.90% |
| Case−Crossover (1:2) | −0.0275 (0.4660) | 0.0000 | 3.30% | 0.47 | 67.49% | 100.00% |
| Case−Crossover (1:3) | −0.0376 (0.4470) | −0.0113 | 3.80% | 0.45 | 73.09% | 100.00% |
| Case−Crossover (1:5) | −0.0456 (0.4261) | −0.0157 | 3.80% | 0.42 | 80.07% | 100.00% |
| Case−Crossover (1:10) | −0.0520 (0.4091) | −0.0158 | 3.00% | 0.41 | 86.48% | 100.00% |
| Case−Crossover (1:25) | −0.0564 (0.4035) | −0.0129 | 3.50% | 0.40 | 88.58% | 100.00% |
| Cox Hazards Model | −0.0595 (0.3788) | −0.0168 | 3.30% | 0.38 | 100.00% | 100.00% |
| **Cohort Size (N=10000 , # event=948)** | | | | | | |
| Case−Crossover (1:1) | 0.0000 (0.3727) | 0.0000 | 4.10% | 0.38 | 50.33% | 100.00% |
| Case−Crossover (1:2) | −0.0083 (0.3266) | 0.0000 | 4.70% | 0.33 | 65.52% | 100.00% |
| Case−Crossover (1:3) | −0.0162 (0.3097) | 0.0070 | 4.90% | 0.31 | 72.69% | 100.00% |
| Case−Crossover (1:5) | −0.0228 (0.2934) | −0.0031 | 4.60% | 0.29 | 80.72% | 100.00% |
| Case−Crossover (1:10) | −0.0295 (0.2822) | −0.0060 | 4.20% | 0.28 | 86.85% | 100.00% |
| Case−Crossover (1:25) | −0.0325 (0.2777) | −0.0067 | 4.20% | 0.28 | 89.41% | 100.00% |
| Cox Hazards Model | −0.0325 (0.2624) | −0.0088 | 4.20% | 0.26 | 100.00% | 100.00% |
| **Cohort Size (N=25000 , # event=2367)** | | | | | | |
| Case−Crossover (1:1) | 0.0063 (0.2280) | 0.0000 | 4.20% | 0.23 | 49.49% | 100.00% |
| Case−Crossover (1:2) | 0.0058 (0.1994) | 0.0071 | 4.00% | 0.20 | 64.73% | 100.00% |
| Case−Crossover (1:3) | 0.0041 (0.1885) | 0.0062 | 4.20% | 0.19 | 72.45% | 100.00% |
| Case−Crossover (1:5) | 0.0015 (0.1764) | 0.0009 | 3.20% | 0.18 | 82.76% | 100.00% |
| Case−Crossover (1:10) | −0.0015 (0.1712) | 0.0007 | 4.00% | 0.18 | 87.85% | 100.00% |
| Case−Crossover (1:25) | −0.0039 (0.1685) | −0.0049 | 3.90% | 0.17 | 90.68% | 100.00% |
| Cox Hazards Model | −0.0045 (0.1604) | 0.0049 | 4.10% | 0.16 | 100.00% | 100.00% |
| **Cohort Size (N=50000 , # event=4738)** | | | | | | |
| Case−Crossover (1:1) | 0.0036 (0.1640) | 0.0061 | 6.20% | 0.16 | 47.47% | 100.00% |
| Case−Crossover (1:2) | 0.0052 (0.1391) | 0.0100 | 3.60% | 0.14 | 65.92% | 100.00% |
| Case−Crossover (1:3) | 0.0040 (0.1295) | 0.0046 | 4.90% | 0.13 | 76.11% | 100.00% |
| Case−Crossover (1:5) | 0.0019 (0.1219) | 0.0078 | 4.50% | 0.13 | 85.88% | 100.00% |
| Case−Crossover (1:10) | 0.0004 (0.1187) | 0.0052 | 4.60% | 0.12 | 90.69% | 100.00% |
| Case−Crossover (1:25) | −0.0004 (0.1168) | 0.0051 | 4.40% | 0.12 | 93.58% | 100.00% |
| Cox Hazards Model | −0.0010 (0.1130) | 0.0095 | 5.00% | 0.11 | 100.00% | 100.00% |
| **Cohort Size (N=100000, # event=9475)** | | | | | | |
| Case−Crossover (1:1) | −0.0009 (0.1143) | 0.0000 | 5.60% | 0.12 | 49.18% | 100.00% |
| Case−Crossover (1:2) | 0.0012 (0.0994) | 0.0075 | 4.40% | 0.10 | 65.02% | 100.00% |
| Case−Crossover (1:3) | −0.0001 (0.0924) | 0.0039 | 3.80% | 0.10 | 75.23% | 100.00% |
| Case−Crossover (1:5) | −0.0009 (0.0876) | −0.0016 | 4.80% | 0.09 | 83.71% | 100.00% |
| Case−Crossover (1:10) | −0.0020 (0.0842) | −0.0003 | 3.90% | 0.09 | 90.59% | 100.00% |
| Case−Crossover (1:25) | −0.0021 (0.0834) | 0.0023 | 4.10% | 0.09 | 92.39% | 100.00% |
| Cox Hazards Model | −0.0026 (0.0801) | 0.0016 | 5.00% | 0.08 | 100.00% | 100.00% |

Table 4.1: Results from simulation of a cohort with an incidence of approximately 10% ( $\lambda = .0025$) incidence and assuming no transient effect ($\beta = .0$). (* Monte Carlo standard deviation of estimates ** Average of the information based standard error)

| | Mean (SD*) | Median | Type I | Avg Std Error** | Relative Efficiency | Converged |
|---|---|---|---|---|---|---|
| **Cohort Size (N=5000 , # event=243)** | | | | | | |
| Case−Crossover (1:1) | 0.0386 (0.8035) | 0.0000 | 1.24% | 0.81 | 49.80% | 96.40% |
| Case−Crossover (1:2) | −0.0248 (0.6906) | 0.0000 | 2.94% | 0.69 | 67.48% | 98.50% |
| Case−Crossover (1:3) | −0.0525 (0.6499) | 0.0000 | 3.55% | 0.65 | 75.80% | 98.50% |
| Case−Crossover (1:5) | −0.0670 (0.6191) | 0.0207 | 3.76% | 0.61 | 83.11% | 98.50% |
| Case−Crossover (1:10) | −0.0823 (0.5966) | 0.0102 | 3.76% | 0.59 | 88.85% | 98.50% |
| Case−Crossover (1:25) | −0.0873 (0.5827) | 0.0024 | 4.16% | 0.58 | 92.82% | 98.50% |
| Cox Hazards Model | −0.1020 (0.5585) | −0.0114 | 3.64% | 0.55 | 100.00% | 98.80% |
| **Cohort Size (N=10000 , # event=486)** | | | | | | |
| Case−Crossover (1:1) | 0.0142 (0.5801) | 0.0000 | 3.90% | 0.54 | 45.34% | 100.00% |
| Case−Crossover (1:2) | −0.0266 (0.4799) | 0.0000 | 3.10% | 0.47 | 66.08% | 100.00% |
| Case−Crossover (1:3) | −0.0374 (0.4511) | −0.0171 | 3.40% | 0.44 | 74.50% | 100.00% |
| Case−Crossover (1:5) | −0.0448 (0.4346) | −0.0058 | 3.20% | 0.42 | 79.98% | 100.00% |
| Case−Crossover (1:10) | −0.0508 (0.4234) | −0.0189 | 3.50% | 0.40 | 83.95% | 100.00% |
| Case−Crossover (1:25) | −0.0517 (0.4171) | −0.0141 | 3.30% | 0.40 | 86.41% | 100.00% |
| Cox Hazards Model | −0.0553 (0.3868) | −0.0210 | 3.80% | 0.37 | 100.00% | 100.00% |
| **Cohort Size (N=25000 , # event=1215)** | | | | | | |
| Case−Crossover (1:1) | −0.0026 (0.3444) | 0.0000 | 5.20% | 0.33 | 41.99% | 100.00% |
| Case−Crossover (1:2) | −0.0115 (0.2905) | 0.0000 | 4.90% | 0.29 | 58.93% | 100.00% |
| Case−Crossover (1:3) | −0.0165 (0.2722) | −0.0107 | 4.70% | 0.27 | 66.97% | 100.00% |
| Case−Crossover (1:5) | −0.0192 (0.2568) | −0.0086 | 4.80% | 0.26 | 75.14% | 100.00% |
| Case−Crossover (1:10) | −0.0202 (0.2475) | −0.0106 | 4.60% | 0.25 | 80.81% | 100.00% |
| Case−Crossover (1:25) | −0.0198 (0.2417) | −0.0029 | 4.60% | 0.24 | 84.70% | 100.00% |
| Cox Hazards Model | −0.0223 (0.2221) | −0.0165 | 4.50% | 0.23 | 100.00% | 100.00% |
| **Cohort Size (N=50000 , # event=2430)** | | | | | | |
| Case−Crossover (1:1) | −0.0082 (0.2324) | 0.0000 | 3.80% | 0.23 | 48.69% | 100.00% |
| Case−Crossover (1:2) | −0.0146 (0.2007) | −0.0143 | 4.90% | 0.20 | 65.02% | 100.00% |
| Case−Crossover (1:3) | −0.0158 (0.1887) | −0.0127 | 4.70% | 0.19 | 73.46% | 100.00% |
| Case−Crossover (1:5) | −0.0178 (0.1792) | −0.0155 | 4.70% | 0.18 | 81.17% | 100.00% |
| Case−Crossover (1:10) | −0.0155 (0.1748) | −0.0137 | 4.70% | 0.17 | 85.54% | 100.00% |
| Case−Crossover (1:25) | −0.0147 (0.1728) | −0.0117 | 5.00% | 0.17 | 87.60% | 100.00% |
| Cox Hazards Model | −0.0170 (0.1614) | −0.0149 | 4.50% | 0.16 | 100.00% | 100.00% |
| **Cohort Size (N=100000, # event=4861)** | | | | | | |
| Case−Crossover (1:1) | 0.0003 (0.1603) | 0.0000 | 4.10% | 0.16 | 49.34% | 100.00% |
| Case−Crossover (1:2) | −0.0034 (0.1421) | −0.0063 | 5.30% | 0.14 | 62.81% | 100.00% |
| Case−Crossover (1:3) | −0.0036 (0.1327) | −0.0055 | 4.90% | 0.13 | 71.98% | 100.00% |
| Case−Crossover (1:5) | −0.0052 (0.1260) | −0.0038 | 4.40% | 0.13 | 79.69% | 100.00% |
| Case−Crossover (1:10) | −0.0045 (0.1227) | 0.0002 | 5.20% | 0.12 | 84.17% | 100.00% |
| Case−Crossover (1:25) | −0.0047 (0.1204) | 0.0009 | 5.30% | 0.12 | 87.40% | 100.00% |
| Cox Hazards Model | −0.0062 (0.1125) | −0.0035 | 4.70% | 0.11 | 100.00% | 100.00% |

Table 4.2: Results from simulation of a cohort with an incidence of approximately 5% ( $\lambda = .00125$) incidence and assuming no transient effect ($\beta = .0$). (* Monte Carlo standard deviation of estimates ** Average of the information based standard error)

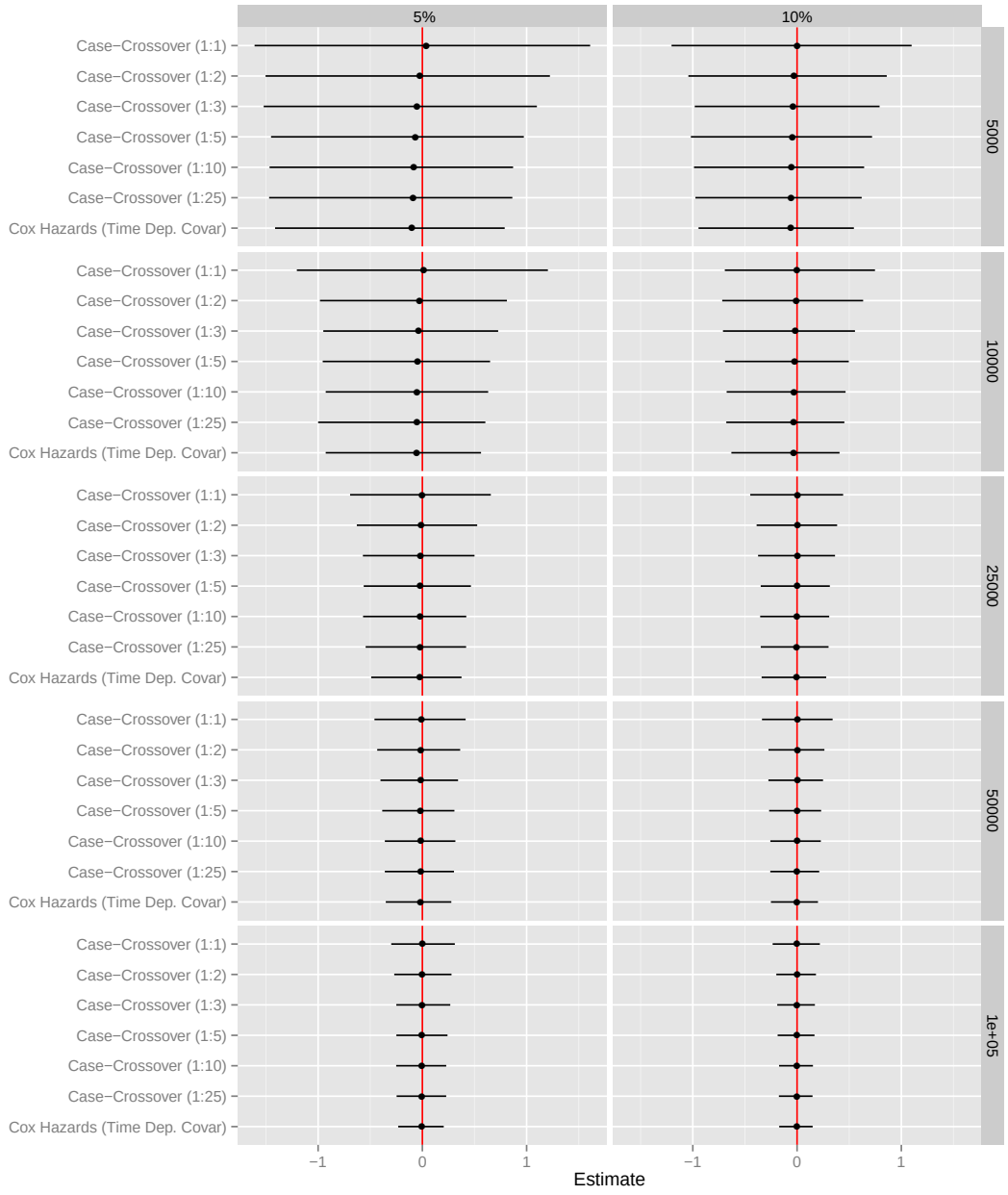


Figure 4.2: Plot of average estimates of log-hazard ratio with 95% percentile intervals by design, cohort size and incidence rate when assuming a log-hazard ratio of $\beta = 0.8$ indicated by the vertical red line.

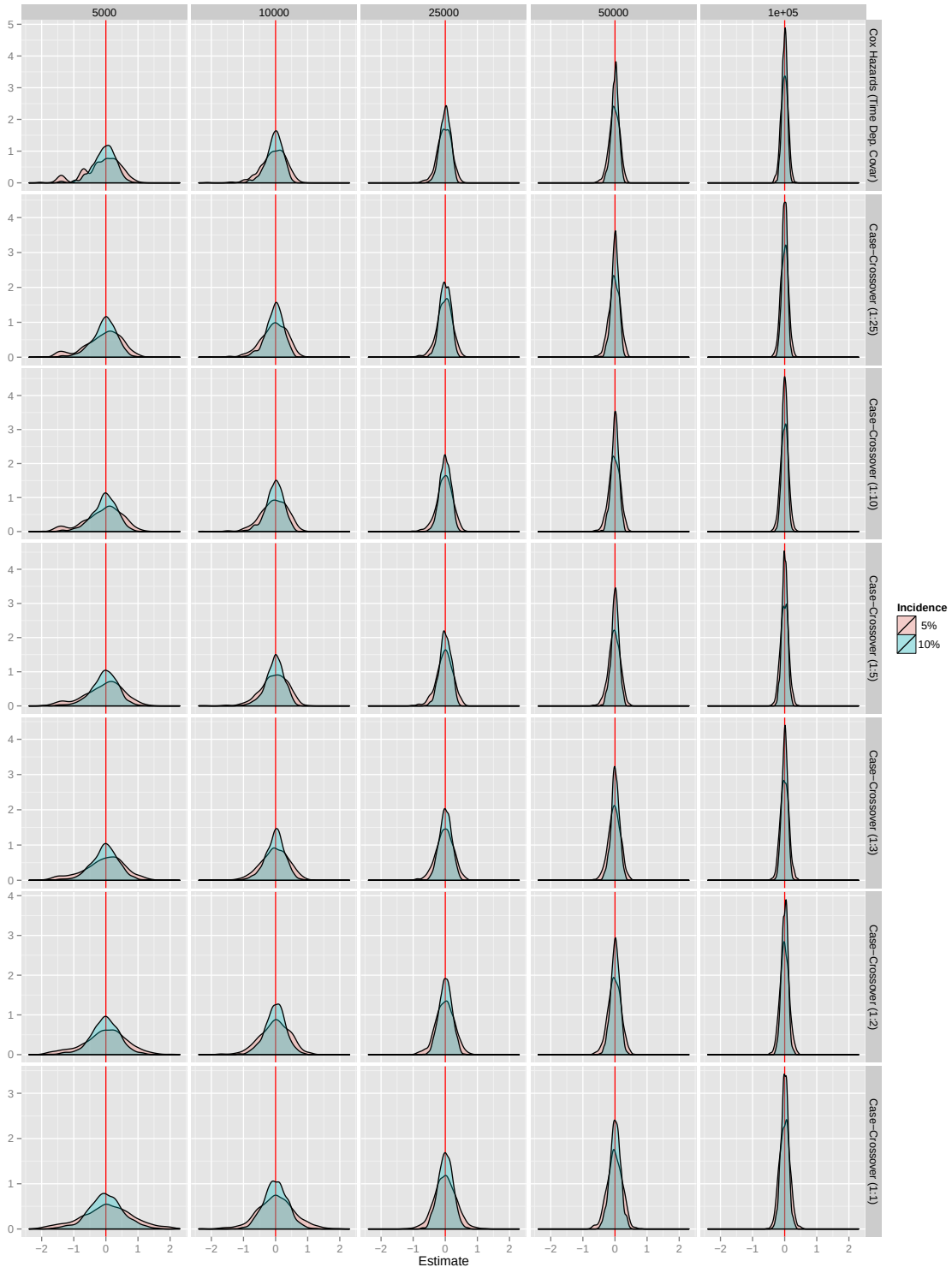


Figure 4.3: Distribution of estimate by design, cohort size and incidence rate when assuming a log-hazard ratio of $\beta = 0$ indicated by the vertical red line.

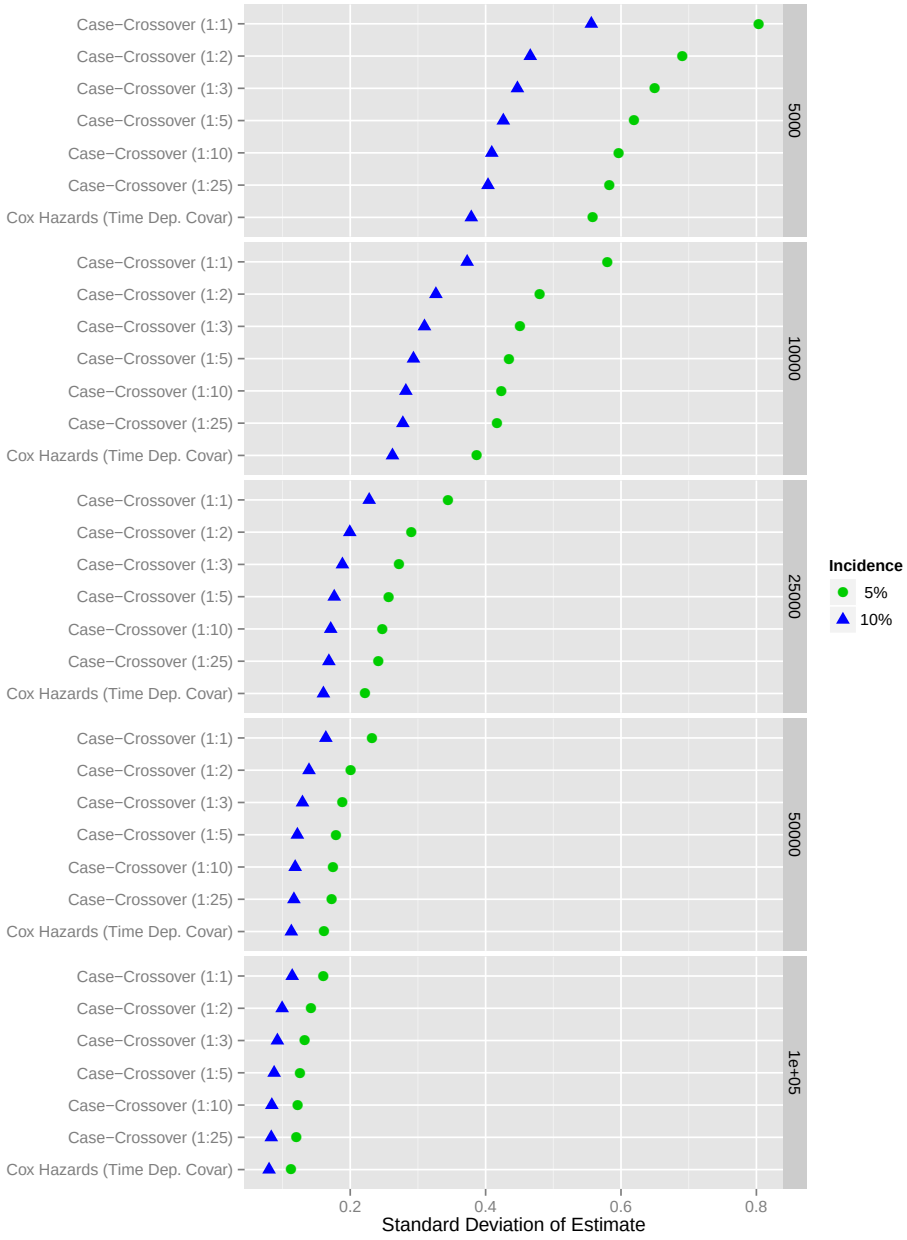


Figure 4.4: Plot of standard deviation of log-hazard ratio estimates by design, cohort size and incidence rate when assuming a log-hazard ratio of $\beta = .0$

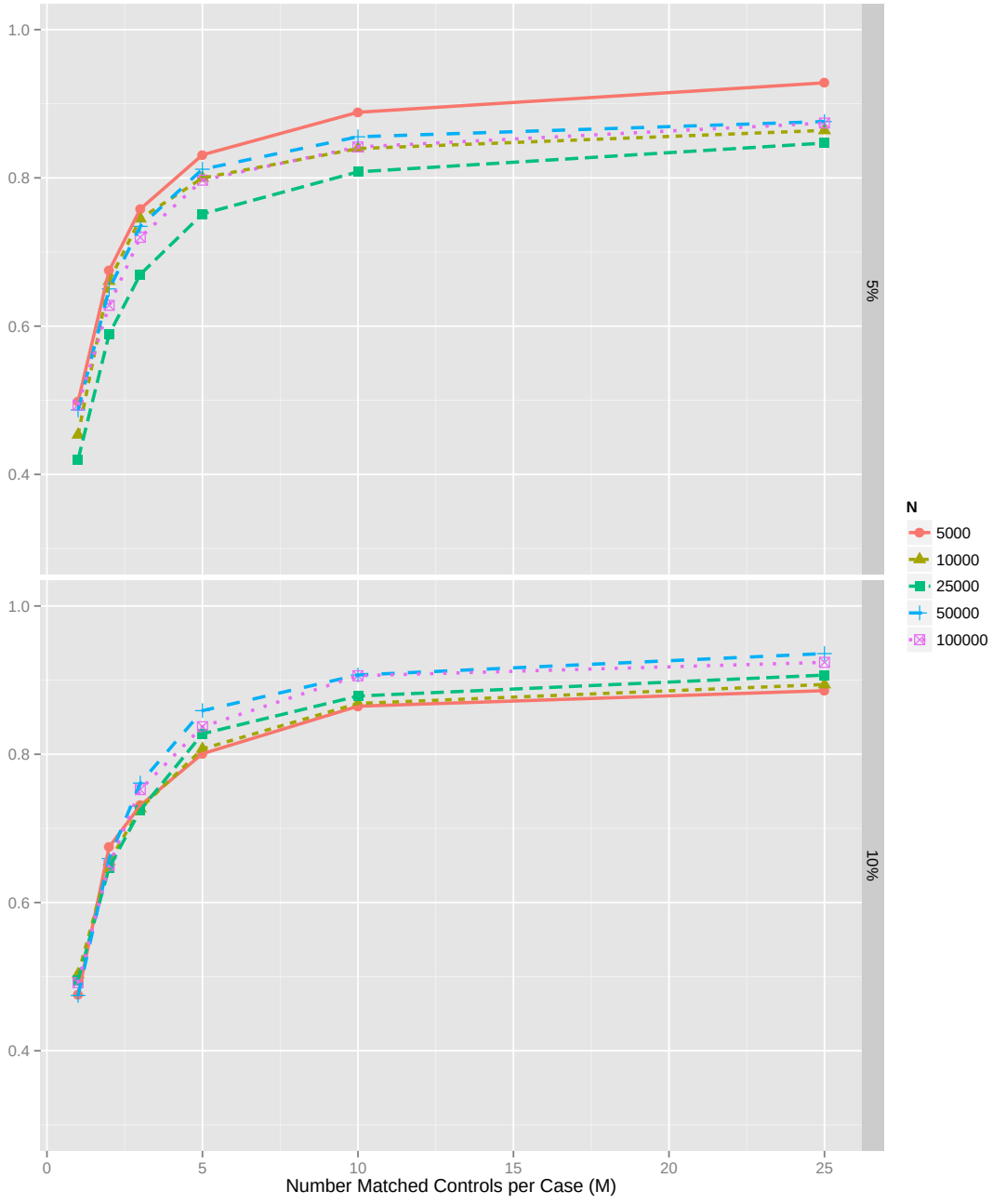


Figure 4.5: Plot of relative efficiency by design, cohort size and incidence rate when assuming a log-hazard ratio of $\beta = 0$

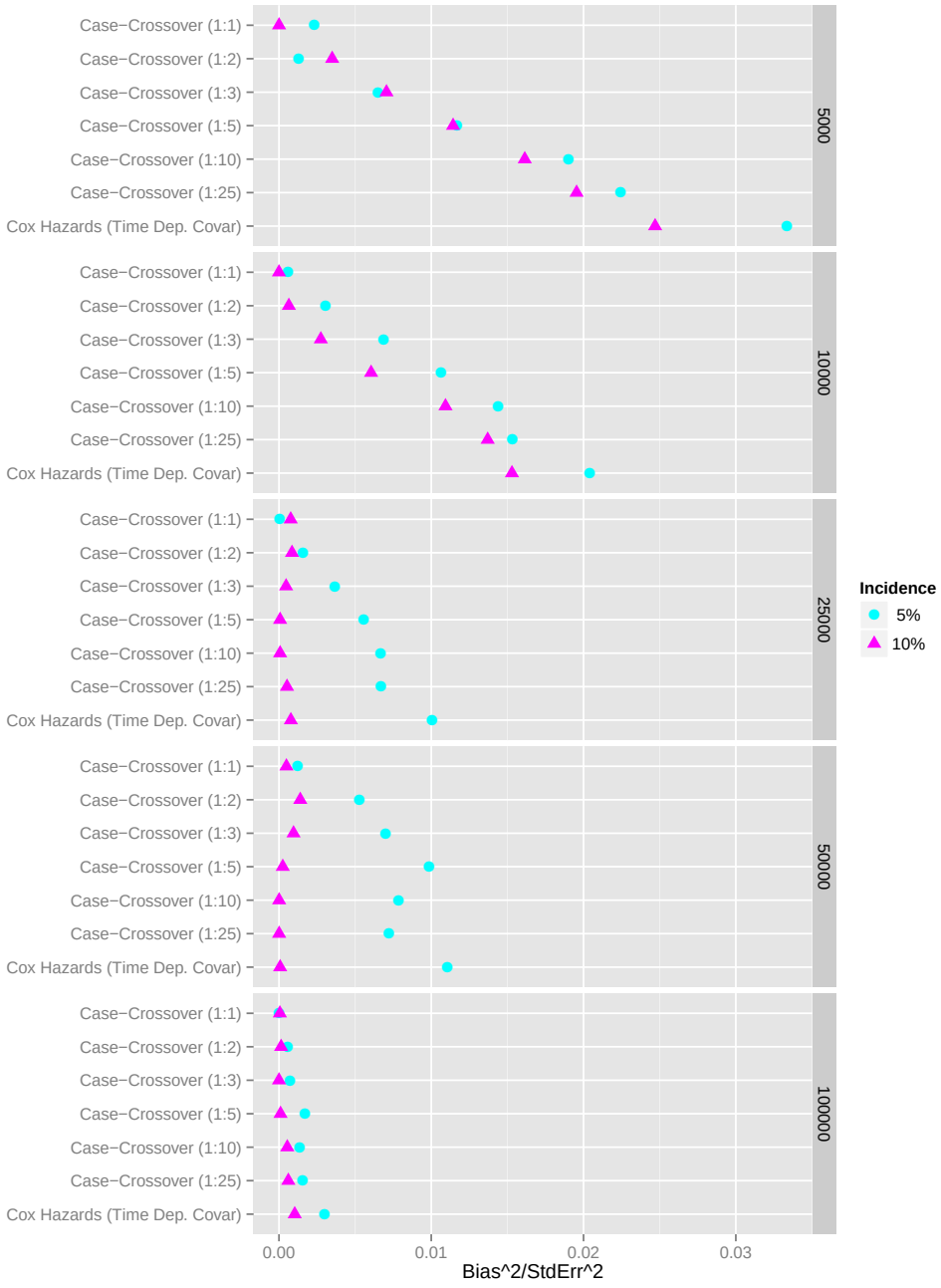


Figure 4.6: Plot of the ratio of $Bias^2$ to $StandardError^2$, or $\frac{(\bar{\hat{\beta}}-\beta)^2}{(SE(\hat{\beta}))^2}$

### 4.1.2 Event is Associated with Exposure

Table 4.3 and 4.4 display the results of the simulation assuming an association of $\beta = 0.8$ for 10% and 5% incidence rates, respectively. A plot of the average estimate, along with 95% percentile intervals is shown in Figure 4.7. From this figure, there appears to be some bias in cohorts of smaller sample size. The 1:1 case-crossover design typically over estimates $\beta$; as high as 7.6% for the cohort of size 10000 with a 10% incidence and does reach less than 1% bias until the sample size for the cohort reaches 100000. The Cox proportional hazards model and case-crossover designs matching 10 or 25 controls underestimate $\beta$. This observed bias is larger for the cohorts with 5% incidence rates compared to 10%, and becomes less than 1% for cohorts of size larger than 10,000 for cohorts with 10% incidence rates and 25,000 for 5%. In the cohorts of small sample size, the 1:3 case-crossover design appears to be optimal in terms of bias. In these smaller samples, larger values of $M$, compared to the 1:3 case-crossover design, increases the degree in which $\beta$ is underestimated.

The percentile intervals, plotted in Figure 4.7, are larger for the cohorts with 5% incidence rates, and narrow with both the size of the cohort and for the number of matched controls for the case-crossover designs. For each cohort size, the estimate and the percentile interval of the case-crossover design becomes very similar to that of the Cox proportional hazards model with time-dependent covariates as $M$ increases. The first 95% percentile interval for the case-crossover design's estimate of $\beta$ that does not overlap zero is the 1:2 reference strategy for the cohort of size 5,000 with a 10% incidence rate, and 10,000 for a 5%; equivalent to 500 events for both incidence rates. The Cox proportional hazards model also requires cohorts of size 5,000 and 10,0000 to observe percentile intervals that do not overlap zero for the respective incidence rates.

The power of each method is calculated as the percentage of simulations that reject the null hypothesis for these studies which assume the data is generated under the alternative hypothesis ($\beta = 0.8$). As expected, the power increases with sample size and exceeds 80% for the Cox proportional hazards model at 500 events and at 1000 events for the case-crossover design. 100% of simulations reject the null hypothesis for cohorts once the

number of events reaches about 2500. Coverage probabilities are higher than expected for smaller cohort sizes, but are very close to 95% once 1000 events are observed.

The full distribution of the estimate is shown in Figure 4.8 and appears to be mostly symmetric. Similar to the percentile intervals, the distribution can be seen to narrow for larger cohort sample size, matched controls, and incidence rates. Figure 4.10 plots the standard deviation of the estimates. From this plot it can be seen that the standard deviation of the estimate decreases with larger cohort sample size, matched controls, and incidence rates. A plot of relative efficiency is shown in Figure 4.10. This plot shows that gains in relative efficiency accrue with larger numbers of matched controls for the case-crossover design, but does not appear to be strongly influenced by cohort size.

### 4.1.3 Baseline Covariate and Exposure are Marginally Associated with Event

The third scenario, depicted in Figure 4.1.c, introduces a baseline covariate, $X_i$, into Equation (4.1). The set of simulation studies in this subsection assume that the baseline covariate and the time-dependent exposure are independent. That is, $P(Z_{i\cdot} = 1) = P(Z_{i\cdot} = 1 \mid X_i = 1)$. To obtain an incidence of approximately 10% and 5% for the event, $\lambda$ is set to values of 0.0018 and 0.0009 respectively. The results for the cohort of 10% incidence is shown in Table 4.5 and in Table 4.6 for 5%.

The mean parameter estimates, 95% percentile intervals, and average information based, standard errors for both the Cox proportional hazards models, assuming the covariate is observed and unobserved, are nearly exactly the same across all simulations for cohorts with sample size greater than 5000. This is expected since $X_i$ is set to be independent of $Z_{i\cdot}$ in these simulations. The over estimation of $\beta$ still exist for the 1:1 case-crossover design, but not to the degree of which it was observed in subsection 4.1.2; a maximum of 4.3% bias for the cohort of size 5000 with a 5% incidence rate. The underestimation of $\beta$, for both the Cox proportional hazards model and case-crossover designs with matching strategies of 1:10 and 1:25, is similar to that in subsection 4.1.2. The Cox proportional hazards model and 1:25 case-crossover design underestimate the log-hazard ratio by up to 8.6% in the cohorts of size 2500 and incidence rate of 10%.

| | Mean (SD*) | Median | Power | Coverage Probability | Bias | Avg Std Error** | Relative Efficiency | Converged |
|---|---|---|---|---|---|---|---|---|
| **Cohort Size (N= 2500, # event=241)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8561 (0.6759) | 0.8109 | 17.85% | 97.64% | 7.01% | 0.69 | 33.13% | 97.50% |
| Case−Crossover (1:2) | 0.8160 (0.5598) | 0.8109 | 34.73% | 97.20% | 2.00% | 0.54 | 48.30% | 99.90% |
| Case−Crossover (1:3) | 0.7854 (0.4912) | 0.8109 | 41.94% | 96.90% | −1.83% | 0.49 | 62.74% | 99.90% |
| Case−Crossover (1:5) | 0.7551 (0.4530) | 0.7828 | 45.35% | 97.10% | −5.61% | 0.44 | 73.77% | 99.90% |
| Case−Crossover (1:10) | 0.7518 (0.4291) | 0.7843 | 50.75% | 96.10% | −6.03% | 0.41 | 82.19% | 99.90% |
| Case−Crossover (1:25) | 0.7470 (0.4146) | 0.7890 | 52.95% | 96.10% | −6.62% | 0.40 | 88.04% | 99.90% |
| Cox Hazards Model | 0.7379 (0.3890) | 0.7946 | 58.26% | 96.40% | −7.76% | 0.36 | 100.00% | 99.90% |
| **Cohort Size (N= 5000, # event=482)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8610 (0.4954) | 0.8109 | 44.14% | 97.00% | 7.62% | 0.47 | 25.68% | 99.90% |
| Case−Crossover (1:2) | 0.8121 (0.3648) | 0.8056 | 60.00% | 96.90% | 1.51% | 0.37 | 47.35% | 100.00% |
| Case−Crossover (1:3) | 0.7955 (0.3272) | 0.8044 | 68.00% | 96.10% | −0.57% | 0.34 | 58.87% | 100.00% |
| Case−Crossover (1:5) | 0.7836 (0.3011) | 0.7948 | 72.90% | 95.50% | −2.05% | 0.31 | 69.50% | 100.00% |
| Case−Crossover (1:10) | 0.7777 (0.2859) | 0.7811 | 76.70% | 95.80% | −2.78% | 0.29 | 77.09% | 100.00% |
| Case−Crossover (1:25) | 0.7752 (0.2721) | 0.7858 | 78.40% | 95.80% | −3.09% | 0.27 | 85.10% | 100.00% |
| Cox Hazards Model | 0.7723 (0.2510) | 0.7890 | 83.60% | 96.00% | −3.47% | 0.25 | 100.00% | 100.00% |
| **Cohort Size (N= 10000, # event=963)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8270 (0.3279) | 0.8183 | 75.50% | 95.80% | 3.38% | 0.32 | 29.13% | 100.00% |
| Case−Crossover (1:2) | 0.8111 (0.2570) | 0.8118 | 88.80% | 95.60% | 1.38% | 0.26 | 47.43% | 100.00% |
| Case−Crossover (1:3) | 0.8042 (0.2322) | 0.8031 | 92.30% | 94.90% | 0.53% | 0.23 | 58.07% | 100.00% |
| Case−Crossover (1:5) | 0.7984 (0.2155) | 0.8062 | 94.20% | 95.10% | −0.20% | 0.21 | 67.44% | 100.00% |
| Case−Crossover (1:10) | 0.7932 (0.2028) | 0.7967 | 95.80% | 94.50% | −0.85% | 0.20 | 76.15% | 100.00% |
| Case−Crossover (1:25) | 0.7922 (0.1956) | 0.8003 | 96.20% | 94.60% | −0.98% | 0.19 | 81.81% | 100.00% |
| Cox Hazards Model | 0.7878 (0.1770) | 0.8007 | 97.70% | 95.00% | −1.52% | 0.17 | 100.00% | 100.00% |
| **Cohort Size (N= 25000, # event=2414)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8197 (0.1990) | 0.8190 | 98.80% | 96.20% | 2.47% | 0.20 | 29.50% | 100.00% |
| Case−Crossover (1:2) | 0.8092 (0.1588) | 0.8155 | 99.80% | 95.50% | 1.15% | 0.16 | 46.30% | 100.00% |
| Case−Crossover (1:3) | 0.8073 (0.1433) | 0.8143 | 99.90% | 95.80% | 0.91% | 0.15 | 56.85% | 100.00% |
| Case−Crossover (1:5) | 0.8058 (0.1326) | 0.8111 | 100.00% | 95.90% | 0.73% | 0.13 | 66.40% | 100.00% |
| Case−Crossover (1:10) | 0.8026 (0.1233) | 0.8089 | 100.00% | 96.10% | 0.33% | 0.13 | 76.85% | 100.00% |
| Case−Crossover (1:25) | 0.8031 (0.1189) | 0.8090 | 100.00% | 95.80% | 0.39% | 0.12 | 82.63% | 100.00% |
| Cox Hazards Model | 0.7988 (0.1081) | 0.8065 | 100.00% | 96.40% | −0.15% | 0.11 | 100.00% | 100.00% |
| **Cohort Size (N= 50000, # event=4827)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8089 (0.1435) | 0.8051 | 100.00% | 94.20% | 1.11% | 0.14 | 29.05% | 100.00% |
| Case−Crossover (1:2) | 0.8039 (0.1107) | 0.8107 | 100.00% | 95.20% | 0.49% | 0.11 | 48.78% | 100.00% |
| Case−Crossover (1:3) | 0.8041 (0.1016) | 0.8086 | 100.00% | 96.00% | 0.51% | 0.10 | 57.89% | 100.00% |
| Case−Crossover (1:5) | 0.8019 (0.0939) | 0.8046 | 100.00% | 95.80% | 0.23% | 0.09 | 67.86% | 100.00% |
| Case−Crossover (1:10) | 0.8001 (0.0888) | 0.8023 | 100.00% | 94.90% | 0.01% | 0.09 | 75.77% | 100.00% |
| Case−Crossover (1:25) | 0.8005 (0.0858) | 0.8032 | 100.00% | 95.50% | 0.06% | 0.09 | 81.26% | 100.00% |
| Cox Hazards Model | 0.7986 (0.0773) | 0.8025 | 100.00% | 95.50% | −0.17% | 0.08 | 100.00% | 100.00% |
| **Cohort Size (N=100000, # event=9659)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8017 (0.0977) | 0.7996 | 100.00% | 95.60% | 0.21% | 0.10 | 30.60% | 100.00% |
| Case−Crossover (1:2) | 0.7978 (0.0778) | 0.8010 | 100.00% | 95.60% | −0.27% | 0.08 | 48.18% | 100.00% |
| Case−Crossover (1:3) | 0.7976 (0.0715) | 0.7992 | 100.00% | 95.60% | −0.30% | 0.07 | 57.16% | 100.00% |
| Case−Crossover (1:5) | 0.7967 (0.0651) | 0.7977 | 100.00% | 96.20% | −0.42% | 0.07 | 68.79% | 100.00% |
| Case−Crossover (1:10) | 0.7961 (0.0616) | 0.7963 | 100.00% | 95.80% | −0.48% | 0.06 | 76.86% | 100.00% |
| Case−Crossover (1:25) | 0.7965 (0.0593) | 0.7984 | 100.00% | 95.60% | −0.44% | 0.06 | 83.05% | 100.00% |
| Cox Hazards Model | 0.7957 (0.0540) | 0.7973 | 100.00% | 95.80% | −0.53% | 0.05 | 100.00% | 100.00% |

Table 4.3: Results from a simulation of a cohort with an incidence of approximately 10% ( $\lambda = .0025$) incidence and assuming a transient effect lasting one time period ($\beta = 0.8$) (* Monte Carlo standard deviation of estimates ** Average of the information based standard error)

| | Mean (SD*) | Median | Power | Coverage Probability | Bias | Avg Std Error** | Relative Efficiency | Converged |
|---|---|---|---|---|---|---|---|---|
| **Cohort Size (N= 5000, # event=248)** | | | | | | | | |
| Case−Crossover (1:1) | 0.7962 (0.6311) | 0.8109 | 15.71% | 97.54% | −0.48% | 0.66 | 34.22% | 97.40% |
| Case−Crossover (1:2) | 0.7867 (0.5526) | 0.7885 | 33.70% | 96.59% | −1.66% | 0.53 | 44.63% | 99.70% |
| Case−Crossover (1:3) | 0.7788 (0.4889) | 0.7802 | 41.40% | 96.70% | −2.65% | 0.48 | 57.01% | 100.00% |
| Case−Crossover (1:5) | 0.7630 (0.4527) | 0.7922 | 46.50% | 96.10% | −4.63% | 0.44 | 66.51% | 100.00% |
| Case−Crossover (1:10) | 0.7468 (0.4160) | 0.7926 | 52.30% | 96.50% | −6.65% | 0.41 | 78.77% | 100.00% |
| Case−Crossover (1:25) | 0.7378 (0.3969) | 0.7775 | 53.30% | 96.70% | −7.77% | 0.39 | 86.54% | 100.00% |
| Cox Hazards Model | 0.7348 (0.3692) | 0.7837 | 58.10% | 96.50% | −8.15% | 0.36 | 100.00% | 100.00% |
| **Cohort Size (N= 10000, # event=496)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8360 (0.4734) | 0.8109 | 45.00% | 96.30% | 4.50% | 0.46 | 28.73% | 100.00% |
| Case−Crossover (1:2) | 0.7988 (0.3721) | 0.8048 | 61.20% | 95.80% | −0.15% | 0.36 | 46.48% | 100.00% |
| Case−Crossover (1:3) | 0.7921 (0.3292) | 0.8019 | 67.10% | 95.90% | −0.98% | 0.33 | 59.39% | 100.00% |
| Case−Crossover (1:5) | 0.7843 (0.3050) | 0.7994 | 72.30% | 95.60% | −1.96% | 0.30 | 69.22% | 100.00% |
| Case−Crossover (1:10) | 0.7777 (0.2867) | 0.8021 | 76.50% | 95.40% | −2.79% | 0.28 | 78.30% | 100.00% |
| Case−Crossover (1:25) | 0.7729 (0.2749) | 0.7943 | 77.70% | 95.40% | −3.39% | 0.27 | 85.20% | 100.00% |
| Cox Hazards Model | 0.7691 (0.2537) | 0.7941 | 82.40% | 95.10% | −3.86% | 0.25 | 100.00% | 100.00% |
| **Cohort Size (N= 25000, # event=1238)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8091 (0.2898) | 0.7985 | 83.60% | 94.90% | 1.14% | 0.28 | 29.07% | 100.00% |
| Case−Crossover (1:2) | 0.7920 (0.2223) | 0.7946 | 93.80% | 96.90% | −1.00% | 0.23 | 49.40% | 100.00% |
| Case−Crossover (1:3) | 0.7890 (0.2034) | 0.7838 | 95.40% | 95.40% | −1.37% | 0.21 | 59.03% | 100.00% |
| Case−Crossover (1:5) | 0.7877 (0.1884) | 0.7890 | 97.10% | 95.50% | −1.54% | 0.19 | 68.81% | 100.00% |
| Case−Crossover (1:10) | 0.7864 (0.1758) | 0.7905 | 98.30% | 95.00% | −1.69% | 0.18 | 79.02% | 100.00% |
| Case−Crossover (1:25) | 0.7849 (0.1687) | 0.7872 | 98.50% | 95.50% | −1.88% | 0.17 | 85.83% | 100.00% |
| Cox Hazards Model | 0.7823 (0.1563) | 0.7898 | 99.30% | 95.10% | −2.21% | 0.15 | 100.00% | 100.00% |
| **Cohort Size (N= 50000, # event=2480)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8019 (0.2057) | 0.7964 | 98.20% | 94.70% | 0.24% | 0.20 | 30.66% | 100.00% |
| Case−Crossover (1:2) | 0.7992 (0.1666) | 0.8008 | 99.90% | 94.00% | −0.10% | 0.16 | 46.69% | 100.00% |
| Case−Crossover (1:3) | 0.7957 (0.1508) | 0.7916 | 100.00% | 94.30% | −0.54% | 0.14 | 57.02% | 100.00% |
| Case−Crossover (1:5) | 0.7950 (0.1390) | 0.7957 | 100.00% | 93.80% | −0.62% | 0.13 | 67.07% | 100.00% |
| Case−Crossover (1:10) | 0.7947 (0.1292) | 0.7967 | 100.00% | 94.20% | −0.67% | 0.12 | 77.70% | 100.00% |
| Case−Crossover (1:25) | 0.7941 (0.1241) | 0.7967 | 100.00% | 94.20% | −0.73% | 0.12 | 84.22% | 100.00% |
| Cox Hazards Model | 0.7911 (0.1139) | 0.7934 | 100.00% | 94.90% | −1.11% | 0.11 | 100.00% | 100.00% |
| **Cohort Size (N=100000, # event=4963)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8024 (0.1410) | 0.7982 | 100.00% | 94.70% | 0.29% | 0.14 | 29.97% | 100.00% |
| Case−Crossover (1:2) | 0.8015 (0.1142) | 0.8011 | 100.00% | 93.80% | 0.19% | 0.11 | 45.67% | 100.00% |
| Case−Crossover (1:3) | 0.7988 (0.1024) | 0.8002 | 100.00% | 94.90% | −0.14% | 0.10 | 56.86% | 100.00% |
| Case−Crossover (1:5) | 0.7989 (0.0945) | 0.7985 | 100.00% | 94.30% | −0.14% | 0.09 | 66.66% | 100.00% |
| Case−Crossover (1:10) | 0.7985 (0.0885) | 0.7982 | 100.00% | 94.80% | −0.19% | 0.09 | 76.02% | 100.00% |
| Case−Crossover (1:25) | 0.7981 (0.0841) | 0.8002 | 100.00% | 95.40% | −0.24% | 0.08 | 84.14% | 100.00% |
| Cox Hazards Model | 0.7962 (0.0772) | 0.7992 | 100.00% | 94.70% | −0.47% | 0.08 | 100.00% | 100.00% |

Table 4.4: Results from a simulation of a cohort with an incidence of approximately 5% ( $\lambda = .00125$) incidence and assuming a transient effect lasting one time period ($\beta = 0.8$) (* Monte Carlo standard deviation of estimates ** Average of the information based standard error)

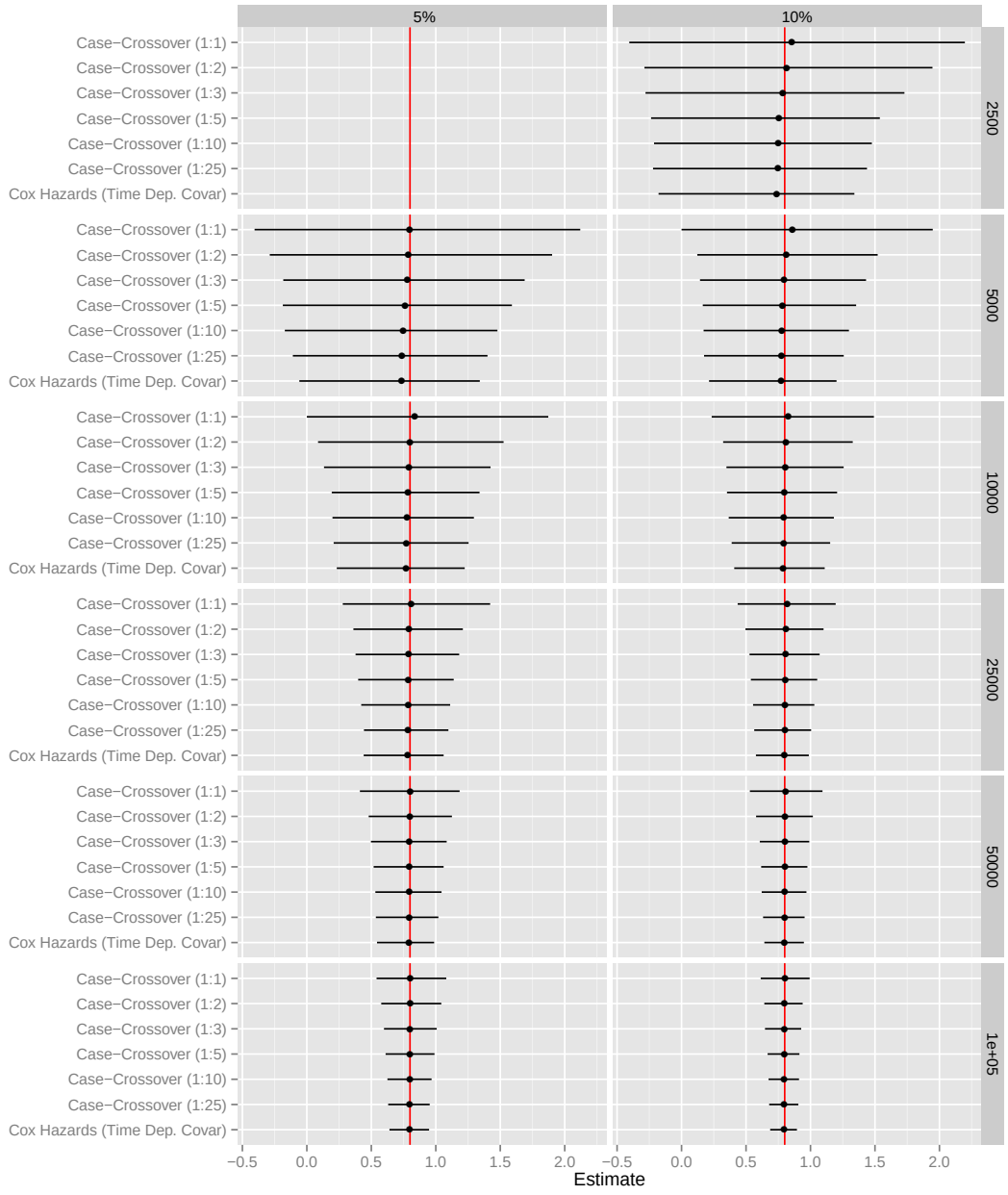


Figure 4.7: Plot of average estimates of log-hazard ratio with 95% percentile intervals by design, cohort size and incidence rate when assuming a log-hazard ratio of $\beta = 0.8$ indicated by the vertical red line.

