PROPENSITY SCORE METHODS FOR COMPETING RISKS

Natnaree Aimyong

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Public Health in the Department of Biostatistics in the Gillings School of Global Public Health.

Chapel Hill
2014

Approved by:

Jason P. Fine

M. Alan Brookhart

Anastasia Ivanova

Chirayath M. Suchindran

Todd A. Schwartz

ABSTRACT

Natnaree Aimyong: Propensity Score Methods for Competing Risks
(Under the direction of Jason P. Fine and M. Alan Brookhart)

Non-experimental studies have increasingly been used to examine the safety and effectiveness of medication. Challenges to this method include confounding, which may cause the estimator to be biased. Propensity score (PS), which is the conditional probability of receiving treatment given all confounders, may be used to control for confounding. Analysis of vulnerable populations may involve competing risks, which may occur before the event of interest. Statistical methods that account for competing risks are needed to obtain valid causal estimate. However, little knowledge attention has been given to this topic in the literature.

The objective of this research was to investigate the performance of estimators under implementation of various PS methods in competing risk survival analyses for estimating marginal and conditional treatment effects. The competing risk models were a cause-specific hazard model and subdistribution hazard model.

According to simulation results, the weighted method produced efficient estimators for marginal treatment effects. However, it leads to an inflated variance when low incidence of event and strong confounder effects. A bootstrapping method can be used to estimate the variance under this scenario. For the conditional treatment effect, PS adjustment in the model performed the best for the null model. Depending on the sample size and the number of confounding variables, the subclassification and matching methods yield best performance under the alternative when treatment effect is non-null.

Heterogeneity of treatment effect associated with statin therapy may be present in el- derly who experience myocardial infarction. Examining treatment effect across age groups and the revascularization procedure illustrated the heterogeneity of statin effects. Statins significantly reduce risks of heart failure among younger age groups. The combination of statins with revascularization procedures presents better treatment effects than occurs with statins alone. Application of propensity score methods to competing risks is illustrated in this study, with the analysis of treatment effects providing an improved understanding of the heterogeneity of the effects of statins therapy.

The efficiency of implementing propensity score method to competing risks is illustrated in this study. Analyzing the treatment effects by subgroup and medical procedure contributes better results for estimating the heterogeneity treatment effect.

I dedicate this work to all mothers, who always support,

and stand behind the success of academia and the researchers.

ACKNOWLEDGMENTS

I wish to express sincere thanks to all who contributed many types of supports especially the following:

First, I would like to thank Royal Thai Government for the scholarship they provided to me and Mahidol University for the full scholarship and the partial scholarship I was awarded. I also appreciate the support from the officers and staff of the Office of Educational Affairs of the Royal Thai Embassy.

I would like to convey my gratefulness to Dr. Jason P. Fine for his excellent academic guidance and inspiration beyond this dissertation. I also express my gratitude to Dr. M. Alan Brookhart for his support and for introducing me to the field of Pharmacoepidemiology. Both of them give me the opportunity to work in professional environment with kindness and patience. I would like to thank my dissertation committee: Dr. Chirayath M. Suchindran, Dr. Richard Bilsborro, Dr. Anastasia Ivanova and Dr. Todd Schwartz, for their valuable advice and comments.

I am very grateful to my parents for their inspiration and courage. I am appreciative to all friends for their help and support.

Last, I would like to express my appreciation to all of my teachers, instructors and professors. They have all contributed to my store of knowledge, especially my kindergarten teacher, who taught me to read and how to write my first letters.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

1.1 Motivating Examples

Cardiovascular disease (CVD) is the leading cause of mortality and is a major cause of morbidity in the United States and worldwide (108). Increasing steadily with age, the prevalence of CVD events is even higher in the population older than 65 years of age, affecting more than 70% of both men and women and 66% of CVD deaths occur in people age above 75 years old (92). As the population is aging, more people are likely to have cardiovascular complications. Increasing attention has been drawn to the management of CVD in the elderly population.