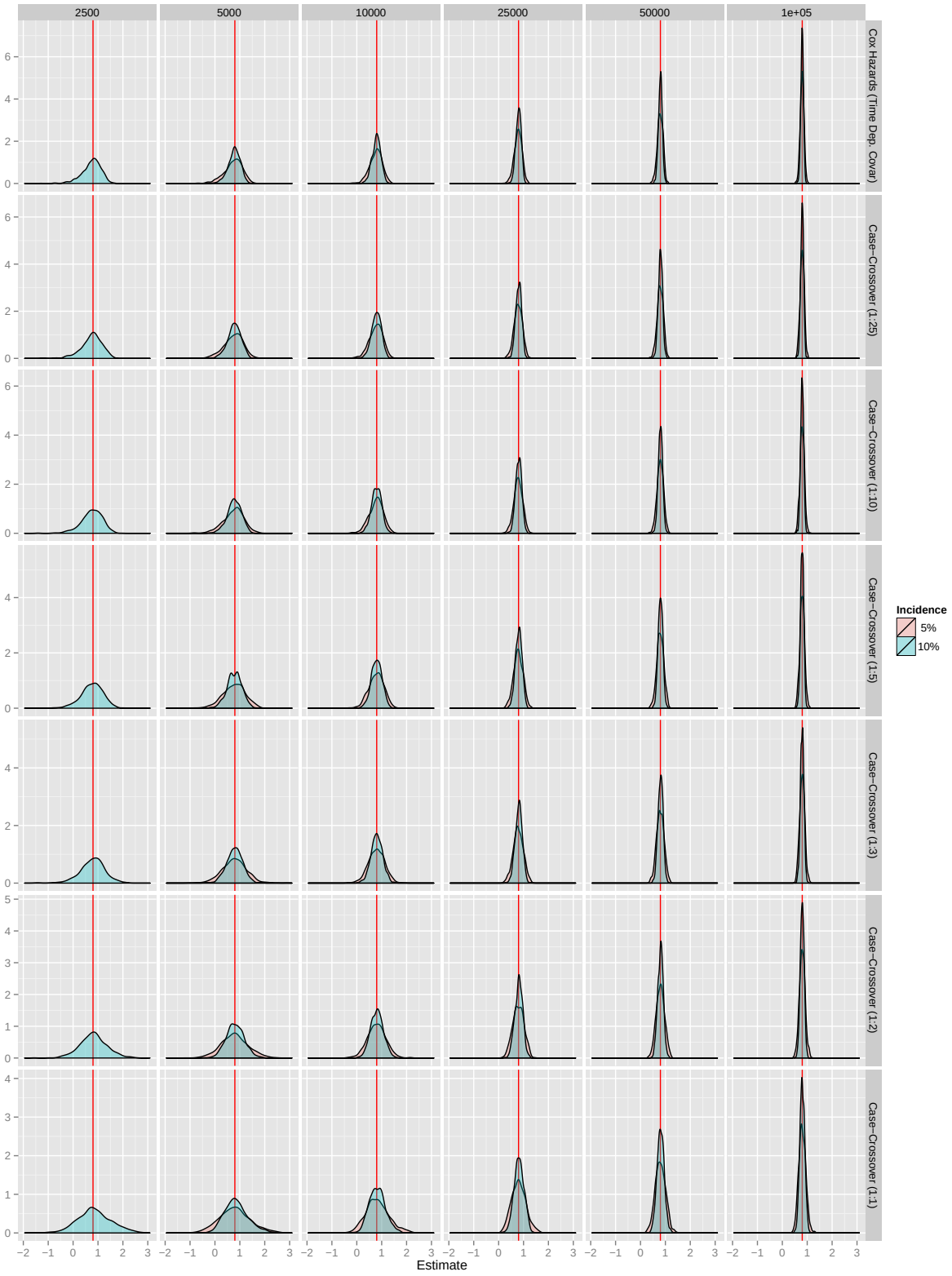


Figure 4.8: Distribution of estimate by design, cohort size and incidence rate when assuming a log-hazard of $\beta = 0.8$ indicated by the vertical red line.

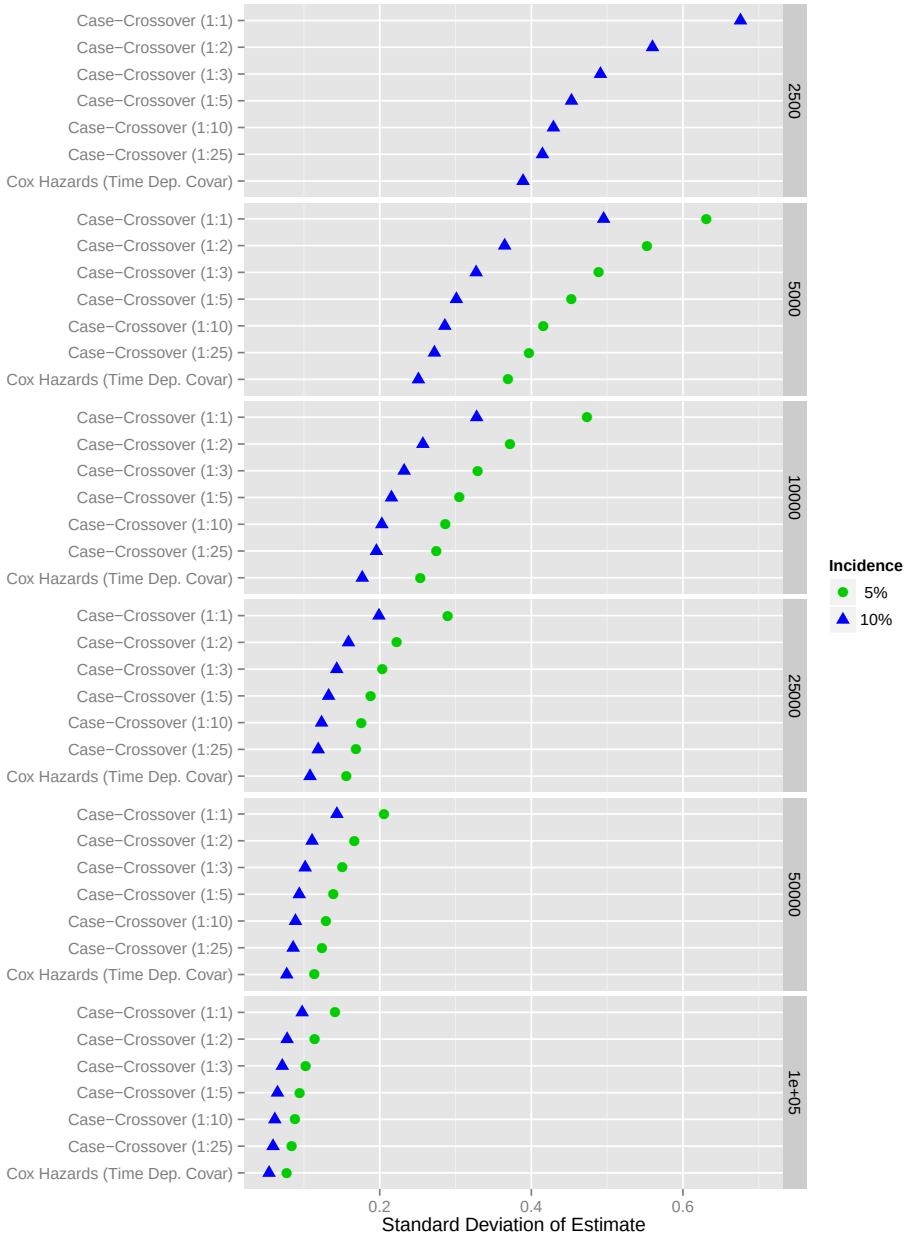


Figure 4.9: Plot of standard deviation of log-hazard ratio estimates by design, cohort size and incidence rate when assuming a log-hazard ratio of $\beta = 0.8$

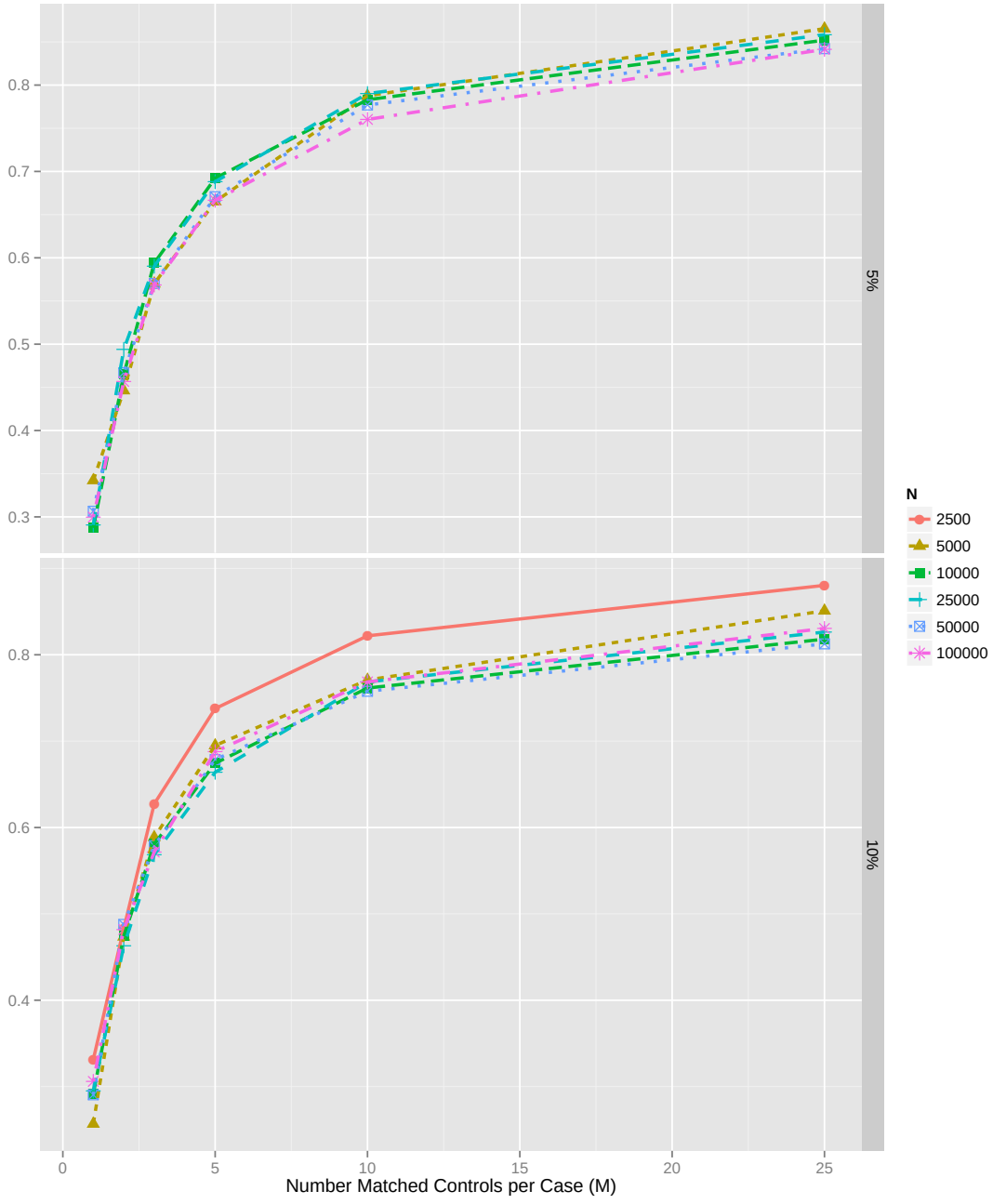


Figure 4.10: Plot of relative efficiency by design, cohort size and incidence rate when assuming a log-hazard ratio of $\beta = 0.8$

The average parameter estimate and respective 95% percentile intervals are plotted in Figure 4.11. Similar to subsection 4.1.2, percentile intervals that do not overlap zero again require the cohort to reach a size 5,000 for a 10% incidence rate and 10,000 for a 5% incidence rate. The 95% percentile intervals decrease with the larger cohort size, matched controls, and incidence rates. 1000 events are required to obtain excellent power for all designs except for the 1:1 case-crossover design which obtains more modest power. A similar relationship is observed in the average standard deviation of the log-hazard ratio estimates in Figure 4.13. The distribution of parameter estimates is displayed in Figure 4.12. The broad distribution of $\beta$ for the 1:1 case-crossover design appears to be right skewed for cohorts of smaller sample size. To a lesser extent, is also observed for 1:2 and 1:3 case-crossover designs.

The relative efficiency of the case-crossover design compared to the Cox proportional hazards model 1 is shown in Figure 4.14. The maximum relative efficiency observed is about 85% and is achieved by the 1:25 case-crossover design on a sample arising from a cohort with a 5% incidence rate. Coverage probabilities are also consistent with the results from subsection subsection 4.1.2.

| | Mean (SD*) | Median | Power | Coverage Probability | Bias | Avg Std Error** | Relative Efficiency | Converged |
|---|---|---|---|---|---|---|---|---|
| **Cohort Size (N= 2500, # event=243)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8056 (0.6687) | 0.6931 | 17.35% | 97.13% | 0.70% | 0.67 | 32.55% | 97.40% |
| Case−Crossover (1:2) | 0.7910 (0.5711) | 0.7968 | 33.23% | 96.10% | −1.12% | 0.54 | 44.63% | 99.90% |
| Case−Crossover (1:3) | 0.7687 (0.5104) | 0.7809 | 40.00% | 95.70% | −3.91% | 0.49 | 55.89% | 100.00% |
| Case−Crossover (1:5) | 0.7466 (0.4688) | 0.7890 | 45.40% | 95.50% | −6.68% | 0.44 | 66.23% | 100.00% |
| Case−Crossover (1:10) | 0.7365 (0.4447) | 0.7861 | 49.80% | 94.40% | −7.94% | 0.41 | 73.61% | 100.00% |
| Case−Crossover (1:25) | 0.7309 (0.4309) | 0.7657 | 50.20% | 94.50% | −8.64% | 0.40 | 78.41% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.7028 (0.1216) | 0.7033 | 100.00% | 96.30% | 0.41% | 0.13 | NA% | 100.00% |
| Z.it | 0.7312 (0.3820) | 0.7694 | 57.10% | 95.60% | −8.60% | 0.36 | 99.76% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.7307 (0.3815) | 0.7715 | 56.90% | 95.80% | −8.66% | 0.36 | 100.00% | 100.00% |
| **Cohort Size (N= 5000, # event=485)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8092 (0.4837) | 0.7732 | 41.44% | 95.80% | 1.15% | 0.46 | 28.68% | 99.90% |
| Case−Crossover (1:2) | 0.7826 (0.3832) | 0.7908 | 58.60% | 95.40% | −2.17% | 0.37 | 45.70% | 100.00% |
| Case−Crossover (1:3) | 0.7732 (0.3469) | 0.7885 | 65.20% | 95.00% | −3.35% | 0.33 | 55.77% | 100.00% |
| Case−Crossover (1:5) | 0.7673 (0.3155) | 0.7792 | 71.00% | 95.10% | −4.09% | 0.31 | 67.39% | 100.00% |
| Case−Crossover (1:10) | 0.7652 (0.2947) | 0.7825 | 74.70% | 95.20% | −4.35% | 0.29 | 77.26% | 100.00% |
| Case−Crossover (1:25) | 0.7601 (0.2856) | 0.7702 | 75.70% | 94.80% | −4.99% | 0.28 | 82.23% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.7017 (0.0900) | 0.7051 | 100.00% | 95.80% | 0.24% | 0.09 | NA% | 100.00% |
| Z.it | 0.7557 (0.2592) | 0.7715 | 82.70% | 95.30% | −5.54% | 0.25 | 99.88% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.7556 (0.2590) | 0.7696 | 82.70% | 95.40% | −5.55% | 0.25 | 100.00% | 100.00% |
| **Cohort Size (N= 10000, # event=970)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8068 (0.3352) | 0.7828 | 73.60% | 95.30% | 0.85% | 0.32 | 28.24% | 100.00% |
| Case−Crossover (1:2) | 0.7956 (0.2627) | 0.8028 | 88.10% | 95.20% | −0.55% | 0.26 | 45.98% | 100.00% |
| Case−Crossover (1:3) | 0.7905 (0.2385) | 0.7971 | 91.20% | 95.70% | −1.18% | 0.23 | 55.79% | 100.00% |
| Case−Crossover (1:5) | 0.7877 (0.2189) | 0.8002 | 93.00% | 95.20% | −1.54% | 0.21 | 66.21% | 100.00% |
| Case−Crossover (1:10) | 0.7876 (0.2036) | 0.7967 | 94.70% | 95.70% | −1.55% | 0.20 | 76.53% | 100.00% |
| Case−Crossover (1:25) | 0.7868 (0.1967) | 0.8002 | 95.60% | 95.00% | −1.65% | 0.19 | 82.04% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.7012 (0.0630) | 0.7005 | 100.00% | 96.10% | 0.17% | 0.06 | NA% | 100.00% |
| Z.it | 0.7818 (0.1782) | 0.7912 | 96.70% | 94.40% | −2.28% | 0.17 | 99.93% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.7816 (0.1781) | 0.7912 | 96.70% | 94.50% | −2.30% | 0.17 | 100.00% | 100.00% |
| **Cohort Size (N= 25000, # event=2424)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8048 (0.2075) | 0.7985 | 98.90% | 93.90% | 0.60% | 0.20 | 27.64% | 100.00% |
| Case−Crossover (1:2) | 0.8001 (0.1616) | 0.8028 | 99.90% | 94.90% | 0.01% | 0.16 | 45.54% | 100.00% |
| Case−Crossover (1:3) | 0.7977 (0.1505) | 0.8003 | 100.00% | 94.00% | −0.29% | 0.15 | 52.50% | 100.00% |
| Case−Crossover (1:5) | 0.7979 (0.1387) | 0.8001 | 100.00% | 94.50% | −0.27% | 0.13 | 61.80% | 100.00% |
| Case−Crossover (1:10) | 0.7980 (0.1275) | 0.8015 | 100.00% | 94.90% | −0.25% | 0.13 | 73.23% | 100.00% |
| Case−Crossover (1:25) | 0.7966 (0.1216) | 0.8005 | 100.00% | 95.00% | −0.42% | 0.12 | 80.45% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.7001 (0.0411) | 0.7001 | 100.00% | 94.40% | 0.01% | 0.04 | NA% | 100.00% |
| Z.it | 0.7920 (0.1091) | 0.7947 | 100.00% | 95.20% | −0.99% | 0.11 | 99.94% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.7920 (0.1091) | 0.7948 | 100.00% | 95.20% | −1.00% | 0.11 | 100.00% | 100.00% |
| **Cohort Size (N= 50000, # event=4850)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8014 (0.1453) | 0.7979 | 100.00% | 94.50% | 0.17% | 0.14 | 28.03% | 100.00% |
| Case−Crossover (1:2) | 0.7994 (0.1105) | 0.7955 | 100.00% | 95.30% | −0.07% | 0.11 | 48.46% | 100.00% |
| Case−Crossover (1:3) | 0.7981 (0.1026) | 0.7996 | 100.00% | 95.60% | −0.24% | 0.10 | 56.25% | 100.00% |
| Case−Crossover (1:5) | 0.7985 (0.0950) | 0.7988 | 100.00% | 95.00% | −0.19% | 0.09 | 65.53% | 100.00% |
| Case−Crossover (1:10) | 0.7997 (0.0888) | 0.7977 | 100.00% | 94.50% | −0.04% | 0.09 | 75.10% | 100.00% |
| Case−Crossover (1:25) | 0.7997 (0.0851) | 0.7979 | 100.00% | 95.30% | −0.04% | 0.09 | 81.71% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.6997 (0.0286) | 0.7000 | 100.00% | 96.10% | −0.04% | 0.03 | NA% | 100.00% |
| Z.it | 0.7968 (0.0770) | 0.7978 | 100.00% | 95.90% | −0.40% | 0.08 | 99.83% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.7967 (0.0769) | 0.7979 | 100.00% | 95.90% | −0.41% | 0.08 | 100.00% | 100.00% |
| **Cohort Size (N=100000, # event=9706)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8033 (0.1024) | 0.8021 | 100.00% | 94.60% | 0.42% | 0.10 | 29.60% | 100.00% |
| Case−Crossover (1:2) | 0.8014 (0.0815) | 0.8026 | 100.00% | 94.60% | 0.18% | 0.08 | 46.67% | 100.00% |
| Case−Crossover (1:3) | 0.8023 (0.0749) | 0.8009 | 100.00% | 94.70% | 0.29% | 0.07 | 55.22% | 100.00% |
| Case−Crossover (1:5) | 0.8022 (0.0682) | 0.8020 | 100.00% | 95.10% | 0.27% | 0.07 | 66.67% | 100.00% |
| Case−Crossover (1:10) | 0.8023 (0.0637) | 0.8023 | 100.00% | 94.40% | 0.28% | 0.06 | 76.47% | 100.00% |
| Case−Crossover (1:25) | 0.8022 (0.0616) | 0.8041 | 100.00% | 94.70% | 0.27% | 0.06 | 81.76% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.6997 (0.0204) | 0.6996 | 100.00% | 95.50% | −0.04% | 0.02 | NA% | 100.00% |
| Z.it | 0.7994 (0.0557) | 0.7993 | 100.00% | 94.30% | −0.08% | 0.05 | 99.92% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.7993 (0.0557) | 0.7996 | 100.00% | 94.20% | −0.09% | 0.05 | 100.00% | 100.00% |

Table 4.5: Results from simulation of a cohort with an incidence of approximately 10% ($\lambda = 0.0009$), where patient characteristic and exposure are independently associated with event, $\gamma = 0.7$ and $\beta = 0.8$ (* Monte Carlo standard deviation of estimates ** Average of the information based standard error)

| | Mean (SD*) | Median | Power | Coverage Probability | Bias | Avg Std Error** | Relative Efficiency | Converged |
|---|---|---|---|---|---|---|---|---|
| **Cohort Size (N= 5000, # event=249)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8344 (0.6773) | 0.8109 | 20.59% | 97.44% | 4.30% | 0.67 | 31.41% | 97.60% |
| Case−Crossover (1:2) | 0.8125 (0.5909) | 0.8109 | 37.44% | 95.20% | 1.56% | 0.53 | 41.26% | 99.90% |
| Case−Crossover (1:3) | 0.7934 (0.5445) | 0.7913 | 43.40% | 94.70% | −0.82% | 0.48 | 48.61% | 100.00% |
| Case−Crossover (1:5) | 0.7805 (0.4837) | 0.8176 | 50.30% | 93.90% | −2.44% | 0.44 | 61.59% | 100.00% |
| Case−Crossover (1:10) | 0.7627 (0.4461) | 0.8106 | 53.00% | 94.70% | −4.67% | 0.40 | 72.41% | 100.00% |
| Case−Crossover (1:25) | 0.7559 (0.4256) | 0.7991 | 55.00% | 95.00% | −5.51% | 0.39 | 79.53% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.7021 (0.1313) | 0.6991 | 100.00% | 94.50% | 0.29% | 0.13 | NA% | 100.00% |
| Z.it | 0.7526 (0.3795) | 0.7972 | 59.90% | 95.00% | −5.92% | 0.36 | 100.04% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.7524 (0.3796) | 0.7958 | 59.50% | 95.20% | −5.95% | 0.36 | 100.00% | 100.00% |
| **Cohort Size (N= 10000, # event=500)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8307 (0.4750) | 0.7985 | 44.90% | 95.20% | 3.83% | 0.45 | 27.75% | 100.00% |
| Case−Crossover (1:2) | 0.8061 (0.3746) | 0.8109 | 64.00% | 95.00% | 0.76% | 0.36 | 44.61% | 100.00% |
| Case−Crossover (1:3) | 0.8033 (0.3426) | 0.7963 | 69.70% | 95.00% | 0.41% | 0.33 | 53.34% | 100.00% |
| Case−Crossover (1:5) | 0.8031 (0.3118) | 0.8102 | 75.40% | 94.10% | 0.39% | 0.30 | 64.38% | 100.00% |
| Case−Crossover (1:10) | 0.7974 (0.2909) | 0.8117 | 79.40% | 93.50% | −0.32% | 0.28 | 73.98% | 100.00% |
| Case−Crossover (1:25) | 0.7935 (0.2782) | 0.8065 | 80.20% | 94.70% | −0.81% | 0.27 | 80.92% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.7006 (0.0888) | 0.7000 | 100.00% | 95.30% | 0.08% | 0.09 | NA% | 100.00% |
| Z.it | 0.7876 (0.2503) | 0.8097 | 84.40% | 95.20% | −1.55% | 0.24 | 99.97% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.7874 (0.2502) | 0.8083 | 84.60% | 95.20% | −1.58% | 0.24 | 100.00% | 100.00% |
| **Cohort Size (N= 25000, # event=1250)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8211 (0.2943) | 0.8109 | 85.60% | 94.50% | 2.64% | 0.28 | 28.47% | 100.00% |
| Case−Crossover (1:2) | 0.8057 (0.2320) | 0.8184 | 93.90% | 95.00% | 0.72% | 0.22 | 45.80% | 100.00% |
| Case−Crossover (1:3) | 0.8029 (0.2092) | 0.8075 | 96.00% | 95.00% | 0.37% | 0.20 | 56.32% | 100.00% |
| Case−Crossover (1:5) | 0.8033 (0.1908) | 0.8091 | 97.50% | 94.90% | 0.41% | 0.19 | 67.75% | 100.00% |
| Case−Crossover (1:10) | 0.7993 (0.1803) | 0.8074 | 98.60% | 94.10% | −0.09% | 0.17 | 75.86% | 100.00% |
| Case−Crossover (1:25) | 0.7964 (0.1732) | 0.8042 | 98.90% | 94.10% | −0.45% | 0.17 | 82.15% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.7002 (0.0556) | 0.6984 | 100.00% | 96.10% | 0.02% | 0.06 | NA% | 100.00% |
| Z.it | 0.7940 (0.1571) | 0.8055 | 99.70% | 95.10% | −0.75% | 0.15 | 99.90% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.7940 (0.1570) | 0.8057 | 99.70% | 95.00% | −0.76% | 0.15 | 100.00% | 100.00% |
| **Cohort Size (N= 50000, # event=2498)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8118 (0.2051) | 0.8011 | 99.00% | 94.90% | 1.48% | 0.20 | 28.18% | 100.00% |
| Case−Crossover (1:2) | 0.8022 (0.1627) | 0.7997 | 99.70% | 94.30% | 0.27% | 0.16 | 44.81% | 100.00% |
| Case−Crossover (1:3) | 0.8006 (0.1463) | 0.7997 | 99.90% | 95.40% | 0.07% | 0.14 | 55.38% | 100.00% |
| Case−Crossover (1:5) | 0.8010 (0.1316) | 0.8046 | 99.90% | 95.50% | 0.13% | 0.13 | 68.47% | 100.00% |
| Case−Crossover (1:10) | 0.7998 (0.1243) | 0.8034 | 100.00% | 94.90% | −0.03% | 0.12 | 76.68% | 100.00% |
| Case−Crossover (1:25) | 0.7979 (0.1200) | 0.8024 | 100.00% | 94.90% | −0.26% | 0.12 | 82.38% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.6997 (0.0394) | 0.6990 | 100.00% | 96.40% | −0.04% | 0.04 | NA% | 100.00% |
| Z.it | 0.7960 (0.1089) | 0.8018 | 100.00% | 95.10% | −0.50% | 0.11 | 100.03% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.7960 (0.1089) | 0.8019 | 100.00% | 95.10% | −0.50% | 0.11 | 100.00% | 100.00% |
| **Cohort Size (N=100000, # event=5000)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8011 (0.1387) | 0.8004 | 100.00% | 95.50% | 0.14% | 0.14 | 32.79% | 100.00% |
| Case−Crossover (1:2) | 0.7986 (0.1142) | 0.7976 | 100.00% | 95.00% | −0.17% | 0.11 | 48.39% | 100.00% |
| Case−Crossover (1:3) | 0.7971 (0.1048) | 0.7952 | 100.00% | 94.50% | −0.37% | 0.10 | 57.40% | 100.00% |
| Case−Crossover (1:5) | 0.7976 (0.0956) | 0.7969 | 100.00% | 94.30% | −0.30% | 0.09 | 68.98% | 100.00% |
| Case−Crossover (1:10) | 0.7971 (0.0901) | 0.7963 | 100.00% | 94.20% | −0.36% | 0.09 | 77.65% | 100.00% |
| Case−Crossover (1:25) | 0.7959 (0.0862) | 0.7985 | 100.00% | 95.10% | −0.51% | 0.08 | 84.86% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.6999 (0.0291) | 0.6991 | 100.00% | 94.50% | −0.01% | 0.03 | NA% | 100.00% |
| Z.it | 0.7955 (0.0795) | 0.7984 | 100.00% | 94.20% | −0.56% | 0.08 | 99.93% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.7955 (0.0794) | 0.7981 | 100.00% | 94.40% | −0.57% | 0.08 | 100.00% | 100.00% |

Table 4.6: Results from simulation of a cohort with an incidence of approximately 5% ($\lambda = 0.0009$), where patient characteristic and exposure are independently associated with event, $\gamma = 0.7$ and $\beta = 0.8$ (* Monte Carlo standard deviation of estimates ** Average of the information based standard error)

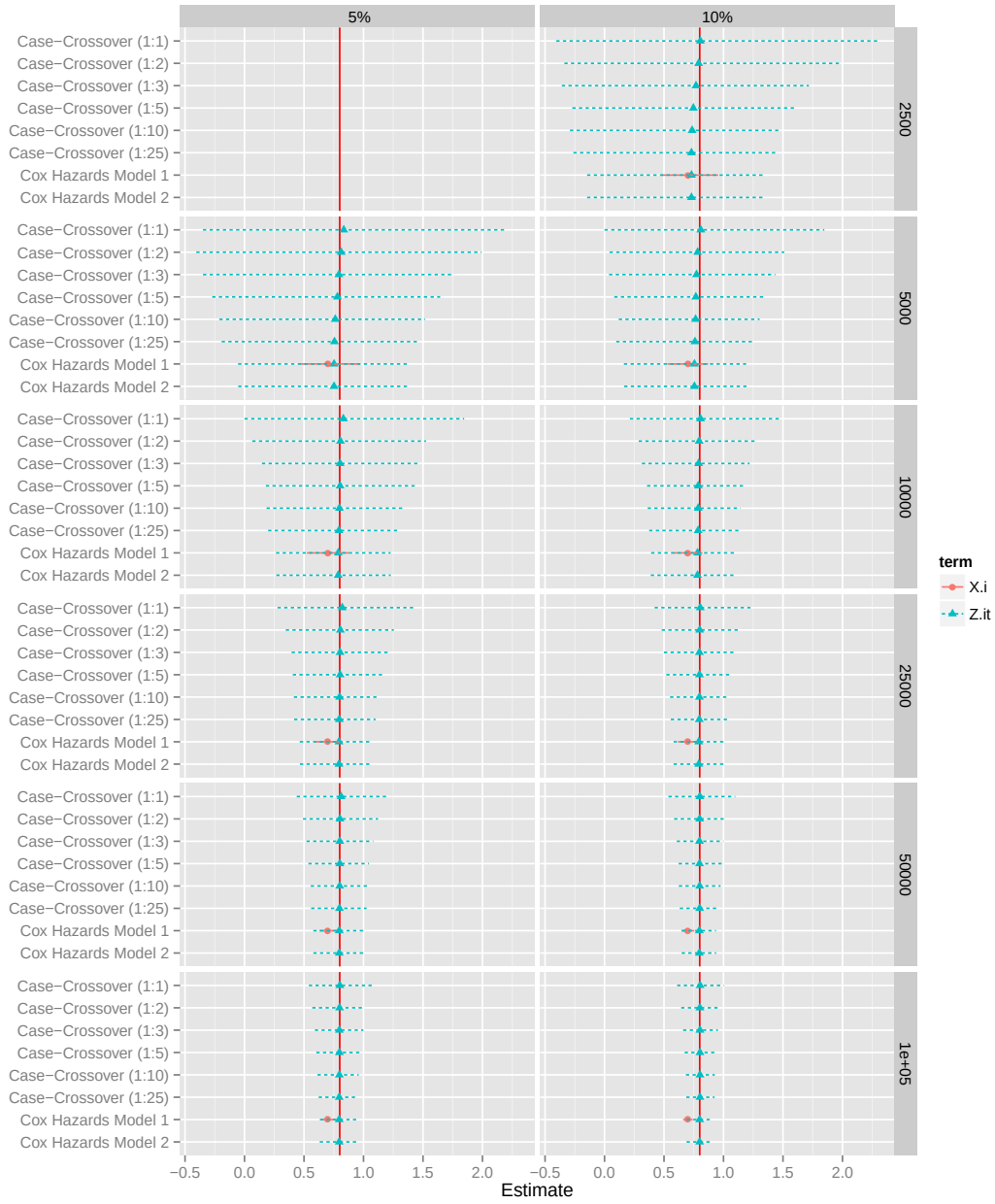


Figure 4.11: Plot of average estimates of log-hazard ratio with 95% percentile intervals by design, cohort size and incidence rate when assuming a log-hazard ratio of $\beta = 0.8$ indicated by the vertical red line.

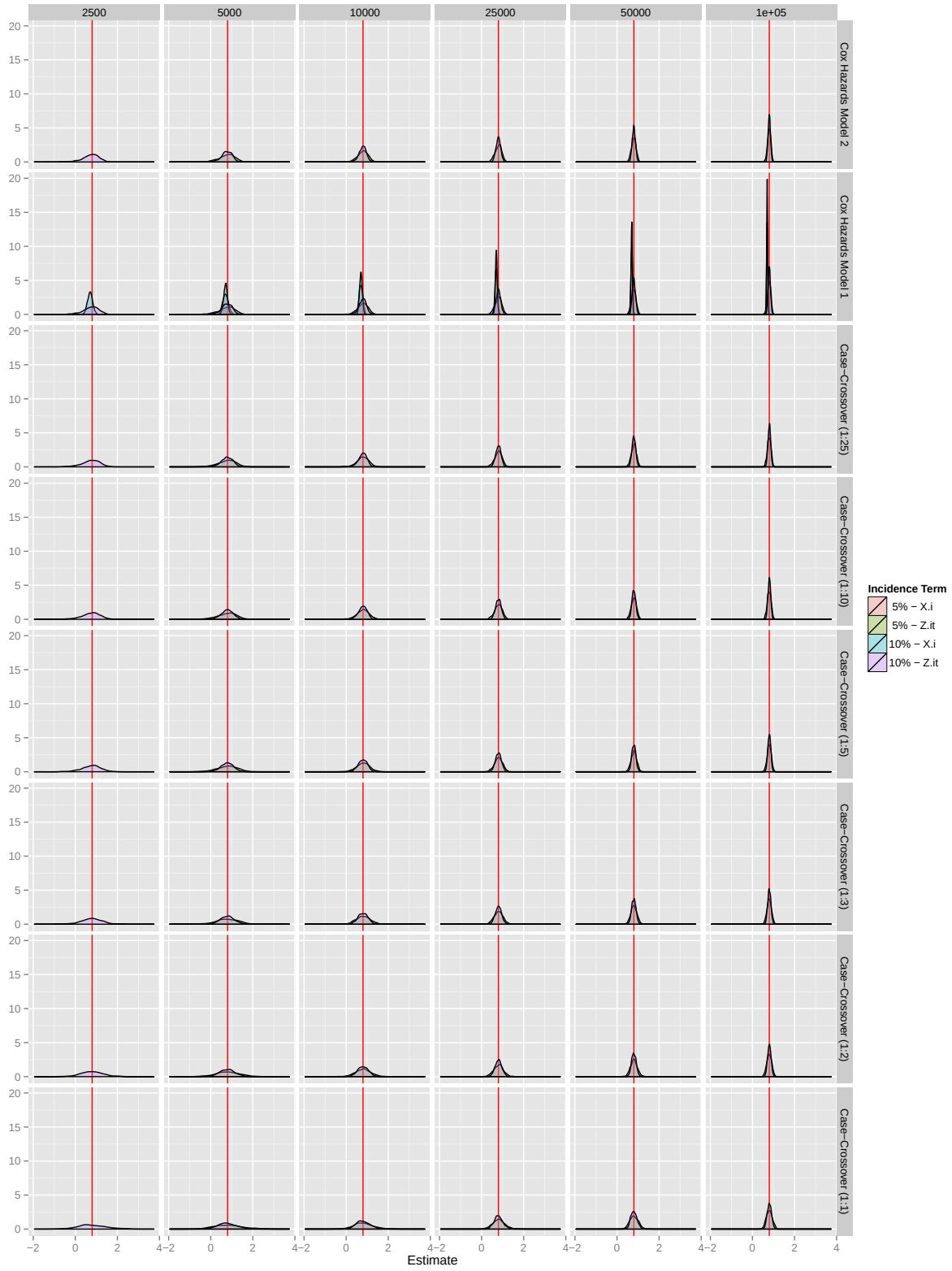


Figure 4.12: Distribution of estimate by design, cohort size and incidence rate when assuming a log-hazard ratio of $\beta = 0.8$ indicated by the vertical red line.

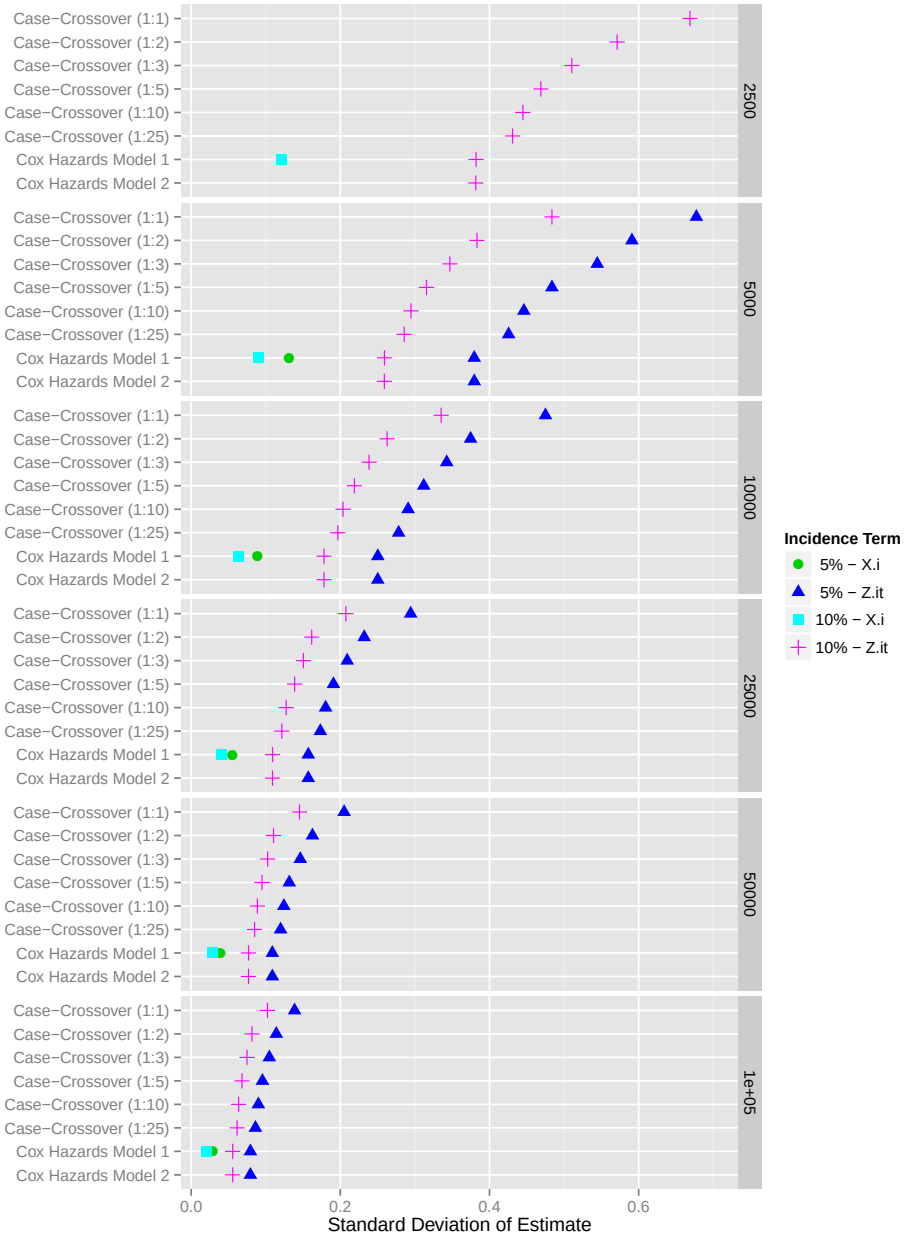


Figure 4.13: Plot of standard deviation of log-hazard ratio estimates by design, cohort size and incidence rate when assuming a log-hazard ratio of $\beta = 0.8$

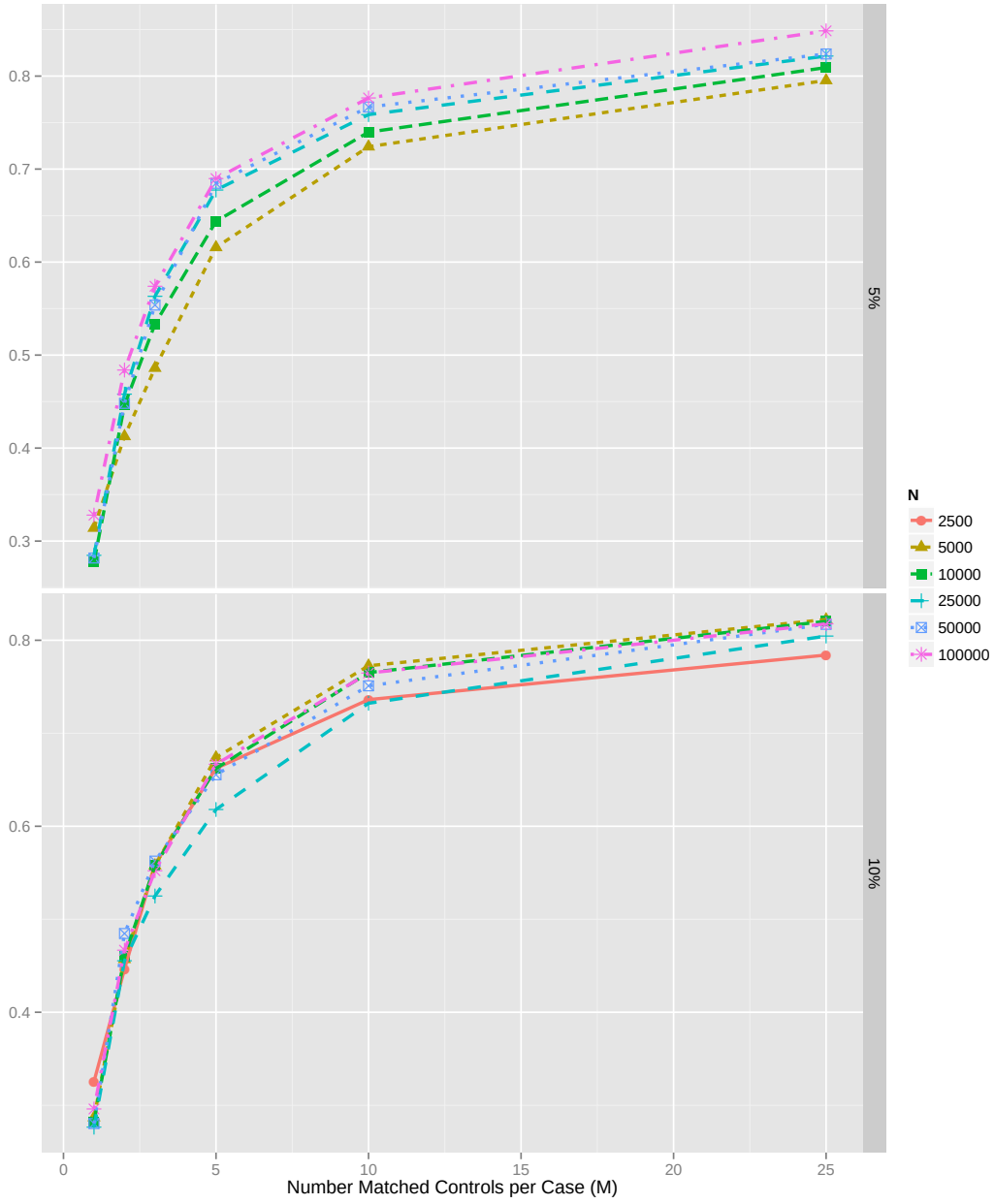


Figure 4.14: Plot of relative efficiency by design, cohort size and incidence rate when assuming a log-hazard ratio of $\beta = 0.8$

### 4.1.4 Exposure Has No Association with Event in Presence of Confounder

The simulations in this section assume that the time-dependent exposure has no effect, or $\beta = 0$, and is illustrated in Figure 4.1.d. Building on the simulations in subsections 4.1.1 and 4.1.3, the baseline covariate is assumed to be associated with the exposure and the event. The marginal probability of observing the change in status of the time-dependent exposure at any time is 0.4, or $P(Z_{i\cdot} = 1) = 0.4$, prior to taking into account the censoring mechanism. For patients with the presence of the baseline covariate, or $X_i = 1$, the probability of observing a status change in the time-dependent covariate at any time is 0.7, denoted $P(Z_{i\cdot} = 1 \mid X_i = 1) = 0.7$.

The results of the simulation studies for cohorts with 10% incidence rates is shown in Table 4.7 and the results for 5% incidence rates is in Table 4.8. The case-crossover design obtains average estimates of approximately 0. There is still some degree of over estimating $\beta$ for the 1:1 case-crossover design and underestimation for larger $M$. The Cox proportional hazards model that assumes the baseline covariate was measured (i.e. model 1) also obtains average estimates very close to the hypothesized value of zero for the time-dependent covariate. Whereas, the Cox proportional hazards model that does not observe the baseline covariate consistently obtains an average estimate ranging from 0.14-0.20; where the degree of bias increases with larger cohort sample sizes.

The average estimates and 95% percentile intervals are plotted in Figure 4.15. The percentile interval of estimates for the Cox proportional hazards model that does not observe the confounder in the cohort assuming a 10% incidence rate in 100000 observations is (-0.02, 0.40). The bias of the Cox hazard model that does not observe the confounder is further illustrated in the top row of Figure 4.16. The distribution of the log-hazard ratio shifts to the right as the sample size of the cohort increase. The distribution on the top, right for the cohort with sample size of 100,000 is almost entirely to the right of the $\beta = 0$ assumption indicated by the vertical red line.

The Type I error for the Cox proportional hazards model, for which the investigator is assumed to not be able to observe the confounder, incorrectly rejects the null hypothesis

in as little as 10.3% of the time in cohorts with 1000 events and as much as 49.3% of the simulations where 10000 events are observed. This means in a similar scenario where the number of events is about 10000, the investigator has a 50% chance of erroneously concluding a significant effect when there is truly no association in the presence of the confounder. The case-crossover design and the Cox proportional hazards model which does observe the confounder, yield Type I error rates around the hypothesized 5% level and largely obtain unbiased estimates of $\beta$.

The relative efficiency of the case-crossover design's estimate of $\beta$ compared to the true Cox proportional hazards model is plotted in Figure 4.18. The 1:5 case-crossover design achieves more than 80% efficiency relative to the true Cox proportional hazards model for all sample sizes with 10% incidence rates, and just less than 80% for 5% incidence rate. The maximum relative efficiency achieved was 92.2% by the 1:25 case-crossover design applied to the cohort of sample size 25,0000. A plot of the standard deviation for these simulations studies is presented in Figure 4.21. The standard deviations for these simulations is larger than any of the respective standard deviations observed in section 4.1.1.

| | Mean (SD*) | Median | Type I | Avg Std Error** | Relative Efficiency | Converged |
|---|---|---|---|---|---|---|
| **Cohort Size (N= 10000, # event=951)** | | | | | | |
| Case−Crossover (1:1) | 0.0055 (0.5290) | 0.0000 | 4.60% | 0.50 | 45.28% | 100.00% |
| Case−Crossover (1:2) | −0.0253 (0.4417) | −0.0264 | 3.60% | 0.43 | 64.96% | 100.00% |
| Case−Crossover (1:3) | −0.0392 (0.4107) | −0.0364 | 4.60% | 0.40 | 75.14% | 100.00% |
| Case−Crossover (1:5) | −0.0436 (0.3987) | −0.0097 | 4.20% | 0.38 | 79.73% | 100.00% |
| Case−Crossover (1:10) | −0.0471 (0.3853) | −0.0241 | 4.90% | 0.37 | 85.35% | 100.00% |
| Case−Crossover (1:25) | −0.0486 (0.3790) | −0.0230 | 4.70% | 0.36 | 88.25% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.7017 (0.0658) | 0.7014 | 100.00% | 0.07 | NA% | 100.00% |
| Z.it | −0.0568 (0.3560) | −0.0368 | 4.90% | 0.34 | 100.00% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.1428 (0.3555) | 0.1613 | 10.30% | 0.34 | 100.30% | 100.00% |
| **Cohort Size (N= 25000, # event=2378)** | | | | | | |
| Case−Crossover (1:1) | −0.0018 (0.3082) | 0.0000 | 4.70% | 0.30 | 52.46% | 100.00% |
| Case−Crossover (1:2) | −0.0128 (0.2621) | 0.0000 | 4.70% | 0.26 | 72.55% | 100.00% |
| Case−Crossover (1:3) | −0.0169 (0.2483) | −0.0088 | 4.70% | 0.25 | 80.82% | 100.00% |
| Case−Crossover (1:5) | −0.0208 (0.2408) | −0.0091 | 4.60% | 0.24 | 85.92% | 100.00% |
| Case−Crossover (1:10) | −0.0258 (0.2349) | −0.0151 | 5.00% | 0.23 | 90.31% | 100.00% |
| Case−Crossover (1:25) | −0.0261 (0.2325) | −0.0122 | 5.00% | 0.22 | 92.17% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.7010 (0.0402) | 0.6995 | 100.00% | 0.04 | NA% | 100.00% |
| Z.it | −0.0284 (0.2232) | −0.0208 | 5.70% | 0.21 | 100.00% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.1712 (0.2230) | 0.1803 | 17.80% | 0.21 | 100.21% | 100.00% |
| **Cohort Size (N= 50000, # event=4760)** | | | | | | |
| Case−Crossover (1:1) | 0.0029 (0.2253) | 0.0000 | 6.60% | 0.21 | 47.83% | 100.00% |
| Case−Crossover (1:2) | −0.0033 (0.1883) | 0.0000 | 4.80% | 0.18 | 68.45% | 100.00% |
| Case−Crossover (1:3) | −0.0061 (0.1787) | 0.0026 | 5.00% | 0.17 | 75.97% | 100.00% |
| Case−Crossover (1:5) | −0.0056 (0.1715) | 0.0028 | 5.40% | 0.17 | 82.51% | 100.00% |
| Case−Crossover (1:10) | −0.0091 (0.1658) | 0.0015 | 6.00% | 0.16 | 88.27% | 100.00% |
| Case−Crossover (1:25) | −0.0091 (0.1638) | 0.0018 | 6.20% | 0.16 | 90.43% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.7012 (0.0284) | 0.7010 | 100.00% | 0.03 | NA% | 100.00% |
| Z.it | −0.0097 (0.1558) | 0.0006 | 5.70% | 0.15 | 100.00% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.1899 (0.1559) | 0.2001 | 29.70% | 0.15 | 99.85% | 100.00% |
| **Cohort Size (N=100000, # event=9520)** | | | | | | |
| Case−Crossover (1:1) | −0.0008 (0.1570) | 0.0000 | 6.10% | 0.15 | 46.65% | 100.00% |
| Case−Crossover (1:2) | −0.0024 (0.1300) | 0.0000 | 4.80% | 0.13 | 68.01% | 100.00% |
| Case−Crossover (1:3) | −0.0034 (0.1248) | 0.0000 | 5.70% | 0.12 | 73.85% | 100.00% |
| Case−Crossover (1:5) | −0.0019 (0.1201) | 0.0002 | 5.50% | 0.12 | 79.68% | 100.00% |
| Case−Crossover (1:10) | −0.0044 (0.1155) | −0.0018 | 6.00% | 0.11 | 86.15% | 100.00% |
| Case−Crossover (1:25) | −0.0048 (0.1136) | −0.0016 | 5.80% | 0.11 | 89.10% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.7001 (0.0197) | 0.6997 | 100.00% | 0.02 | NA% | 100.00% |
| Z.it | −0.0031 (0.1072) | 0.0019 | 5.90% | 0.10 | 100.00% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.1962 (0.1070) | 0.2009 | 49.30% | 0.10 | 100.47% | 100.00% |

Table 4.7: results from simulation of a cohort with an incidence of approximately 10% ($\lambda = 0.0018$), where the patient characteristic is associated with the exposure and the event, but the event is independent of the exposure ($\gamma = .7$ and $\beta = 0$) (* Monte Carlo standard deviation of estimates ** Average of the information based standard error)

| | Mean (SD*) | Median | Type I | Avg Std Error** | Relative Efficiency | Converged |
|---|---|---|---|---|---|---|
| **Cohort Size (N= 25000, # event=1227)** | | | | | | |
| Case−Crossover (1:1) | 0.0052 (0.4447) | 0.0000 | 5.20% | 0.43 | 44.27% | 100.00% |
| Case−Crossover (1:2) | −0.0284 (0.3702) | −0.0264 | 4.20% | 0.37 | 63.87% | 100.00% |
| Case−Crossover (1:3) | −0.0293 (0.3513) | −0.0328 | 3.90% | 0.35 | 70.94% | 100.00% |
| Case−Crossover (1:5) | −0.0329 (0.3386) | −0.0301 | 4.50% | 0.33 | 76.36% | 100.00% |
| Case−Crossover (1:10) | −0.0359 (0.3229) | −0.0215 | 3.90% | 0.32 | 83.94% | 100.00% |
| Case−Crossover (1:25) | −0.0374 (0.3157) | −0.0286 | 4.70% | 0.32 | 87.81% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.6991 (0.0569) | 0.6997 | 100.00% | 0.06 | NA% | 100.00% |
| Z.it | −0.0367 (0.2959) | −0.0151 | 4.20% | 0.30 | 100.00% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.1616 (0.2947) | 0.1809 | 13.20% | 0.30 | 100.76% | 100.00% |
| **Cohort Size (N= 50000, # event=2455)** | | | | | | |
| Case−Crossover (1:1) | 0.0085 (0.3006) | 0.0000 | 4.00% | 0.30 | 46.85% | 100.00% |
| Case−Crossover (1:2) | −0.0109 (0.2533) | −0.0114 | 3.80% | 0.26 | 65.96% | 100.00% |
| Case−Crossover (1:3) | −0.0113 (0.2433) | −0.0045 | 4.70% | 0.24 | 71.51% | 100.00% |
| Case−Crossover (1:5) | −0.0144 (0.2320) | −0.0055 | 3.70% | 0.23 | 78.62% | 100.00% |
| Case−Crossover (1:10) | −0.0159 (0.2229) | −0.0111 | 4.50% | 0.22 | 85.18% | 100.00% |
| Case−Crossover (1:25) | −0.0146 (0.2182) | −0.0123 | 4.50% | 0.22 | 88.86% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.6999 (0.0409) | 0.6996 | 100.00% | 0.04 | NA% | 100.00% |
| Z.it | −0.0131 (0.2057) | −0.0049 | 4.20% | 0.21 | 100.00% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.1854 (0.2045) | 0.1942 | 19.10% | 0.21 | 101.16% | 100.00% |
| **Cohort Size (N=100000, # event=4910)** | | | | | | |
| Case−Crossover (1:1) | 0.0113 (0.2107) | 0.0000 | 5.00% | 0.21 | 48.99% | 100.00% |
| Case−Crossover (1:2) | 0.0006 (0.1825) | 0.0000 | 4.70% | 0.18 | 65.28% | 100.00% |
| Case−Crossover (1:3) | 0.0000 (0.1733) | 0.0089 | 5.30% | 0.17 | 72.40% | 100.00% |
| Case−Crossover (1:5) | −0.0013 (0.1671) | 0.0043 | 5.10% | 0.16 | 77.85% | 100.00% |
| Case−Crossover (1:10) | −0.0031 (0.1591) | 0.0060 | 5.20% | 0.16 | 85.84% | 100.00% |
| Case−Crossover (1:25) | −0.0029 (0.1564) | 0.0051 | 4.90% | 0.15 | 88.91% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.7007 (0.0286) | 0.7006 | 100.00% | 0.03 | NA% | 100.00% |
| Z.it | −0.0037 (0.1475) | 0.0027 | 6.40% | 0.14 | 100.00% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.1949 (0.1469) | 0.2011 | 30.30% | 0.14 | 100.80% | 100.00% |

Table 4.8: Results from simulation of a cohort with an incidence of approximately 5% ($\lambda = 0.0009$), where the patient characteristic is associated with the exposure and the event, but the event is independent of the exposure ($\gamma = .7$ and $\beta = 0$) (* Monte Carlo standard deviation of estimates ** Average of the information based standard error)

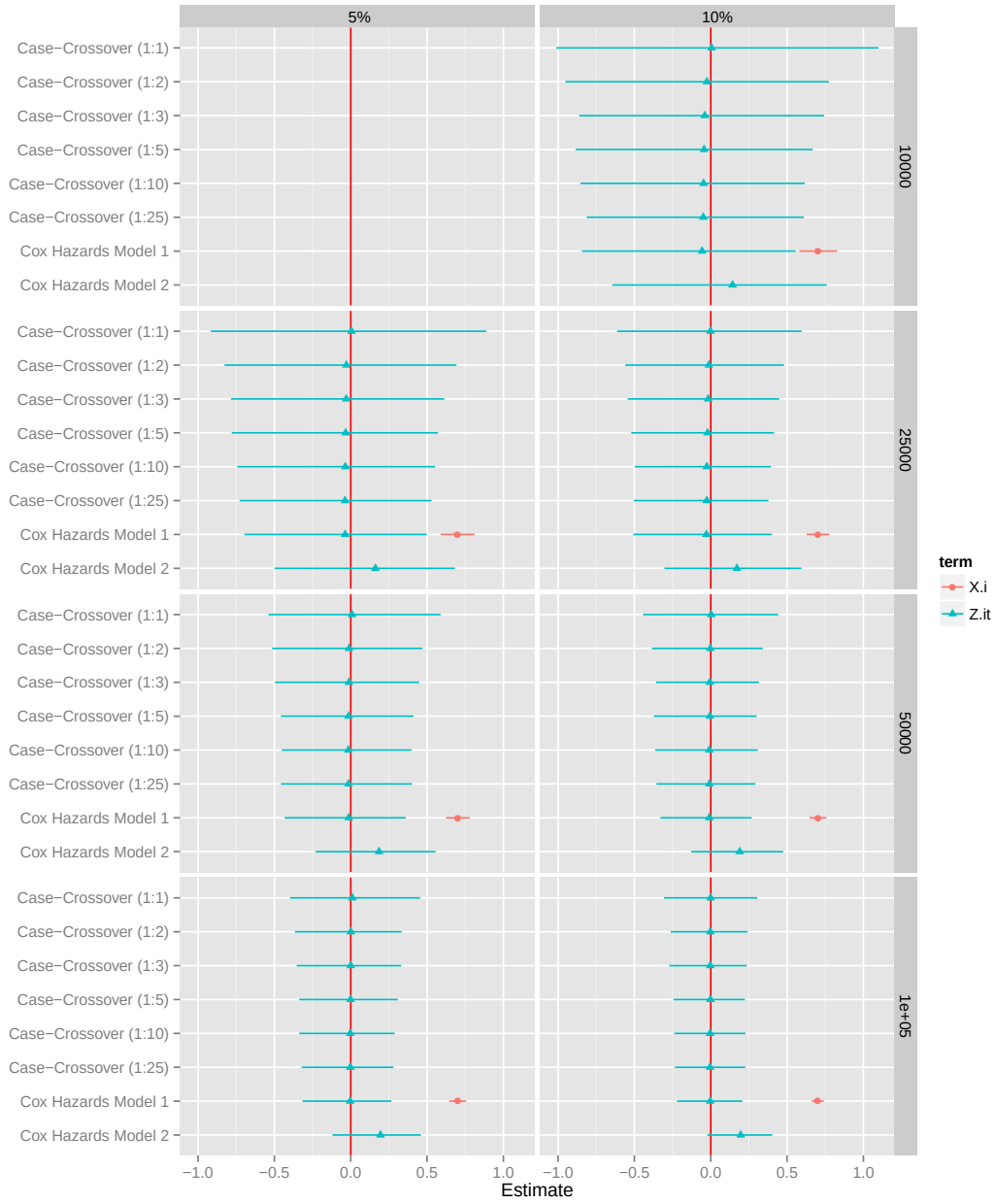


Figure 4.15: Plot of average estimates of the log-hazard ratio with 95% percentile intervals by design, cohort size and incidence rate when assuming a log-hazard ratio of $\beta = 0$ indicated by the vertical red line.

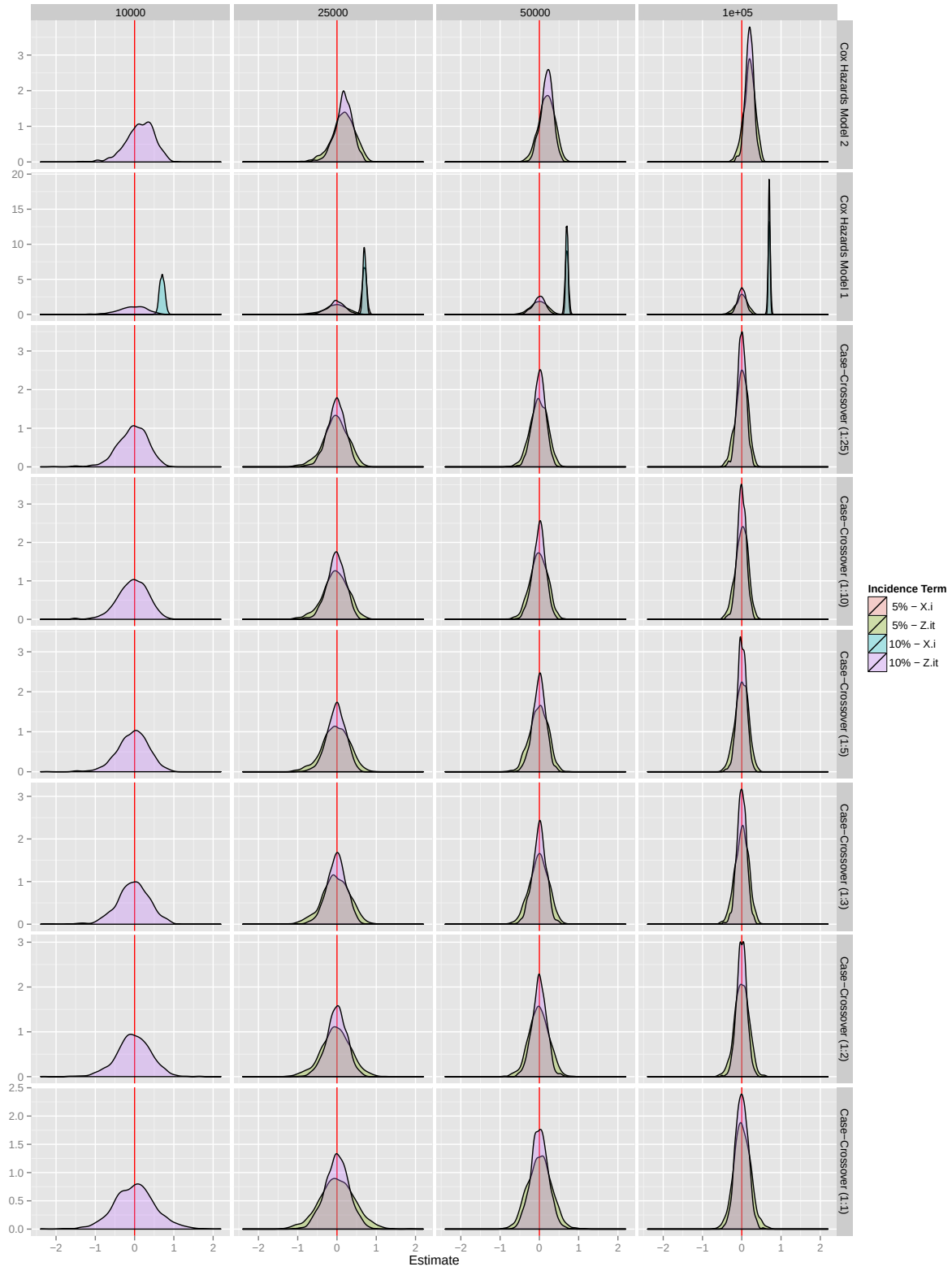


Figure 4.16: Distribution of estimate by design, cohort size and incidence rate when assuming a log-hazard ratio of $\beta = 0$ indicated by the vertical red line.

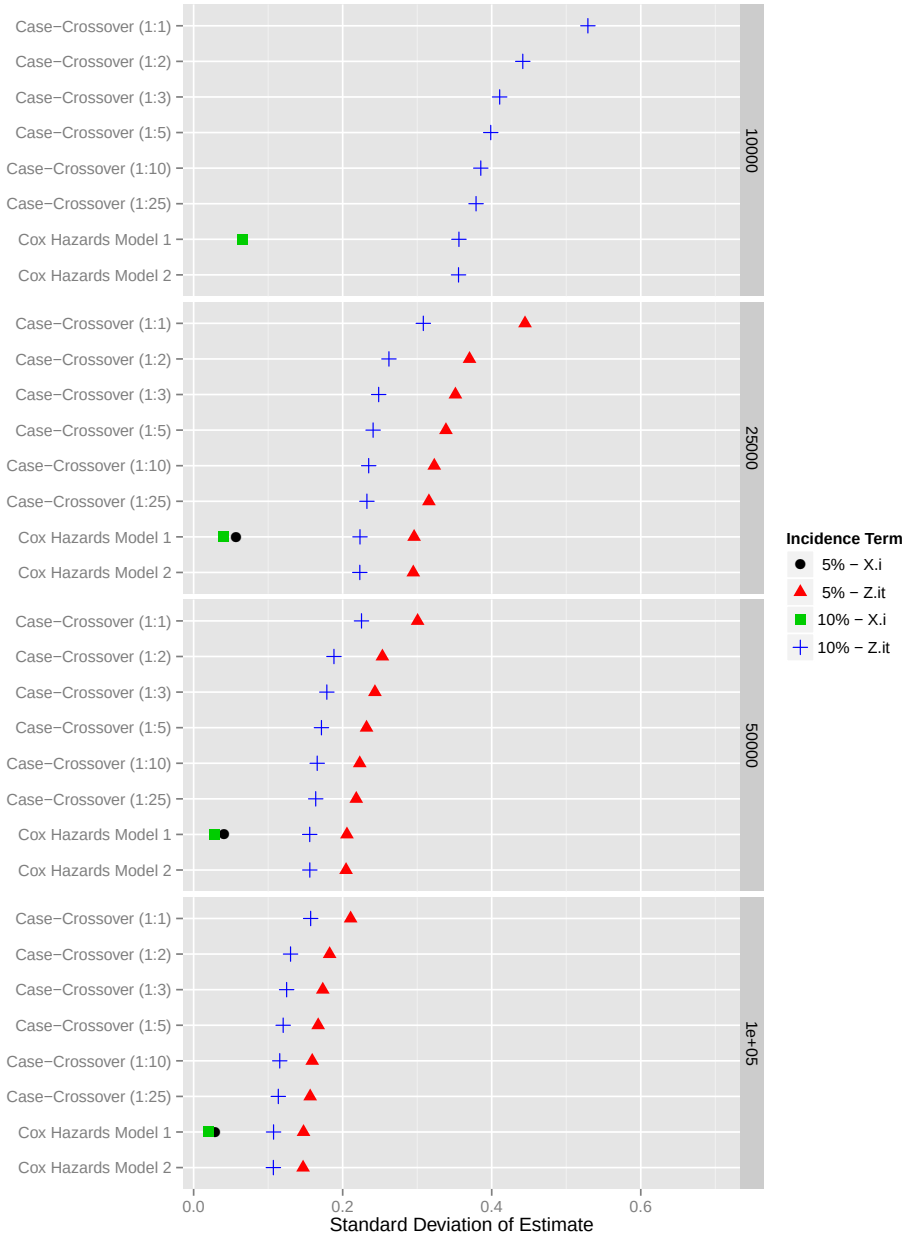


Figure 4.17: Plot of standard deviation of log-hazard ratio estimates by design, cohort size and incidence rate when assuming a log-hazard ratio of $\beta = .0$

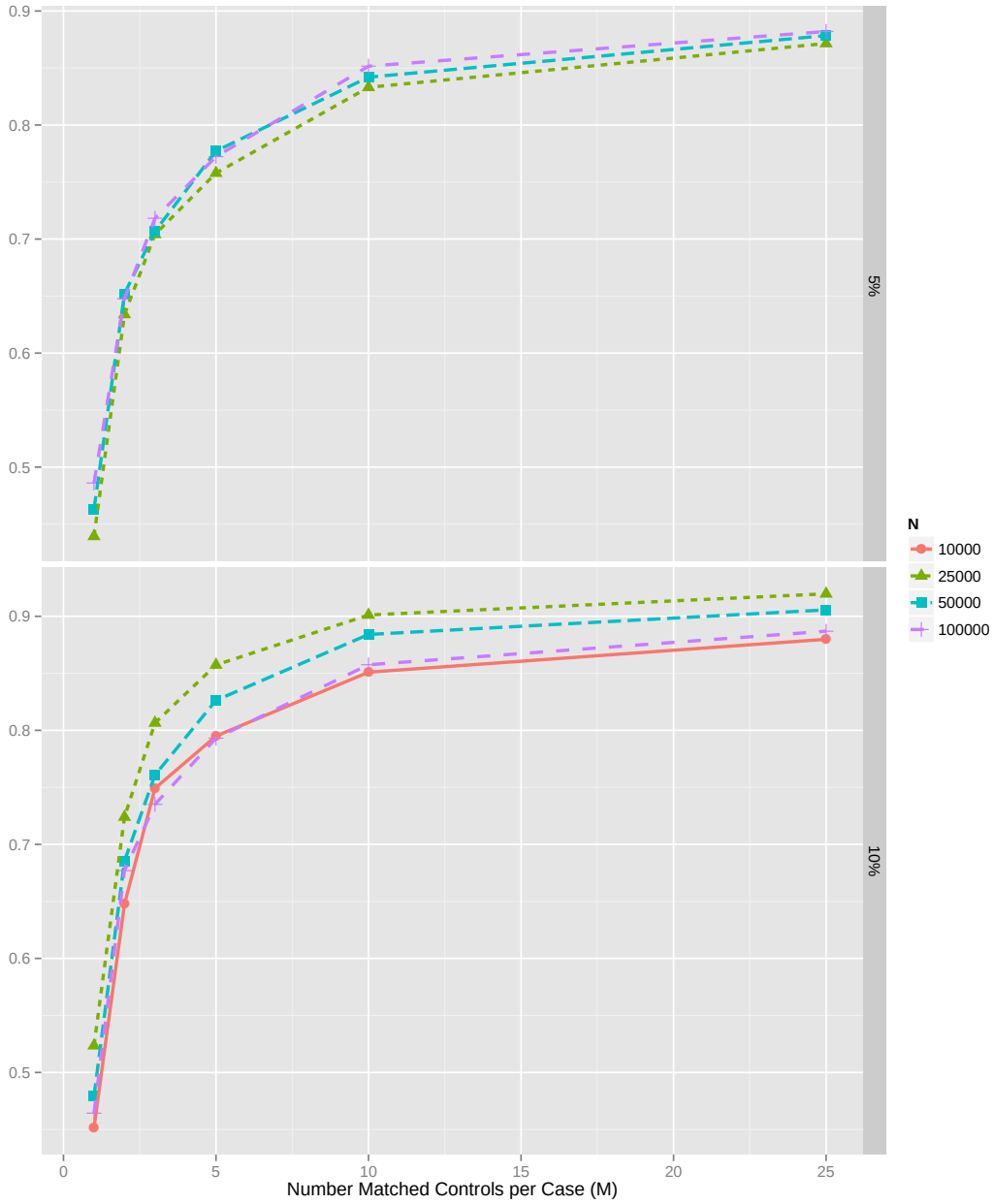


Figure 4.18: Plot of relative efficiency by design, cohort size and incidence rate when assuming a log-hazard ratio of $\beta = 0$

### 4.1.5 Exposure Has Positive Association in Presence of a Confounder

The final scenario, highlighted in Figure 4.1.e, assumes that the time-dependent exposure increases the risk of observing the event in the presence of a confounder by a log-hazard ratio of 0.8. Similar to section 4.1.3, the simulations in this scenario assume $P(Z_{i\cdot} = 0.4$ and $P(Z_{i\cdot} = 1 \mid X_i = 1) = 0.7$. The results of the simulations in this subsection are displayed in Table 4.9 and 4.10 for cohorts with incidence rates of 10% and 5%, respectively.

The average parameter estimates of the log-hazard ratio for the time-dependent exposure are generally close to the assumed effect of $\beta = 0.8$ for the case-crossover designs. Similar patterns of bias exist for the case-crossover design in smaller cohorts; the 1:1 design over estimates the log-hazard ratio while 1:10 and 1:25 underestimate. The Cox proportional hazards model which assumes to have observed the baseline confounder still underestimates the log-hazard ratio for $\beta$ by almost 7% for cohorts of size 5,000 with a 10% incidence ratio and of size 10,000 with a 5% incidence ratio. The Cox proportional hazards model that does not observe the baseline confounder over estimates $\beta$, with average estimates ranging from 0.94-1.00 for the various sample size. This is likely the justification for the lower coverage probabilities since the distribution of estimates for $\beta$ have been shifted up by the bias introduced by the unobserved confounder.

Figure 4.19 contains a plot of the average log-hazard ratio estimates with corresponding 95% percentile intervals. The degree of which not controlling for the confounder in Cox proportional hazards model 2 is apparent in that the 95% confidence interval does not contain the hypothesized value for cohorts of size 100,000 with 5% incidence rates and greater than 50,000 for cohorts with 10% incidence rates. The entire distribution of $\beta$ when the confounder is not observed are shown to lie almost entirely above the hypothesized value shown in the top right of Figure 4.20.

Compared to the simulations presented in sections 4.1.1-4.1.4, the relative efficiency seems to be generally smaller. This is highlighted by the 1:5 case-crossover design never exceeding 70% relative efficiency. A plot of relative efficiency is displayed in Figure 4.19. The maximum relative efficiency observed was 86% for the 1:25 case-crossover design applied to cohorts of size 100,000 with 5% incidence rate. The standard deviation of these

estimates are shown in Figure 4.21, and follow the same relationships already observed in previous simulations. The standard deviation decreases as the sample size, number of matched controls and incidence rate increase.

| | Mean (SD*) | Median | Power | Coverage Probability | Bias | Avg Std Error** | Relative Efficiency | Converged |
|---|---|---|---|---|---|---|---|---|
| **Cohort Size (N= 5000, # event=482)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8291 (0.6290) | 0.8109 | 23.24% | 97.48% | 3.64% | 0.61 | 32.18% | 99.40% |
| Case−Crossover (1:2) | 0.7940 (0.5153) | 0.8109 | 42.24% | 96.70% | −0.75% | 0.49 | 47.96% | 99.90% |
| Case−Crossover (1:3) | 0.7824 (0.4697) | 0.8145 | 49.10% | 95.20% | −2.19% | 0.44 | 57.71% | 100.00% |
| Case−Crossover (1:5) | 0.7602 (0.4316) | 0.7880 | 52.40% | 94.20% | −4.98% | 0.40 | 68.37% | 100.00% |
| Case−Crossover (1:10) | 0.7594 (0.4079) | 0.7986 | 59.50% | 94.70% | −5.07% | 0.37 | 76.55% | 100.00% |
| Case−Crossover (1:25) | 0.7573 (0.3953) | 0.8134 | 61.40% | 94.20% | −5.34% | 0.36 | 81.50% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.6996 (0.0896) | 0.7037 | 100.00% | 95.50% | −0.06% | 0.09 | NA% | 100.00% |
| Z.it | 0.7448 (0.3577) | 0.8045 | 65.70% | 94.70% | −6.90% | 0.33 | 99.55% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.9437 (0.3569) | 0.9989 | 78.30% | 87.50% | 17.96% | 0.33 | 100.00% | 100.00% |
| **Cohort Size (N= 10000, # event=963)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8060 (0.4342) | 0.7621 | 49.90% | 96.30% | 0.74% | 0.42 | 27.45% | 100.00% |
| Case−Crossover (1:2) | 0.7872 (0.3463) | 0.7800 | 66.10% | 96.00% | −1.60% | 0.33 | 43.14% | 100.00% |
| Case−Crossover (1:3) | 0.7842 (0.3132) | 0.7804 | 72.10% | 95.70% | −1.97% | 0.30 | 52.75% | 100.00% |
| Case−Crossover (1:5) | 0.7751 (0.2795) | 0.7939 | 77.20% | 95.80% | −3.11% | 0.28 | 66.22% | 100.00% |
| Case−Crossover (1:10) | 0.7773 (0.2611) | 0.7915 | 80.30% | 95.60% | −2.84% | 0.26 | 75.92% | 100.00% |
| Case−Crossover (1:25) | 0.7792 (0.2526) | 0.8031 | 83.40% | 96.00% | −2.60% | 0.25 | 81.10% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.6990 (0.0654) | 0.6983 | 100.00% | 95.40% | −0.14% | 0.07 | NA% | 100.00% |
| Z.it | 0.7708 (0.2285) | 0.7876 | 87.70% | 95.60% | −3.65% | 0.23 | 99.08% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.9698 (0.2275) | 0.9848 | 95.60% | 83.80% | 21.22% | 0.23 | 100.00% | 100.00% |
| **Cohort Size (N= 25000, # event=2411)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8031 (0.2682) | 0.7885 | 88.40% | 93.90% | 0.39% | 0.26 | 28.71% | 100.00% |
| Case−Crossover (1:2) | 0.7977 (0.2102) | 0.7950 | 97.60% | 96.10% | −0.28% | 0.21 | 46.75% | 100.00% |
| Case−Crossover (1:3) | 0.7982 (0.1912) | 0.7970 | 98.40% | 95.70% | −0.22% | 0.19 | 56.53% | 100.00% |
| Case−Crossover (1:5) | 0.7937 (0.1719) | 0.7957 | 99.00% | 95.60% | −0.79% | 0.17 | 69.92% | 100.00% |
| Case−Crossover (1:10) | 0.7963 (0.1628) | 0.8005 | 99.20% | 95.30% | −0.46% | 0.16 | 77.96% | 100.00% |
| Case−Crossover (1:25) | 0.7958 (0.1570) | 0.7977 | 99.70% | 95.20% | −0.53% | 0.16 | 83.76% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.6998 (0.0407) | 0.7008 | 100.00% | 95.10% | −0.03% | 0.04 | NA% | 100.00% |
| Z.it | 0.7886 (0.1448) | 0.7948 | 99.70% | 94.90% | −1.42% | 0.14 | 98.53% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.9878 (0.1437) | 0.9906 | 100.00% | 69.30% | 23.48% | 0.14 | 100.00% | 100.00% |
| **Cohort Size (N= 50000, # event=4819)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8073 (0.1875) | 0.7961 | 99.80% | 94.90% | 0.91% | 0.18 | 28.47% | 100.00% |
| Case−Crossover (1:2) | 0.8075 (0.1457) | 0.8079 | 100.00% | 94.90% | 0.94% | 0.15 | 47.13% | 100.00% |
| Case−Crossover (1:3) | 0.8061 (0.1315) | 0.8072 | 100.00% | 95.70% | 0.76% | 0.13 | 57.89% | 100.00% |
| Case−Crossover (1:5) | 0.8028 (0.1199) | 0.8052 | 100.00% | 96.20% | 0.35% | 0.12 | 69.54% | 100.00% |
| Case−Crossover (1:10) | 0.8042 (0.1141) | 0.8087 | 100.00% | 95.60% | 0.53% | 0.11 | 76.86% | 100.00% |
| Case−Crossover (1:25) | 0.8027 (0.1101) | 0.8082 | 100.00% | 95.50% | 0.33% | 0.11 | 82.59% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.6992 (0.0290) | 0.6998 | 100.00% | 94.40% | −0.11% | 0.03 | NA% | 100.00% |
| Z.it | 0.7987 (0.1003) | 0.8064 | 100.00% | 96.00% | −0.16% | 0.10 | 99.50% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.9978 (0.1000) | 1.0043 | 100.00% | 45.80% | 24.73% | 0.10 | 100.00% | 100.00% |
| **Cohort Size (N=100000, # event=9632)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8026 (0.1275) | 0.7982 | 100.00% | 95.30% | 0.33% | 0.13 | 29.29% | 100.00% |
| Case−Crossover (1:2) | 0.8026 (0.1026) | 0.8051 | 100.00% | 95.70% | 0.32% | 0.10 | 45.26% | 100.00% |
| Case−Crossover (1:3) | 0.8025 (0.0917) | 0.8043 | 100.00% | 95.80% | 0.31% | 0.09 | 56.60% | 100.00% |
| Case−Crossover (1:5) | 0.8003 (0.0855) | 0.8042 | 100.00% | 95.70% | 0.04% | 0.09 | 65.22% | 100.00% |
| Case−Crossover (1:10) | 0.8010 (0.0797) | 0.8050 | 100.00% | 95.60% | 0.12% | 0.08 | 75.01% | 100.00% |
| Case−Crossover (1:25) | 0.8003 (0.0773) | 0.8025 | 100.00% | 95.40% | 0.03% | 0.08 | 79.76% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.7002 (0.0198) | 0.7006 | 100.00% | 95.70% | 0.03% | 0.02 | NA% | 100.00% |
| Z.it | 0.7978 (0.0694) | 0.8000 | 100.00% | 95.50% | −0.28% | 0.07 | 98.91% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.9972 (0.0690) | 1.0005 | 100.00% | 20.70% | 24.65% | 0.07 | 100.00% | 100.00% |

Table 4.9: Results from simulation of a cohort with an incidence of approximately 10% ($\lambda = 0.0018$), where the patient characteristic is confounding with the exposure and the event ($\gamma = .7$ and $\beta = 0.8$), and $P(Z_{i\cdot} = 1) = .4$ and $P(Z_{i\cdot} = 1 \mid X_i = 1) = .7$ (* Monte Carlo standard deviation of estimates ** Average of the information based standard error)

| | Mean (SD*) | Median | Power | Coverage Probability | Bias | Avg Std Error** | Relative Efficiency | Converged |
|---|---|---|---|---|---|---|---|---|
| **Cohort Size (N= 10000, # event=496)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8462 (0.6132) | 0.8109 | 26.63% | 97.29% | 5.77% | 0.61 | 30.82% | 99.50% |
| Case−Crossover (1:2) | 0.8012 (0.5046) | 0.7885 | 41.80% | 95.60% | 0.15% | 0.48 | 45.52% | 100.00% |
| Case−Crossover (1:3) | 0.7712 (0.4530) | 0.7945 | 48.70% | 95.70% | −3.60% | 0.43 | 56.49% | 100.00% |
| Case−Crossover (1:5) | 0.7588 (0.4173) | 0.7919 | 53.80% | 95.50% | −5.15% | 0.39 | 66.56% | 100.00% |
| Case−Crossover (1:10) | 0.7547 (0.3894) | 0.7839 | 57.30% | 95.70% | −5.67% | 0.37 | 76.46% | 100.00% |
| Case−Crossover (1:25) | 0.7517 (0.3751) | 0.7939 | 60.90% | 95.20% | −6.03% | 0.35 | 82.39% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.6941 (0.0925) | 0.6953 | 100.00% | 94.60% | −0.85% | 0.09 | NA% | 100.00% |
| Z.it | 0.7464 (0.3415) | 0.7942 | 65.30% | 95.00% | −6.70% | 0.32 | 99.40% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.9432 (0.3405) | 0.9845 | 79.00% | 88.40% | 17.90% | 0.32 | 100.00% | 100.00% |
| **Cohort Size (N= 25000, # event=1240)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8217 (0.3797) | 0.8023 | 64.50% | 94.80% | 2.72% | 0.36 | 29.27% | 100.00% |
| Case−Crossover (1:2) | 0.8031 (0.3079) | 0.8044 | 78.30% | 93.80% | 0.39% | 0.29 | 44.51% | 100.00% |
| Case−Crossover (1:3) | 0.7953 (0.2803) | 0.8034 | 82.70% | 94.40% | −0.59% | 0.27 | 53.70% | 100.00% |
| Case−Crossover (1:5) | 0.7896 (0.2613) | 0.8133 | 87.00% | 93.50% | −1.30% | 0.24 | 61.82% | 100.00% |
| Case−Crossover (1:10) | 0.7886 (0.2395) | 0.8049 | 89.60% | 93.80% | −1.43% | 0.23 | 73.56% | 100.00% |
| Case−Crossover (1:25) | 0.7863 (0.2285) | 0.8017 | 90.00% | 94.70% | −1.72% | 0.22 | 80.87% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.6964 (0.0566) | 0.6968 | 100.00% | 94.90% | −0.52% | 0.06 | NA% | 100.00% |
| Z.it | 0.7811 (0.2066) | 0.7989 | 92.80% | 94.30% | −2.36% | 0.20 | 98.85% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.9786 (0.2054) | 0.9951 | 98.00% | 80.10% | 22.32% | 0.20 | 100.00% | 100.00% |
| **Cohort Size (N= 50000, # event=2479)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8025 (0.2619) | 0.7922 | 90.80% | 94.10% | 0.31% | 0.25 | 28.88% | 100.00% |
| Case−Crossover (1:2) | 0.7945 (0.2115) | 0.7985 | 97.00% | 94.70% | −0.69% | 0.20 | 44.32% | 100.00% |
| Case−Crossover (1:3) | 0.7920 (0.1928) | 0.7972 | 97.70% | 94.90% | −1.00% | 0.19 | 53.32% | 100.00% |
| Case−Crossover (1:5) | 0.7902 (0.1767) | 0.7924 | 98.70% | 94.60% | −1.23% | 0.17 | 63.43% | 100.00% |
| Case−Crossover (1:10) | 0.7921 (0.1629) | 0.7998 | 99.00% | 94.20% | −0.98% | 0.16 | 74.72% | 100.00% |
| Case−Crossover (1:25) | 0.7926 (0.1553) | 0.7978 | 99.20% | 94.80% | −0.93% | 0.15 | 82.18% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.6993 (0.0403) | 0.6974 | 100.00% | 95.00% | −0.10% | 0.04 | NA% | 100.00% |
| Z.it | 0.7882 (0.1415) | 0.7984 | 99.70% | 94.20% | −1.47% | 0.14 | 98.95% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.9865 (0.1408) | 0.9955 | 100.00% | 70.50% | 23.31% | 0.14 | 100.00% | 100.00% |
| **Cohort Size (N=100000, # event=4962)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8045 (0.1778) | 0.7985 | 99.80% | 95.00% | 0.56% | 0.18 | 30.14% | 100.00% |
| Case−Crossover (1:2) | 0.8021 (0.1478) | 0.7965 | 100.00% | 94.70% | 0.26% | 0.14 | 43.63% | 100.00% |
| Case−Crossover (1:3) | 0.8011 (0.1310) | 0.8011 | 100.00% | 95.90% | 0.14% | 0.13 | 55.52% | 100.00% |
| Case−Crossover (1:5) | 0.7997 (0.1204) | 0.8027 | 100.00% | 95.10% | −0.04% | 0.12 | 65.78% | 100.00% |
| Case−Crossover (1:10) | 0.8005 (0.1110) | 0.7995 | 100.00% | 95.30% | 0.06% | 0.11 | 77.38% | 100.00% |
| Case−Crossover (1:25) | 0.8000 (0.1062) | 0.8024 | 100.00% | 95.30% | 0.00% | 0.11 | 84.51% | 100.00% |
| Cox Hazards Model 1 − X.i | 0.6988 (0.0281) | 0.6992 | 100.00% | 94.90% | −0.17% | 0.03 | NA% | 100.00% |
| Z.it | 0.7968 (0.0985) | 0.7941 | 100.00% | 95.10% | −0.41% | 0.10 | 98.23% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.9949 (0.0976) | 0.9939 | 100.00% | 48.80% | 24.36% | 0.10 | 100.00% | 100.00% |

Table 4.10: Results from simulation of a cohort with an incidence of approximately 5% ($\lambda = 0.0018$), where the patient characteristic is confounding with the exposure and the event ($\gamma = .7$ and $\beta = 0.8$), and $P(Z_{i\cdot} = 1) = .4$ and $P(Z_{i\cdot} = 1 \mid X_i = 1) = .7$ (* Monte Carlo standard deviation of estimates ** Average of the information based standard error)

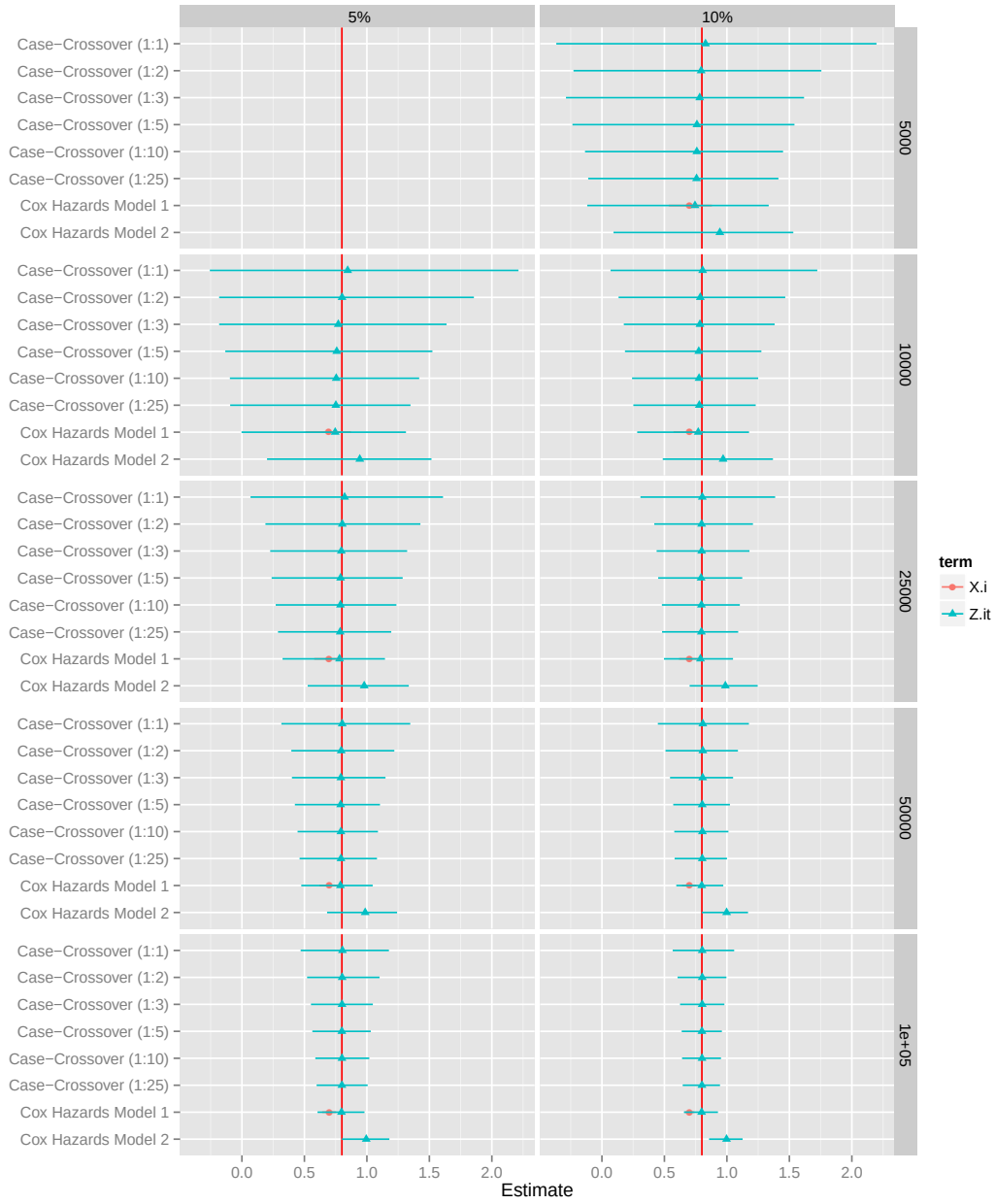


Figure 4.19: Plot of average estimates of the log-hazard ratio with 95% percentile intervals by design, cohort size and incidence rate when assuming a log-hazard ratio of $\beta = 0.8$ indicated by the vertical red line.