Statins, a class of 3-hydroxy-3-methlglutaryl coenzyme A reductase inhibitors, are used to lower low-density lipoprotein cholesterol (LDL-c) levels to prevent CVD events (100). LDL-c has been recognized as an established risk factor for CVD events, and remains the primary focus for evaluation of pharmacologic effectiveness based on treatment target goals. Clinical trials have demonstrated that statins are highly effective at lowering LDL-c and thus help to reduce the incidence of cardiovascular events(32, 90, 112). Post-myocardial infarction (MI) patients with high normal or only mildly elevated lipid levels also benefit from statin treatment (85, 23). However, older populations were not well represented in most statin trials or age-specific results were not reported. As a result, little is known about the benefits of treatment with statins in the elderly, especially among the elderly patients who just had a MI.

Non-experimental studies are increasingly used to evaluate the safety and effectiveness of medications when used in usual clinical care (11). It is an appealing design for investigation of statin treatment effects among post MI patients. However, two major challenges are present. The first challenge is presence of possible confounding systematic differences in prognosis between patients treated with statins and those untreated. If left uncontrolled, the observed difference in outcome risks cannot be interpreted as a causal effect solely due to statins. Another challenge is the presence of competing risks, as the occurrence of some events may precede and thus preclude the occurrence of the event of interest. In the elderly population after an acute MI, multiple comorbidities and worsened health status put them at a higher risk of mortality. Death may thus preclude the occurrence of the event of interest, (such as another MI) and make the evaluation of the effect of statins on MI difficult. Sophisticated methods such as competing risks survival analyses are needed in this setting (7, 64). Statistical methods that can account for both competing risks and confounding are needed to obtain a valid causal estimate.

A competing risks survival analysis is a method to address the presence of multiple events in a survival analysis of the time between the start of follow-up and either the occurrence of the event(s) of interest or a censoring event (51). The outcomes of competing risk model are three mutually exclusive events including occurrence of the event of interest, occurrence of a competing event, and a censoring event (lost to follow-up, end of follow-up etc.). The regression models in competing risks may be used to estimate the treatment effects.

The classical approach for confounding control is based on multivariable regression models. This approach may be heavily model dependent, with different models potentially giving very different results. Moreover, the correct outcome model specification can be challenging, especially in the settings of rare outcomes, many confounders, and/or treatment heterogeneity (11, 14). Matching and stratification methods are popular tools used to control confounding factors. However, these methods have difficulties when applied to a large number of confounders. Propensity score (PS) methods, which estimate

treatment effects without relying on modeling the outcome, are an approach being increasingly used in non-experimental studies (96).

Competing risk survival analyses with PS methods to control for confounding can then be used to investigate effect of statin treatment on cardiovascular events among post MI patients. However, this approach has not been extensively explored, and little is known about the performance of this approach compared to the existing standard methods. It is of great interest to investigate the performance of competing risk models using propensity score methods to control confounding for estimating marginal and conditional treatment effects.

1.2 Causal Inference

Causal inference is the process of drawing a conclusion about whether a causal relationship does in fact exist (43). Let $Y_i(1)$ and $Y_i(0)$ be the potential outcomes that would have been observed for individual i in the treatment group or the control (untreated) group, respectively. Individual causal effects can then be obtained by contrasting the values of the two potential outcomes. However, only one of the potential outcomes is observed for each individual, thus, in general, individual causal effects cannot be obtained.

Since it is generally impossible to identify individual causal effects, an aggregated causal effect, the average causal effect in a population of individuals, becomes the focus of interest. Let $P[Y_z = 1]$ be the counterfactual risk of outcome Y_z, or the risk of developing outcome Y in all subjects in the population receiving counterfactual treatment z. An average causal treatment effect is present when

$$P[Y_{z=1} = 1] \neq P[Y_{z=0} = 1].$$

The assumptions of causal inference are the stable unit treatment value assumption (SUTVA) and strong ignorability. SUTVA means the potential outcome of individual i is independent of potential outcome of individual j, $i \neq j$. The strong ignorability

assumption is the independence of treatment assignment and outcomes conditional on the covariates.