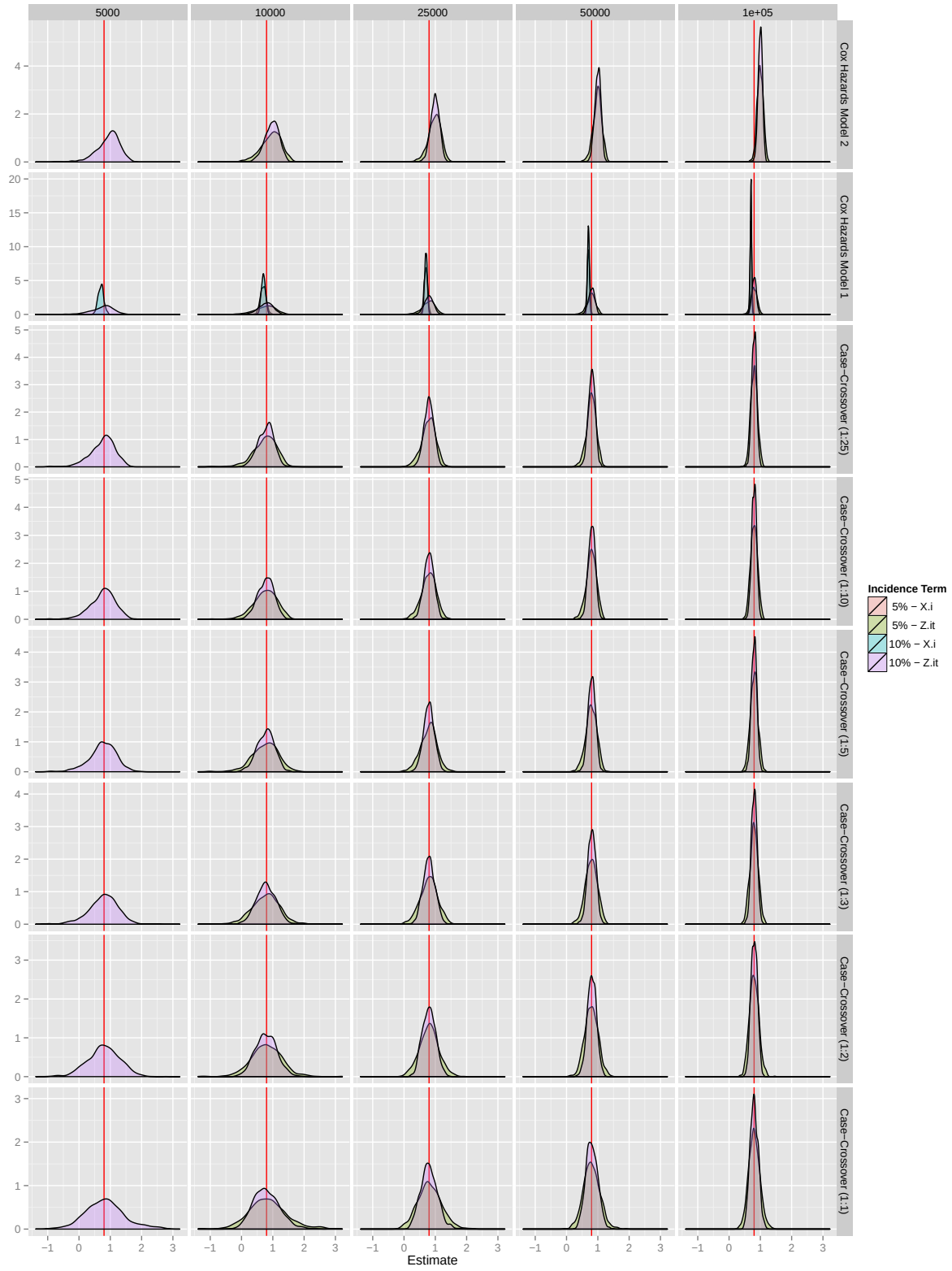


Figure 4.20: Distribution of estimate by design, cohort size and incidence rate when assuming a log-hazard ratio of $\beta = 0.8$ indicated by the vertical red line.

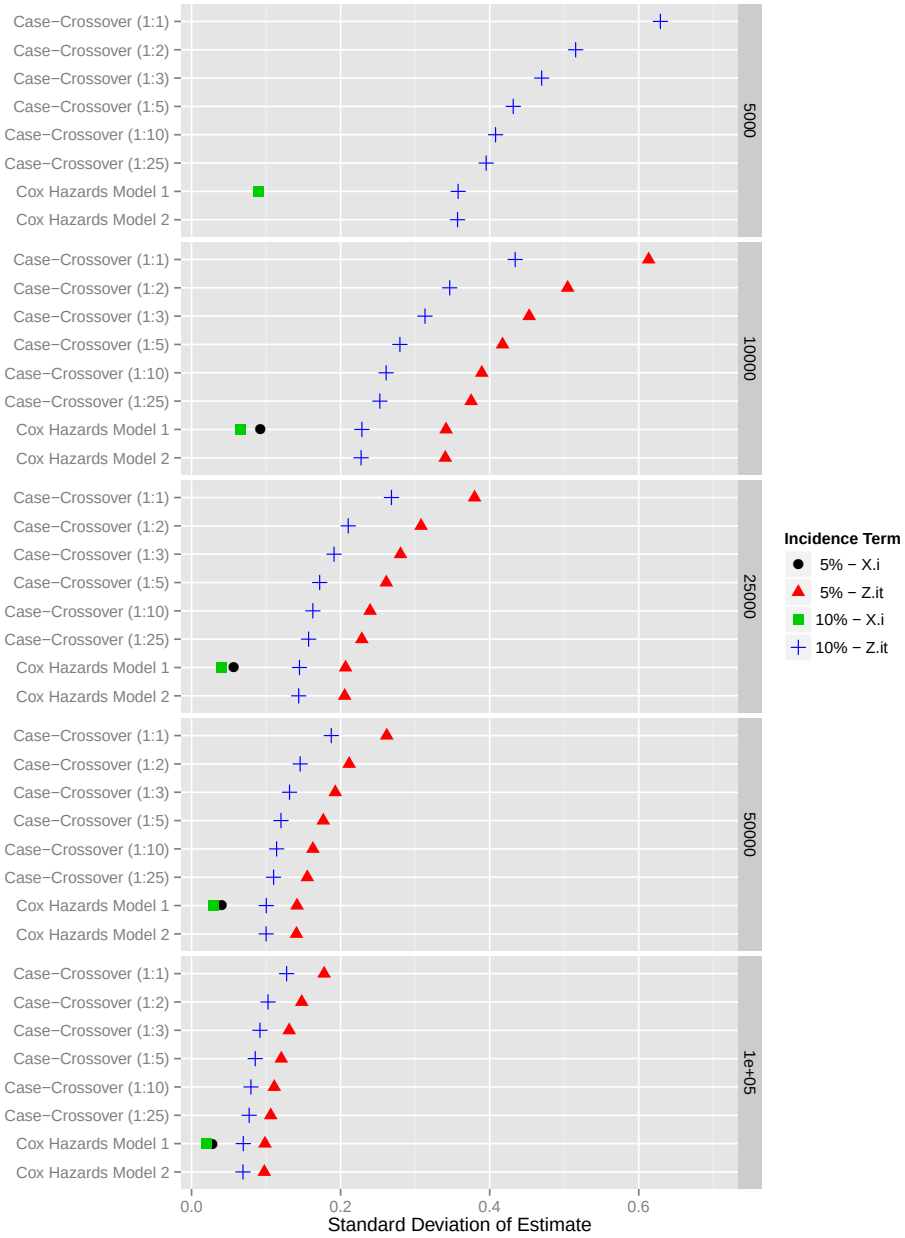


Figure 4.21: Plot of standard deviation of log-hazard ratio estimates by design, cohort size and incidence rate when assuming a log-hazard ratio of $\beta = 0.8$

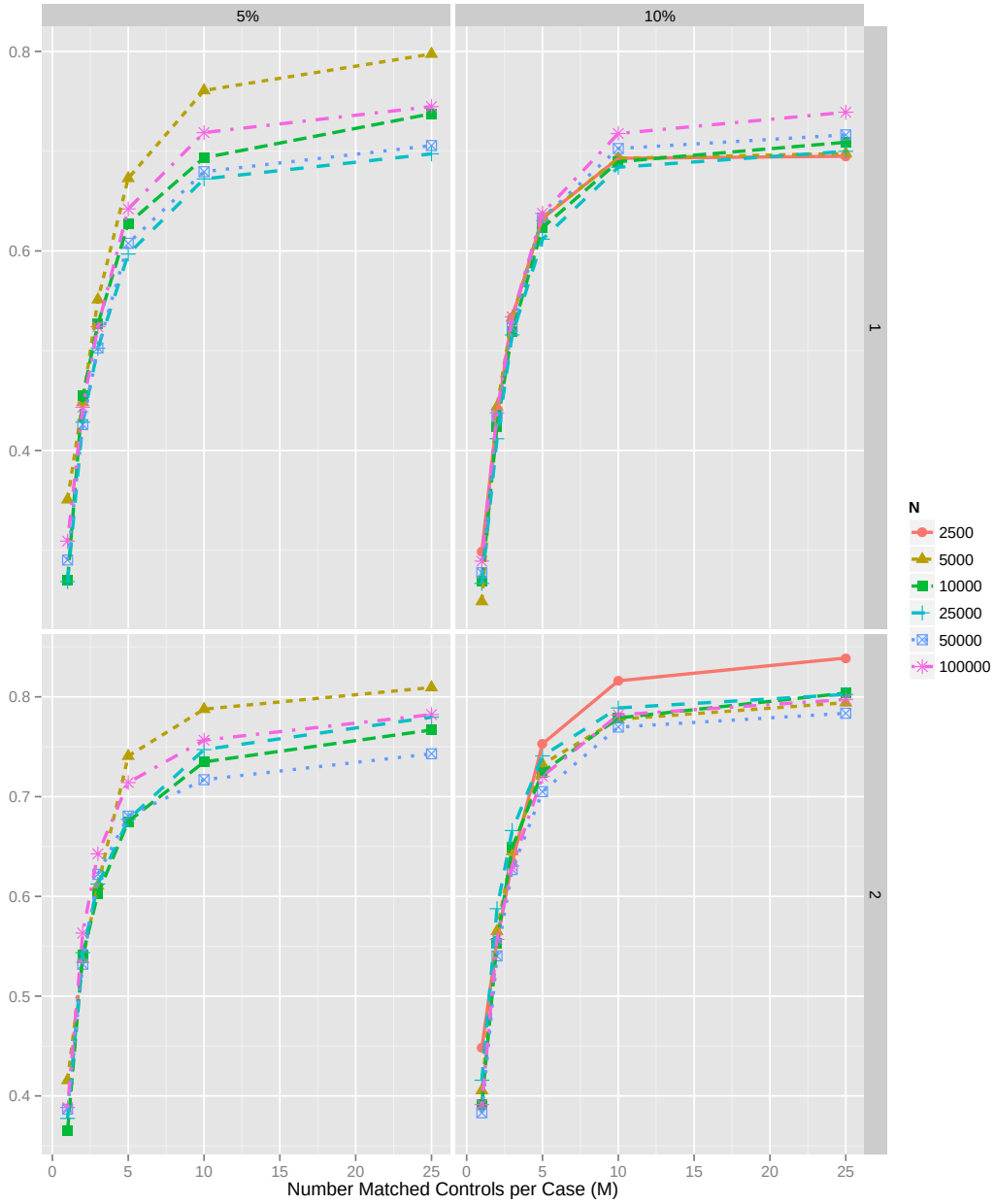


Figure 4.22: Plot of relative efficiency by design, cohort size and incidence rate when assuming a log-hazard ratio of $\beta = 0.8$

## 4.2 Distributed Lag

This section expands upon the simulations outlined in section 4.1. The simulations in this section assume that the transient effect resulting from a change in time-dependent exposure has an effect that lasts more than one time interval. Survival times are assumed to arise from the hazard function specified in Equation (3.17). The derivations yielding the function to generate survival times in Equation (4.2) is extended to generate survival times assuming a distributed lag of effect from the inverse cumulative hazard function given by

$$\tilde{T} = \begin{cases} \frac{-log(u)}{\lambda e^{\gamma X_i}}, & -log(u) \in R_1 \\ \frac{-log(u)+\lambda e^{\gamma X_i}\left(e^{\beta_1}\tau_i - \tau_i\right)}{\lambda e^{\gamma X_i+\beta_1}}, & -log(u) \in R_2 \\ \frac{-log(u)+\lambda e^{\gamma X_i}\left(e^{\beta_2}(\tau_i+1) - \tau_i - e^{\beta_1}\right)}{\lambda e^{\gamma X_i+\beta_2}}, & -log(u) \in R_3 \\ \vdots & \\ \frac{-log(u)+\lambda e^{\gamma X_i}\left(e^{\beta_l}(\tau_i+l) - \tau_i - \sum_{j=1}^{l-1} e^{\beta_j}\right)}{\lambda e^{\gamma X_i+\beta_l}}, & -log(u) \in R_{l+1} \\ \frac{-log(u)+\lambda e^{\gamma X_i}\left((\tau_i+l) - \tau_i - \sum_{j=1}^{l} e^{\beta_j}\right)}{\lambda e^{\gamma X_i}}, & -log(u) \in R_{l+2} \end{cases} \tag{4.4}$$

where

$$\begin{aligned} R_1 &= [0, \lambda e^{\gamma X_i}\tau_i) \\ R_2 &= \left[\lambda e^{\gamma X_i}\tau_i, \lambda e^{\gamma X_i}(\tau_i + e_1^{\beta})\right) \\ R_3 &= \left[\lambda e^{\gamma X_i}(\tau_i + e_1^{\beta}), \lambda e^{\gamma X_i}(\tau_i + e_1^{\beta} + e_2^{\beta})\right) \\ &\vdots \\ R_{1+2} &= \left[\lambda e^{\gamma X_i}(\tau_i + \sum_{j=1}^{l-1} e_j^{\beta}), \lambda e^{\gamma X_i}(\tau_i + \sum_{j=1}^{l} e_j^{\beta})\right) \end{aligned} \tag{4.5}$$

All simulations in this section assume that $\beta_l$ are constrained by the function $\Theta_l(\beta, \theta) = \beta\theta^{(l-1)}$. The first simulation demonstrates how the simple transient effect reviewed in section 4.1 can be extended to have an effect described by two parameters. The second simulation assumes the effect is distributed over a lag of seven time periods.

### 4.2.1 Two Lagged Variables

This first simulation assumes that the lag of the effect lasts over two time periods such that the hazard function is

$$\lambda_i(t) = \lambda_0(t)e^{\beta_1 Z_{it} + \beta_2 Z_{it-1}}$$

The log-hazard ratios over the two time intervals are assumed to be $\beta_1 = 0.8$ and $\beta_2 = 0.4$. The results from these simulated studies are displayed in Table 4.11 for cohorts with an assumed incidence rate of 10% and in Table 4.12 for cohorts assuming an incidence rate of 5%.

Average estimates for both $\beta_1$ and $\beta_2$ both tend to yield results close to the hypothesized values of 0.8 and 0.4 respectively. The patterns of bias still appear to exist for this simulation: the 1:1 case-crossover overestimates $\beta_1$ and $\beta_2$ and the Cox proportional hazards model, 1:10 and 1:25 matched case-crossover designs underestimate $\beta_1$ and $\beta_2$ for cohorts of smaller sample size. Differing from earlier simulations, the presence of bias exceeding 1% still exists in cohorts of size 100,000 for $\beta_2$. The average estimates of $\beta_1$ and $\beta_2$ with corresponding percentile intervals is plotted in Figure 4.23. In this plot there are two dark red lines that indicate the hypothesized values of 0.8 and 0.4. For smaller sample sizes, the 1:3 case-crossover designs tend to provide the least amount of bias and as $M$ increases the magnitude of underestimating $\beta_1$ and $\beta_2$ increases. The level of bias intervals decrease with the higher incidence rate and larger cohort sample sizes. The negative bias observed in some of these designs may be a result of the skewed left distributions observed in Figure 4.24.

Similar to the simulations in previous sections, 1000 events are required to observe power exceeding 80% for $\beta_1$ assuming a type I error rate of 5%. The ability to detect a difference for the second lagged parameter is more difficult, requiring about 5000 observations are required to observe power exceeding 80%. Coverage probabilities for both variables are consistent with previous simulations.

A plot of relative efficiency is depicted in Figure 4.26. This plot shows similar patterns to previous simulations: relative efficiency increases with the number of matched controls and sample size of the cohort. The second lagged parameter has consistently higher relative

| | Mean (SD*) | Median | Power | Coverage Probability | Bias | Avg Std Error** | Relative Efficiency | Converged |
|---|---|---|---|---|---|---|---|---|
| **Cohort Size (N= 2500, # event=244)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8377 (0.6922) | 0.8109 | 16.67% | 97.33% | 4.71% | 0.70 | 29.85% | 93.60% |
| Case−Crossover (1:1) | 0.3966 (0.7427) | 0.4055 | 3.31% | 98.40% | −0.84% | 0.77 | 44.80% | 93.60% |
| **Cohort Size (N= 2500, # event=243)** | | | | | | | | |
| Case−Crossover (1:2) | 0.8222 (0.5744) | 0.8267 | 32.90% | 96.97% | 2.78% | 0.56 | 43.35% | 99.10% |
| Case−Crossover (1:2) | 0.3874 (0.6669) | 0.3925 | 10.29% | 97.38% | −3.14% | 0.64 | 55.56% | 99.10% |
| Case−Crossover (1:3) | 0.7949 (0.5180) | 0.7952 | 39.60% | 96.48% | −0.63% | 0.51 | 53.31% | 99.50% |
| Case−Crossover (1:3) | 0.3606 (0.6242) | 0.3939 | 11.86% | 96.58% | −9.85% | 0.59 | 63.41% | 99.50% |
| Case−Crossover (1:5) | 0.7829 (0.4754) | 0.8105 | 45.93% | 95.88% | −2.13% | 0.46 | 63.27% | 99.50% |
| Case−Crossover (1:5) | 0.3301 (0.5729) | 0.3819 | 15.38% | 97.19% | −17.48% | 0.55 | 75.29% | 99.50% |
| Case−Crossover (1:10) | 0.7663 (0.4543) | 0.8132 | 49.45% | 95.98% | −4.22% | 0.44 | 69.29% | 99.50% |
| Case−Crossover (1:10) | 0.3197 (0.5503) | 0.3820 | 15.08% | 97.09% | −20.09% | 0.53 | 81.61% | 99.50% |
| Case−Crossover (1:25) | 0.7634 (0.4537) | 0.8127 | 51.46% | 95.68% | −4.58% | 0.43 | 69.49% | 99.50% |
| Case−Crossover (1:25) | 0.3175 (0.5427) | 0.3874 | 15.58% | 96.88% | −20.64% | 0.52 | 83.90% | 99.50% |
| Cox Hazards Model 2 − Z.it | 0.7557 (0.3782) | 0.7953 | 59.78% | 95.69% | −5.54% | 0.36 | 100.00% | 99.70% |
| Cox Hazards Model 2 − Z.it | 0.2997 (0.4971) | 0.3667 | 18.25% | 96.29% | −25.07% | 0.46 | 100.00% | 99.70% |
| **Cohort Size (N= 5000, # event=487)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8669 (0.5256) | 0.8329 | 45.29% | 95.29% | 8.36% | 0.48 | 24.88% | 99.80% |
| Case−Crossover (1:1) | 0.4087 (0.5212) | 0.4055 | 10.02% | 97.19% | 2.18% | 0.51 | 40.55% | 99.80% |
| Case−Crossover (1:2) | 0.8193 (0.3935) | 0.8277 | 59.90% | 95.90% | 2.41% | 0.38 | 44.39% | 100.00% |
| Case−Crossover (1:2) | 0.3995 (0.4415) | 0.4004 | 17.50% | 96.30% | −0.12% | 0.43 | 56.51% | 100.00% |
| Case−Crossover (1:3) | 0.7998 (0.3591) | 0.8113 | 66.00% | 95.20% | −0.03% | 0.35 | 53.29% | 100.00% |
| Case−Crossover (1:3) | 0.3808 (0.4132) | 0.3807 | 20.10% | 95.30% | −4.80% | 0.40 | 64.51% | 100.00% |
| Case−Crossover (1:5) | 0.7961 (0.3294) | 0.7990 | 70.90% | 95.50% | −0.49% | 0.32 | 63.35% | 100.00% |
| Case−Crossover (1:5) | 0.3698 (0.3878) | 0.3771 | 21.80% | 94.60% | −7.56% | 0.37 | 73.25% | 100.00% |
| Case−Crossover (1:10) | 0.7912 (0.3150) | 0.8120 | 72.40% | 94.60% | −1.10% | 0.30 | 69.29% | 100.00% |
| Case−Crossover (1:10) | 0.3676 (0.3764) | 0.3881 | 23.40% | 94.20% | −8.11% | 0.36 | 77.76% | 100.00% |
| Case−Crossover (1:25) | 0.7903 (0.3139) | 0.8166 | 73.40% | 94.30% | −1.22% | 0.30 | 69.77% | 100.00% |
| Case−Crossover (1:25) | 0.3669 (0.3724) | 0.3819 | 24.70% | 94.40% | −8.27% | 0.36 | 79.42% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.7890 (0.2622) | 0.8071 | 83.90% | 94.00% | −1.38% | 0.25 | 100.00% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.3544 (0.3319) | 0.3666 | 28.90% | 94.30% | −11.39% | 0.31 | 100.00% | 100.00% |
| **Cohort Size (N= 10000, # event=974)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8352 (0.3310) | 0.8333 | 75.50% | 96.00% | 4.40% | 0.33 | 26.94% | 100.00% |
| Case−Crossover (1:1) | 0.4083 (0.3474) | 0.4055 | 19.30% | 95.80% | 2.06% | 0.35 | 39.11% | 100.00% |
| Case−Crossover (1:2) | 0.8190 (0.2638) | 0.8165 | 88.00% | 95.10% | 2.37% | 0.26 | 42.43% | 100.00% |
| Case−Crossover (1:2) | 0.4030 (0.2921) | 0.4016 | 27.30% | 96.40% | 0.74% | 0.30 | 55.32% | 100.00% |
| Case−Crossover (1:3) | 0.8093 (0.2385) | 0.8125 | 90.80% | 96.50% | 1.16% | 0.24 | 51.88% | 100.00% |
| Case−Crossover (1:3) | 0.4003 (0.2697) | 0.3956 | 32.80% | 96.60% | 0.06% | 0.27 | 64.89% | 100.00% |
| Case−Crossover (1:5) | 0.8048 (0.2176) | 0.8160 | 93.00% | 96.30% | 0.61% | 0.22 | 62.34% | 100.00% |
| Case−Crossover (1:5) | 0.3975 (0.2552) | 0.4005 | 36.20% | 95.80% | −0.63% | 0.26 | 72.46% | 100.00% |
| Case−Crossover (1:10) | 0.8029 (0.2069) | 0.8151 | 94.80% | 95.80% | 0.36% | 0.21 | 68.95% | 100.00% |
| Case−Crossover (1:10) | 0.3952 (0.2462) | 0.4091 | 39.40% | 95.10% | −1.19% | 0.25 | 77.88% | 100.00% |
| Case−Crossover (1:25) | 0.8025 (0.2041) | 0.8169 | 94.70% | 95.30% | 0.31% | 0.21 | 70.89% | 100.00% |
| Case−Crossover (1:25) | 0.3949 (0.2423) | 0.4052 | 40.20% | 95.50% | −1.26% | 0.25 | 80.38% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.8010 (0.1718) | 0.8074 | 97.80% | 95.00% | 0.13% | 0.17 | 100.00% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.3857 (0.2173) | 0.3984 | 47.40% | 95.20% | −3.59% | 0.22 | 100.00% | 100.00% |

Table 4.11: Results from a simulation of a cohort with an incidence of approximately 10% ( $\lambda = .0025$) and effect deteriorating over two time periods ($\beta_1 = .8, \beta_2 = .4$) (* Monte Carlo standard deviation of estimates ** Average of the information based standard error)

efficiency compared to the first lagged covariate. The relative efficiency is consistently at or just beneath 80%. The first lagged factor is typically around 70%. The standard deviation of estimates is plotted in Figure 4.25. Patterns of the standard deviation for each coefficient are consistent with the earlier simulation studies. For both incidence rates, the standard deviation for the second lagged parameter is always larger than the first.

| | Mean (SD*) | Median | Power | Coverage Probability | Bias | Avg Std Error** | Relative Efficiency | Converged |
|---|---|---|---|---|---|---|---|---|
| **Cohort Size (N= 25000, # event=2434)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8204 (0.2041) | 0.8154 | 98.90% | 95.90% | 2.55% | 0.20 | 26.68% | 100.00% |
| Case−Crossover (1:1) | 0.4038 (0.2139) | 0.3891 | 43.70% | 96.00% | 0.94% | 0.22 | 41.57% | 100.00% |
| Case−Crossover (1:2) | 0.8120 (0.1643) | 0.8202 | 99.70% | 95.20% | 1.50% | 0.17 | 41.18% | 100.00% |
| Case−Crossover (1:2) | 0.3995 (0.1799) | 0.3942 | 57.90% | 95.50% | −0.12% | 0.18 | 58.76% | 100.00% |
| Case−Crossover (1:3) | 0.8058 (0.1468) | 0.8107 | 100.00% | 95.90% | 0.73% | 0.15 | 51.59% | 100.00% |
| Case−Crossover (1:3) | 0.4014 (0.1690) | 0.4049 | 64.80% | 95.40% | 0.35% | 0.17 | 66.60% | 100.00% |
| Case−Crossover (1:5) | 0.8053 (0.1348) | 0.8063 | 100.00% | 96.30% | 0.66% | 0.14 | 61.18% | 100.00% |
| Case−Crossover (1:5) | 0.4018 (0.1602) | 0.4065 | 69.90% | 95.40% | 0.45% | 0.16 | 74.08% | 100.00% |
| Case−Crossover (1:10) | 0.8042 (0.1275) | 0.8086 | 100.00% | 95.80% | 0.52% | 0.13 | 68.39% | 100.00% |
| Case−Crossover (1:10) | 0.4032 (0.1552) | 0.4106 | 72.40% | 94.80% | 0.80% | 0.16 | 78.89% | 100.00% |
| Case−Crossover (1:25) | 0.8040 (0.1260) | 0.8022 | 100.00% | 95.20% | 0.50% | 0.13 | 69.99% | 100.00% |
| Case−Crossover (1:25) | 0.4029 (0.1539) | 0.4093 | 73.00% | 95.10% | 0.72% | 0.15 | 80.24% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.8013 (0.1054) | 0.7992 | 100.00% | 96.20% | 0.16% | 0.11 | 100.00% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.3989 (0.1379) | 0.4021 | 81.00% | 94.40% | −0.26% | 0.14 | 100.00% | 100.00% |
| **Cohort Size (N= 50000, # event=4866)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8119 (0.1418) | 0.8102 | 100.00% | 95.70% | 1.48% | 0.14 | 27.75% | 100.00% |
| Case−Crossover (1:1) | 0.4060 (0.1510) | 0.4055 | 77.20% | 94.70% | 1.50% | 0.15 | 38.31% | 100.00% |
| Case−Crossover (1:2) | 0.8087 (0.1134) | 0.8094 | 100.00% | 96.00% | 1.09% | 0.12 | 43.40% | 100.00% |
| Case−Crossover (1:2) | 0.4036 (0.1271) | 0.4034 | 87.70% | 94.50% | 0.89% | 0.13 | 54.03% | 100.00% |
| Case−Crossover (1:3) | 0.8059 (0.1032) | 0.8029 | 100.00% | 96.00% | 0.74% | 0.11 | 52.34% | 100.00% |
| Case−Crossover (1:3) | 0.4038 (0.1180) | 0.4091 | 90.60% | 96.20% | 0.96% | 0.12 | 62.69% | 100.00% |
| Case−Crossover (1:5) | 0.8050 (0.0939) | 0.8075 | 100.00% | 96.80% | 0.62% | 0.10 | 63.31% | 100.00% |
| Case−Crossover (1:5) | 0.4030 (0.1113) | 0.4067 | 93.40% | 95.80% | 0.76% | 0.11 | 70.50% | 100.00% |
| Case−Crossover (1:10) | 0.8038 (0.0891) | 0.8047 | 100.00% | 96.10% | 0.47% | 0.09 | 70.27% | 100.00% |
| Case−Crossover (1:10) | 0.4040 (0.1065) | 0.4096 | 94.80% | 95.30% | 1.00% | 0.11 | 76.97% | 100.00% |
| Case−Crossover (1:25) | 0.8035 (0.0883) | 0.8028 | 100.00% | 96.20% | 0.43% | 0.09 | 71.64% | 100.00% |
| Case−Crossover (1:25) | 0.4044 (0.1056) | 0.4084 | 95.20% | 95.40% | 1.10% | 0.11 | 78.35% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.8013 (0.0747) | 0.8035 | 100.00% | 95.90% | 0.16% | 0.08 | 100.00% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.4023 (0.0934) | 0.4067 | 97.40% | 95.20% | 0.56% | 0.10 | 100.00% | 100.00% |
| **Cohort Size (N=100000, # event=9731)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8073 (0.1018) | 0.8027 | 100.00% | 94.50% | 0.91% | 0.10 | 28.91% | 100.00% |
| Case−Crossover (1:1) | 0.4026 (0.1040) | 0.4055 | 97.00% | 97.50% | 0.66% | 0.11 | 39.13% | 100.00% |
| Case−Crossover (1:2) | 0.8053 (0.0827) | 0.8067 | 100.00% | 95.20% | 0.67% | 0.08 | 43.78% | 100.00% |
| Case−Crossover (1:2) | 0.4028 (0.0872) | 0.4039 | 99.40% | 96.30% | 0.69% | 0.09 | 55.70% | 100.00% |
| Case−Crossover (1:3) | 0.8039 (0.0749) | 0.8029 | 100.00% | 94.80% | 0.49% | 0.08 | 53.40% | 100.00% |
| Case−Crossover (1:3) | 0.4040 (0.0820) | 0.4047 | 99.70% | 95.50% | 1.01% | 0.09 | 63.03% | 100.00% |
| Case−Crossover (1:5) | 0.8041 (0.0686) | 0.8039 | 100.00% | 96.00% | 0.51% | 0.07 | 63.77% | 100.00% |
| Case−Crossover (1:5) | 0.4035 (0.0767) | 0.4017 | 99.80% | 95.70% | 0.87% | 0.08 | 72.03% | 100.00% |
| Case−Crossover (1:10) | 0.8030 (0.0646) | 0.8033 | 100.00% | 95.70% | 0.37% | 0.07 | 71.77% | 100.00% |
| Case−Crossover (1:10) | 0.4042 (0.0736) | 0.4034 | 99.90% | 95.70% | 1.06% | 0.08 | 78.19% | 100.00% |
| Case−Crossover (1:25) | 0.8024 (0.0637) | 0.8041 | 100.00% | 95.90% | 0.30% | 0.07 | 73.91% | 100.00% |
| Case−Crossover (1:25) | 0.4047 (0.0729) | 0.4035 | 100.00% | 95.70% | 1.16% | 0.08 | 79.74% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.8002 (0.0547) | 0.8001 | 100.00% | 95.80% | 0.03% | 0.05 | 100.00% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.4033 (0.0651) | 0.4059 | 99.90% | 95.80% | 0.83% | 0.07 | 100.00% | 100.00% |

Table 4.11 (contd.): Results from a simulation of a cohort with an incidence of approximately 10% ( $\lambda = .0025$) and effect deteriorating over two time periods ($\beta_1 = .8, \beta_2 = .4$) (* Monte Carlo standard deviation of estimates ** Average of the information based standard error)

| | Mean (SD*) | Median | Power | Coverage Probability | Bias | Avg Std Error** | Relative Efficiency | Converged |
|---|---|---|---|---|---|---|---|---|
| **Cohort Size (N= 5000, # event=250)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8458 (0.6805) | 0.8473 | 17.02% | 97.13% | 5.72% | 0.69 | 35.08% | 94.00% |
| Case−Crossover (1:1) | 0.3694 (0.7458) | 0.3365 | 3.19% | 98.19% | −7.64% | 0.76 | 41.56% | 94.00% |
| Case−Crossover (1:2) | 0.8239 (0.6019) | 0.8267 | 34.94% | 96.39% | 2.98% | 0.56 | 44.84% | 99.60% |
| Case−Crossover (1:2) | 0.3516 (0.6561) | 0.3567 | 10.74% | 97.29% | −12.11% | 0.63 | 53.70% | 99.60% |
| Case−Crossover (1:3) | 0.7941 (0.5429) | 0.8216 | 40.72% | 95.09% | −0.73% | 0.50 | 55.12% | 99.70% |
| Case−Crossover (1:3) | 0.3267 (0.6152) | 0.3384 | 10.73% | 96.89% | −18.33% | 0.58 | 61.07% | 99.70% |
| Case−Crossover (1:5) | 0.7790 (0.4914) | 0.8071 | 45.44% | 95.79% | −2.63% | 0.46 | 67.28% | 99.70% |
| Case−Crossover (1:5) | 0.3123 (0.5586) | 0.3660 | 13.04% | 96.99% | −21.94% | 0.55 | 74.06% | 99.70% |
| Case−Crossover (1:10) | 0.7594 (0.4620) | 0.7974 | 48.14% | 94.98% | −5.08% | 0.43 | 76.09% | 99.70% |
| Case−Crossover (1:10) | 0.3020 (0.5417) | 0.3740 | 13.24% | 97.19% | −24.50% | 0.52 | 78.77% | 99.70% |
| Case−Crossover (1:25) | 0.7598 (0.4513) | 0.7957 | 49.95% | 95.99% | −5.02% | 0.43 | 79.74% | 99.70% |
| Case−Crossover (1:25) | 0.2978 (0.5344) | 0.3628 | 13.74% | 96.89% | −25.55% | 0.52 | 80.94% | 99.70% |
| Cox Hazards Model 2 − Z.it | 0.7414 (0.4030) | 0.7941 | 60.82% | 95.99% | −7.33% | 0.36 | 100.00% | 99.80% |
| Cox Hazards Model 2 − Z.it | 0.2832 (0.4808) | 0.3493 | 17.94% | 97.49% | −29.19% | 0.46 | 100.00% | 99.80% |
| **Cohort Size (N= 10000, # event=500)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8521 (0.4983) | 0.8109 | 42.73% | 96.59% | 6.51% | 0.48 | 27.02% | 99.70% |
| Case−Crossover (1:1) | 0.4182 (0.5308) | 0.4055 | 10.23% | 95.99% | 4.55% | 0.51 | 36.50% | 99.70% |
| Case−Crossover (1:2) | 0.8069 (0.3842) | 0.8052 | 58.30% | 96.50% | 0.87% | 0.38 | 45.45% | 100.00% |
| Case−Crossover (1:2) | 0.3849 (0.4360) | 0.3993 | 16.20% | 95.70% | −3.78% | 0.42 | 54.11% | 100.00% |
| Case−Crossover (1:3) | 0.7914 (0.3569) | 0.8006 | 64.00% | 95.50% | −1.07% | 0.34 | 52.67% | 100.00% |
| Case−Crossover (1:3) | 0.3700 (0.4131) | 0.3836 | 18.60% | 94.70% | −7.49% | 0.39 | 60.26% | 100.00% |
| Case−Crossover (1:5) | 0.7838 (0.3270) | 0.7804 | 69.40% | 95.00% | −2.02% | 0.32 | 62.75% | 100.00% |
| Case−Crossover (1:5) | 0.3624 (0.3903) | 0.3766 | 20.80% | 94.10% | −9.40% | 0.37 | 67.50% | 100.00% |
| Case−Crossover (1:10) | 0.7740 (0.3110) | 0.7715 | 72.10% | 94.40% | −3.24% | 0.30 | 69.34% | 100.00% |
| Case−Crossover (1:10) | 0.3556 (0.3741) | 0.3699 | 21.70% | 94.60% | −11.10% | 0.35 | 73.48% | 100.00% |
| Case−Crossover (1:25) | 0.7735 (0.3016) | 0.7816 | 73.40% | 95.00% | −3.31% | 0.29 | 73.76% | 100.00% |
| Case−Crossover (1:25) | 0.3517 (0.3662) | 0.3682 | 22.20% | 94.60% | −12.07% | 0.35 | 76.68% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.7602 (0.2590) | 0.7720 | 82.90% | 94.40% | −4.97% | 0.25 | 100.00% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.3477 (0.3207) | 0.3890 | 26.20% | 95.30% | −13.07% | 0.31 | 100.00% | 100.00% |
| **Cohort Size (N= 25000, # event=1251)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8229 (0.2986) | 0.7941 | 84.50% | 95.30% | 2.87% | 0.29 | 26.85% | 100.00% |
| Case−Crossover (1:1) | 0.4045 (0.3175) | 0.4055 | 25.80% | 95.30% | 1.13% | 0.31 | 37.75% | 100.00% |
| Case−Crossover (1:2) | 0.8044 (0.2364) | 0.8059 | 92.70% | 96.00% | 0.55% | 0.23 | 42.84% | 100.00% |
| Case−Crossover (1:2) | 0.3955 (0.2646) | 0.4027 | 33.60% | 95.00% | −1.11% | 0.26 | 54.35% | 100.00% |
| Case−Crossover (1:3) | 0.8007 (0.2183) | 0.8025 | 95.30% | 95.10% | 0.08% | 0.21 | 50.25% | 100.00% |
| Case−Crossover (1:3) | 0.3905 (0.2493) | 0.3883 | 40.20% | 95.20% | −2.38% | 0.24 | 61.24% | 100.00% |
| Case−Crossover (1:5) | 0.7930 (0.2003) | 0.7903 | 96.90% | 95.10% | −0.87% | 0.20 | 59.69% | 100.00% |
| Case−Crossover (1:5) | 0.3873 (0.2371) | 0.3831 | 42.70% | 94.30% | −3.18% | 0.23 | 67.71% | 100.00% |
| Case−Crossover (1:10) | 0.7891 (0.1887) | 0.7893 | 97.50% | 94.80% | −1.36% | 0.19 | 67.22% | 100.00% |
| Case−Crossover (1:10) | 0.3847 (0.2257) | 0.3901 | 44.10% | 94.30% | −3.82% | 0.22 | 74.71% | 100.00% |
| Case−Crossover (1:25) | 0.7893 (0.1853) | 0.7908 | 97.70% | 95.20% | −1.34% | 0.18 | 69.73% | 100.00% |
| Case−Crossover (1:25) | 0.3837 (0.2209) | 0.3877 | 44.70% | 94.40% | −4.09% | 0.22 | 77.97% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.7899 (0.1547) | 0.7985 | 99.60% | 95.00% | −1.27% | 0.15 | 100.00% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.3822 (0.1951) | 0.3934 | 53.00% | 94.50% | −4.45% | 0.19 | 100.00% | 100.00% |

Table 4.12: Results from a simulation of a cohort with an incidence of approximately 5% ( $\lambda = .00125$) and effect deteriorating over two time periods ($\beta_1 = .8, \beta_2 = .4$) (* Monte Carlo standard deviation of estimates ** Average of the information based standard error)

| | Mean (SD*) | Median | Power | Coverage Probability | Bias | Avg Std Error** | Relative Efficiency | Converged |
|---|---|---|---|---|---|---|---|---|
| **Cohort Size (N= 50000, # event=2502)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8185 (0.2042) | 0.8081 | 99.40% | 95.10% | 2.31% | 0.20 | 29.02% | 100.00% |
| Case−Crossover (1:1) | 0.4006 (0.2200) | 0.3956 | 46.10% | 95.60% | 0.14% | 0.22 | 38.67% | 100.00% |
| Case−Crossover (1:2) | 0.8088 (0.1686) | 0.8030 | 99.60% | 95.50% | 1.10% | 0.16 | 42.59% | 100.00% |
| Case−Crossover (1:2) | 0.3998 (0.1875) | 0.3986 | 58.80% | 94.90% | −0.06% | 0.18 | 53.20% | 100.00% |
| Case−Crossover (1:3) | 0.8067 (0.1552) | 0.8067 | 99.80% | 94.20% | 0.84% | 0.15 | 50.28% | 100.00% |
| Case−Crossover (1:3) | 0.3968 (0.1734) | 0.4028 | 65.10% | 94.10% | −0.79% | 0.17 | 62.21% | 100.00% |
| Case−Crossover (1:5) | 0.8029 (0.1411) | 0.8025 | 99.90% | 94.00% | 0.36% | 0.14 | 60.77% | 100.00% |
| Case−Crossover (1:5) | 0.3924 (0.1658) | 0.4020 | 67.30% | 94.50% | −1.91% | 0.16 | 68.05% | 100.00% |
| Case−Crossover (1:10) | 0.8017 (0.1335) | 0.8022 | 99.90% | 94.10% | 0.21% | 0.13 | 67.94% | 100.00% |
| Case−Crossover (1:10) | 0.3928 (0.1615) | 0.3990 | 70.80% | 93.40% | −1.79% | 0.15 | 71.70% | 100.00% |
| Case−Crossover (1:25) | 0.8023 (0.1310) | 0.8042 | 99.90% | 93.70% | 0.28% | 0.13 | 70.57% | 100.00% |
| Case−Crossover (1:25) | 0.3936 (0.1587) | 0.3989 | 72.50% | 93.50% | −1.60% | 0.15 | 74.29% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.7982 (0.1100) | 0.7995 | 100.00% | 94.40% | −0.23% | 0.11 | 100.00% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.3905 (0.1368) | 0.3930 | 80.50% | 94.10% | −2.37% | 0.13 | 100.00% | 100.00% |
| **Cohort Size (N=100000, # event=5004)** | | | | | | | | |
| Case−Crossover (1:1) | 0.8104 (0.1428) | 0.8074 | 100.00% | 95.70% | 1.29% | 0.14 | 30.91% | 100.00% |
| Case−Crossover (1:1) | 0.3963 (0.1541) | 0.3969 | 73.60% | 95.90% | −0.93% | 0.15 | 38.84% | 100.00% |
| Case−Crossover (1:2) | 0.8081 (0.1192) | 0.8042 | 100.00% | 93.60% | 1.02% | 0.12 | 44.34% | 100.00% |
| Case−Crossover (1:2) | 0.3969 (0.1279) | 0.3945 | 87.00% | 96.10% | −0.78% | 0.13 | 56.32% | 100.00% |
| Case−Crossover (1:3) | 0.8068 (0.1096) | 0.8053 | 100.00% | 93.90% | 0.85% | 0.11 | 52.41% | 100.00% |
| Case−Crossover (1:3) | 0.3962 (0.1198) | 0.3945 | 90.90% | 96.20% | −0.95% | 0.12 | 64.26% | 100.00% |
| Case−Crossover (1:5) | 0.8072 (0.0991) | 0.8054 | 100.00% | 95.00% | 0.90% | 0.10 | 64.22% | 100.00% |
| Case−Crossover (1:5) | 0.3956 (0.1136) | 0.3942 | 92.70% | 95.30% | −1.11% | 0.11 | 71.41% | 100.00% |
| Case−Crossover (1:10) | 0.8055 (0.0936) | 0.8075 | 100.00% | 95.50% | 0.69% | 0.09 | 71.87% | 100.00% |
| Case−Crossover (1:10) | 0.3945 (0.1104) | 0.3992 | 93.60% | 95.40% | −1.38% | 0.11 | 75.66% | 100.00% |
| Case−Crossover (1:25) | 0.8059 (0.0920) | 0.8082 | 100.00% | 95.10% | 0.73% | 0.09 | 74.48% | 100.00% |
| Case−Crossover (1:25) | 0.3949 (0.1086) | 0.3986 | 94.80% | 96.00% | −1.27% | 0.11 | 78.22% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.8020 (0.0794) | 0.8014 | 100.00% | 94.10% | 0.25% | 0.08 | 100.00% | 100.00% |
| Cox Hazards Model 2 − Z.it | 0.3933 (0.0960) | 0.3993 | 97.60% | 95.80% | −1.67% | 0.09 | 100.00% | 100.00% |

Table 4.12 (contd.): Results from a simulation of a cohort with an incidence of approximately 5% ( $\lambda = .00125$) and effect deteriorating over two time periods ($\beta_1 = .8, \beta_2 = .4$) (* Monte Carlo standard deviation of estimates ** Average of the information based standard error)

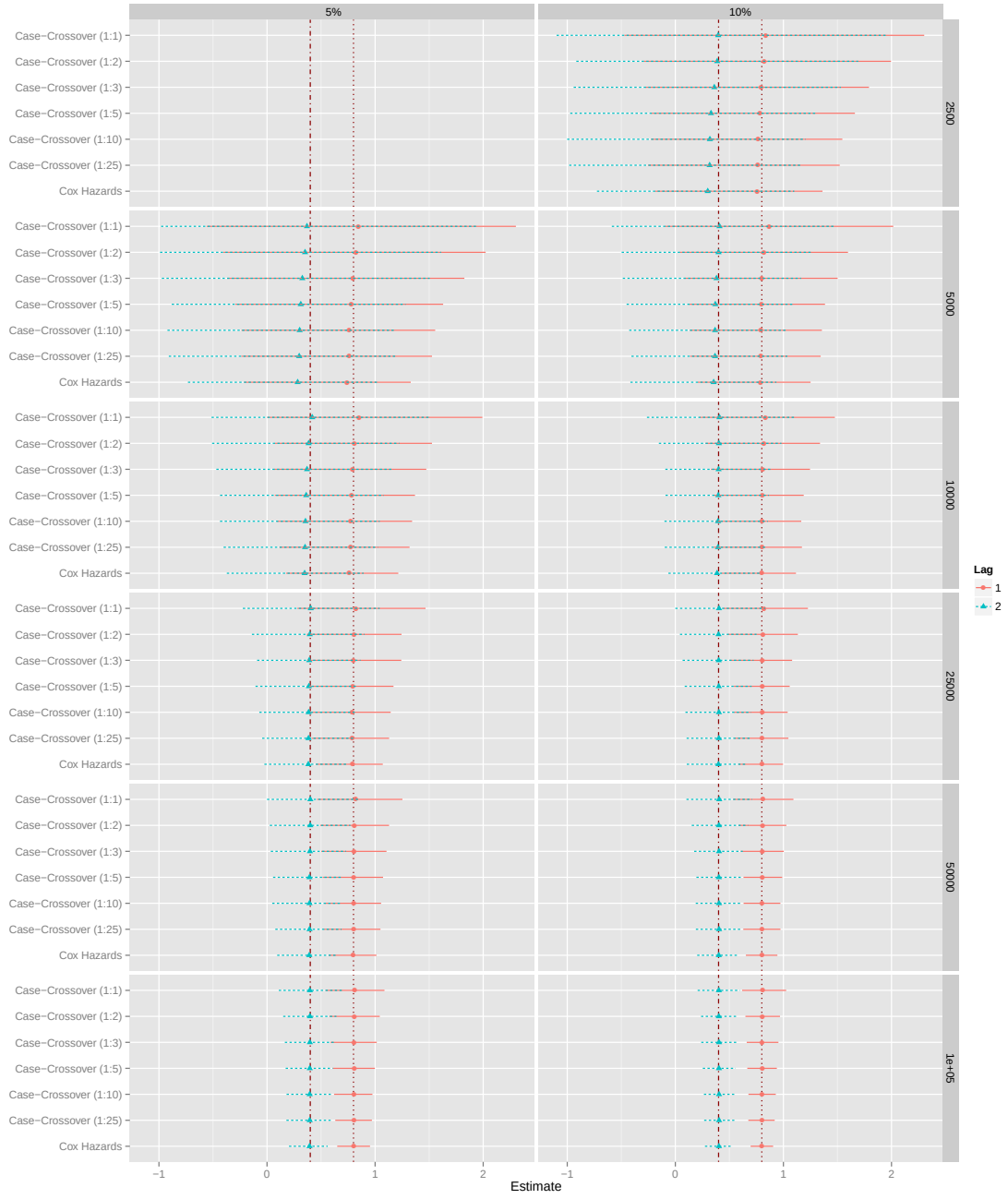


Figure 4.23: Plot of average estimates of the log-hazard ratio with 95% percentile intervals by parameter, design, cohort size and incidence rate when assuming a log-hazard ratios of $\beta_1 = 0.8, \beta_2 = 0.4$ indicated by the vertical dark red lines.

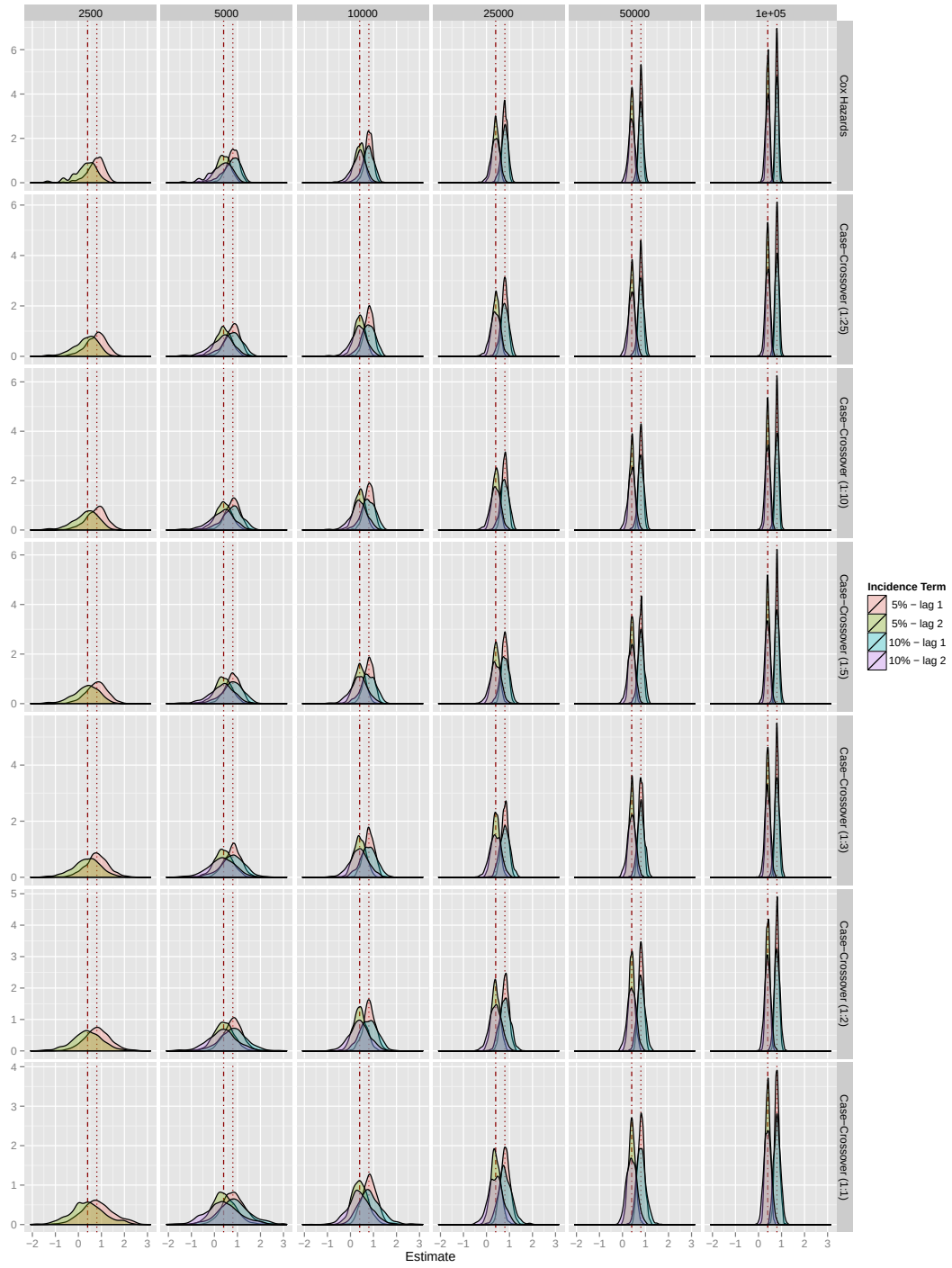


Figure 4.24: Distribution of estimate by parameter, design, cohort size and incidence rate when assuming a log-hazard ratios of $\beta_1 = 0.8, \beta_2 = 0.4$ indicated by the vertical dark red lines.

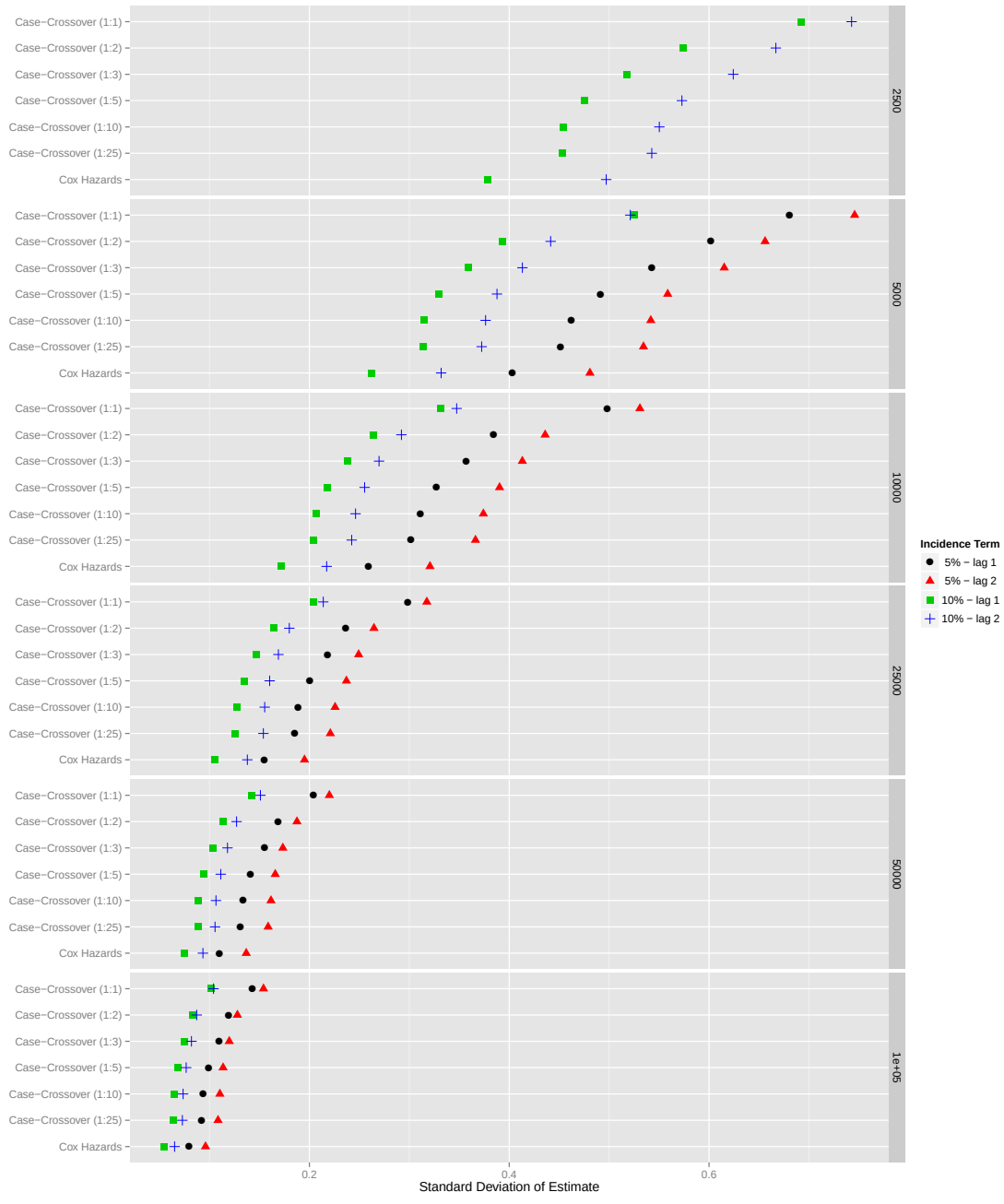


Figure 4.25: Plot of standard deviation of log-hazard ratio estimates by parameter, design, cohort size and incidence rate when assuming a log-hazard ratio of $\beta_1 = 0.8, \beta_2 = 0.4$

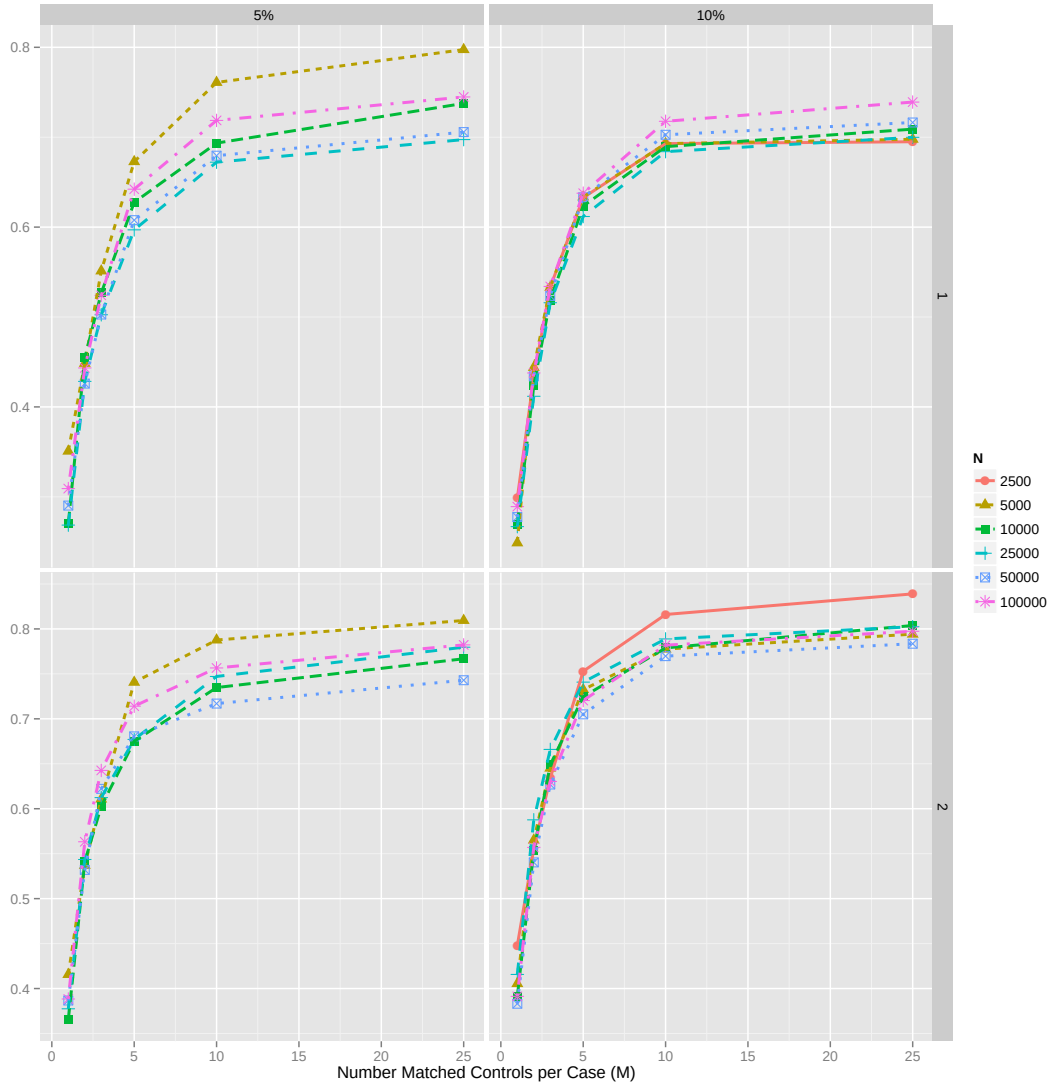


Figure 4.26: Plot of relative efficiency by parameter, design, cohort size and incidence rate when assuming a log-hazard ratios of $\beta_1 = 0.8, \beta_2 = 0.4$

### 4.2.2 Seven Lagged Variables

The final simulation assumes that the effect of a change in the time-dependent exposure is distributed over seven time intervals such that the hazard function is given by

$$\lambda_i(t) = \lambda e^{\beta_1 Z_{it}+\beta_2 Z_{it-1}+\beta_3 Z_{it-2}+\beta_4 Z_{it-3}+\beta_5 Z_{it-4}+\beta_6 Z_{it-5}+\beta_7 Z_{it-6}} \tag{4.6}$$

This hazard will be referred to as the full model since a parameter is used to estimate the effect at each time interval. If the constraint $\Theta_l(\beta, \theta) = \beta\theta^{(l-1)}$ is applied to this equation then the hazard takes the form

$$\lambda_i(t) = \lambda e^{\beta Z_{it}+\beta\theta Z_{it-1}+\beta\theta^2 Z_{it-2}+\beta\theta^3 Z_{it-3}+\beta\theta^4 Z_{it-4}+\beta\theta^5 Z_{it-5}+\beta\theta^6 Z_{it-6}} \tag{4.7}$$

Since this function constrains the seven parameters with a function requiring only two parameters, it will therefore be referred to as the two parameter model. For this simulation $\beta$ is assumed to have an initial effect with a log-hazard of 0.8, and then deteriorate by 0.5 over each time interval so that the assumed hazard is

$$\lambda_i(t) = \lambda e^{0.8Z_{it}+0.4Z_{it-1}+0.2Z_{it-2}+0.1Z_{it-3}+0.05Z_{it-4}+0.025Z_{it-5}+0.0125Z_{it-6}} \tag{4.8}$$

This hazard function is used to generate the survival times using Equation (4.4). In addition to the hazards functions in Equations (4.6) and (4.7), a naive model is also fit to each simulation which assumes the duration of the effect can be estiamted by one parameter for all seven days after the exposure change, or $\beta_i = \beta_j$ for all $i \neq j$. This naive model can be written as

$$\lambda_i(t)_{naive} = \lambda e^{\beta_n Z_{it}+\beta_n Z_{it-1}+\beta_n Z_{it-2}+\beta_n Z_{it-3}+\beta_n Z_{it-4}+\beta_n Z_{it-5}+\beta_n Z_{it-6}} \tag{4.9}$$

which can be written as

$$\lambda_i(t)_{naive} = \lambda e^{\beta_n Z_{it}} \tag{4.10}$$

where

$$Z_{it} = \begin{cases} 1 & t \in (\tau_i, \tau_i + 7] \\ 0 & t \notin (\tau_i, \tau_i + 7] \end{cases} \tag{4.11}$$

The results of estimating parameters assuming the full, two parameter, and naive model are presented in Tables 4.13 and 4.14 for cohorts of sample size 50,000 assuming 10% and

5% incidence rates respectively. The amount of bias for the first lag in the full model are on par with earlier simulations results for the cohorts with 25,000. The 1:1 case-crossover design over estimates the first lag by 2.4%, the second lag at 5.0% and then underestimates the remaining five lags. The first lag is the only parameter estimate from the full model with 95% percentile intervals not overlapping zero for either incidence rate. This is further illustrated by plotting the full distribution of estimates by the lag which is shown in 4.27. The distributions appear to be mostly symmetric. The distribution for the first lag is almost entirely above zero, while the other distributions are centered around the hypothesized value: which approaches zero with larger lags. A significant issue with the full model is that there is insufficient power to detect differences for lagged parameter of order higher than two.

The $\beta$ coefficient in the two parameter model is over estimated by a range of 2.2 to 4% for different $M$ matched case-crossover designs and the $\theta$ coefficient is typically underestimated by approximately 8%. The distribution of the parameters estimated in the two parameter model are displayed in Figure 4.28. The distribution for $\beta$ is mostly symmetrical around 0.8. The estimates for the cohort assuming a 10% incidence is more narrow than the cohort with a 5% incidence rate. There is some indication that the distribution for the 5% estimates is bimodal. The column on the right shows the distribution of $\theta$ in Equation (4.7). The distribution seems to be skewed to the left which may explain whhy the average estimate of $\theta$ is biased towards underestimating the parameter. As $M$ increases for the case-crossover design, the distribution of $\theta$ appears to also become bimodal.

To select which of the three model specifications fit the data best, likelihood ratio tests and AIC are used for assessment. The two parameter model is a special case of the full model since $\Theta_l(\beta, \theta) = \beta\theta^{(l-1)}$. Since the two parameter model is nested within the full model a likelihood ratio test can be conducted where the difference in the degrees of freedom is five. Table 4.16 shows the results of this test. For both 10% and 5% incidence rates, the full model is rejected about 5% of the time; as expected the two-parameter model is a sufficient model compared to full model. A second likelihood ratio test is calculated for each simulation comparing the full and naive models. Equations and demonstrate how the naive model is nested within the full model with a difference in six degrees of freedom.

For the 10% incidence rate, the full model is preferred in at least 97% of the simulations, and by 72% of simulations assuming a 5% incidence rate. A final likelihood ratio test is conducted comparing the two parameter model to the naive model; the two parameter model is the same as the naive model when $\theta = 1$. This likelihood ratio tests prefers the two parameter model in nearly all of the simulations assuming a 10% incidence and more than 92% of the simulations assuming a 5% incidence.

An assessment of AIC for each model is also displayed in Table 4.16. AIC assesses model fit of the model while penalizing the model for using more degrees of freedom. The average AIC is smallest for the two parameter model for all case-crossover designs, and the full model is smaller than the naive model for all case-crossover designs as well. In fact, the two parameter model is selected as the best model in at least 92% of simulations with a 10% incidence, and the full model comes in second being chosen as the best model in about 7% of simulations. The naive model is not observed to be the best model in any of the 1000 simulations assuming a 10% incidence rate. The two parameter model is still preferred in about 90% of the 5% incidence rate simulations, and the full model is in second but the naive model has captured best model in almost 3% of the simulations.

For this simulation a Cox proportional hazards model was fit only for the full model assumption. The Cox proportional hazards model yielded non-overlapping 95% percentile intervals for both the first and second lags, but each estimate tends to underestimate the true value. The distribution of the seventh lagged covariate, also shown in Figure 4.28, may also be bimodal.

| | Mean (SD*) | Median | Power | Coverage Probability | Bias | Avg Std Error** |
|---|---|---|---|---|---|---|
| **Case−Crossover (1:1) − Naive** | | | | | | |
| lag [1−7] | 0.2705 (0.0698) | 0.2685 | 97.10% | NA% | NA% | 0.07 |
| **Case−Crossover (1:3) − Naive** | | | | | | |
| lag [1−7] | 0.2702 (0.0595) | 0.2694 | 99.20% | NA% | NA% | 0.06 |
| **Case−Crossover (1:5) − Naive** | | | | | | |
| lag [1−7] | 0.2701 (0.0581) | 0.2693 | 99.40% | NA% | NA% | 0.06 |
| **Case−Crossover (1:10) − Naive** | | | | | | |
| lag [1−7] | 0.2702 (0.0582) | 0.2693 | 99.40% | NA% | NA% | 0.06 |
| **Case−Crossover (1:3) − full** | | | | | | |
| lag 1 | 0.8096 (0.1427) | 0.8104 | 100.00% | 94.20% | 1.20% | 0.14 |
| lag 2 | 0.4011 (0.1543) | 0.4070 | 74.50% | 95.20% | 0.28% | 0.15 |
| lag 3 | 0.2017 (0.1644) | 0.1995 | 24.90% | 94.70% | 0.85% | 0.16 |
| lag 4 | 0.0990 (0.1681) | 0.1085 | 8.00% | 95.60% | −1.02% | 0.17 |
| lag 5 | 0.0495 (0.1733) | 0.0601 | 6.00% | 95.80% | −1.04% | 0.17 |
| lag 6 | 0.0233 (0.1736) | 0.0302 | 4.90% | 95.40% | −6.61% | 0.17 |
| lag 7 | 0.0083 (0.1738) | 0.0097 | 5.10% | 95.30% | −33.54% | 0.17 |
| **Case−Crossover (1:5) − full** | | | | | | |
| lag 1 | 0.8083 (0.1373) | 0.8054 | 100.00% | 94.30% | 1.04% | 0.13 |
| lag 2 | 0.3992 (0.1507) | 0.4010 | 75.70% | 95.70% | −0.19% | 0.15 |
| lag 3 | 0.2005 (0.1605) | 0.2018 | 24.30% | 94.90% | 0.24% | 0.16 |
| lag 4 | 0.0982 (0.1661) | 0.1090 | 8.80% | 95.60% | −1.80% | 0.17 |
| lag 5 | 0.0481 (0.1702) | 0.0567 | 5.80% | 95.90% | −3.74% | 0.17 |
| lag 6 | 0.0238 (0.1699) | 0.0300 | 5.10% | 96.00% | −4.86% | 0.17 |
| lag 7 | 0.0092 (0.1711) | 0.0084 | 5.90% | 94.80% | −26.76% | 0.17 |
| **Case−Crossover (1:10) − full** | | | | | | |
| lag 1 | 0.8086 (0.1370) | 0.8058 | 100.00% | 94.60% | 1.08% | 0.13 |
| lag 2 | 0.3992 (0.1508) | 0.3998 | 75.50% | 95.30% | −0.19% | 0.15 |
| lag 3 | 0.2006 (0.1604) | 0.2045 | 24.40% | 94.90% | 0.29% | 0.16 |
| lag 4 | 0.0983 (0.1662) | 0.1064 | 8.90% | 95.60% | −1.73% | 0.17 |
| lag 5 | 0.0481 (0.1702) | 0.0568 | 6.00% | 95.70% | −3.77% | 0.17 |
| lag 6 | 0.0243 (0.1701) | 0.0281 | 5.00% | 96.10% | −2.78% | 0.17 |
| lag 7 | 0.0091 (0.1709) | 0.0084 | 5.90% | 95.10% | −27.50% | 0.17 |
| **Cox Hazards (Time Dep. Covar) − full** | | | | | | |
| lag 1 | 0.7984 (0.0792) | 0.8003 | 100.00% | 94.90% | −0.20% | 0.08 |
| lag 2 | 0.3977 (0.0959) | 0.4011 | 97.40% | 95.10% | −0.56% | 0.10 |
| lag 3 | 0.1998 (0.1066) | 0.2013 | 47.90% | 95.00% | −0.10% | 0.11 |
| lag 4 | 0.0946 (0.1171) | 0.0990 | 15.70% | 94.80% | −5.38% | 0.11 |
| lag 5 | 0.0456 (0.1221) | 0.0460 | 9.10% | 95.20% | −8.89% | 0.12 |
| lag 6 | 0.0180 (0.1262) | 0.0259 | 7.00% | 94.70% | −27.87% | 0.12 |
| lag 7 | 0.0090 (0.1259) | 0.0166 | 6.00% | 94.80% | −27.78% | 0.12 |
| **Case−Crossover (1:1) − 2 parameter** | | | | | | |
| beta | 0.8206 (0.1670) | 0.8183 | 100.00% | 92.00% | 2.58% | 0.15 |
| theta | 0.4812 (0.1703) | 0.5011 | 84.00% | 88.40% | −3.77% | 0.14 |
| **Case−Crossover (1:3) − 2 parameter** | | | | | | |
| beta | 0.8131 (0.1389) | 0.8085 | 100.00% | 91.50% | 1.64% | 0.12 |
| theta | 0.4883 (0.1419) | 0.5024 | 89.70% | 89.80% | −2.34% | 0.12 |
| **Case−Crossover (1:5) − 2 parameter** | | | | | | |
| beta | 0.8111 (0.1340) | 0.8060 | 100.00% | 91.50% | 1.38% | 0.12 |
| theta | 0.4887 (0.1408) | 0.4987 | 91.60% | 89.60% | −2.27% | 0.12 |
| **Case−Crossover (1:10) − 2 parameter** | | | | | | |
| beta | 0.8113 (0.1337) | 0.8059 | 100.00% | 91.50% | 1.41% | 0.12 |
| theta | 0.4894 (0.1374) | 0.5005 | 91.70% | 89.80% | −2.11% | 0.12 |

Table 4.13: Results from simulation of a cohort with an incidence of approximately 10% ( $\lambda = .0025$) incidence and assuming a deteriorating effect over seven periods starting at $\beta = .8$ and geometrical decreasing at a rate of $\theta = .5$ (* Monte Carlo standard deviation of estimates ** Average of the information based standard error)

| | Mean (SD*) | Median | Power | Coverage Probability | Bias | Avg Std Error** |
|---|---|---|---|---|---|---|
| **Case−Crossover (1:1) − Naive** | | | | | | |
| lag [1−7] | 0.2628 (0.0951) | 0.2626 | 78.30% | NA% | NA% | 0.10 |
| **Case−Crossover (1:3) − Naive** | | | | | | |
| lag [1−7] | 0.2613 (0.0809) | 0.2610 | 88.50% | NA% | NA% | 0.08 |
| **Case−Crossover (1:5) − Naive** | | | | | | |
| lag [1−7] | 0.2610 (0.0796) | 0.2626 | 89.50% | NA% | NA% | 0.08 |
| **Case−Crossover (1:10) − Naive** | | | | | | |
| lag [1−7] | 0.2610 (0.0795) | 0.2627 | 89.10% | NA% | NA% | 0.08 |
| **Case−Crossover (1:3) − full** | | | | | | |
| lag 1 | 0.8013 (0.1961) | 0.8024 | 98.30% | 95.20% | 0.16% | 0.19 |
| lag 2 | 0.3819 (0.2131) | 0.3813 | 44.10% | 95.60% | −4.52% | 0.21 |
| lag 3 | 0.1830 (0.2281) | 0.1922 | 13.70% | 96.00% | −8.50% | 0.23 |
| lag 4 | 0.0952 (0.2418) | 0.1094 | 7.60% | 94.70% | −4.81% | 0.24 |
| lag 5 | 0.0425 (0.2373) | 0.0447 | 4.90% | 95.90% | −15.03% | 0.24 |
| lag 6 | 0.0046 (0.2435) | 0.0121 | 5.40% | 94.80% | −81.59% | 0.24 |
| lag 7 | −0.0025 (0.2555) | 0.0000 | 5.20% | 94.70% | −119.71% | 0.24 |
| **Case−Crossover (1:5) − full** | | | | | | |
| lag 1 | 0.7999 (0.1919) | 0.8012 | 98.20% | 94.70% | −0.01% | 0.19 |
| lag 2 | 0.3813 (0.2084) | 0.3836 | 45.90% | 95.20% | −4.68% | 0.21 |
| lag 3 | 0.1821 (0.2245) | 0.1930 | 14.60% | 95.80% | −8.94% | 0.22 |
| lag 4 | 0.0946 (0.2362) | 0.1069 | 7.40% | 94.60% | −5.44% | 0.23 |
| lag 5 | 0.0409 (0.2326) | 0.0522 | 4.40% | 96.30% | −18.21% | 0.24 |
| lag 6 | 0.0019 (0.2383) | 0.0081 | 5.00% | 95.00% | −92.50% | 0.24 |
| lag 7 | −0.0050 (0.2496) | 0.0028 | 5.50% | 94.20% | −139.97% | 0.24 |
| **Case−Crossover (1:10) − full** | | | | | | |
| lag 1 | 0.7994 (0.1908) | 0.8020 | 98.40% | 94.80% | −0.07% | 0.19 |
| lag 2 | 0.3819 (0.2080) | 0.3840 | 45.30% | 95.30% | −4.53% | 0.21 |
| lag 3 | 0.1822 (0.2242) | 0.1908 | 14.70% | 95.90% | −8.90% | 0.22 |
| lag 4 | 0.0946 (0.2353) | 0.1041 | 7.60% | 95.10% | −5.36% | 0.23 |
| lag 5 | 0.0406 (0.2327) | 0.0519 | 4.60% | 96.50% | −18.71% | 0.24 |
| lag 6 | 0.0012 (0.2376) | 0.0044 | 4.90% | 95.20% | −95.36% | 0.24 |
| lag 7 | −0.0051 (0.2490) | 0.0023 | 5.70% | 94.50% | −140.51% | 0.24 |
| **Cox Hazards (Time Dep. Covar) − full** | | | | | | |
| lag 1 | 0.7960 (0.1079) | 0.7962 | 100.00% | 95.30% | −0.50% | 0.11 |
| lag 2 | 0.3875 (0.1340) | 0.3935 | 79.60% | 95.00% | −3.12% | 0.13 |
| lag 3 | 0.1859 (0.1465) | 0.1909 | 26.50% | 95.70% | −7.04% | 0.15 |
| lag 4 | 0.0881 (0.1591) | 0.0991 | 9.40% | 96.30% | −11.90% | 0.16 |
| lag 5 | 0.0373 (0.1600) | 0.0450 | 5.60% | 96.10% | −25.44% | 0.17 |
| lag 6 | 0.0015 (0.1744) | 0.0108 | 5.00% | 95.30% | −93.89% | 0.17 |
| lag 7 | −0.0080 (0.1794) | 0.0023 | 4.30% | 96.00% | −163.79% | 0.18 |
| **Case−Crossover (1:1) − 2 parameter** | | | | | | |
| beta | 0.8268 (0.2389) | 0.8252 | 99.40% | 91.10% | 3.35% | 0.21 |
| theta | 0.4560 (0.2346) | 0.4830 | 64.60% | 87.80% | −8.80% | 0.19 |
| **Case−Crossover (1:3) − 2 parameter** | | | | | | |
| beta | 0.8066 (0.1918) | 0.8080 | 99.90% | 91.20% | 0.83% | 0.17 |
| theta | 0.4666 (0.2074) | 0.4831 | 72.00% | 86.70% | −6.67% | 0.17 |
| **Case−Crossover (1:5) − 2 parameter** | | | | | | |
| beta | 0.8054 (0.1872) | 0.8066 | 99.90% | 91.00% | 0.67% | 0.16 |
| theta | 0.4660 (0.2062) | 0.4809 | 72.90% | 86.10% | −6.80% | 0.17 |
| **Case−Crossover (1:10) − 2 parameter** | | | | | | |
| beta | 0.8052 (0.1861) | 0.8089 | 99.90% | 91.00% | 0.65% | 0.16 |
| theta | 0.4660 (0.2057) | 0.4820 | 72.80% | 86.50% | −6.80% | 0.17 |

Table 4.14: Results from simulation of a cohort with an incidence of approximately 5% ( $\lambda = .00125$) incidence and assuming a deteriorating effect over seven periods starting at $\beta = .8$ and geometrical decreasing at a rate of $\theta = .5$ (* Monte Carlo standard deviation of estimates ** Average of the information based standard error)

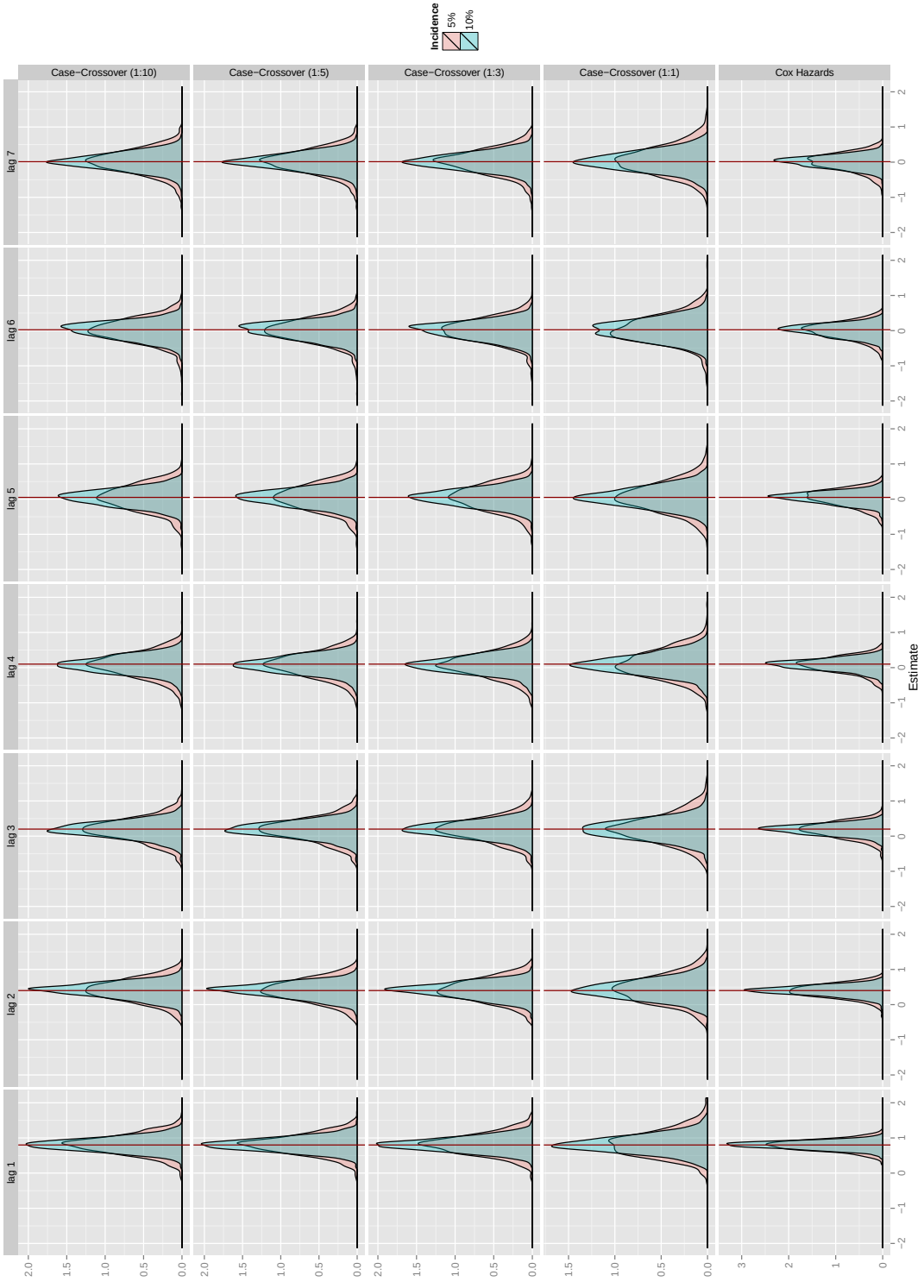


Figure 4.27: Distribution of estimate from full model displayed parameter, design, cohort size and incidence rate when assuming a distributed lag effect with constrained parameters $\beta = 0.8, \theta = 0.5$ indicated by the vertical dark red lines.

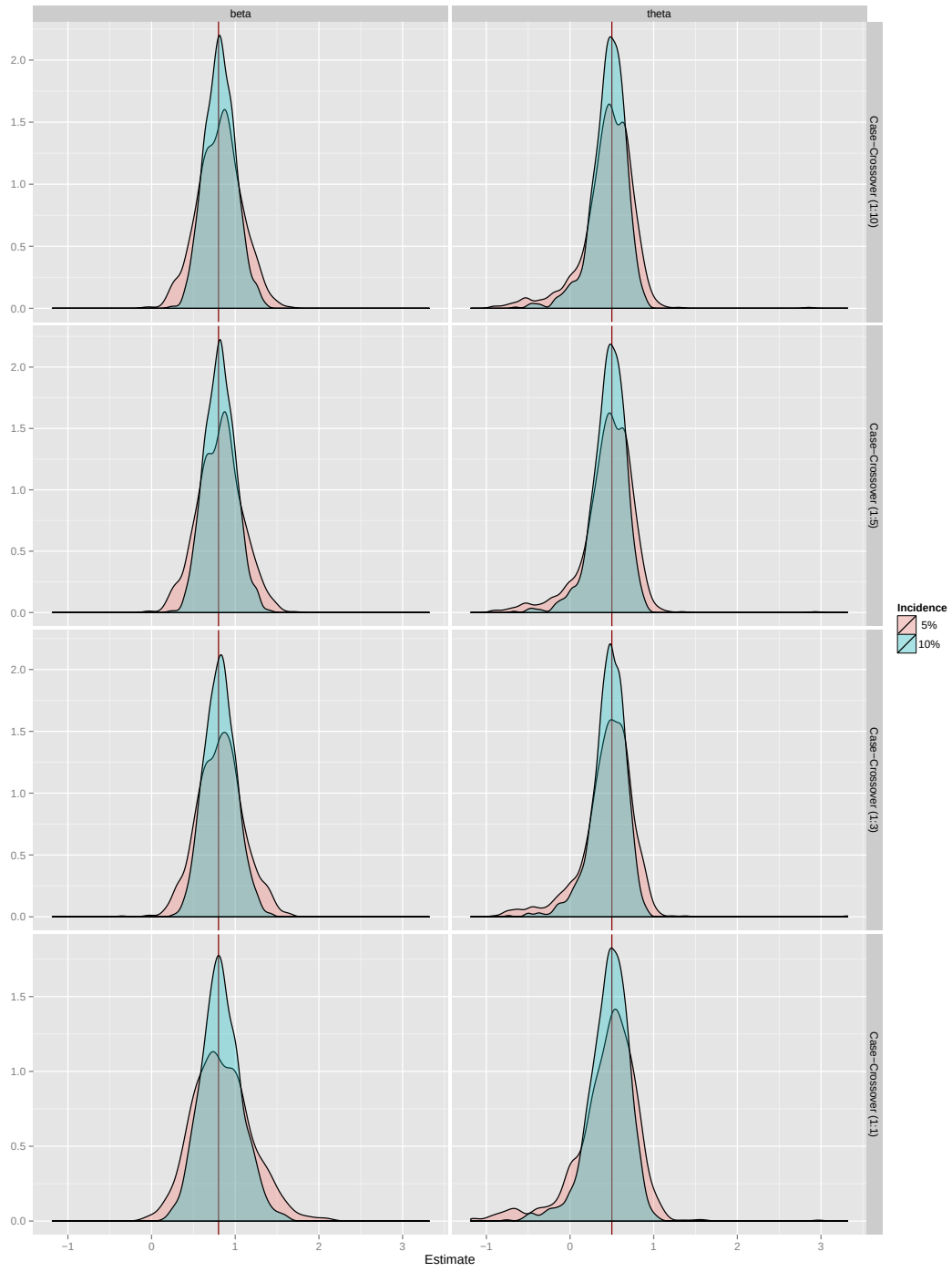


Figure 4.28: Distribution of estimates from two parameter model displayed by parameter, design, cohort size and incidence rate when assuming a distributed lag effect with constrained parameters $\beta = 0.8, \theta = 0.5$ indicated by the vertical dark red lines.

| | Likelihood Ratio Test (% Reject Null) | | | Average AIC | | | Preferred Model via AIC | | |
|---|---|---|---|---|---|---|---|---|---|
| | Full vs 2−parm | Full vs Naive | 2−parm vs Naive | Full | 2−parm | Naive | Full | 2−parm | Naive |
| **10%** | | | | | | | | | |
| Case−Crossover (1:3) | 5.20% | 97.30% | 99.90% | 6974 | 6969 | 6992 | 7.50% | 92.50% | 0.00% |
| Case−Crossover (1:5) | 4.80% | 97.70% | 99.90% | 7545 | 7540 | 7563 | 7.00% | 93.00% | 0.00% |
| Case−Crossover (1:10) | 4.70% | 97.90% | 99.90% | 7604 | 7599 | 7622 | 7.30% | 92.70% | 0.00% |
| **5%** | | | | | | | | | |
| Case−Crossover (1:3) | 4.80% | 72.40% | 92.20% | 3633 | 3628 | 3639 | 7.40% | 89.80% | 2.80% |
| Case−Crossover (1:5) | 4.40% | 73.10% | 92.30% | 3935 | 3930 | 3941 | 6.80% | 90.60% | 2.60% |
| Case−Crossover (1:10) | 4.50% | 73.40% | 92.40% | 3966 | 3961 | 3973 | 6.80% | 90.90% | 2.30% |

Table 4.15: Results of likelihood ratio tests and assessment of AIC from simulations of cohorts with incidence rates of approximately 10% ( $\lambda = .0025$) and 5% ( $\lambda = .00125$)

| | Mean (SD) | Median | Coverage Probability | Bias |
|---|---|---|---|---|
| **Case−Crossover (1:5) − 2 parameter** | | | | |
| beta | 0.8111 (0.1341) | 0.8062 | 91.40% | 1.39% |
| theta | 0.4900 (0.1348) | 0.4989 | 90.00% | −2.00% |

Table 4.16: Results applying the profile likelihood approach from simulations of a cohort with an incidence of approximately 10% ( $\lambda = .0025$) incidence and assuming a deteriorating effect over seven periods starting at $\beta = .8$ and geometrical decreasing at a rate of $\theta = .5$ (* Monte Carlo standard deviation of estimates ** Average of the information based standard error)

# Chapter 5

# Analysis of the VNSNY Medication and Fall Data

This chapter presents an analysis of the VNSNY Medication and Fall data. This thesis focuses on an exposure which indicates an increase in the number of medications classified as either an antidepressant or pyschostimulant. The first section provides an overview of the VNSNY Medication and Fall data and the home health care setting. Section 5.2 conducts an empirical exploration of the patient characteristics present at baseline that predict the risk of falling during home health care. The factors identified from this section are treated as confounders in the cohort designs in the following sections.

The remainder of the chapter explores the transient effect of falling resulting from an increase in antidepressants or pyschostimulants by applying the case-crossover and cohort designs. The initial approach to this data analysis is based on previous studies which explored how changes in medications may trigger a fall by deploying the case-crossover designs [71, 87, 88]. These studies reported several associations for various medication classifications, but limited their analysis by assuming the effect lasted no more than four days. This chapter initially follows the methodology of these studies, but then extends the duration for up to 28 days, and concludes with an exploration of possible non-linear deteriorating effect.

## 5.1 Setting

From 2010 to 2011 there were 192,438 admissions into Visiting Nurse Service of New York's (VNSNY) certified home health care program (CHHA). Of these admissions, 4.0% (7840), displayed in Figure 5.1, were documented in the electronic health record to have fallen at least once while receiving home health care services. The date of the first fall is recorded in agency systems and is the primary outcome in this study.

Two Kaplan Meier survival curves are plotted in Figure 5.2: the time to first fall and time to discharge. From this plot, it can be seen that a majority of patients are discharged by 60 days from admission. In addition, there are two noticeable points in the time to discharge curve, these occur at 60 and 120 days. These shifts can be explained by the Medicare and Medicaid reimbursement rules for home health care services. When a patient is admitted into home health care, the Center for Medicare and Medicaid Services requires a registered nurse or physical therapist complete a clinical assessment, known as the Outcomes and Assessment Information Set (OASIS), and have physician orders justifying that home health care services are necessary for the patient [16]. These requirements certify a 60 day home health care episode [26]. When the 60 days near completion, the nurse or therapist must again recertify that the patient requires another 60 days of home health care services. As a result, many patients are discharged from home health care in the final days of each sixty day interval.

Figure 5.3 overlays parametric fits to the time to first event on top of the Kaplan Meier curve. From this plot, the exponential distribution appears to not fit the survival curve over the entire period. However, nearly 90% of patients are discharged from home health care by 60 days, and during this period the exponential distribution does appear to fit the Kaplan Meier curve. The case-crossover design is appropriate since the assumption of constant or small hazard is valid.

The OASIS assessment comprises of much of the information that is known about the patient. The OASIS requires the clinician to asses the patient status in several domains such as: demographics, patient history, home health care diagnosis, living arrangements, sensory, integumentary, respiratory, cardiac status, elimination, neuro/emotional/behavior, activities of daily living, instrumental activities of daily living, medications, and care management. Many of these factors could confound a relationship between a medication change and fall. The results from the OASIS are stored in the agencies electronic health records and are stored in an Oracle data warehouse.

The nurse or therapist must also review the medication regimen at start of care and document into the plan of care. The plan of care requires approval from a medical doctor in order for the home health care episode to be valid. Any time the home health care clinician

observes a change in the medication regimen after admission, he or she must document the updates into the electronic health record. The home health care patient's time-dependent, medication regimen is also stored in a data warehouse and is the source for defining the exposure of interest in this dissertation.

The last component of the electronic health record utilized in this study is the identification of falls. The nurse or therapist provide home visits throughout the episode of care. In any visit, if the patient, care giver, or home health aide inform the nurse or therapist that the patient has fallen since the last visit, the clinician documents the date of fall into the patient electronic health record.

The VNSNY CHHA population is located in the New York City greater metropolitan area. This includes the five boroughs of New York City (i.e Bronx, Brooklyn, Manhattan, Staten Island, Queens) and Westchester and Nassau counties. The population has a diverse racial, linguistic, and minority profile. These demographics and other characteristics about the sample can be observed in Table 5.1. Since this population is home bound and receiving post acute services, a majority of the population is older than 65. In addition to patient demographics, Table 5.1 includes factors that are known to increase the risk of falling for the patient such as risk of falls, frailty, ability to ambulate, transfer, and use the toilet, etc. Nearly all factors in Table 5.1 are derived from the OASIS , and no time-varying measures are included in this table. The diagnosis classifications used in this analysis are based on the chronic conditions in home care work by Murtaugh, Peng and colleagues [68]. Table 5.2 contains descriptive statistics of the medication regimen present at admission into home health care. These measures are derived from First Databank's therapeutic classifications.

Tables 5.1 and 5.2 often serve as a first step in a cohort design to describe the characteristics associated with patients who experienced the event; results from this table are consistent with the research, outlined in Section 2.3, identifying patient characteristics associated with falling. For example, of those patients who fell while receiving home health care services, 32.9% were taking medications classified as psychostimulants/antidepressants at admission compared to 19.0% of those who did not fall. Section 5.2 seeks to identify which of these baseline factors are associated with a fall.

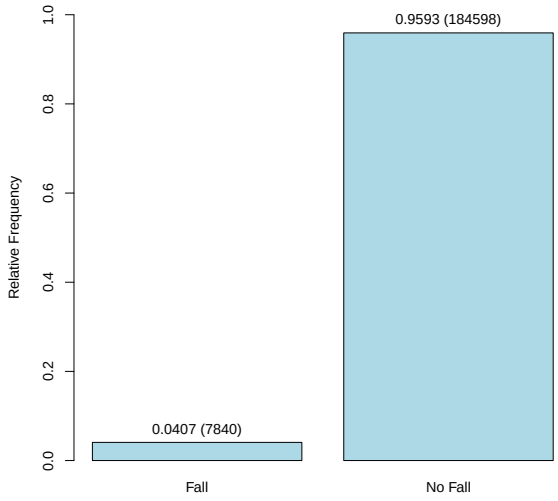


Figure 5.1: Barplot of the Incidence of Falling in the VNSNY Medication and Fall Data

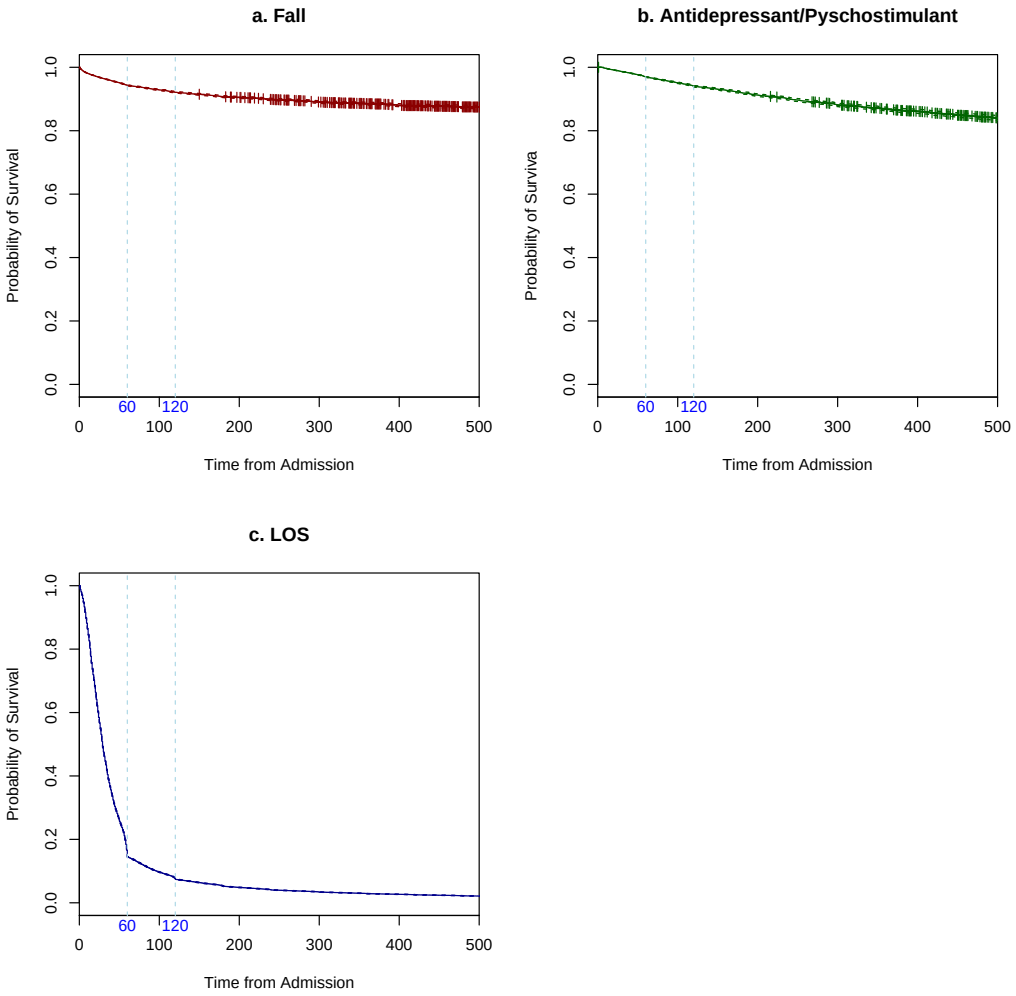


Figure 5.2: Kaplan Meier plots for time to: a.) first fall, b.) increase in antidepressants/pyschostimlatns, c.) discharge (i.e. length of service)

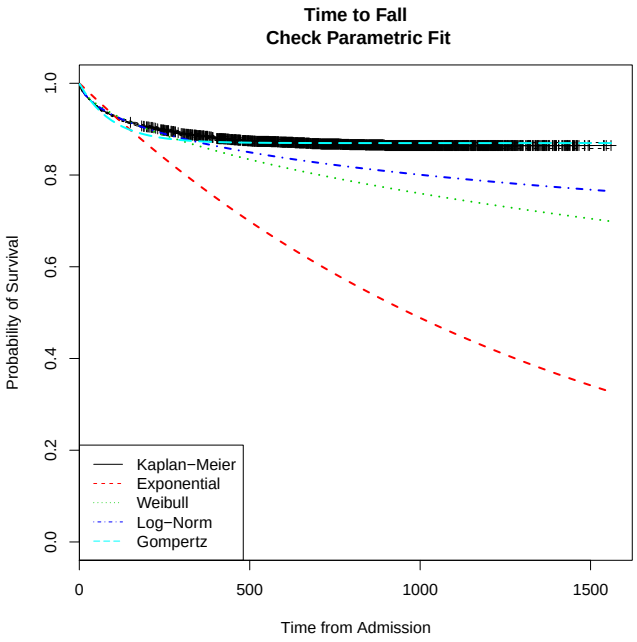


Figure 5.3: Kaplan Meier plot of time to fall with parametric fit overlays

| | No Fall | Fall |
|---|---|---|
| **Age Group** | | |
| 18 to 54 | 16.69% ( 30807) | 10.34% ( 811) |
| 55 to 64 | 16.16% ( 29835) | 12.07% ( 946) |
| 65 to 74 | 20.32% ( 37519) | 18.32% (1436) |
| 75 to 84 | 25.92% ( 47842) | 30.24% (2371) |
| 85 + | 20.90% ( 38587) | 29.03% (2276) |
| missing | 0.00% ( 8) | 0.00% ( 0) |
| **Demographics** | | |
| Female | 61.49% (113511) | 61.68% (4836) |
| **Borough** | | |
| Bronx | 15.42% ( 28471) | 15.22% (1193) |
| Brooklyn | 20.02% ( 36960) | 13.48% (1057) |
| Manhattan | 25.25% ( 46612) | 28.78% (2256) |
| Nassau | 6.96% ( 12840) | 10.27% ( 805) |
| Queens | 23.07% ( 42581) | 21.91% (1718) |
| Staten Is | 5.21% ( 9613) | 4.62% ( 362) |
| Westchester | 4.07% ( 7521) | 5.73% ( 449) |
| **Race** | | |
| American Indian or Alaska Native | 0.34% ( 634) | 0.28% ( 22) |
| Asian | 5.72% ( 10565) | 4.35% ( 341) |
| Black | 24.49% ( 45206) | 18.69% (1465) |
| Hispanic | 22.58% ( 41691) | 17.92% (1405) |
| Pacific Islander | 0.55% ( 1012) | 0.41% ( 32) |
| White | 46.64% ( 86090) | 58.64% (4597) |
| Unkown | 0.00% ( 5) | 0.00% ( 0) |
| **Language Spoken** | | |
| English | 66.13% (122068) | 76.05% (5962) |
| Spanish | 16.19% ( 29878) | 13.42% (1052) |
| Other | 17.69% ( 32652) | 10.54% ( 826) |
| **Prior Conditions** | | |
| Urinary Incontinence | 14.13% ( 26087) | 22.51% (1765) |
| Idwelling/Suprapubic Catheter | 1.19% ( 2188) | 1.35% ( 106) |
| Intractable Pain | 8.27% ( 15272) | 7.54% ( 591) |
| Impaired Decision Making | 7.69% ( 14193) | 10.42% ( 817) |
| Disruptive Behavior | 0.84% ( 1557) | 1.15% ( 90) |
| Memory Loss | 5.34% ( 9860) | 7.65% ( 600) |
| **Home Therapies** | | |
| Enteral Nutrition | 1.36% ( 2510) | 0.91% ( 71) |
| Parenteral Nutrition | 0.12% ( 214) | 0.06% ( 5) |
| Intravenous | 1.32% ( 2428) | 1.05% ( 82) |
| **Risk for Hospitalization** | | |
| Decline in Mental, Emotional, Behavioral | 11.52% ( 21267) | 18.53% (1453) |
| Multi Hospitalization | 27.98% ( 51643) | 34.85% (2732) |
| History of Falls | 13.65% ( 25197) | 33.70% (2642) |
| Five or More Medications | 64.06% (118249) | 72.61% (5693) |
| Frailty Indicators | 23.28% ( 42983) | 31.94% (2504) |
| **Overall Status** | | |
| 0−Stable | 22.62% ( 41756) | 17.36% (1361) |
| 1−Likely to Stable | 58.42% (107836) | 57.77% (4529) |
| 2−Fragile | 17.01% ( 31402) | 22.46% (1761) |
| 3−Serious | 1.30% ( 2399) | 1.53% ( 120) |
| Unknown | 0.65% ( 1205) | 0.88% ( 69) |
| **Risk Factors** | | |
| Smoking | 6.27% ( 11573) | 6.89% ( 540) |
| Obesity | 12.86% ( 23738) | 12.67% ( 993) |
| Alcohol | 1.39% ( 2563) | 1.67% ( 131) |
| Drug | 0.96% ( 1770) | 1.19% ( 93) |
| **Lives With** | | |
| Alone | 24.97% ( 46102) | 29.30% (2297) |
| Family | 31.40% ( 57962) | 30.13% (2362) |
| Other | 3.38% ( 6233) | 3.95% ( 310) |
| Significant Other | 0.77% ( 1419) | 0.61% ( 48) |
| Spouse | 17.26% ( 31857) | 18.24% (1430) |
| Unknown | 22.22% ( 41025) | 17.77% (1393) |

Table 5.1: Descriptives of Baseline Patient Characteristics

| | No Fall | Fall |
|---|---|---|
| **Availabiity of Assistance** | | |
| 1−Around Clock | 65.50% (120909) | 65.60% (5143) |
| 2−Regular Day | 9.80% ( 18092) | 9.41% ( 738) |
| 3−Regular Night | 9.18% ( 16952) | 9.23% ( 724) |
| 4−Occassional | 13.11% ( 24209) | 13.57% (1064) |
| 5−None Available | 2.40% ( 4436) | 2.18% ( 171) |
| **Vision Impairment (0−2)** | | |
| 0−Normal | 81.46% (150365) | 77.26% (6057) |
| 1−Partially | 17.13% ( 31618) | 21.22% (1664) |
| 2−Severely | 1.42% ( 2615) | 1.52% ( 119) |
| **Hearing Impairment (0−4)** | | |
| 0−Adequate | 81.73% (150880) | 74.83% (5867) |
| 1−Mild | 17.03% ( 31437) | 24.09% (1889) |
| 2−Severely | 0.91% ( 1674) | 1.03% ( 81) |
| Unable | 0.33% ( 607) | 0.04% ( 3) |
| **Understand Verbal Content (0−3)** | | |
| 0−Understands | 78.71% (145295) | 72.97% (5721) |
| 1−Usually | 17.06% ( 31501) | 22.92% (1797) |
| 2−Sometimes | 3.17% ( 5853) | 3.64% ( 285) |
| 3−Rarely | 0.40% ( 733) | 0.28% ( 22) |
| Unknown | 0.66% ( 1216) | 0.19% ( 15) |
| **Speech Impairment (0−5)** | | |
| 0−No Impairment | 73.61% (135889) | 66.49% (5213) |
| 1−Minimal | 18.64% ( 34414) | 24.59% (1928) |
| 2−Moderdate | 4.42% ( 8151) | 6.30% ( 494) |
| 3−Severe | 1.78% ( 3285) | 1.81% ( 142) |
| 4−Unable | 0.96% ( 1772) | 0.59% ( 46) |
| 5−Nonresponsive | 0.59% ( 1087) | 0.22% ( 17) |
| **Frequency of Pain (0−4)** | | |
| 0−None | 36.53% ( 67440) | 36.56% (2866) |
| 1−No Interference Activity | 11.12% ( 20532) | 10.32% ( 809) |
| 2−Less than Daily | 10.77% ( 19889) | 11.01% ( 863) |
| 3−Daily Not Constant | 37.82% ( 69813) | 37.32% (2926) |
| 4−All the Time | 3.75% ( 6924) | 4.80% ( 376) |
| **Stage of Most Problematic Pressure Ulcer (1−4)** | | |
| 0−None | 94.18% (173851) | 91.56% (7178) |
| 1−Stage I | 1.21% ( 2228) | 2.10% ( 165) |
| 2−Stage II | 2.68% ( 4942) | 3.66% ( 287) |
| 3−Stage III | 0.90% ( 1659) | 1.08% ( 85) |
| 4−Stage IV | 0.45% ( 822) | 0.56% ( 44) |
| Unstageable | 0.59% ( 1096) | 1.03% ( 81) |
| **Status of Most Problematic Stasis Ulcer (0−3)** | | |
| 0−Newly Epithelialized | 0.04% ( 76) | 0.03% ( 2) |
| 1−Fully Granulating | 0.10% ( 182) | 0.14% ( 11) |
| 2−Early/Partial Granulation | 0.37% ( 679) | 0.59% ( 46) |
| 3−Not Healing | 1.28% ( 2368) | 1.91% ( 150) |
| None | 98.04% (180974) | 97.16% (7617) |
| Not Observ | 0.17% ( 319) | 0.18% ( 14) |
| **Status of Most Problematic Surgical Wound (0−3)** | | |
| 0−Newly Epithelialized | 5.30% ( 9792) | 3.52% ( 276) |
| 1−Fully Granulating | 3.82% ( 7054) | 2.03% ( 159) |
| 2−Early/Partial Granulation | 6.94% ( 12802) | 3.16% ( 248) |
| 3−Not Healing | 10.68% ( 19711) | 5.68% ( 445) |
| None | 70.05% (129319) | 83.86% (6575) |
| Not Observ | 3.21% ( 5920) | 1.75% ( 137) |
| **Dyspnea (0−4)** | | |
| 0−Never | 56.52% (104341) | 51.30% (4022) |
| 1−Walk 20ft/Stairs | 24.90% ( 45960) | 26.19% (2053) |
| 2−Moderate Exertion | 13.78% ( 25431) | 16.56% (1298) |
| 3−Minimal Exertion | 3.74% ( 6899) | 4.78% ( 375) |
| 4−At Rest | 1.07% ( 1967) | 1.17% ( 92) |
| **Respiratory Treatments** | | |
| Oxygen | 5.79% ( 10681) | 6.28% ( 492) |
| Ventilator | 0.09% ( 174) | 0.06% ( 5) |
| Airway Pressure | 0.57% ( 1049) | 0.50% ( 39) |

Table 5.1 (contd.): Descriptives of Baseline Patient Characteristics

| | No Fall | Fall |
|---|---|---|
| **Urinary Tract Infection** | | |
| Infection | 5.61% ( 10351) | 7.56% ( 593) |
| None | 93.72% (173003) | 91.65% (7185) |
| Prophylatctic Trt | 0.22% ( 400) | 0.27% ( 21) |
| Unknown | 0.46% ( 844) | 0.52% ( 41) |
| **When Urinary Incontinence Occurs** | | |
| 0−Time−Voiding | 8.44% ( 15581) | 12.55% ( 984) |
| 1−Occasional Stress | 3.74% ( 6898) | 5.60% ( 439) |
| 2−Night Only | 1.44% ( 2662) | 2.27% ( 178) |
| 3−Day Only | 0.28% ( 526) | 0.38% ( 30) |
| 4−Day and Night | 14.80% ( 27323) | 22.47% (1762) |
| No Incontinence | 68.36% (126196) | 53.53% (4197) |
| Urinary Catheter | 2.93% ( 5412) | 3.19% ( 250) |
| **Bowel Incontinence Frequency** | | |
| 0−Never/Rarely | 86.35% (159400) | 81.86% (6418) |
| 1−Less than 1/week | 2.40% ( 4432) | 4.53% ( 355) |
| 2−1to3/week | 2.93% ( 5400) | 5.13% ( 402) |
| 3−4to6/week | 1.56% ( 2877) | 2.18% ( 171) |
| 4−Daily | 3.83% ( 7069) | 4.23% ( 332) |
| 5−More than 1 Daily | 0.97% ( 1782) | 0.96% ( 75) |
| Bowel Ostomy | 1.95% ( 3593) | 1.08% ( 85) |
| Unknown | 0.02% ( 45) | 0.03% ( 2) |
| **Cognitive Functioning (0−4)** | | |
| 0−Alert | 70.10% (129410) | 59.55% (4669) |
| 1−Prompting | 20.84% ( 38465) | 29.52% (2314) |
| 2−Some Asst | 6.08% ( 11228) | 8.53% ( 669) |
| 3−Consderable Asst | 2.16% ( 3995) | 2.14% ( 168) |
| 4−Dependent | 0.81% ( 1500) | 0.26% ( 20) |
| **Confusion (0−4)** | | |
| 0−Never | 58.75% (108458) | 47.73% (3742) |
| 1−New Situation | 32.92% ( 60762) | 41.35% (3242) |
| 2−On Awakening | 0.74% ( 1363) | 1.34% ( 105) |
| 3−Day and Evening | 5.18% ( 9565) | 7.60% ( 596) |
| 4−Constantly | 2.08% ( 3848) | 1.90% ( 149) |
| Nonresponsive | 0.33% ( 602) | 0.08% ( 6) |
| **Anxious (0−3)** | | |
| 0−None | 67.81% (125173) | 61.56% (4826) |
| 1−Less than Daily | 18.57% ( 34288) | 20.69% (1622) |
| 2−Daily, not Constant | 11.99% ( 22137) | 15.64% (1226) |
| 3−All of the Time | 1.13% ( 2080) | 1.91% ( 150) |
| Nonresponsive | 0.50% ( 920) | 0.20% ( 16) |
| **Cognitive, Behavioral, Psychiatric Symptoms** | | |
| Memory Deficity | 6.45% ( 11911) | 9.76% ( 765) |
| Impaired Decision Making | 9.50% ( 17542) | 12.65% ( 992) |
| Verbal Disruption | 0.69% ( 1273) | 1.17% ( 92) |
| Physical Aggression | 0.32% ( 594) | 0.42% ( 33) |
| Disruptive | 0.47% ( 867) | 0.59% ( 46) |
| Delusional | 0.55% ( 1016) | 0.61% ( 48) |
| **Frequency of Disruptive Behavior** | | |
| 0−Never | 95.70% (176668) | 93.57% (7336) |
| 1−Less than 1/month | 1.33% ( 2447) | 1.68% ( 132) |
| 2−1/month | 0.16% ( 298) | 0.37% ( 29) |
| 3−Several/month | 0.55% ( 1020) | 0.70% ( 55) |
| 4−Several/week | 0.85% ( 1561) | 1.35% ( 106) |
| 5−Daily | 1.41% ( 2604) | 2.32% ( 182) |
| **Grooming (0−3)** | | |
| 0−Able | 34.20% ( 63134) | 21.91% (1718) |
| 1−Utensil Asst | 34.93% ( 64489) | 36.57% (2867) |
| 2−Assistance | 23.39% ( 43169) | 33.71% (2643) |
| 3−Dependent | 7.48% ( 13806) | 7.81% ( 612) |
| **Dress Upper Body (0−3)** | | |
| 0−Able | 24.66% ( 45531) | 12.79% (1003) |
| 1−Clothing Laid Out | 33.42% ( 61691) | 31.25% (2450) |
| 2−Assistance | 33.27% ( 61419) | 46.34% (3633) |
| 3−Dependent | 8.64% ( 15957) | 9.62% ( 754) |

Table 5.1 (contd.): Descriptives of Baseline Patient Characteristics

| | No Fall | Fall |
|---|---|---|
| **Dress Lower Body (0−3)** | | |
| 0−Able | 18.87% ( 34831) | 8.53% ( 669) |
| 1−Clothing Laid Out | 18.62% ( 34365) | 15.37% (1205) |
| 2−Assistance | 48.35% ( 89261) | 57.33% (4495) |
| 3−Dependent | 14.16% ( 26141) | 18.76% (1471) |
| **Bathing (0−6)** | | |
| 0−Able | 9.39% ( 17341) | 2.36% ( 185) |
| 1−Use of Device | 10.79% ( 19912) | 6.16% ( 483) |
| 2−Intermittent Asst | 26.13% ( 48227) | 22.96% (1800) |
| 3−Requires Presence | 30.99% ( 57202) | 44.29% (3472) |
| 4−Bedside/Sink | 5.89% ( 10878) | 4.85% ( 380) |
| 5−Bedside/Sink with Asst | 9.84% ( 18169) | 12.40% ( 972) |
| 6−Dependent | 6.97% ( 12869) | 6.99% ( 548) |
| **Toilet Transferring (0−4)** | | |
| 0−Able | 53.29% ( 98376) | 38.98% (3056) |
| 1−Assistance | 34.36% ( 63434) | 45.62% (3577) |
| 2−Bedside | 4.72% ( 8715) | 7.77% ( 609) |
| 3−Bedpan | 1.02% ( 1882) | 1.80% ( 141) |
| 4−Dependent | 6.60% ( 12191) | 5.83% ( 457) |
| **Toilet Hygiene (0−3)** | | |
| 0−Able | 44.68% ( 82470) | 31.96% (2506) |
| 1−Supplies | 29.55% ( 54549) | 31.93% (2503) |
| 2−Assistance | 18.02% ( 33260) | 28.32% (2220) |
| 3−Dependent | 7.76% ( 14319) | 7.79% ( 611) |
| **Transferring (0−5)** | | |
| 0−Able | 23.26% ( 42940) | 10.48% ( 822) |
| 1−Minimum Asst | 62.32% (115034) | 69.53% (5451) |
| 2−Bear Weight No Pivot | 8.38% ( 15464) | 14.46% (1134) |
| 3−No Pivot with Asst | 3.29% ( 6075) | 4.23% ( 332) |
| 4−Bedfast, Turn | 0.79% ( 1463) | 0.66% ( 52) |
| 5−Bedfast | 1.96% ( 3622) | 0.62% ( 49) |
| **Ambulation (0−6)** | | |
| 0−Able | 14.14% ( 26094) | 3.35% ( 263) |
| 1−One Hand Device | 23.98% ( 44267) | 17.55% (1376) |
| 2−Two Hand Device | 36.73% ( 67806) | 43.32% (3396) |
| 3−Assistance | 17.10% ( 31560) | 27.87% (2185) |
| 4−Chairfast, Wheel | 2.24% ( 4133) | 2.88% ( 226) |
| 5−Chairfast, no Wheel | 4.45% ( 8210) | 4.48% ( 351) |
| 6−Bedfast | 1.37% ( 2528) | 0.55% ( 43) |
| **Feeding (0−5)** | | |
| 0−Able | 43.75% ( 80769) | 34.31% (2690) |
| 1−Some Asst | 49.49% ( 91355) | 59.07% (4631) |
| 2−Assistance | 5.39% ( 9944) | 5.82% ( 456) |
| 3−Orally and Nasog Tube/Gastr | 0.42% ( 778) | 0.31% ( 24) |
| 4−Nasog Tube/Gastrostomy | 0.79% ( 1464) | 0.42% ( 33) |
| 5−Unable | 0.16% ( 288) | 0.08% ( 6) |
| **Ability to Plan and Prepare Light Meals** | | |
| 0−Able | 16.01% ( 29558) | 9.01% ( 706) |
| 1−Not Regular Basis | 40.95% ( 75587) | 37.13% (2911) |
| 2−Unable | 43.04% ( 79453) | 53.86% (4223) |
| **Ability to Use Telephone** | | |
| 0−Fully Capable | 77.08% (142296) | 69.86% (5477) |
| 1−Special Phone | 6.45% ( 11915) | 8.15% ( 639) |
| 2−Placing Calls Diff | 5.22% ( 9635) | 7.87% ( 617) |
| 3−Limited | 4.32% ( 7982) | 6.96% ( 546) |
| 4−Listen with Asst | 2.69% ( 4973) | 3.37% ( 264) |
| 5−Unable | 3.75% ( 6918) | 3.42% ( 268) |
| No Phone | 0.48% ( 879) | 0.37% ( 29) |
| **Oral Medications (0−3)** | | |
| 0−Able | 50.33% ( 92911) | 36.39% (2853) |
| 1−Preparation | 27.76% ( 51239) | 34.55% (2709) |
| 2−Reminders | 10.07% ( 18597) | 14.53% (1139) |
| 3−Unable | 10.85% ( 20028) | 13.75% (1078) |
| No Meds | 0.99% ( 1823) | 0.78% ( 61) |

Table 5.1 (contd.): Descriptives of Baseline Patient Characteristics

| | No Fall | Fall |
|---|---|---|
| **Injectable Medications (0−3)** | | |
| 0−Able | 9.78% ( 18046) | 8.33% ( 653) |
| 1−Preparation | 3.16% ( 5841) | 3.34% ( 262) |
| 2−Reminders | 1.31% ( 2411) | 1.67% ( 131) |
| 3−Unable | 7.37% ( 13605) | 8.06% ( 632) |
| No Meds | 78.38% (144695) | 78.60% (6162) |
| **ADL Assistance** | | |
| 0−No Assistance Needed | 20.76% ( 38328) | 9.35% ( 733) |
| 1−Caregiver Currently Providing | 59.61% (110044) | 63.35% (4967) |
| 2−Caregiver Needs Training | 6.17% ( 11395) | 9.13% ( 716) |
| 3−Caregiver Unlikely | 4.05% ( 7467) | 5.19% ( 407) |
| 4−Unclear | 2.93% ( 5417) | 3.97% ( 311) |
| 5−No Caregiver | 6.47% ( 11947) | 9.01% ( 706) |
| **IADL Assistance** | | |
| 0−No Assistance Needed | 7.93% ( 14631) | 3.04% ( 238) |
| 1−Caregiver Currently Providing | 74.92% (138302) | 75.13% (5890) |
| 2−Caregiver Needs Training | 4.84% ( 8932) | 6.56% ( 514) |
| 3−Caregiver Unlikely | 2.90% ( 5346) | 3.24% ( 254) |
| 4−Unclear | 2.83% ( 5217) | 3.43% ( 269) |
| 5−No Caregiver | 6.59% ( 12170) | 8.61% ( 675) |
| **Medication Administration** | | |
| 0−No Assistance Needed | 45.66% ( 84288) | 33.55% (2630) |
| 1−Caregiver Currently Providing | 45.71% ( 84380) | 55.66% (4364) |
| 2−Caregiver Needs Training | 3.50% ( 6463) | 4.59% ( 360) |
| 3−Caregiver Unlikely | 1.58% ( 2912) | 1.63% ( 128) |
| 4−Unclear | 1.28% ( 2366) | 1.81% ( 142) |
| 5−No Caregiver | 2.27% ( 4189) | 2.76% ( 216) |
| **Medical Procedure Treatments** | | |
| 0−No Assistance Needed | 69.37% (128055) | 71.88% (5635) |
| 1−Caregiver Currently Providing | 15.25% ( 28157) | 13.12% (1029) |
| 2−Caregiver Needs Training | 5.83% ( 10767) | 5.23% ( 410) |
| 3−Caregiver Unlikely | 3.62% ( 6675) | 3.71% ( 291) |
| 4−Unclear | 2.68% ( 4948) | 2.49% ( 195) |
| 5−No Caregiver | 3.25% ( 5996) | 3.57% ( 280) |
| **Management of Equipment** | | |
| 0−No Assistance Needed | 89.29% (164833) | 88.33% (6925) |
| 1−Caregiver Currently Providing | 8.26% ( 15256) | 8.75% ( 686) |
| 2−Caregiver Needs Training | 1.24% ( 2290) | 1.48% ( 116) |
| 3−Caregiver Unlikely | 0.39% ( 729) | 0.37% ( 29) |
| 4−Unclear | 0.36% ( 664) | 0.54% ( 42) |
| 5−No Caregiver | 0.45% ( 826) | 0.54% ( 42) |
| **Supervision and Safety** | | |
| 0−No Assistance Needed | 59.21% (109297) | 47.60% (3732) |
| 1−Caregiver Currently Providing | 34.78% ( 64201) | 43.47% (3408) |
| 2−Caregiver Needs Training | 2.73% ( 5041) | 4.15% ( 325) |
| 3−Caregiver Unlikely | 0.72% ( 1330) | 0.97% ( 76) |
| 4−Unclear | 0.99% ( 1835) | 1.45% ( 114) |
| 5−No Caregiver | 1.57% ( 2894) | 2.36% ( 185) |
| **Advocacy or Facilitation** | | |
| 0−No Assistance Needed | 27.07% ( 49966) | 20.88% (1637) |
| 1−Caregiver Currently Providing | 65.62% (121130) | 69.89% (5479) |
| 2−Caregiver Needs Training | 2.44% ( 4502) | 3.10% ( 243) |
| 3−Caregiver Unlikely | 0.87% ( 1613) | 0.94% ( 74) |
| 4−Unclear | 1.65% ( 3055) | 2.32% ( 182) |
| 5−No Caregiver | 2.35% ( 4332) | 2.87% ( 225) |
| **Skin Ulcer** | | |
| 0−None/Asymptomatic | 91.30% (168535) | 87.98% (6898) |
| 1−Well Controlled | 0.11% ( 206) | 0.18% ( 14) |
| 2−Controlled with Difficulty | 5.58% ( 10304) | 7.67% ( 601) |
| 3−Poorly Controlled | 3.01% ( 5553) | 4.17% ( 327) |
| **Hypertension** | | |
| 0−None/Asymptomatic | 33.23% ( 61350) | 31.58% (2476) |
| 1−Well Controlled | 10.30% ( 19017) | 11.07% ( 868) |
| 2−Controlled with Difficulty | 46.73% ( 86267) | 46.76% (3666) |
| 3−Poorly Controlled | 9.73% ( 17964) | 10.59% ( 830) |

Table 5.1 (contd.): Descriptives of Baseline Patient Characteristics

| | No Fall | Fall |
|---|---|---|
| **Diabetes** | | |
| 0−None/Asymptomatic | 65.80% (121461) | 64.38% (5047) |
| 1−Well Controlled | 3.61% ( 6669) | 3.80% ( 298) |
| 2−Controlled with Difficulty | 23.90% ( 44125) | 24.87% (1950) |
| 3−Poorly Controlled | 6.69% ( 12343) | 6.95% ( 545) |
| **Arthritis and Musculoskeletal Diseases** | | |
| 0−None/Asymptomatic | 80.91% (149351) | 83.10% (6515) |
| 1−Well Controlled | 1.64% ( 3023) | 2.14% ( 168) |
| 2−Controlled with Difficulty | 15.46% ( 28530) | 11.98% ( 939) |
| 3−Poorly Controlled | 2.00% ( 3694) | 2.78% ( 218) |
| **Heart Failure** | | |
| 0−None/Asymptomatic | 85.78% (158353) | 83.66% (6559) |
| 1−Well Controlled | 0.83% ( 1536) | 1.06% ( 83) |
| 2−Controlled with Difficulty | 9.11% ( 16825) | 10.40% ( 815) |
| 3−Poorly Controlled | 4.27% ( 7884) | 4.89% ( 383) |
| **Chronic Pulmonary Disease** | | |
| 0−None/Asymptomatic | 84.26% (155535) | 83.23% (6525) |
| 1−Well Controlled | 1.75% ( 3239) | 1.72% ( 135) |
| 2−Controlled with Difficulty | 10.80% ( 19945) | 11.86% ( 930) |
| 3−Poorly Controlled | 3.18% ( 5879) | 3.19% ( 250) |
| **Acute Myocardial Infarction; CIHD** | | |
| 0−None/Asymptomatic | 81.86% (151121) | 80.45% (6307) |
| 1−Well Controlled | 2.04% ( 3772) | 2.50% ( 196) |
| 2−Controlled with Difficulty | 12.91% ( 23838) | 13.43% (1053) |
| 3−Poorly Controlled | 3.18% ( 5867) | 3.62% ( 284) |
| **Cardiac Dysrhythmia** | | |
| 0−None/Asymptomatic | 88.49% (163354) | 86.58% (6788) |
| 1−Well Controlled | 1.60% ( 2957) | 1.68% ( 132) |
| 2−Controlled with Difficulty | 7.91% ( 14596) | 9.30% ( 729) |
| 3−Poorly Controlled | 2.00% ( 3691) | 2.44% ( 191) |
| **Stroke or Late Effects of CVA** | | |
| 0−None/Asymptomatic | 91.49% (168880) | 87.97% (6897) |
| 1−Well Controlled | 0.48% ( 877) | 0.54% ( 42) |
| 2−Controlled with Difficulty | 6.11% ( 11270) | 8.42% ( 660) |
| 3−Poorly Controlled | 1.93% ( 3571) | 3.07% ( 241) |
| **Dementia** | | |
| 0−None/Asymptomatic | 92.26% (170314) | 89.91% (7049) |
| 1−Well Controlled | 1.17% ( 2164) | 1.70% ( 133) |
| 2−Controlled with Difficulty | 5.32% ( 9822) | 6.58% ( 516) |
| 3−Poorly Controlled | 1.24% ( 2298) | 1.81% ( 142) |
| **Neurological Diseases Other than Alzheimer's** | | |
| 0−None/Asymptomatic | 95.54% (176369) | 91.26% (7155) |
| 1−Well Controlled | 0.37% ( 681) | 0.60% ( 47) |
| 2−Controlled with Difficulty | 2.93% ( 5412) | 5.46% ( 428) |
| 3−Poorly Controlled | 1.16% ( 2136) | 2.68% ( 210) |
| **Alzheimer's or Other Cerebral Degeneration** | | |
| 0−None/Asymptomatic | 97.36% (179717) | 96.70% (7581) |
| 1−Well Controlled | 0.34% ( 626) | 0.32% ( 25) |
| 2−Controlled with Difficulty | 1.80% ( 3319) | 2.26% ( 177) |
| 3−Poorly Controlled | 0.51% ( 936) | 0.73% ( 57) |
| **Cancer** | | |
| 0−None/Asymptomatic | 93.45% (172500) | 92.81% (7276) |
| 1−Well Controlled | 0.23% ( 429) | 0.37% ( 29) |
| 2−Controlled with Difficulty | 3.90% ( 7206) | 4.03% ( 316) |
| 3−Poorly Controlled | 2.42% ( 4463) | 2.79% ( 219) |
| **Depression** | | |
| 0−None/Asymptomatic | 91.72% (169317) | 87.02% (6822) |
| 1−Well Controlled | 1.84% ( 3397) | 3.00% ( 235) |
| 2−Controlled with Difficulty | 5.53% ( 10212) | 8.58% ( 673) |
| 3−Poorly Controlled | 0.91% ( 1672) | 1.40% ( 110) |
| **Peripheral Vascular Disease** | | |
| 0−None/Asymptomatic | 96.64% (178388) | 95.92% (7520) |
| 1−Well Controlled | 0.32% ( 595) | 0.42% ( 33) |
| 2−Controlled with Difficulty | 2.35% ( 4345) | 2.82% ( 221) |
| 3−Poorly Controlled | 0.69% ( 1270) | 0.84% ( 66) |

Table 5.1 (contd.): Descriptives of Baseline Patient Characteristics

| | No Fall | Fall |
|---|---|---|
| **Chronic Hepatic Renal Disease** | | |
| 0−None/Asymptomatic | 91.64% (169174) | 90.27% (7077) |
| 1−Well Controlled | 0.78% ( 1437) | 0.93% ( 73) |
| 2−Controlled with Difficulty | 5.87% ( 10838) | 6.63% ( 520) |
| 3−Poorly Controlled | 1.71% ( 3149) | 2.17% ( 170) |
| **AIDS/HIV** | | |
| 0−None/Asymptomatic | 99.02% (182795) | 99.11% (7770) |
| 1−Well Controlled | 0.08% ( 153) | 0.08% ( 6) |
| 2−Controlled with Difficulty | 0.70% ( 1296) | 0.51% ( 40) |
| 3−Poorly Controlled | 0.19% ( 354) | 0.31% ( 24) |

Table 5.1 (contd.): Descriptives of Baseline Patient Characteristics

| | No Fall | Fall |
|---|---|---|
| **Medication Regimen** | | |
| Anti−Ulcer And Other Gi Drugs | 37.52% ( 69266) | 42.54% (3335) |
| Anti−Diarrheals | 0.89% ( 1634) | 1.36% ( 107) |
| Antispasmodics, Anticholinergics | 3.60% ( 6649) | 5.89% ( 462) |
| Bile Therapy | 0.43% ( 803) | 0.50% ( 39) |
| Laxatives | 31.67% ( 58454) | 33.00% (2587) |
| Psychotropics | 12.10% ( 22345) | 17.33% (1359) |
| Muscle Relaxants | 2.77% ( 5107) | 4.16% ( 326) |
| Antiparkinson Agents | 3.07% ( 5662) | 6.16% ( 483) |
| Cns Stimulants | 0.68% ( 1249) | 0.99% ( 78) |
| Psychostimulants/ Antidepressants | 18.98% ( 35030) | 32.93% (2582) |
| Amphetamine Preparations | 0.10% ( 188) | 0.18% ( 14) |
| Anti−Obesity Preparations, All Others | 0.02% ( 37) | 0.04% ( 3) |
| Antihistamines | 4.38% ( 8079) | 4.86% ( 381) |
| Bronchodilators | 15.51% ( 28632) | 17.41% (1365) |
| Antitussives−Expectorants | 1.12% ( 2062) | 1.19% ( 93) |
| Cough And Cold Preparations | 0.55% ( 1011) | 0.59% ( 46) |
| Adrenergics | 0.28% ( 519) | 0.64% ( 50) |
| Nasal And Otic Preparations | 1.88% ( 3468) | 2.16% ( 169) |
| Opthalmic Preparations | 6.39% ( 11787) | 7.58% ( 594) |
| Tetracyclines | 0.82% ( 1509) | 0.85% ( 67) |
| Penicillins | 3.73% ( 6891) | 2.77% ( 217) |
| Strptomycins | 1.02% ( 1879) | 1.24% ( 97) |
| Sulfonamides | 3.25% ( 5996) | 3.39% ( 266) |
| Erthromycins | 1.30% ( 2408) | 1.29% ( 101) |
| Cephalasporins | 4.03% ( 7440) | 3.09% ( 242) |
| Antibiotics, Other | 6.56% ( 12116) | 6.34% ( 497) |
| Antibacterials, Urinary | 2.58% ( 4763) | 2.86% ( 224) |
| Antineoplastics | 2.68% ( 4952) | 3.38% ( 265) |
| Antiparasitics | 2.26% ( 4171) | 2.40% ( 188) |
| Antimalarials | 0.52% ( 966) | 0.66% ( 52) |
| Antivirals | 3.27% ( 6043) | 3.06% ( 240) |
| Tb Preparations | 0.41% ( 760) | 0.20% ( 16) |
| Trimethoprim | 0.03% ( 53) | 0.08% ( 6) |
| Vaginal Cleansers | 0.00% ( 1) | 0.00% ( 0) |
| Antibacterials And Antiseptics, General | 0.08% ( 149) | 0.06% ( 5) |
| Diagnostics | 0.01% ( 10) | 0.00% ( 0) |
| Analgesics, Narcotic | 34.06% ( 62870) | 28.78% (2256) |
| Analgesics, Non−Narcotic, General | 46.17% ( 85237) | 49.80% (3904) |
| Antiarthritics | 11.39% ( 21034) | 12.65% ( 992) |
| Anesthetics Gen Inject | 0.00% ( 2) | 0.00% ( 0) |
| Anesthetics Local/ Topical | 2.10% ( 3882) | 2.86% ( 224) |
| Sedative, Barbituate | 0.21% ( 386) | 0.54% ( 42) |
| Sedative, Non−Barbituate | 6.80% ( 12552) | 9.82% ( 770) |
| Anticonvulsants | 15.87% ( 29304) | 24.54% (1924) |
| Antinauseants | 5.86% ( 10825) | 6.96% ( 546) |
| Corticotropins | 0.00% ( 4) | 0.01% ( 1) |
| Glucocorticoids | 10.63% ( 19623) | 12.67% ( 993) |
| Mineralocorticoids | 0.32% ( 582) | 0.85% ( 67) |
| Aldosterone Antagonists | 3.22% ( 5945) | 4.31% ( 338) |
| Antidotes | 0.09% ( 175) | 0.29% ( 23) |
| Thyroid Preparations | 12.20% ( 22523) | 16.03% (1257) |
| Anti−Thyroid Preparations | 0.40% ( 743) | 0.43% ( 34) |
| Iodine Therapy | 0.01% ( 10) | 0.00% ( 0) |
| Diabetic Therapy | 29.39% ( 54255) | 31.22% (2448) |
| Anabolics | 0.01% ( 24) | 0.03% ( 2) |
| Androgens | 0.11% ( 196) | 0.20% ( 16) |
| Estrogens | 0.32% ( 596) | 0.40% ( 31) |
| Progesterone | 0.08% ( 141) | 0.11% ( 9) |
| Oral Contraceptives | 0.06% ( 107) | 0.03% ( 2) |
| Other Hormones | 0.48% ( 892) | 0.73% ( 57) |
| Lipotropics | 44.29% ( 81756) | 48.97% (3839) |
| Cholesterol Reducers | 0.34% ( 630) | 0.47% ( 37) |
| Digestants | 0.05% ( 101) | 0.03% ( 2) |
| Protein Lysates | 0.08% ( 152) | 0.08% ( 6) |
| Enzymes | 0.60% ( 1114) | 0.82% ( 64) |

Table 5.2: Descriptives of Medications Present at Baseline: First Databank's Therapeutic Classification

| | No Fall | Fall |
|---|---|---|
| Rauwolfia Preparations | 0.00% ( 2) | 0.00% ( 0) |
| Other Hypotensives | 43.10% ( 79568) | 44.27% (3471) |
| Vasodilators: Coronary | 6.36% ( 11738) | 7.92% ( 621) |
| Vasodilators: Peripheral | 0.01% ( 10) | 0.00% ( 0) |
| Digitalis Preparations | 5.03% ( 9278) | 6.53% ( 512) |
| Xanthene Derivatives | 0.28% ( 523) | 0.27% ( 21) |
| Cardiovascular Preparations, Other | 55.86% (103114) | 62.12% (4870) |
| Anticoagulants | 29.95% ( 55286) | 31.85% (2497) |
| Hemostatics | 0.01% ( 25) | 0.01% ( 1) |
| Diuretics | 29.29% ( 54071) | 33.75% (2646) |
| Vitamins, Fat Soluble | 7.19% ( 13270) | 9.54% ( 748) |
| Vitamins, Water Soluble | 9.70% ( 17911) | 12.82% (1005) |
| Multivitamins | 14.28% ( 26353) | 16.73% (1312) |
| Folic Acid Preparations | 7.83% ( 14456) | 9.92% ( 778) |
| B Complex With Vitamin C | 0.12% ( 222) | 0.20% ( 16) |
| Vitamin K Preparations | 0.30% ( 551) | 0.15% ( 12) |
| Infant Formulas | 0.00% ( 8) | 0.01% ( 1) |
| Electrolytes And Misc | 19.98% ( 36889) | 24.15% (1893) |
| Hematinics | 13.14% ( 24247) | 14.27% (1119) |
| Biologicals | 0.02% ( 35) | 0.03% ( 2) |
| Coal Tar | 0.00% ( 3) | 0.00% ( 0) |
| Emollients Protectives | 0.64% ( 1190) | 0.89% ( 70) |
| Fungicides | 3.09% ( 5713) | 3.30% ( 259) |
| Dermatologicals, All Others | 0.40% ( 737) | 0.74% ( 58) |
| Hemorrhoidal Preparations | 0.10% ( 187) | 0.11% ( 9) |
| Oxytocics | 0.00% ( 0) | 0.01% ( 1) |
| Parasympathetic Agents | 5.79% ( 10692) | 8.67% ( 680) |
| Unclassified Drug Products | 21.63% ( 39934) | 27.68% (2170) |

Table 5.2 (contd.): Descriptives of Medications Present at Baseline: First Databank's Therapeutic Classification

## 5.2 Identifying Baseline Risk Factors in Full Cohort Setting

Unlike the case-crossover design, the score function for the Cox partial likelihood in Equation (3.7) contains the matrix of characteristics present at baseline, which are denoted by $\boldsymbol{X_i}$. All confounders must be controlled for in the Cox proportional hazards model in order to obtain an unbiased estimate of $\beta$. While it is possible that unmeasured confounders exist, the OASIS assessment provides a comprehensive list of factors that describe the patients health status at admission into home health care. The inclusion of all observed baseline factors would result in a complicated Cox proportional hazards model with a large number of degrees of freedom. Therefore, a more parsimonious model is sought through a variable selection analysis resulting in a subset of baseline factors that are associated with falling. The analysis in this section follows previous work, on the same population, by Rosati and colleagues to identify the factors associated with patients who are hospitalized [84].

All factors displayed in the descriptive analysis displayed in Table 5.1 and the therapeutic medication classifications in Table 5.2 are assumed to be candidate predictors of falling. The VNSNY Medication and Fall data set is randomly split into two data sets. 70% of patient are allocated to a training data set for which variable selection and parameter estimation will be conducted. The remaining 30% of the sample is held out to assess the model fit.

A Cox proportional hazards model, predicting the time to the first fall during home health care, is developed by using the elastic net shrinkage penalty to select and estimate the coefficients of predictors [112]. Shrinkage methods have been shown to have superior performance over traditional stepwise methods in both simulations and empirical data and are underutilized in applied settings [67, 90]. Shrinkage methods alleviate two issues that occur in traditional maximum likelihood methods: estimation and variable selection in the presence of multi-collinearity, and over fitting the model to the training data set. Parameter estimates can still be obtained and are interpreted just as the coefficients in a Cox hazards model via maximum partial likelihood. The elastic net mixing parameter is fixed at 0.5 and the complexity parameter considers range of possible values. The complexity parameter

was chosen that yielded the model within one standard error of the model that achieved the minimum partial likelihood deviance. Standard errors are estimated from 10-fold cross-validation. This analysis was completed using the **R** function `glmnet.cv` from the glmnet package in **R** [45, 73]. The factors identified as predictors of falls experienced during home health care by the elastic net were also used in a Cox proportional hazards model where parameter estimates were approximated using traditional (i.e. unpenalized) partial likelihood using `coxph` function in the survival package in **R** [97]. Results from the Cox proportional hazards model estimated from the elastic net and traditional partial likelihood are both displayed in Table 5.3.

Using parameter estimates obtained from the training data set, predictions are made on the hold-out data set and compared to the actual outcome. Figure 5.4 displays time-dependent Receiver Operating Characteristic (ROC) curves of both models extrapolated on the 30% hold out data. Time points of 4, 13, 34, and 60 are fixed to obtain area under the curve (AUC) estimates of 0.694, 0.694, 0.693, and 0.691 for the elastic net and 0.702, 0.699, 0.696, and 0.697 for the traditional estimation respectively. Times 4, 13, 34 are the inter-quartile range and median times of falling, and 60 marks the end of the first CMS home health care episode. Slight improvements in the AUC for the traditional estimation over the elastic net can be explained by the additional degrees of freedom allowed to enter into the model.

The strongest predictor of a fall during home health care, as expected, is the history of falling. The elastic net estimates the hazard of falling is 2.05 higher for patients who have a history of falling compared to those who don not. There are several other notable, strong, positive predictors of falling. Caucasians are at 1.36 times the risk of falling as patients of race other than Caucasian or African Americans. An indicator that no surgical wound is present increases the hazard of falling by 33%. Patients who are admitted into home health care after a surgical procedure are considered post-acute restorative and are less likely to be frail [68]. Patients that require a two hand device or assistance to ambulate are at 21% or 31% higher risk than those who do not. Patients with a home health care diagnosis classified as a neurological disease, other than Alzheimer's, with symptoms that are considered to be poorly controlled increase the baseline hazard by 1.36 fold.

The therapeutic medication classifications which yield the strongest hazards ratios include psychostimulants/antidepressants, barbituates, anticonvulsants, mineralocorticoids, and antidotes which have hazard ratios of 1.36, 1.34, 1.25, 1.47, and 1.62 respectively. These associations are describing the effect of the presence of a medication at admission and not the transient effect of a medication change.

| | Elastic Net | | Traditional Partial Likelihood | | |
|---|---|---|---|---|---|
| | Log−Hazard Ratio | Hazard Ratio | log−hazard (Std.Error) | Hazard Ratio (95% CI) | p−value |
| **Language Spoken** | | | | | |
| Spanish (ref) | | | | | |
| English | 0.21 | 1.24 | 0.27 (0.06) | (1.17, 1.47) | 0.0000 |
| Other | −0.16 | 0.85 | −0.20 (0.06) | (0.72, 0.93) | 0.0020 |
| **Age Group** | | | | | |
| 18 to 54 (ref) | | | | | |
| 55 to 64 | | | −0.09 (0.06) | (0.82, 1.03) | 0.1363 |
| 65 to 74 | | | −0.03 (0.06) | (0.87, 1.08) | 0.5878 |
| 75 to 84 | 0.00 | 1.00 | 0.05 (0.05) | (0.94, 1.16) | 0.4030 |
| 85 + | | | 0.05 (0.06) | (0.94, 1.18) | 0.3831 |
| **Race** | | | | | |
| Other (ref) | | | | | |
| Black | −0.02 | 0.98 | −0.11 (0.05) | (0.80, 0.99) | 0.0364 |
| White | 0.31 | 1.36 | 0.29 (0.05) | (1.22, 1.47) | 0.0000 |
| **Prior Conditions** | | | | | |
| Urinary Incontinence | 0.09 | 1.09 | 0.11 (0.04) | (1.04, 1.20) | 0.0017 |
| **Risk for Hospitalization** | | | | | |
| Decline in Mental, Emotional, Behavioral | 0.15 | 1.17 | 0.23 (0.04) | (1.16, 1.35) | 0.0000 |
| Multi Hospitalization | 0.05 | 1.05 | 0.06 (0.03) | (1.00, 1.13) | 0.0387 |
| History of Falls | 0.72 | 2.05 | 0.70 (0.03) | (1.89, 2.13) | 0.0000 |
| Frailty Indicators | 0.09 | 1.09 | 0.09 (0.03) | (1.03, 1.16) | 0.0054 |
| **Understand Verbal Content** | | | | | |
| 0−Understands (ref) | | | | | |
| 1−Usually | | | −0.07 (0.04) | (0.87, 1.00) | 0.0637 |
| 2−Sometimes | | | −0.11 (0.08) | (0.76, 1.05) | 0.1724 |
| 3−Rarely | | | 0.02 (0.27) | (0.60, 1.74) | 0.9503 |
| Unknown | −0.05 | 0.95 | −0.74 (0.35) | (0.24, 0.95) | 0.0344 |
| **Stage of Most Problematic Pressure Ulcer** | | | | | |
| 0−None (ref) | | | | | |
| 1−Stage I | 0.06 | 1.06 | 0.13 (0.10) | (0.93, 1.38) | 0.2010 |
| 2−Stage II | | | −0.02 (0.09) | (0.82, 1.18) | 0.8544 |
| 3−Stage III | | | −0.18 (0.15) | (0.63, 1.11) | 0.2146 |
| 4−Stage IV | | | −0.08 (0.19) | (0.64, 1.34) | 0.6662 |
| Unstageable | | | 0.12 (0.14) | (0.85, 1.50) | 0.3873 |
| **Status of Most Problematic Surgical Wound** | | | | | |
| 0−Newly Epithelialized (ref) | | | | | |
| 1−Fully Granulating | | | −0.07 (0.12) | (0.73, 1.18) | 0.5383 |
| 2−Early/Partial Granulation | | | −0.06 (0.11) | (0.77, 1.16) | 0.5659 |
| 3−Not Healing | | | −0.09 (0.09) | (0.76, 1.10) | 0.3276 |
| None | 0.29 | 1.33 | 0.30 (0.08) | (1.17, 1.57) | 0.0001 |
| Not Observ | | | 0.00 (0.12) | (0.79, 1.27) | 0.9896 |
| **Bowel Incontinence Frequency** | | | | | |
| 0−Rare or Never (ref) | | | 0.04 (0.05) | (0.94, 1.14) | 0.4705 |
| 1−Less than 1 to 3/wk | | | 0.04 (0.05) | (0.94, 1.14) | 0.4705 |
| 2−4to6/week | | | 0.00 (0.10) | (0.82, 1.22) | 0.9757 |
| 3−Daily | | | 0.11 (0.08) | (0.96, 1.31) | 0.1570 |
| 4−Bowel Ostomy | −0.13 | 0.87 | −0.44 (0.13) | (0.49, 0.84) | 0.0010 |
| 5−More than 1 Daily | | | 0.11 (0.14) | (0.84, 1.49) | 0.4364 |
| **Cognitive Functioning** | | | | | |
| 0−Alert (ref) | | | | | |
| 1−Prompting | 0.04 | 1.04 | 0.07 (0.04) | (1.00, 1.15) | 0.0576 |
| 2−Some Asst | | | −0.02 (0.06) | (0.87, 1.10) | 0.7228 |
| 3−Consderable Asst | −0.01 | 0.99 | −0.14 (0.12) | (0.69, 1.09) | 0.2260 |
| 4−Dependent | −0.19 | 0.82 | −0.76 (0.30) | (0.26, 0.85) | 0.0125 |
| **Anxious** | | | | | |
| 0−None (ref) | | | | | |
| 1−Less than Daily | | | −0.08 (0.04) | (0.86, 0.99) | 0.0238 |
| 2−Daily, not Constant | | | 0.01 (0.04) | (0.93, 1.10) | 0.7633 |
| 3−All of the Time | 0.07 | 1.07 | 0.23 (0.10) | (1.03, 1.53) | 0.0246 |
| Nonresponsive | | | −0.06 (0.32) | (0.50, 1.76) | 0.8457 |
| **Little Interest/Pleasure** | | | | | |
| 0−Not at all (0−1) (ref) | | | | | |
| 1−Several days (2−6) | | | −0.11 (0.05) | (0.81, 0.99) | 0.0336 |
| 2−More than half (7−11) | 0.17 | 1.19 | 0.18 (0.10) | (0.97, 1.47) | 0.0894 |
| 3−Nearly every day (12−14) | | | 0.01 (0.15) | (0.76, 1.35) | 0.9480 |
| 8−Other Asmnt−Further Eval | | | −0.21 (0.13) | (0.63, 1.05) | 0.1141 |
| 9−Other Asmnt−No Further Eval | | | 0.01 (0.07) | (0.88, 1.16) | 0.8706 |
| No Screening | | | 0.25 (1.51) | (0.07, 24.85) | 0.8692 |
| Nonresponsive | | | −0.21 (0.37) | (0.40, 1.67) | 0.5732 |

Table 5.3: The first set of columns use elastic net penalty, the second contains traditional Cox partial likelihood estimates. References for non-mutually exclusive conditions is the absence of that particular condition, these items are categorized under: prior conditions, risk for hospitalization, cognitive, behavioral, psychiatric symptoms, medication regimen

| | Elastic Net | | Traditional Partial Likelihood | | |
|---|---|---|---|---|---|
| | Log−Hazard Ratio | Hazard Ratio | log−hazard (Std.Error) | Hazard Ratio (95% CI) | p−value |
| **Down, Depressed Hopeless** | | | | | |
| 0−Not at all (0−1) (ref) | | | | | |
| 1−Several days (2−6) | 0.07 | 1.08 | 0.19 (0.05) | (1.10, 1.34) | 0.0001 |
| 2−More than half (7−11) | | | 0.10 (0.10) | (0.91, 1.35) | 0.3011 |
| 3−Nearly every day (12−14) | | | 0.12 (0.14) | (0.85, 1.48) | 0.4062 |
| 8−Other Asmnt−Further Eval | | | | | |
| 9−Other Asmnt−No Further Eval | | | | | |
| No Screening | | | −0.25 (1.51) | (0.04, 15.04) | 0.8670 |
| Nonresponsive | | | 0.43 (0.36) | (0.75, 3.14) | 0.2413 |
| **Cognitive, Behavioral, Psychiatric Symptoms** | | | | | |
| Delusional | −0.08 | 0.93 | −0.33 (0.18) | (0.50, 1.03) | 0.0696 |
| **Grooming** | | | | | |
| 0−Able (ref) | | | | | |
| 1−Utensil Asst | 0.02 | 1.02 | 0.05 (0.04) | (0.96, 1.14) | 0.2674 |
| 2−Assistance | | | 0.00 (0.05) | (0.91, 1.10) | 0.9715 |
| 3−Dependent | | | −0.02 (0.08) | (0.84, 1.16) | 0.8325 |
| **Bathing** | | | | | |
| 0−Able (ref) | | | | | |
| 1−Use of Device | | | 0.24 (0.11) | (1.03, 1.58) | 0.0259 |
| 2−Intermittent Asst | | | 0.26 (0.11) | (1.05, 1.61) | 0.0144 |
| 3−Requires Presence | 0.11 | 1.12 | 0.38 (0.11) | (1.19, 1.82) | 0.0005 |
| 4−Bedside/Sink | | | 0.31 (0.12) | (1.08, 1.73) | 0.0101 |
| 5−Bedside/Sink with Asst | 0.16 | 1.17 | 0.48 (0.12) | (1.29, 2.03) | 0.0000 |
| 6−Dependent | | | 0.36 (0.13) | (1.11, 1.86) | 0.0058 |
| **Toilet Transferring** | | | | | |
| 0−Able (ref) | | | | | |
| 1−Assistance | 0.02 | 1.02 | 0.05 (0.04) | (0.97, 1.14) | 0.2260 |
| 2−Bedside | 0.03 | 1.03 | 0.09 (0.07) | (0.96, 1.24) | 0.1986 |
| 3−Bedpan | 0.14 | 1.15 | 0.30 (0.11) | (1.08, 1.68) | 0.0089 |
| 4−Dependent | −0.17 | 0.84 | −0.24 (0.11) | (0.64, 0.96) | 0.0212 |
| **Toilet Hygiene** | | | | | |
| 0−Able (ref) | | | | | |
| 1−Supplies | | | −0.06 (0.04) | (0.87, 1.02) | 0.1397 |
| 2−Assistance | 0.03 | 1.03 | −0.06 (0.05) | (0.86, 1.05) | 0.2818 |
| 3−Dependent | −0.01 | 0.99 | −0.20 (0.10) | (0.68, 0.99) | 0.0402 |
| **Transferring** | | | | | |
| 0−Able (ref) | | | | | |
| 1−Minimum Asst | 0.09 | 1.10 | 0.09 (0.05) | (0.99, 1.21) | 0.0920 |
| 2−Bear Weight No Pivot | 0.20 | 1.23 | 0.25 (0.07) | (1.12, 1.47) | 0.0004 |
| 3−No Pivot with Asst | | | 0.15 (0.10) | (0.95, 1.42) | 0.1493 |
| 4−Bedfast, Turn | | | −0.07 (0.20) | (0.64, 1.37) | 0.7300 |
| 5−Bedfast | −0.42 | 0.66 | −0.82 (0.23) | (0.28, 0.69) | 0.0003 |
| **Ambulation** | | | | | |
| 0−Able (ref) | | | | | |
| 1−One Hand Device | | | 0.45 (0.09) | (1.32, 1.87) | 0.0000 |
| 2−Two Hand Device | 0.19 | 1.21 | 0.63 (0.09) | (1.57, 2.22) | 0.0000 |
| 3−Assistance | 0.27 | 1.31 | 0.70 (0.09) | (1.68, 2.41) | 0.0000 |
| 4−Chairfast, Wheel | | | 0.33 (0.12) | (1.09, 1.78) | 0.0078 |
| 5−Chairfast, no Wheel | | | 0.44 (0.12) | (1.21, 1.98) | 0.0005 |
| 6−Bedfast | | | 0.46 (0.24) | (0.99, 2.56) | 0.0571 |
| **Feeding** | | | | | |
| 0−Able (ref) | | | | | |
| 1−Some Asst | | | −0.09 (0.03) | (0.86, 0.97) | 0.0049 |
| 2−Assistance | −0.01 | 0.99 | −0.21 (0.08) | (0.70, 0.94) | 0.0053 |
| 3−Orally and Nasog Tube/Gastr | | | −0.16 (0.24) | (0.53, 1.36) | 0.4929 |
| 4−Nasog Tube/Gastrostomy | | | −0.41 (0.21) | (0.44, 1.01) | 0.0536 |
| 5−Unable | | | −0.36 (0.50) | (0.26, 1.87) | 0.4750 |
| **Oral Medications** | | | | | |
| 0−Able (ref) | | | | | |
| 1−Preparation | 0.05 | 1.05 | 0.02 (0.04) | (0.94, 1.12) | 0.5933 |
| 2−Reminders | | | −0.08 (0.06) | (0.83, 1.03) | 0.1675 |
| 3−Unable | | | −0.06 (0.06) | (0.83, 1.06) | 0.3215 |
| No Meds | | | 0.05 (0.16) | (0.76, 1.44) | 0.7824 |
| **ADL Assistance** | | | | | |
| 0−No Assistance Needed (ref) | | | 0.13 (0.06) | (1.01, 1.28) | 0.0309 |
| 1−Caregiver Currently Providing | | | 0.13 (0.06) | (1.01, 1.28) | 0.0309 |
| 2−Caregiver Needs Training/Unclear | 0.01 | 1.01 | 0.20 (0.07) | (1.06, 1.40) | 0.0063 |
| 3−Caregiver Unlikely/No Caregiver | | | 0.21 (0.07) | (1.08, 1.41) | 0.0019 |

Table 5.3 (contd.):

| | Elastic Net | | Traditional Partial Likelihood | | |
|---|---|---|---|---|---|
| | Log−Hazard Ratio | Hazard Ratio | log−hazard (Std.Error) | Hazard Ratio (95% CI) | p−value |
| **Medication Administration** | | | | | |
| 0−No Assistance Needed (ref) | 0.02 | 1.02 | 0.06 (0.04) | (0.98, 1.16) | 0.1642 |
| 1−Caregiver Currently Providing | 0.02 | 1.02 | 0.06 (0.04) | (0.98, 1.16) | 0.1642 |
| 2−Caregiver Needs Training/Unclear | | | 0.05 (0.07) | (0.92, 1.21) | 0.4638 |
| 3−Caregiver Unlikely/No Caregiver | −0.15 | 0.86 | −0.30 (0.08) | (0.63, 0.86) | 0.0001 |
| **AIDS/HIV** | | | | | |
| 0−None/Asymptomatic (ref) | | | | | |
| 1−Well Controlled | | | 0.03 (0.45) | (0.43, 2.49) | 0.9412 |
| 2−Controlled with Difficulty | −0.44 | 0.64 | −0.87 (0.20) | (0.28, 0.63) | 0.0000 |
| 3−Poorly Controlled | | | 0.16 (0.25) | (0.73, 1.90) | 0.5116 |
| **Cancer** | | | | | |
| 0−None/Asymptomatic (ref) | | | | | |
| 1−Well Controlled | | | 0.38 (0.22) | (0.95, 2.25) | 0.0823 |
| 2−Controlled with Difficulty | 0.12 | 1.13 | 0.32 (0.07) | (1.21, 1.57) | 0.0000 |
| 3−Poorly Controlled | | | 0.16 (0.09) | (0.99, 1.39) | 0.0713 |
| **Depression** | | | | | |
| 0−None/Asymptomatic (ref) | | | | | |
| 1−Well Controlled | 0.01 | 1.01 | 0.14 (0.08) | (0.98, 1.35) | 0.0868 |
| 2−Controlled with Difficulty | | | 0.04 (0.05) | (0.94, 1.15) | 0.4447 |
| 3−Poorly Controlled | | | −0.24 (0.13) | (0.61, 1.00) | 0.0537 |
| **Neurological Diseases Other than Alzheimer's** | | | | | |
| 0−None/Asymptomatic (ref) | | | | | |
| 1−Well Controlled | | | 0.23 (0.18) | (0.89, 1.78) | 0.1997 |
| 2−Controlled with Difficulty | 0.09 | 1.09 | 0.22 (0.07) | (1.08, 1.44) | 0.0023 |
| 3−Poorly Controlled | 0.31 | 1.36 | 0.49 (0.09) | (1.36, 1.97) | 0.0000 |
| **Stroke or Late Effects of CVA** | | | | | |
| 0−None/Asymptomatic (ref) | | | | | |
| 1−Well Controlled | | | 0.05 (0.18) | (0.74, 1.50) | 0.7811 |
| 2−Controlled with Difficulty | 0.04 | 1.04 | 0.17 (0.05) | (1.07, 1.31) | 0.0008 |
| 3−Poorly Controlled | 0.01 | 1.01 | 0.21 (0.08) | (1.05, 1.45) | 0.0106 |
| **Chronic Hepatic Renal Disease** | | | | | |
| 0−None/Asymptomatic (ref) | | | | | |
| 1−Well Controlled | | | 0.36 (0.14) | (1.10, 1.87) | 0.0086 |
| 2−Controlled with Difficulty | 0.04 | 1.04 | 0.19 (0.06) | (1.09, 1.35) | 0.0005 |
| 3−Poorly Controlled | 0.07 | 1.07 | 0.31 (0.09) | (1.13, 1.64) | 0.0009 |
| **Skin Ulcer** | | | | | |
| 0−None/Asymptomatic (ref) | | | | | |
| 1−Well Controlled | | | 0.14 (0.34) | (0.59, 2.22) | 0.6887 |
| 2−Controlled with Difficulty | 0.06 | 1.06 | 0.19 (0.07) | (1.06, 1.37) | 0.0037 |
| 3−Poorly Controlled | | | 0.13 (0.08) | (0.97, 1.34) | 0.1066 |
| **Arthritis and Musculoskeletal Diseases** | | | | | |
| 0−None/Asymptomatic (ref) | | | | | |
| 1−Well Controlled | | | 0.04 (0.10) | (0.86, 1.26) | 0.6909 |
| 2−Controlled with Difficulty | −0.13 | 0.88 | −0.22 (0.04) | (0.74, 0.87) | 0.0000 |
| 3−Poorly Controlled | | | 0.04 (0.09) | (0.88, 1.23) | 0.6120 |
| **Medication Regimen** | | | | | |
| Antispasmodics, Anticholinergics | 0.03 | 1.03 | 0.11 (0.06) | (1.00, 1.25) | 0.0587 |
| Antiparkinson Agents | 0.03 | 1.03 | 0.04 (0.07) | (0.90, 1.20) | 0.5961 |
| Psychostimulants/ Antidepressants | 0.30 | 1.36 | 0.32 (0.03) | (1.29, 1.47) | 0.0000 |
| Strptomycins | 0.01 | 1.01 | 0.28 (0.12) | (1.04, 1.67) | 0.0222 |
| Sedative, Barbituate | 0.29 | 1.34 | 0.66 (0.18) | (1.36, 2.75) | 0.0002 |
| Sedative, Non−Barbituate | 0.07 | 1.07 | 0.14 (0.05) | (1.05, 1.26) | 0.0027 |
| Anticonvulsants | 0.22 | 1.25 | 0.26 (0.03) | (1.22, 1.39) | 0.0000 |
| Glucocorticoids | 0.03 | 1.03 | 0.09 (0.04) | (1.01, 1.18) | 0.0343 |
| Mineralocorticoids | 0.39 | 1.47 | 0.56 (0.14) | (1.32, 2.34) | 0.0001 |
| Aldosterone Antagonists | 0.02 | 1.02 | 0.14 (0.07) | (1.01, 1.32) | 0.0342 |
| Antidotes | 0.48 | 1.62 | 0.84 (0.24) | (1.45, 3.69) | 0.0005 |
| Cardiovascular Preparations, Other | 0.00 | 1.00 | 0.03 (0.03) | (0.98, 1.10) | 0.2578 |
| Vitamins, Water Soluble | 0.01 | 1.01 | 0.06 (0.04) | (0.98, 1.15) | 0.1354 |
| Folic Acid Preparations | 0.00 | 1.00 | 0.06 (0.05) | (0.97, 1.17) | 0.1769 |
| Dermatologicals, All Others | 0.09 | 1.10 | 0.33 (0.15) | (1.03, 1.88) | 0.0309 |
| Unclassified Drug Products | 0.00 | 1.00 | 0.03 (0.03) | (0.97, 1.09) | 0.3790 |

Table 5.3 (contd.):

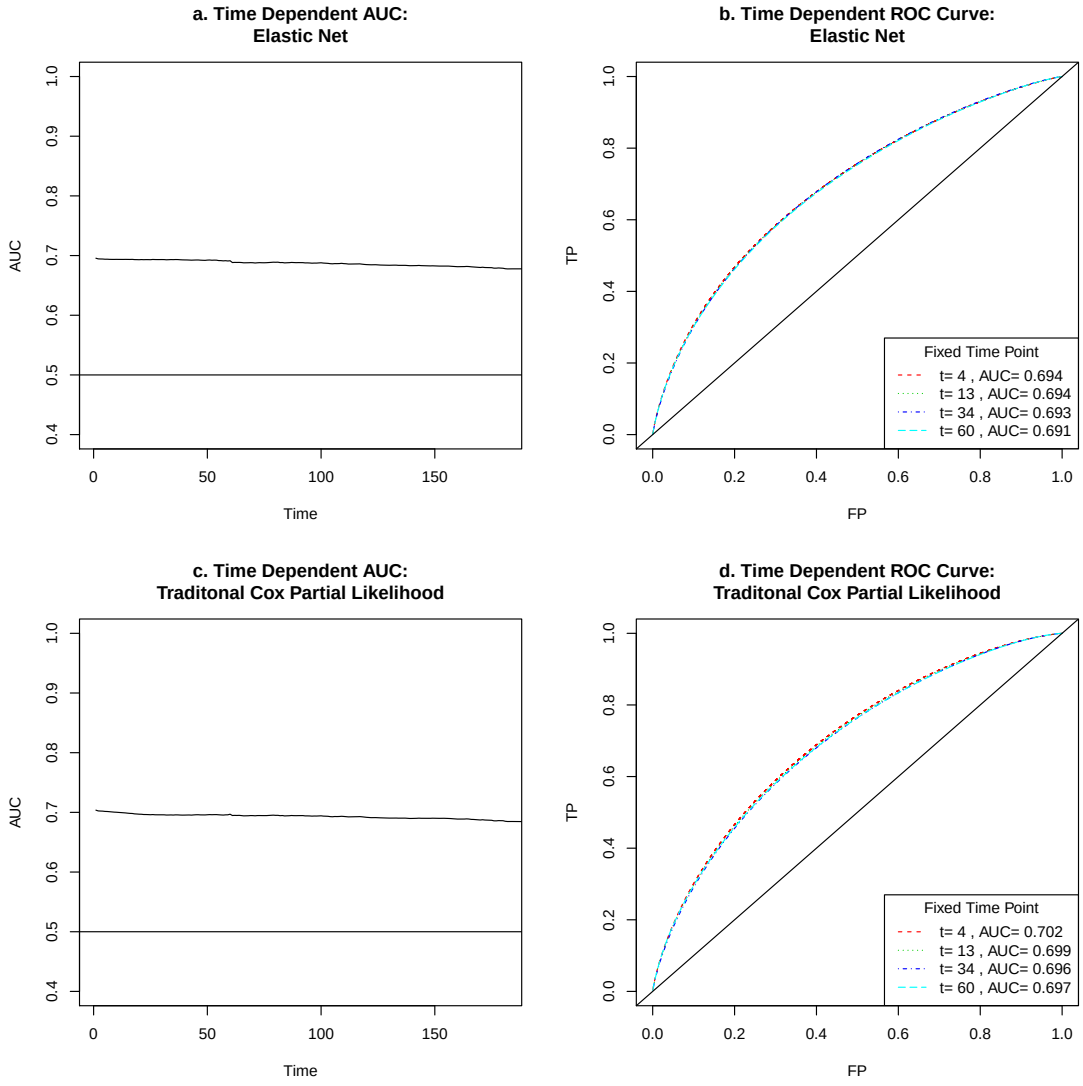


Figure 5.4: Time-dependent ROC curve analysis of model performance on the validation data set. Plots (a) and (b) are from automated variable selection using elastic net shrinkage, and plots (c) and (d) use the variables identified in the elastic net but estimates are obtained via traditional Cox partial likelihood . Plots (a) and (c) plot the area under the curve versus time, while plots (b) and (d) plot the time-dependent ROC curve.

## 5.3 Design Comparison

This study explores the potential for the risk of falling to elevate over the days following an increase in the number of psychostimulant or antidepressant. This section compares results from the case-crossover design and the Cox proportional hazards model in the cohort design to the VNSNY Medication and Fall data. The time-dependent covariate in the full cohort design is defined as an indication when the number of medications classified as a psychostimulants or antidepressants increases. This indicator requires information on the number of medications present on the prior day, therefore, time of entry for this study occurs on day two of home health care. As a result, 1318 day one falls are eliminated from both designs.

Both designs explore several durations in which the effect is assumed to last. This covariate is specified in Equation (5.1), and varying durations of the effect are explored at $\Delta = 1, 2, 3, 4, 5, 7, 14, 21, 28$ for both designs.

$$Z_{it} = \begin{cases} 1 & t \in (\tau_i, \tau_i + \Delta] \\ 0 & t \notin (\tau_i, \tau_i + \Delta] \end{cases} \tag{5.1}$$

The results of the case-crossover design and the cohort design applying a Cox proportional hazards model are displayed in Table 5.4. Both designs consider the following hazard function

$$\lambda(t) = \lambda_0(t) e^{\boldsymbol{X}\boldsymbol{\gamma} + \beta Z_{it}} \tag{5.2}$$

where $\boldsymbol{X}$ contains the baseline factors identified by Section 5.2. These factors are treated as covariates in the specification of the Cox proportional hazards model and are implicitly controlled for in the case-crossover design.

The case-crossover design applies a 1:5 matching scheme, where a minimum of one control is required to be included. For assumed durations of the effect of 1,2,3,4,5,7, the first day of the first control reference window is forced to be seven days apart from the first day of the case window. This imposes a same day of the week comparison, and eliminating any seasonality that may exists on different days of the week [71]. The case-crossover design is forced to drop falls when atleast one entire control-time is not available after admission. An illustration of the case-crossover design reference strategies is displayed in Figure 5.5.

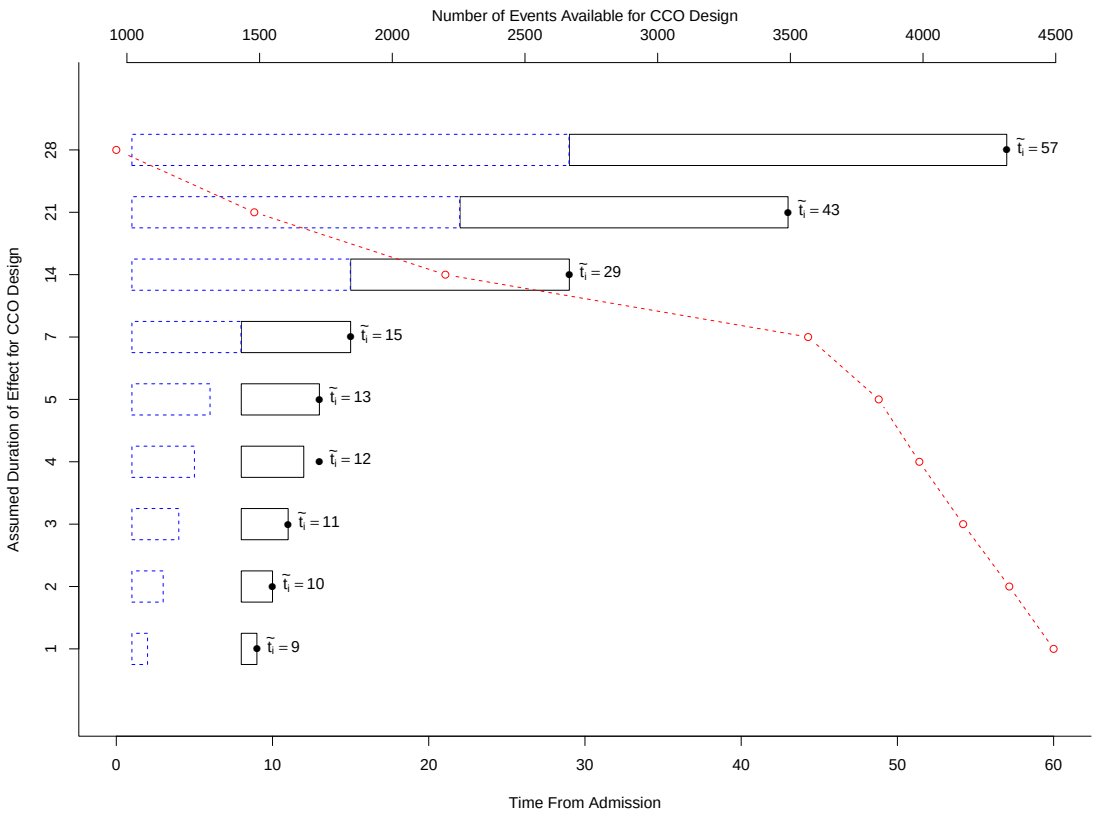


Figure 5.5: Dual axis plot illustrating the application of several case-crossover design reference strategies applied to VNSNY medication and falls data. The minimum time of fall is plotted with the corresponding reference strategy assuming varying durations of the effect. The red dotted line plots the number of events available for each case-crossover design

This plot shows the minimum required time for a fall to be included in the analysis for each reference strategy. The minimum number of days to be included in the analysis is 9, 10, 11, 12, 13, 15, 29, 43, 57 for $\Delta$ = 1,2,3,4,5,7,14,21,28, respectively. For each increase in length of the assumed effect, the more falls are dropped from the analysis. This is also plotted on the second x-axis in Figure 5.5 for the corresponding design. Only 68.9% of the 6522 falls available for the cohort design remain for the assumed duration of one day, and as little as 14.7% of the falls for a duration of 28 days since the minimum time for inclusion of a fall is 57 days from admission. Choosing $\Delta$ is difficult in the case-crossover design. The varying number of observations available changes with delta and makes goodness of fit measures incomparable (e.g. AIC, BIC, etc.).

The case-crossover design provides no evidence to support an association between an

| Duration | Case−Crossover Design | | | | Cohort Design with Cox Hazards Model | | | | Relative |
|---|---|---|---|---|---|---|---|---|---|
| (days) | log−hazard | Hazard Ratio(95% CI) | p−value | Falls Remaining | log−hazard | Hazard Ratio(95% CI) | p−value | AIC | Efficiency |
| 1 | 0.54 (0.44) | 1.71 (0.72, 4.06) | 0.2205 | 68.87% (4492) | 0.68 (0.38) | 1.97 (0.94, 4.14) | 0.0722 | 145044 | 74.05 |
| 2 | 0.45 (0.34) | 1.57 (0.81, 3.05) | 0.1814 | 66.31% (4325) | 0.63 (0.28) | 1.87 (1.09, 3.23) | 0.0237 | 145043 | 67.58 |
| 3 | 0.59 (0.27) | 1.80 (1.06, 3.05) | 0.0288 | 63.65% (4151) | 0.72 (0.22) | 2.06 (1.34, 3.17) | 0.0009 | 145039 | 66.21 |
| 4 | 0.60 (0.23) | 1.82 (1.16, 2.86) | 0.0093 | 61.12% (3986) | 0.85 (0.18) | 2.33 (1.64, 3.32) | 0.0000 | 145031 | 61.50 |
| 5 | 0.51 (0.22) | 1.67 (1.09, 2.55) | 0.0185 | 58.77% (3833) | 0.79 (0.17) | 2.21 (1.59, 3.06) | 0.0000 | 145031 | 59.43 |
| 7 | 0.53 (0.20) | 1.70 (1.15, 2.50) | 0.0076 | 54.69% (3567) | 0.78 (0.15) | 2.18 (1.64, 2.90) | 0.0000 | 145027 | 53.89 |
| 14 | 0.41 (0.20) | 1.51 (1.01, 2.26) | 0.0427 | 33.73% (2200) | 0.64 (0.12) | 1.89 (1.50, 2.38) | 0.0000 | 145029 | 33.00 |
| 21 | 0.36 (0.21) | 1.43 (0.94, 2.18) | 0.0914 | 22.69% (1480) | 0.54 (0.11) | 1.72 (1.40, 2.12) | 0.0000 | 145034 | 24.88 |
| 28 | 0.34 (0.24) | 1.40 (0.88, 2.23) | 0.1555 | 14.72% ( 960) | 0.60 (0.10) | 1.82 (1.51, 2.19) | 0.0000 | 145026 | 15.97 |

Table 5.4: Analysis of an increase in psychostimulants/antidepressants and falls where the assumed duration in which the effect is specified by one time-dependent variable. Case-crossover design analysis of durations 1,2,3,4,5,7 fix the case- and control times to be separated by one week. The Cox proportional hazards model includes all covariates selected in by the elastic net procedure shown in Table 5.3, log-hazard estimates for these parameters are suppressed to isolate the effect of increase in medication. The estimates from Cox proportional hazards model includes 6522 falls.

increase in the number of psychostimulants or antidepressants and the risk of falling assuming a window for the effect of length one or two. The p-values for assumed durations of 3, 4, 5, 7 suggest there is strong evidence to reject the null hypothesis assuming no effect. There is still moderate evidence (p-value=0.0427) that there is an increased effect when the duration is extended to a 14 day window, but only 33.7% of the original number of falls are used in the conditional likelihood for this estimate. When the effect is extended to 21 and 28 days both estimate a hazard ratio of approximately 1.4, however, both fail to reject the null hypothesis of no effect assuming a type I error rate of 5%. In order to allow enough time for a reference strategy for 21 and 28 time intervals, the number of falls used in the designs drops to less than 22.7% of the falls used by the cohort design.

The same assumed durations of the effect are explored for the Cox proportional hazards model with the time dependent covariate, defined in Equation (5.1), and patient characteristics $X$ from Equation (5.2). All factors identified by the elastic net model shown in Table 5.3 are included in the model, but parameter estimates are suppressed to ease presentation in Table 5.4 by isolating the time-dependent covariate. There is weak evidence to support the hypothesis that an increase in the number of psychostimulants or antidepressants increases the risk of falling during the following day. For every other assumed duration the Cox proportional hazards model finds strong evidence to support a hazard ratio with point estimates ranging from 1.87 to 2.33. Since each Cox proportional hazards model uses

the same number of falls when constructing the partial likelihood, the Akaike Information Criterion (AIC) can be used to compare models by selecting the model that minimizes the AIC. The AIC is minimized for the assumed duration of seven days, which estimates the effect of an increase in the number of psychostimulants or antidepressants as 2.18 (95% CI; 1.64, 2.90).

Throughout the analysis the Cox proportional hazards model provides much smaller confidence intervals; due to smaller standard errors resulting from a larger risk set. The smaller standard errors are consistent with the findings from the simulation studies. The relative efficiency of the Case-crossover design, assuming a duration of seven days, is 56% of that obtained in the Cox proportional hazards model.

The estimates from the cohort design consistently estimate a larger log-hazard than the case-crossover design. The closest estimates occurs at a duration of three days, where the cohort design point estimate is 23% larger than the case-crossover design. The largest observed difference occurs at 28 days, where the cohort design is 76% higher, but at this duration the case-crossover design is only using 14.7% of the cohort falls. At durations of 2,4,5,7,14,21, the cohort design estimate is 40%-54% larger than the case-crossover design. There are several possible explanations for why the Cox proportional hazards model estimates are different than the case-crossover design. First, the case-crossover design disregarded 42.7% to 88.8%.

One explanation could be due to the association between the increase in medications and the risk of falling is different during early days of home health care. One way to explore the impact of removing earlier falls in the case-crossover design is to also remove the same falls from the Cox proportional hazards model in the cohort design. The results of dropping the same falls is shown in the left set of columns in Table 5.5. Dropping falls in the Cox proportional hazards model actually increases estimates of the log-hazard for all assumed durations of the effect. Hence, contrary to explaining why the Cox proportional hazards model in 5.4 are higher than the case-crossover design.

Another explanation is the possible presence of an unmeasured confounder in the full cohort design that is not observed in the VNSNY Medication and Fall data. It is difficult to assess whether an unmeasured confounder can explain for the difference in log-hazard

| | Cohort Design | | | | | |
|---|---|---|---|---|---|---|
| | Drop Early Falls | | | Time Dependent Covariate Only | | |
| Duration (days) | log−hazard | Hazard Ratio(95% CI) | p−value | log−hazard | Hazard Ratio(95% CI) | p−value |
| 1 | 0.92 (0.38) | 2.50 (1.19, 5.25) | 0.0155 | 0.87 (0.38) | 2.38 (1.13, 4.99) | 0.0220 |
| 2 | 0.80 (0.29) | 2.22 (1.26, 3.91) | 0.0058 | 0.81 (0.28) | 2.26 (1.31, 3.89) | 0.0033 |
| 3 | 0.95 (0.22) | 2.58 (1.66, 4.00) | 0.0000 | 0.91 (0.22) | 2.48 (1.62, 3.81) | 0.0000 |
| 5 | 0.98 (0.18) | 2.67 (1.89, 3.79) | 0.0000 | 0.97 (0.17) | 2.65 (1.91, 3.67) | 0.0000 |
| 7 | 0.99 (0.16) | 2.70 (1.99, 3.67) | 0.0000 | 0.96 (0.15) | 2.61 (1.96, 3.47) | 0.0000 |
| 14 | 0.71 (0.16) | 2.04 (1.49, 2.79) | 0.0000 | 0.81 (0.12) | 2.24 (1.78, 2.82) | 0.0000 |
| 21 | 0.68 (0.16) | 1.97 (1.43, 2.72) | 0.0000 | 0.71 (0.11) | 2.04 (1.66, 2.51) | 0.0000 |
| 28 | 0.59 (0.19) | 1.80 (1.24, 2.62) | 0.0020 | 0.76 (0.09) | 2.13 (1.77, 2.57) | 0.0000 |

Table 5.5: Sensitivity analysis of cohort design. The first series of analysis drops the same falls as the corresponding case-crossover design assuming the same duration of the effect. The series of analysis on the right do not control for any baseline covariates and only includes the time-dependent exposure

estimates between the two designs. It is possible to examine the impact of including the baseline covariates in the Cox proportional hazards model. The right set of columns in Table 5.5 does not include any baseline covariates. The impact of only having the time-dependent covariate for the a change in antideperessant or pychostimulant medications results in larger hazard ratio estimates. These larger estimates likely suggest that some of the factors identified by the elastic net regression are positively correlated with the increase in antideperessant or pychostimulant medications and the probability of falling. An unmeasured confounder that is positively correlated with both the event and the medication increase could account for the difference in log-hazard estimates observed in Table 5.4.

The final and most likely explanation for the difference in estimates of the hazard ratio between the two designs may be explained by the reference strategy of the case-crossover design. The Cox-hazards model observes strong evidence that the duration of the effect may last up to or possibly exceed 28 days. The cohort design identified the seven day window as the the best model, but the other models may indicate that the transient effect may dissipate over the following weeks. The case-crossover design examining the effect over seven days fixes the control time to the immediately preceding seven day period. If the true effect of the medication lasts longer than seven days, then by defining the control window during a time where a carry over effect exists will bias estimates towards the null hypothesis. The difficulty in accounting for carry over effect for this length of time in the case-crossover design is the loss of an extraordinary amount of falls from the analysis given

the short time line of a typical 60-day home health care episode.

## 5.4 Exploration of a Nonlinear Effect

The cohort design in Section 5.3 identified that analysis assuming an effect lasting for seven days was the best model according to AIC. The case-crossover design also found strong evidence that the seven day window was associated with the event while using 54.7% of the falls. To further explore the transient effect in the first seven days after the medication change in more detail the following model is assumed

$$\lambda(t) = \lambda e^{\boldsymbol{X}\gamma+\beta_1 Z_{it}+\beta_2 Z_{it-1}+\beta_3 Z_{it-2}+\beta_4 Z_{it-3}+\beta_5 Z_{it-4}+\beta_6 Z_{it-5}+\beta_7 Z_{it-6}} \tag{5.3}$$

Results from this model are shown in the column labeled Model 7 in Table 5.6. This table shows that there is no evidence to support $\beta_5, \beta_6, \beta_7$ are not equal to zero. These parameters are removed one-by-one until model four is selected since it minimizes the AIC. There is some, but weak evidence (p-value=0.09) to support an association at a lag of four days after the medication change. There is strong evidence (p-value=0.01) that a change in medications increase the log-hazard ratio for falling by 1.19 fold three days after the medication change. The first two days after the medication changed yielded small estimates of the hazard ratio with no evidence to support rejecting the null hypothesis (p-value = 0.75 and 0.74 respectively). This may suggest that there is an incubation period for the effect that does not manifest until three days later.

A likelihood ratio test is constructed testing the hypothesis

$$\begin{aligned} H_o : \lambda(t) &= \lambda_0(t) e^{\boldsymbol{X}\boldsymbol{\gamma}+\beta Z_{it}} \\ H_A : \lambda(t) &= \lambda e^{\boldsymbol{X}\gamma+\beta_1 Z_{it}+\beta_2 Z_{it-1}+\beta_3 Z_{it-2}+\beta_4 Z_{it-3}+\beta_5 Z_{it-4}+\beta_6 Z_{it-5}+\beta_7 Z_{it-6}} \end{aligned} \tag{5.4}$$

where the null hypothesis assumes a constant effect lasting for seven days. This test yields a p-value=0.8438, which fails to reject the null hypothesis. Therefore, the case-crossover design concludes that the effect of the medication increase is constant over the seven days after the medication change. For illustration purposes, the constraint $\Theta_l(\beta, \theta) = \beta\theta^{(l-1)}$ to Equation (5.3) confirms the results of the likelihood ratio test. This analysis estimates the log-hazard estimate $\hat{\beta} = 0.53$ (95%CI; 0.48, 1.02) and $\hat{\theta} = 0.95$ (95%CI; 0.56, 1.36). The

| | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | | Model 6 | | Model 7 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Lag | Estimate(Std.Err) | p−value | Estimate(Std.Err) | p−value | Estimate(Std.Err) | p−value | Estimate(Std.Err) | p−value | Estimate(Std.Err) | p−value | Estimate(Std.Err) | p−value | Estimate(Std.Err) | p−value |
| 1 | 0.16 (0.50) | 0.75 | 0.16 (0.50) | 0.75 | 0.16 (0.50) | 0.75 | 0.16 (0.50) | 0.75 | 0.16 (0.50) | 0.75 | 0.16 (0.50) | 0.75 | 0.16 (0.50) | 0.75 |
| 2 | | | 0.19 (0.58) | 0.74 | 0.19 (0.58) | 0.74 | 0.19 (0.58) | 0.74 | 0.19 (0.58) | 0.74 | 0.19 (0.58) | 0.74 | 0.19 (0.58) | 0.74 |
| 3 | | | | | 1.19 (0.49) | 0.01 | 1.19 (0.49) | 0.01 | 1.19 (0.49) | 0.01 | 1.19 (0.49) | 0.01 | 1.19 (0.49) | 0.01 |
| 4 | | | | | | | 0.79 (0.46) | 0.09 | 0.79 (0.46) | 0.09 | 0.79 (0.46) | 0.09 | 0.79 (0.46) | 0.09 |
| 5 | | | | | | | | | 0.22 (0.59) | 0.71 | 0.22 (0.59) | 0.71 | 0.22 (0.59) | 0.71 |
| 6 | | | | | | | | | | | 0.91 (0.61) | 0.14 | 0.91 (0.61) | 0.14 |
| 7 | | | | | | | | | | | | | 0.18 (0.56) | 0.74 |

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 |
|---|---|---|---|---|---|---|---|
| AIC | 9801 | 9802 | 9799 | 9798 | 9800 | 9800 | 9802 |

Table 5.6: Results from seven day distributed lag case-crossover design

| Lag | Case−Crossover | | | Cox Hazards | | | Relative |
|---|---|---|---|---|---|---|---|
| (days) | Estimate | Hazard Ratio | p−value | Cox Hazards | Hazard Ratio | p−value | Efficiency |
| 1−7 | 0.45 (0.35) | 1.57 (0.78, 3.14) | 0.2047 | 0.79 (0.15) | 2.19 (1.65, 2.91) | 0.0000 | 16.82% |
| 8−14 | 0.88 (0.43) | 2.42 (1.05, 5.56) | 0.0381 | 0.41 (0.20) | 1.51 (1.03, 2.22) | 0.0359 | 21.46% |
| 15−21 | 0.10 (0.42) | 1.11 (0.49, 2.51) | 0.8097 | 0.21 (0.24) | 1.24 (0.77, 1.99) | 0.3851 | 34.01% |

Table 5.7: Results from analysis with three lagged covariates defined as seven day intervals. Case-crossover design has allowed for one additional week for possible carry-over effect

confidence interval for $\hat{\theta}$ implies that $\theta$ can be assumed to equal one, and suggests that the effect is constant and $\hat{\beta}$ is consistent with the results form Table 5.4.

The results from Table 5.4 suggests that the effect may last longer than seven days. A final analysis considers the hazard in Equation (5.5)

$$\lambda(t) = \lambda e^{\boldsymbol{X}\gamma+\beta_1 Z_{i,t^*}+\beta_2 Z_{i,t^*-1}+\beta_3 Z_{i,t^*-2}} \tag{5.5}$$

where the special notation $t^*$ is introduced since the lagged covariates are now expressed as seven day intervals of time such that

$$\begin{aligned}
Z_{it^*} &= \begin{cases} 0 & t \leq \tau_i \\ 1 & t \in (\tau_i, \tau_i + 7] \\ 0 & t > \tau_i + 7 \end{cases} \\
Z_{i,t^*-1} &= \begin{cases} 0 & t \leq \tau_i + 7 \\ 1 & t \in (\tau_i + 7, \tau_i + 14] \\ 0 & t > \tau_i + 14 \end{cases} \\
Z_{i,t^*-2} &= \begin{cases} 0 & t \leq \tau_i + 14 \\ 1 & t \in (\tau_i + 14, \tau_i + 21] \\ 0 & t > \tau_i + 21 \end{cases}
\end{aligned} \tag{5.6}$$

The pattern of the lagged seven day covariates in Table 5.7 for the Cox proportional hazards model appear to follow the geometric lag relationship. The hazard of falling increases by 2.2 (95% CI; 1.65, 2.91) in the first week after observing an increase in antidepressants or psychostimulants medications. The risk subsides some what during the second week after the increase, but the risk of falling is still elevated by 1.51 (95% CI; 1.03, 2.22) times. During the third week following the medication increase there is very weak evidence (p-value 0.39) that there may still be some elevated hazard of falling with a hazard ratio of

1.24 (95% CI; 0.77, 1.99), but with the 95% confidence interval overlapping one the effect is minimum.

Due to the design reference strategy, the case-crossover design in Table 5.7 only has 18.8% (1223) of the falls available in the cohort design. The effect estimated by the case-crossover design does not follow the same deteriorating effect that the Cox proportional hazards model observed. A likelihood ratio test comparing a naive model over $\Delta = 21$ days, to the three variable, seven day interval model such that

$$\begin{aligned} H_o : \lambda(t) &= \lambda_0(t) e^{\boldsymbol{X}\boldsymbol{\gamma} + \beta Z_{it^*}} \\ H_A : \lambda(t) &= \lambda e^{\boldsymbol{X}\gamma + \beta_1 Z_{i,t^*} + \beta_2 Z_{i,t^*-1} + \beta_3 Z_{i,t^*-2}} \end{aligned} \tag{5.7}$$

fails to reject the null hypothesis (p-value=0.64). Because this case-crossover design is using so few falls and since a robust set of baseline factors are controlled for in the Cox proportional hazards model, the effect described by the cohort design is believed to most accurately describe the effect of the medication change.

# Chapter 6

# Summary and Discussion

This thesis establishes an important connection between the "prospective" cohort design and the case-crossover design. By making this connection, this thesis demonstrates that the concept of distributed lag models can be applied to the case-crossover design to estimate non-linear effects. It also shows that the translation of distributed lag variables from the "prospective" cohort design to the case-crossover design provides a mirror image interpretation of the association. Lagged covariates from the case-crossover design can be "reflected" for interpretation under the "prospective" paradigm. Finally, this thesis introduces a geometric, distributed lag, estimation method to the case-crossover design, which provides a parsimonious, two-parameter estimation in the presence of a non-linear deteriorating effect.

The simulations prove that the geometric, two-parameter, model provides unbiased estimates, but coverage probabilities are less than desirable due to underestimated standard errors. Other studies have found that distributed lag models underestimate standard errors [61, 94]. Investigators are recommended to report bootstrapped standard errors to avoided inflated type I errors [94]. Furthermore, it has been shown that the likelihood ratio tests or AIC and BIC criteria can be used to assess competing model specifications, but likelihood ratio tests may favor the more constrained model [29, 94]. The results of the distributed lag analysis for the two competing designs in the VNSNY Medication and Fall data provide some what different results. The Cox proportional hazards model results are consistent with a geometric lag over three weeks, but the unconstrained distributed lag for the case-crossover design suggests something more attune to a quadratic relationship. Almon also presented the polynomial distributed lag model, which could be considered in future work [2]. In this analysis, the results from the case-crossover design are not

considered informative or representative of the true effect due to the excessive loss of data.

Simulations and the VNSNY Medication and Fall data analysis both assume that the effect deteriorated or can be described by unconstrained distributed lag model with $L$ lagged variables. In all instances, $L$ is fixed by the investigator. In practice, however, the investigator should consider other methods, such as the use splines, for exploring the non-linear relationship. Let the interval $[\xi_{i(M)b}, \xi_{i(M)e}]$ denote the beginning and end times of the $M^{th}$ strata from the case-crossover design, and let the time from medication change in reference to the event, or end of the control interval, be defined as

$$\mu_{i(M)} = \begin{cases} \xi_{i(M)e} - \tau_i, & \tau_i \in [\xi_{i(M)b}, \xi_{i(M)e}] \\ 0 & \tau_i \notin [\xi_{i(M)b}, \xi_{i(M)e}] \end{cases} \tag{6.1}$$

Then, a conditional likelihood for a linear spline which divides the time from medication change into $k$ knots at $a, b, c, ...$ is constructed as

$$L_{cco}(\beta_k) = \prod_{i \in \{i; \delta_i = 1\}} \frac{e^{\beta_1 (\mu_{i(0)} - a)_+ + \beta_2 (\mu_{i(0)} - b)_+ + \beta_3 (\mu_{i(0)} - c)_+ + ...}}{\sum_{m=0}^{M} e^{\beta_1 (\mu_{i(M)} - a)_+ + \beta_2 (\mu_{i(M)} - b)_+ + \beta_3 (\mu_{i(M)} - c)_+ + ...}} \tag{6.2}$$

where

$$(u)_+ = \begin{cases} u & u > 0 \\ 0 & u \leq 0 \end{cases} \tag{6.3}$$

Gasparrini has recently presented a general statistical framework for modeling distributed lag non-linear models [29]. He uses the notation $s(z, t)$ to specify an exposure-response curve in terms of exposure history of $z$ evaluated at time $t$ such that

$$s(z, t) = \sum_{l=l_o}^{L} f(x_{t-l}) * w(l) \tag{6.4}$$

He has demonstrated this framework in Cox proportional hazards models, but can be applied to various study designs and regression models and has code available in the **R** package dlnm [28]. The next steps following the work of this thesis is to incorporate Gasparrini framework into the case-crossover design in future work.

The simulation studies in this thesis are believed to be the first to evaluate the case-crossover design using a time-to-event data and under the case paradigm; the population giving rise to the cases are also included in the comparative cohort design. This thesis shows

that the case-crossover and the "prospective" cohort design obtain unbiased estimates of the log-hazard ratio when the sample size is sufficient. In smaller sample sizes, however, the 1:1 and 1:2 case-crossover designs tend to over estimate the magnitude of the log-hazard ratio when an effect truly does exist, while the Cox-hazards model and the case-crossover design, matching 10 or 25 controls, typically underestimates the log-hazard ratio. In the simulations presented in this thesis, the 1:3 case-crossover design results in estimates with the least amount of bias when the number of events is 500 or less. Further analysis of the simulations show that the number of discordant pairs for the conditional logistic regression is very small, yielding only about 6 per simulation for 5000 cohort and 10%.

Bias in conditional logistic regression has been thoroughly documented for sparse data sets [32, 33, 38, 92]. Sun et al. also observed bias in simulations applying conditional maximum likelihood to 1:1 and 1:2 matched case-control studies with small sample size [92]. The authors have attributed the overestimation to the asymptotic properties of conditional maximum likelihood and have suggested corrective methods to obtain unbiased estimates in small sample sizes [92]. Heinze and Puhr have observed bias of similar magnitude for 1:4 matched-case-control data [38]. The simulations in this dissertation are based on Austin's guidelines generating survival times to simulate Cox proportional hazards models with time-varying covariates [3]. Austin and colleagues have recently published a series of simulation studies similar to those in this dissertation. He observed bias for a time-varying variable exceeding -20% for cohorts with 5% incidence and a treatment prevalence of 10% [4]. Furthermore, Austin also observed that this negative bias was still present but converged to 0 with larger treatment prevalence and incidence rates.

This thesis has also provides evidence to conclude that the case-crossover design is preferred over the Cox proportional hazards model when unmeasured confounders are present at baseline. When a baseline factor is present and correlated with the exposure and the event, the Cox hazards model over-estimates the log-hazard ratio by up to 25% compared to the true value. When the exposure is assumed to be independent of event, the bias inflates empirical Type I errors reaching up to 50% when the nominal type I error is 5%. Erroneously concluding a drug has a side effect, such as falling, could have serious implications. Such type I errors may result in physicians and patients being reluctant to

take a drug because of potential side effects even though it has real therapeutic value.

In settings where the investigator is able to observe all confounders, the Cox proportional hazards model approach is preferred because of its superior efficiency compared to the case-crossover design. When $\beta = 0$, the relative efficiency achieved in the case-crossover design compared to the cohort design follows the relative efficiency $= M/(M+1)$ rule established by Brewslow and Day for matched case-control designs [11, 12]. A 1:5 case-crossover design is at the limits of what the authors consider feasible in practice. In addition, the larger the number of matched controls can expose the case-crossover design to other forms of bias associated with increasing the within subject time frame for reference strategy. When $M$ increases, the effect of capturing trends in the exposure or non-constant hazards will be magnified when otherwise may be assumed to have minimal effect. Even though the Cox proportional hazards model is consistently preferred in terms of efficiency, an assessment of relative efficiency should take into consideration the cost associated with obtaining data on patients who did not experience the event.

A major disadvantage to the case-crossover design in VNSNY Medication and Fall data set is the amount of data discarded due to limitations of the reference strategy. On top of losing efficiency, loss of data presents complications when assessing competing models or exploring the duration of the effect. By increasing the assumed duration, the length of the case- and control- windows must be increased, and as a result additional observations are discarded. Goodness of fit criteria, such as AIC or BIC, are incomparable between case-crossover designs when number of events used in the conditional likelihood are different. The duration of the effect observed in the analysis of the VNSNY Medication and Fall data was unanticipated and presents another issue when applying the case-crossover design to this data. Previous case-crossover design studies have only considered transient effects resulting from a medication change that lasted days, and the effect observed by the Cox hazards analysis suggested the effect may last as long as three weeks after the medication change. Selecting control-times and not appropriately accounting for the carry-over effect will bias results towards the null. If the full cohort is available in practice as it is here, the Cox hazard model may serve as a diagnostic method for determining how long the carry over effect may last. The combination of the relative short follow-up time and the duration

of the effect for this particular medication classification suggests that the case-crossover design is not applicable to the VNSNY Medication and Fall data.

The Cox proportional hazards model was preferred in a majority of the simulation studies and in the VNSNY Medication and Fall data. Still, there are several reasons why the investigator may still want to consider the case-crossover design. First, it may still be preferable in some settings since it is able to implicitly control for confounders that do not change over time. Obtaining data on potential confounders may not be as convenient and complete in other applications. It could also be preferable when there is a substantial cost to obtaining data. The investigator may also want to avoid applying the Cox proportional hazards model because it requires checking the proportional hazards assumption for each individual covariate. This assumption could be violated, or the exercise of checking each covariate may be daunting when a study is controlling for a larger number of covariates. Finally, the time of origin may be difficult to define in many applications. The Cox hazards model is not the only approach to exploring the type of associations with medication changes and the risk of falling. Many researchers may wish to use other methods such as a longitudinal logistic regression with time-varying variables [83, 86].

## 6.1 Limitations and Future Work

Future work may consider studying the effects of violating several assumptions made about the time-to-event data generated in the simulation studies. First, the exposure distribution in these studies were assumed to be independent and constant. The case-crossover design's estimates will be biased when the exposure of interest tends to either increase or decrease for all patients in the study [85] . The self-controlled, case series method, described in Chapter 2, is another epidemiological design methodology proposed as an alternative to the case-crossover design and is able to account for secular trends in the exposure [20, 21, 22]. Future work may consider introducing a trend to the exposure and include the case series analysis in the time-to-event simulations. If the trend is consistent in both patients who experience the event and those who do not, it is hypothesized that the Cox proportional hazards model should be able to take trends into account while yielding superior efficiency over the case series design.

Future work may also explore the effect on the case-crossover design when censoring is informative. In addition, these studies only considered the time to the first fall even though patients can experience more than one fall. Recurrent events within the case-crossover design has already been explored by [58], but future work may compare their methodology to how the Cox proportional hazards model can take within subject correlations into effect.

# Appendix

| | |
|---|---|
| DATE: | March 1, 2013 |
| TO: | Carlin Brickner, M.S. |
| FROM: | Visiting Nurse Service of New York Institutional Review Board (IRB) |
| STUDY TITLE: | [419844-1] Medication Changes and Falls: An Investigation on the Performance of the Case-Crossover Design in the Home Health Care Setting |
| IRB REFERENCE #: | I13-002 |
| SUBMISSION TYPE: | New Project |
| ACTION: | APPROVED |
| APPROVAL DATE: | February 28, 2013 |
| EXPIRATION DATE: | February 27, 2014 |
| REVIEW TYPE: | Expedited Review |
| REVIEW CATEGORY: | Minimal Risk |

Thank you for your submission of New Project materials for this research study. Visiting Nurse Service of New York Institutional Review Board (IRB) has APPROVED your submission. This approval is based on an appropriate risk/benefit ratio and a study design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on the applicable federal regulation.

Please remember that informed consent is a process beginning with a description of the study and insurance of participant understanding followed by a signed consent form. Informed consent must continue throughout the study via a dialogue between the researcher and research participant. Federal regulations require each participant receive a copy of the signed consent document.

Please note that any revision to previously approved materials must be approved by this office prior to initiation. Please use the appropriate revision forms for this procedure.

All SERIOUS and UNEXPECTED adverse events must be reported to this office. Please use the appropriate adverse event forms for this procedure. All FDA and sponsor reporting requirements should also be followed.

Please report all NON-COMPLIANCE issues or COMPLAINTS regarding this study to this office.

Please note that all research records must be retained for a minimum of three years.

Based on the risks, this project requires Continuing Review by this office on an annual basis. Please use the appropriate renewal forms for this procedure.

If you have any questions, please contact Lori King at (212) 609-5766 or lori.king@vnsny.org. Please include your study title and reference number in all correspondence with this office.

For the Institutional Review Board,

Susan Regan, IRB Chair

## Waiver of Patient Authorization

Study Title: [419844-1] Medication Changes and Falls: An Investigation on the Performance of the Case-Crossover Design in the Home Health Care Setting
PI Name: Carlin Brickner, M.S.

## Documentation of Waiver Approval

***The research proposal listed above meets the following criteria:***

1. The research could not practicably be conducted without the waiver
2. The research could not practicably be conducted without access to and use of the protected health information described within the research project proposal
3. The use or disclosure of protected health information involves no more than minimal risk to the privacy of individuals, based on, at least, the presence of the following elements:

   a. An adequate plan to protect the identifiers from improper use and disclosure;
   b. An adequate plan to destroy the identifiers at the earliest opportunity consistent with conduct of the research, unless there is health or research justification for retaining the identifiers, or such retention is otherwise required by law; and
   c. Adequate written assurances that the protected health information will not be reused or disclosed to any other person or entity, except as required by law, for authorized oversight of the research project, or for other research for which the use or disclosure of protected health information would be permitted by this subpart

***A brief description of the protected health information being requested for use or access in the research project listed above has been reviewed, and has been determined to be necessary in conducting the research. All of the necessary criteria required under the Privacy Rule (45 CFR §164.512) have been satisfied for approval of waiver of authorization. The Institutional Review Board of VNSNY has reviewed and approved the waiver of authorization under normal or expedited review procedures as outlined in the Common Rule.***

Susan Regan, IRB Chair

***Protected health information** refers to individually identifiable health information maintained or transmitted electronically or in any other form or medium, including written and oral methods

| | |
|---|---|
| DATE: | January 28, 2014 |
| TO: | Carlin Brickner, M.S. |
| FROM: | Visiting Nurse Service of New York Institutional Review Board (IRB) |
| STUDY TITLE: | [419844-2] Medication Changes and Falls: An Investigation on the Performance of the Case-Crossover Design in the Home Health Care Setting |
| IRB REFERENCE #: | I13-002 |
| SUBMISSION TYPE: | Continuing Review/Progress Report |
| ACTION: | APPROVED |
| APPROVAL DATE: | January 28, 2014 |
| EXPIRATION DATE: | January 27, 2015 |
| REVIEW TYPE: | Expedited Review |
| REVIEW CATEGORY: | Minimal Risk |

Thank you for your submission of Continuing Review/Progress Report materials for this research study. Visiting Nurse Service of New York Institutional Review Board (IRB) has APPROVED your submission. This approval is based on an appropriate risk/benefit ratio and a study design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on the applicable federal regulation.

Please remember that informed consent is a process beginning with a description of the study and insurance of participant understanding followed by a signed consent form. Informed consent must continue throughout the study via a dialogue between the researcher and research participant. Federal regulations require each participant receive a copy of the signed consent document.

Please note that any revision to previously approved materials must be approved by this office prior to initiation. Please use the appropriate revision forms for this procedure.

All SERIOUS and UNEXPECTED adverse events must be reported to this office. Please use the appropriate adverse event forms for this procedure. All FDA and sponsor reporting requirements should also be followed.

Please report all NON-COMPLIANCE issues or COMPLAINTS regarding this study to this office.

Please note that all research records must be retained for a minimum of three years.

Based on the risks, this project requires Continuing Review by this office on an annual basis. Please use the appropriate renewal forms for this procedure.

If you have any questions, please contact Lori King at (212) 609-5766 or lori.king@vnsny.org. Please include your study title and reference number in all correspondence with this office.

For the Institutional Review Board,

Susan Regan, IRB Chair

RUTGERS
THE STATE UNIVERSITY OF NEW JERSEY

** This is an auto-generated email. Please do not reply to this email message.
The originating e-mail account is not monitored.
If you have questions, please contact your local IRB office or log into eIRB.Rutgers.edu **

**DHHS Federal Wide Assurance Identifier:** FWA00003913
**IRB Chair Person:** Nancy Fiedler
**IRB Director:** Donna Hoagland
**Effective Date:** 4/23/2014

## eIRB Notice of Approval

STUDY PROFILE

**Study ID:** Pro2013003710

**Title:** Medication Changes and Falls: An Investigation on the Performance of the Case-Crossover Design in the Home Health Care Setting

| | | | |
|---|---|---|---|
| **Principal Investigator:** | Carlin Brickner | **Study Coordinator:** | |
| **Co-Investigator(s):** | Shou-En Lu<br>Dirk Moore | **Other Study Staff:** | There are no items to display |
| **Sponsor:** | There are no items to display | **Approval Cycle:** | Twelve Months |
| **Risk Determination:** | Minimal Risk | **Device Determination:** | Not Applicable |

| | | | | | |
|---|---|---|---|---|---|
| **Review Type:** | Expedited | **Expedited Category:** | There are no items to display | **Exempt Category:** | There are no items to display |
| **Subjects:** | 0 | **Specimens:** | | **Records:** | |

CURRENT SUBMISSION STATUS

| | | | |
|---|---|---|---|
| **Submission Type:** | Facilitated Review or NCI-CIRB Independent Review | **Submission Status:** | Approved |
| **Approval Date:** | 4/23/2014 | **Expiration Date:** | 4/22/2015 |

| | | | | | |
|---|---|---|---|---|---|
| **Pregnancy Code:** | No Pregnant Women as Subjects<br>Not Applicable | **Pediatric Code:** | No Children As Subjects | **Prisoner Code:** | There are no items to display |

| | | | | | |
|---|---|---|---|---|---|
| **Protocol:** | protocol_brickner20140113 | **Consent:** | There are no items to display | **Other Materials:** | There are no items to display |

* IRB APPROVAL IS GRANTED SUBJECT TO THE STIPULATION(S) THAT:

*** Study Performance Sites:**

There are no items to display

| | |
|---|---|
| Visiting Nurse Service of New York | The Center for Home Care Policy & Research 1250 Broadway - 20th Floor New York, NY 10001 |

**ALL APPROVED INVESTIGATOR(S) MUST COMPLY WITH THE FOLLOWING:**

1. Conduct the research in accordance with the protocol, applicable laws and regulations, and the principles of research ethics as set forth in the Belmont Report.

2. **Continuing Review:** Approval is valid until the protocol expiration date shown above. To avoid lapses in approval, submit a continuation application at least eight weeks before the study expiration date.

3. **Expiration of IRB Approval:** If IRB approval expires, effective the date of expiration and until the continuing review approval is issued: **All research activities must stop unless the IRB finds that it is in the best interest of individual subjects to continue. (This determination shall be based on a separate written request from the PI to the IRB.) No new subjects may be enrolled and no samples/charts/surveys may be collected, reviewed, and/or analyzed.**

4. **Amendments/Modifications/Revisions**: If you wish to change any aspect of this study, including but not limited to, study procedures, consent form(s), investigators, advertisements, the protocol document, investigator drug brochure, or accrual goals, you are required to obtain IRB review and approval prior to implementation of these changes unless necessary to eliminate apparent immediate hazards to subjects.

5. **Unanticipated Problems**: Unanticipated problems involving risk to subjects or others must be reported to the IRB Office (45 CFR 46, 21 CFR 312, 812) as required, in the appropriate time as specified in the attachment online at: http://rbhs.rutgers.edu/hsweb

6. **Protocol Deviations and Violations**: Deviations from/violations of the approved study protocol must be reported to the IRB Office (45 CFR 46, 21 CFR 312, 812) as required, in the appropriate time as specified in the attachment online at: http://rbhs.rutgers.edu/hsweb

7. **Consent/Assent**: The IRB has reviewed and approved the consent and/or assent process, waiver and/or alteration described in this protocol as required by 45 CFR 46 and 21 CFR 50, 56, (if FDA regulated research). Only the versions of the documents included in the approved process may be used to document informed consent and/or assent of study subjects; each subject must receive a copy of the approved form(s); and a copy of each signed form must be filed in a secure place in

the subject's medical/patient/research record.

8. **Completion of Study:** Notify the IRB when your study has been stopped for any reason. Neither study closure by the sponsor or the investigator removes the obligation for submission of timely continuing review application or final report.

9. The Investigator(s) did not participate in the review, discussion, or vote of this protocol.

10. Letter Comments: *There are no additional comments.*

# References


[1] P. D. Allison and N. A. Christakis. Fixed-effects methods for the analysis of nonrepeated events. *Sociological Methodology*, 36(1):155–172, 2006.

[2] S. Almon. The distributed lag between capital appropriations and expenditures. *Econometrica*, 33(1):pp. 178–196, 1965.

[3] P. C. Austin. Generating survival times to simulate cox proportional hazards models with time-varying covariates. *Statistics in Medicine*, 31(29):3946–3958, Dec 2012.

[4] P. C. Austin, G. M. Anderson, C. Cigsar, and A. Gruneir. Comparing the cohort design and the nested case–control design in the presence of both time-invariant and time-dependent treatment and competing risks: bias and precision. *Pharmacoepidemiology and Drug Safety*, 21(7):714–724, 2012.

[5] M. Avalos, Y. Grandvalet, N. D. Adroher, L. Orriols, and E. Lagarde. Analysis of multiple exposures in the case-crossover design via sparse conditional likelihood. *Statistics in Medicine*, 31(21):2290–2302, Sep 2012.

[6] J. C. Ayus, A. L. Negri, K. Kalantar-Zadeh, and M. L. Moritz. Is chronic hyponatremia a novel risk factor for hip fracture in the elderly? *Nephrol Dial Transplant*, 27(10):3725–3731, Oct 2012.

[7] T. F. Bateson and J. Schwartz. Selection bias and confounding in case-crossover analyses of environmental time-series data. *Epidemiology*, 12(6):654–661, Nov 2001.

[8] H. D. Berlie and C. L. Garwood. Diabetes medications related to an increased risk of falls and fall-related morbidity in the elderly. *Annals of Pharmacotherapy*, 44(4):712–717, Apr 2010.

[9] C. E. Bishop, D. Gilden, J. Blom, J. Kubisiak, R. Hakim, A. Lee, and D. W. Garnick. Medicare spending for injured elders: are there opportunities for savings? *Health Affairs*, 21(6):215–223, 2002.

[10] J. M. Bolton, C. Metge, L. Lix, H. Prior, J. Sareen, and W. D. Leslie. Fracture risk from psychotropic medications: a population-based analysis. *Journal of Clinical Psychopharmacology*, 28(4):384–391, Aug 2008.

[11] N. E. Breslow and N. E. Day. Statistical methods in cancer research. volume i - the analysis of case-control studies. *IARC Scientific Publications*, (32):5–338, 1980.

[12] N. E. Breslow and N. E. Day. Statistical methods in cancer research. volume ii - the design and analysis of cohort studies. *IARC Scientific Publications*, (32):5–338, 1980.

[13] D. A. Butt, M. Mamdani, P. C. Austin, K. Tu, T. Gomes, and R. H. Glazier. The risk of hip fracture after initiating antihypertensive drugs in the elderly. *Archives of Internal Medicine*, 172(22):1739–1744, Dec 2012.

[14] E. Carracedo-Martinez, C. Sanchez, M. Taracido, M. Saez, V. Jato, and A. Figueiras. Effect of short-term exposure to air pollution and pollen on medical emergency calls: a case-crossover study in spain. *Allergy*, 63(3):347–353, 2008.

[15] E. Carracedo-Martnez, M. Taracido, A. Tobias, M. Saez, and A. Figueiras. Case-crossover analysis of air pollution health effects: a systematic review of methodology and application. *Environ Health Perspect*, 118(8):1173–1182, Aug 2010.

[16] Center for Medicare and Medicaid Services. OASIS C. Online, *http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HomeHealthQualityInits/OASISC.html*, Accessed August 2014.

[17] D. Chew, S. Mattschoss, M. Horsfall, C. Astley, J. Vaile, and M. Joseph. Patterns of inflammatory activation associated with precipitants of acute coronary syndromes: a case-crossover study. *Internal medicine journal*, 42(10):1096–1103, 2012.

[18] G. P. Consiglio, A. M. Burden, M. Maclure, L. McCarthy, and S. M. Cadarette. Case-crossover study design in pharmacoepidemiology: systematic review and recommendations. *Pharmacoepidemiol Drug Saf*, Sep 2013.

[19] S. J. Diem, T. L. Blackwell, K. L. Stone, K. Yaffe, E. M. Haney, M. M. Bliziotes, and K. E. Ensrud. Use of antidepressants and rates of hip bone loss in older women: the study of osteoporotic fractures. *Archives of Internal Medicine*, 167(12):1240–1245, Jun 2007.

[20] C. Farrington, J. Nash, and E. Miller. Case series analysis of adverse reactions to vaccines: a comparative evaluation. *American journal of epidemiology*, 143(11):1165–1173, 1996.

[21] C. P. Farrington. Relative incidence estimation from case series for vaccine safety evaluation. *Biometrics*, 51(1):228–235, Mar 1995.

[22] C. P. Farrington. Control without separate controls: evaluation of vaccine safety using case-only methods. *Vaccine*, 22(15-16):2064–2070, May 2004.

[23] A. Figueiras, E. Carracedo-Martnez, M. Saez, and M. Taracido. Analysis of case-crossover designs using longitudinal approaches: a simulation study. *Epidemiology*, 16(2):239–246, Mar 2005.

[24] L. Filleul, V. Rondeau, A. Cantagrel, J.-F. Dartigues, and J.-F. Tessier. Do subject characteristics modify the effects of particulate air pollution on daily mortality among the elderly? *Journal of occupational and environmental medicine*, 46(11):1115–1122, 2004.

[25] C. for Disease Control, N. C. f. I. P. Prevention, and Control. Injury Prevention and Control: Data and Statistics. Online, *http://www.cdc.gov/injury/wisqars/*, Accessed March 2013.

[26] C. for Medicare and M. Servicess. Home Health PPS. Online, *http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/HomeHealthPPS/*, Accessed August 2014.

[27] F. Forastiere, M. Stafoggia, G. Berti, L. Bisanti, A. Cernigliaro, M. Chiusolo, S. Mallone, R. Miglio, P. Pandolfi, M. Rognoni, et al. Particulate matter and daily mortality: a case-crossover analysis of individual effect modifiers. *Epidemiology*, 19(4):571–580, 2008.

[28] A. Gasparrini. Distributed lag linear and non-linear models in R: the package dlnm. *Journal of Statistical Software*, 43(8):1–20, 2011.

[29] A. Gasparrini. Modeling exposure-lag-response associations with distributed lag non-linear models. *Statistics in Medicine*, 33(5):881–899, Feb 2014.

[30] S. Greenland. Confounding and exposure trends in case-crossover and case-time-control designs. *Epidemiology*, 7(3):231–239, May 1996.

[31] S. Greenland. A unified approach to the analysis of case-distribution (case-only) studies. *Statistics in medicine*, 18(1):1–15, 1999.

[32] S. Greenland. Small-sample bias and corrections for conditional maximum-likelihood odds-ratio estimators. *Biostatistics*, 1(1):113–122, 2000.

[33] S. Greenland, J. A. Schwartzbaum, and W. D. Finkle. Problems due to small samples and sparse data in conditional logistic regression analysis. *American Journal of Epidemiology*, 151(5):531–539, 2000.

[34] J. Gribbin, R. Hubbard, J. Gladman, C. Smith, and S. Lewis. Risk of falls associated with antihypertensive medication: self-controlled case series. *Pharmacoepidemiol Drug Saf*, 20(8):879–884, Aug 2011.

[35] S. J. Hambidge, J. M. Glanz, E. K. France, D. McClure, S. Xu, K. Yamasaki, L. Jackson, J. P. Mullooly, K. M. Zangwill, S. M. Marcy, S. B. Black, E. M. Lewis, H. R. Shinefield, E. Belongia, J. Nordin, R. T. Chen, D. K. Shay, R. L. Davis, F. DeStefano, and V. S. D. T. . Safety of trivalent inactivated influenza vaccine in children 6 to 23 months old. *JAMA*, 296(16):1990–1997, Oct 2006.

[36] E. M. Haney, B. K. S. Chan, S. J. Diem, K. E. Ensrud, J. A. Cauley, E. Barrett-Connor, E. Orwoll, M. M. Bliziotes, and O. F. i. M. S. G. . Association of low bone mineral density with selective serotonin reuptake inhibitor use by older men. *Archives of Internal Medicine*, 167(12):1246–1251, Jun 2007.

[37] J. M. Hausdorff, D. A. Rios, and H. K. Edelberg. Gait variability and fall risk in community-living older adults: a 1-year prospective study. *Archives of Physical Medicine and Rehabilitation*, 82(8):1050–1056, Aug 2001.

[38] G. Heinze and R. Puhr. Bias-reduced and separation-proof conditional logistic regression with small or sparse data sets. *Statistics in medicine*, 29(7-8):770–777, 2010.

[39] S. Hernndez-Daz, M. A. Hernn, K. Meyer, M. M. Werler, and A. A. Mitchell. Case-crossover and case-time-control designs in birth defects epidemiology. *American Journal of Epidemiology*, 158(4):385–391, Aug 2003.

[40] E. B. Hoffman, P. K. Sen, and C. R. Weinberg. Within-cluster resampling. *Biometrika*, 88:1121–1134, 2001.

[41] M. C. Hornbrook, V. J. Stevens, D. J. Wingfield, J. F. Hollis, M. R. Greenlick, and M. G. Ory. Preventing falls among community-dwelling older persons: results from a randomized trial. *Gerontologist*, 34(1):16–23, Feb 1994.

[42] D. W. Hosmer, S. Lemeshow, and S. May. *Applied Survival Analysis*. Wiley, 2nd edition, 2008.

[43] S. Hugonnet, A. Villaveces, and D. Pittet. Nurse staffing level and nosocomial infections: empirical evaluation of the case-crossover and case-time-control designs. *American Journal of Epidemiology*, 165(11):1321–1327, Jun 2007.

[44] H. Janes, L. Sheppard, and T. Lumley. Overlap bias in the case-crossover design, with application to air pollution exposures. *Statistics in Medicine*, 24(2):285–300, Jan 2005.

[45] R. T. Jerome Friedman, Trevor Hastie. Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1):1–22, 2010.

[46] S. M. Lavsa, T. J. Fabian, M. I. Saul, S. L. Corman, and K. C. Coley. Influence of medications and diagnoses on fall risk in psychiatric inpatients. *American Journal of Health-System Pharmacy*, 67(15):1274–1280, Aug 2010.

[47] J.-T. Lee and J. Schwartz. Reanalysis of the effects of air pollution on daily mortality in seoul, korea: A case-crossover design. *Environmental Health Perspectives*, 107(8):633, 1999.

[48] R. M. Leipzig, R. G. Cumming, and M. E. Tinetti. Drugs and falls in older people: a systematic review and meta-analysis: I. psychotropic drugs. *Journal of the American Geriatrics Society*, 47(1):30–39, Jan 1999.

[49] R. M. Leipzig, R. G. Cumming, and M. E. Tinetti. Drugs and falls in older people: a systematic review and meta-analysis: Ii. cardiac and analgesic drugs. *Journal of the American Geriatrics Society*, 47(1):40–50, Jan 1999.

[50] F. Leisch, A. Weingessel, and K. Hornik. *bindata: Generation of Artificial Binary Data*, 2012. R package version 0.9-19.

[51] J. Lepeule, V. Rondeau, L. Filleul, and J.-F. Dartigues. Survival analysis to estimate association between short-term mortality and air pollution. *Environ Health Perspect*, 114(2):242–247, Feb 2006.

[52] D. Levy, L. Sheppard, H. Checkoway, J. Kaufman, T. Lumley, J. Koenig, and D. Siscovick. A case-crossover analysis of particulate matter air pollution and out-of-hospital primary cardiac arrest. *Epidemiology*, 12(2):193–199, 2001.

[53] M. Lin, Y. Chen, R. T. Burnett, P. J. Villeneuve, and D. Krewski. The influence of ambient coarse particulate matter on asthma hospitalization in children: case-crossover and time-series analyses. *Environmental health perspectives*, 110(6):575, 2002.

[54] M. Lin, D. M. Stieb, and Y. Chen. Coarse particulate matter and hospitalization for respiratory infections in children younger than 15 years in toronto: a case-crossover analysis. *Pediatrics*, 116(2):e235–e240, 2005.

[55] Y. Lu, J. M. Symons, A. S. Geyh, and S. L. Zeger. An approach to checking case-crossover analyses based on equivalence with time-series methods. *Epidemiology*, 19(2):169–175, Mar 2008.

[56] Y. Lu and S. L. Zeger. On the equivalence of case-crossover and time series methods in environmental epidemiology. *Biostatistics*, 8(2):337–344, Apr 2007.

[57] T. Lumley and D. Levy. Bias in the case-crossover design: implications for studies of air pollution. *Environmetrics*, 11(6):689–704, 2000.

[58] X. Luo and G. S. Sorock. Analysis of recurrent event data under the case-crossover design with applications to elderly falls. *Statistics in Medicine*, 27(15):2890–2901, Jul 2008.

[59] M. Maclure. The case-crossover design: a method for studying transient effects on the risk of acute events. *American Journal of Epidemiology*, 133(2):144–153, Jan 1991.

[60] M. Maclure and M. A. Mittleman. Should we use a case-crossover design? *Annual Review of Public Health*, 21:193–221, 2000.

[61] M. Mahmud, M. Abrahamowicz, K. Leffondré, and Y. P. Chaubey. Selecting the optimal transformation of a continuous covariate in cox's regression: Implications for hypothesis testing. *Communications in Statistics?Simulation and Computation®*, 35(1):27–45, 2006.

[62] R. J. Marshall and R. T. Jackson. Analysis of case-crossover designs. *Statistics in Medicine*, 12(24):2333–2341, Dec 1993.

[63] S. P. McEvoy, M. R. Stevenson, A. T. McCartt, M. Woodward, C. Haworth, P. Palamara, and R. Cercarelli. Role of mobile phones in motor vehicle crashes resulting in hospital attendance: a case-crossover study. *BMJ*, 331(7514):428, Aug 2005.

[64] L. B. Meuleners, J. Duke, A. H. Lee, P. Palamara, J. Hildebrand, and J. Q. Ng. Psychoactive medications and crash involvement requiring hospitalization for older drivers: a population-based study. *Journal of the American Geriatrics Society*, 59(9):1575–1580, Sep 2011.

[65] M. A. Mittleman, M. Maclure, J. B. Sherwood, R. P. Mulry, G. H. Tofler, S. C. Jacobs, R. Friedman, H. Benson, and J. E. Muller. Triggering of acute myocardial infarction onset by episodes of anger. determinants of myocardial infarction onset study investigators. *Circulation*, 92(7):1720–1725, Oct 1995.

[66] J. Mller, J. Hallqvist, L. Laflamme, F. Mattsson, S. Ponzer, S. Sadigh, and K. Engstrm. Emotional stress as a trigger of falls leading to hip or pelvic fracture. results from the tofa study - a case-crossover study among elderly people in stockholm, sweden. *BMC Geriatrics*, 9:7, 2009.

[67] K. G. M. Moons, A. R. T. Donders, E. W. Steyerberg, and F. E. Harrell. Penalized maximum likelihood estimation to directly adjust diagnostic and prognostic prediction models for overoptimism: a clinical example. *Journal of Clinical Epidemiology*, 57(12):1262–1270, Dec 2004.

[68] C. Murtaugh, T. Peng, A. Totten, B. Costello, S. Moore, and H. Aykan. Complexity in geriatric home healthcare. *Journal for Healthcare Quality*, 31(2):34–43, 2009.

[69] W. Navidi. Bidirectional case-crossover designs for exposures with time trends. *Biometrics*, 54(2):596–605, Jun 1998.

[70] W. Navidi and E. Weinhandl. Risk set sampling for case-crossover designs. *Epidemiology*, 13(1):100–105, Jan 2002.

[71] C. I. Neutel, S. Perry, and C. Maxwell. Medication use and risk of falls. *Pharmacoepidemiol Drug Saf*, 11(2):97–104, Mar 2002.

[72] M. C. Nevitt, S. R. Cummings, S. Kidd, and D. Black. Risk factors for recurrent nonsyncopal falls. a prospective study. *JAMA*, 261(18):2663–2668, May 1989.

[73] T. H. R. T. Noah Simon, Jerome Friedman. Regularization paths for cox's proportional hazards model via coordinate descent. *Journal of Statistical Software*, 39(5):1–13, 2011.

[74] J. Olazarn, D. Valle, J. A. Serra, P. Cano, and R. Mu?iz. Psychotropic medications and falls in nursing homes: A cross-sectional study. *Journal of the American Medical Directors Association*, Dec 2012.

[75] U. Olsson Mller, P. Midlv, J. Kristensson, C. Ekdahl, J. Berglund, and U. Jakobsson. Prevalence and predictors of falls and dizziness in people younger and older than 80 years of age–a longitudinal cohort study. *Archives of Gerontology and Geriatrics*, 56(1):160–168, 2013.

[76] T. Park, M. Ki, and S.-G. Yi. Statistical analysis of mmr vaccine adverse events on aseptic meningitis using the case cross-over design. *Statistics in medicine*, 23(12):1871–1883, 2004.

[77] A. Peters, S. von Klot, N. Berglind, A. Hörmann, H. Löwel, F. Nyberg, J. Pekkanen, C. A. Perucci, M. Stafoggia, J. Sunyer, et al. Comparison of different methods in analyzing short-term air pollution effects in a cohort study of susceptible individuals. *Epidemiologic Perspectives & Innovations*, 3(1):10, 2006.

[78] C. A. Pope III, D. G. Renlund, A. G. Kfoury, H. T. May, and B. D. Horne. Relation of heart failure hospitalization to exposure to fine particulate air pollution. *The American journal of cardiology*, 102(9):1230–1234, 2008.

[79] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012. ISBN 3-900051-07-0.

[80] D. A. Redelmeier and R. J. Tibshirani. Interpretation and bias in case-crossover studies. *Journal of Clinical Epidemiology*, 50(11):1281–1287, Nov 1997.

[81] B. Renneboog, W. Musch, X. Vandemergel, M. U. Manto, and G. Decaux. Mild chronic hyponatremia is associated with falls, unsteadiness, and attention deficits. *American Journal of Medicine*, 119(1):71.e1–71.e8, Jan 2006.

[82] R. H. Rieger and C. R. Weinberg. Analysis of clustered binary outcomes using within-cluster paired resampling. *Biometrics*, 58(2):332–341, Jun 2002.

[83] J. M. Robins, S. Greenland, and F.-C. Hu. Estimation of the causal effect of a time-varying exposure on the marginal mean of a repeated binary outcome. *Journal of the American Statistical Association*, 94(447):pp. 687–700, 1999.

[84] R. J. Rosati, L. Huang, M. Navaie-Waliser, and P. H. Feldman. Risk factors for repeated hospitalizations among home healthcare recipients. *Journal for Healthcare Quality*, 25(2):4–11, 2003.

[85] Sammy Suissa. The case-time-control design. *Epidemiology*, 6(3):248–253, May 1995.

[86] J. S. Schildcrout and P. J. Heagerty. Regression analysis of longitudinal binary data with time-dependent environmental covariates: bias and efficiency. *Biostatistics*, 6(4):633–652, 2005.

[87] H. Shuto, O. Imakyure, J. Matsumoto, T. Egawa, Y. Jiang, M. Hirakawa, Y. Kataoka, and T. Yanagawa. Medication use as a risk factor for inpatient falls in an acute care hospital: a case-crossover study. *British Journal of Clinical Pharmacology*, 69(5):535–542, May 2010.

[88] G. S. Sorock, a. A. Quigley, M. K. Rutledge, J. Taylor, X. Luo, P. Foulis, M.-C. Wang, R. Varadhan, M. Bellantoni, and S. P. Baker. Central nervous system medication changes and falls in nursing home residents. *Geriatric Nursing*, 30(5):334–340, 2009.

[89] J. A. Stevens, P. S. Corso, E. A. Finkelstein, and T. R. Miller. The costs of fatal and non-fatal falls among older adults. *Injury Prevention*, 12(5):290–295, Oct 2006.

[90] E. W. Steyerberg, M. J. Eijkemans, F. Harrell, Jr, and J. D. Habbema. Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets. *Statistics in Medicine*, 19(8):1059–1079, Apr 2000.

[91] S. Suissa. The case-time-control design: further assumptions and conditions. *Epidemiology*, 9(4):441–445, Jul 1998.

[92] J. X. Sun, S. Sinha, S. Wang, and T. Maiti. Bias reduction in conditional logistic regression. *Statistics in medicine*, 30(4):348–355, 2011.

[93] J. Sunyer, J. Schwartz, A. Tobías, D. Macfarlane, J. Garcia, and J. M. Antó. Patients with chronic obstructive pulmonary disease are at increased risk of death associated with urban particle air pollution: a case-crossover analysis. *American Journal of Epidemiology*, 151(1):50–56, 2000.

[94] M.-P. Sylvestre and M. Abrahamowicz. Flexible modeling of the cumulative effects of time-dependent exposures on the hazard. *Statistics in Medicine*, 28(27):3437–3453, Nov 2009.

[95] J. Symons, L. Wang, E. Guallar, E. Howell, F. Dominici, M. Schwab, B. Ange, J. Samet, J. Ondov, D. Harrison, et al. A case-crossover study of fine particulate matter air pollution and onset of congestive heart failure symptom exacerbation leading to hospitalization. *American journal of epidemiology*, 164(5):421–433, 2006.

[96] L. H. Tecer, O. Alagha, F. Karaca, G. Tuncel, and N. Eldes. Particulate matter (pm2. 5, pm10-2.5, and pm10) and children's hospital admissions for asthma and respiratory diseases: A bidirectional case-crossover study. *Journal of Toxicology and Environmental Health, Part A*, 71(8):512–520, 2008.

[97] T. Therneau. *A Package for Survival Analysis in S*, 2012. R package version 2.36-14.

[98] M. E. Tinetti, J. Doucette, E. Claus, and R. Marottoli. Risk factors for serious injury during falls by older persons in the community. *Journal of the American Geriatrics Society*, 43(11):1214–1221, Nov 1995.

[99] M. E. Tinetti, W. L. Liu, and E. B. Claus. Predictors and prognosis of inability to get up after falls among elderly persons. *JAMA*, 269(1):65–70, Jan 1993.

[100] M. E. Tinetti, M. Speechley, and S. F. Ginter. Risk factors for falls among elderly persons living in the community. *New England Journal of Medicine*, 319(26):1701–1707, Dec 1988.

[101] C. Viboud, P. Y. Bolle, J. Kelly, A. Auquier, J. Schlingmann, J. C. Roujeau, and A. Flahault. Comparison of the statistical efficiency of case-crossover and case-control designs: application to severe cutaneous adverse reactions. *Journal of Clinical Epidemiology*, 54(12):1218–1227, Dec 2001.

[102] S. K. Vines and C. P. Farrington. Within-subject exposure dependency in case-crossover studies. *Statistics in Medicine*, 20(20):3039–3049, Oct 2001.

[103] S. Wang, C. Linkletter, M. Maclure, D. Dore, V. Mor, S. Buka, and G. A. Wellenius. Future cases as present controls to adjust for exposure trend bias in case-only studies. *Epidemiology*, 22(4):568–574, Jul 2011.

[104] S. V. Wang, J. J. Gagne, R. J. Glynn, and S. Schneeweiss. Case-crossover studies of therapeutics: design approaches to addressing time-varying prognosis in elderly populations. *Epidemiology*, 24(3):375–378, May 2013.

[105] L. Warner, M. Macaluso, H. D. Austin, D. K. Kleinbaum, L. Artz, M. E. Fleenor, I. Brill, D. R. Newman, and E. W. Hook. Application of the case-crossover design to reduce unmeasured confounding in studies of condom effectiveness. *American Journal of Epidemiology*, 161(8):765–773, Apr 2005.

[106] W. Whang, J. E. Manson, F. B. Hu, C. U. Chae, K. M. Rexrode, W. C. Willett, M. J. Stampfer, and C. M. Albert. Physical exertion, exercise, and sudden cardiac death in women. *JAMA*, 295(12):1399–1403, Mar 2006.

[107] H. J. Whitaker, C. P. Farrington, B. Spiessens, and P. Musonda. Tutorial in biostatistics: the self-controlled case series method. *Statistics in Medicine*, 25(10):1768–1797, May 2006.

[108] J. Whitney, J. C. T. Close, S. H. D. Jackson, and S. R. Lord. Understanding risk of falls in people with cognitive impairment living in residential care. *Journal of the American Medical Directors Association*, 13(6):535–540, Jul 2012.

[109] J. M. Williamson, S. Datta, and G. A. Satten. Marginal analyses of clustered data when cluster size is informative. *Biometrics*, 59(1):36–42, Mar 2003.

[110] Y. Zhang, R. Woods, C. E. Chaisson, T. Neogi, J. Niu, T. E. McAlindon, and D. Hunter. Alcohol consumption as a trigger of recurrent gout attacks. *American Journal of Medicine*, 119(9):800.e13–800.e18, Sep 2006.

[111] G. Ziere, J. P. Dieleman, T. J. M. van der Cammen, A. Hofman, H. A. P. Pols, and B. H. C. Stricker. Selective serotonin reuptake inhibiting antidepressants are associated with an increased risk of nonvertebral fractures. *Journal of Clinical Psychopharmacology*, 28(4):411–417, Aug 2008.

[112] H. Zou and T. Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320, 2005.

# Curriculum Vitae

## CARLIN BRICKNER

4339 41st St Apt 1D ⋄ Sunnyside, NY 11104

### EDUCATION

**Rutgers University** *2007–2015*
Doctorate of Public Health in Biostatistics

**The University of Akron** *2004–2005*
Master of Science in Statistics

**The Pennsylvania State University** *2000–2004*
Bachelor of Science in Mathematics
Minor in Statistics

### WORK EXPERIENCE

**Visiting Nurse Service of New York** New York, NY
*Associate Director of Biostatistics* *2014– Present*
*Senior Statistical Analyst* *2010–2014*
*Statistical Analyst* *2006–2010*

**The Reputation Institute** New York, NY
*Consultant* *Dec 2005–Apr 2006*

### TEACHING EXPERIENCE

**Rutgers University** Piscataway, NJ
*Teaching Assistant* *Jan 2012–May 2012*
*Longitudinal Data Analysis*

**Rutgers University** Piscataway, NJ
*Teaching Assistant* *Jun 2010–Aug 2010*
*Regression Methods for Public Health Studies*

**The University of Akron** Akron, OH
*Teaching Assistant* *Aug 2004–Aug 2005*
*Introduction to Statistics*

### PUBLICATIONS

Rosati, R. J., Russell, D., Peng, T., **Brickner, Carlin**, Kurowski, D., Christopher, M. A., and Sheehan, K. M. (2014). Medicare home health payment reform may jeopardize access for clinically complex and socially vulnerable patients. *Health Aff (Millwood)*, 33(6):946–956.