In randomized trials, only one potential outcome is observed for each individual. However, the randomization process ensures that the balancing between the background variables between treated and untreated groups. In other words, the randomization process ensures the independent predictors of the outcome are equally distributed between the treated and untreated groups. The treated and untreated groups are exchangeable, and the exchangeability thus implies lack of confounding (65).

In non-experimental studies, treatment is not randomly assigned and the reason for receiving treatment is likely to be associated with some predictors of the outcome. Thus, exchangeability is not guaranteed. However, in weaker conditions, conditioning on many pre-exposure covariates of the treated and untreated groups is often reasonable to allow exchangeability. Let L be the confounding or background covariates. The conditional exchangeability implies

$$P[Y = 1|L = 1, Z = z] = P[Y_z = 1|L = l]$$

PS is the conceptual tool used to achieve accurate causal inference by balancing or conditioning on the background variables of treated and untreated subjects (57).

1.3 Propensity score and non-experimental studies

A major challenge for non-experimental studies is confounding caused by systematic differences in prognosis between patients exposed to intervention of interest and the comparison group. Sources of confounding in non-experimental studies of medications can arise from multiple areas. Physicians may tend to prescribe medications to patients who are most likely to benefit from them, which can cause the intervention to appear to cause events that they actually prevent (91). Patients initiating a preventive medication may be more aware of their health status and more likely to engage in other healthy, prevention-oriented behaviors leading to the bias known as healthy user bias (11). The

validity of the treatment effect estimated from non-experimental studies is a concern if these biases are left uncontrolled.

Propensity score methods have emerged as a useful method to control for confounding in the non-experimental setting. It is widely used in a variety of areas, including medical (95), economic (42) and social research (67). The methods were formalized by Rosenbaum and Rubin (80) and were shown to be able to control confounding. At the beginning, the propensity score was developed to estimate the causal effect of binary treatment (48). A propensity score is the conditional probability of receiving treatment given all confounders. Among patients with the same PS, treatment is unrelated to confounders. Therefore, the two treatment groups have the same distribution of measured confounders.

The true propensity score of non-randomization studies is unknown. However, it can be estimated from observed data. Statistical methods such as logistic regression and other discriminant models can be used to estimate PS. Multivariable logistic regression is the most widely used method for PS estimation. The estimation of the propensity score can be implemented as a continuous variable via standard approaches including matching, subclassification (stratification) by propensity score, propensity score adjustment to the multivariate model, and weighting (96).

1.3.1 Matching

Matching on certain covariates to remove confounding by the matching variables has been used extensively in cohort studies (96). Matching methods attempt to choose a single or multiple patients from the untreated group with the same values of the matching variables for each patient in the treated group. It can be easily implemented when matching only needs to be done for a small number of covariates. A large number of confounders in a non-experimental study makes matching on all confounders difficult. PS as a summary score reduces these multidimensional confounders to one dimension, which helps overcome this limitation (19).

Rosenbuam and Rubin recommended three multivariate matching criteria (82). The

first option is nearest available matching on the estimated PS. The treated and untreated patients are first randomly ordered, and the first treated patient is matched with the first untreated patient who has the nearest PS. The second option is Mahalanobis metric matching (84). The Mahalanobis distance is the distance between two dimensional points scaled by the statistical variation in each component of the point. Treated and untreated patients are first randomly ordered, and then each treated patient is pair-matched with the first untreated patient who has the closest Mahalanobis distance. The third alternative is nearest available Mahalanobis metric matching within calipers defined by PS. The treated group is first randomly ordered. For each treated patient, a group of untreated patients are chosen given that the difference of PS between the treated patient and the untreated patient is less than the caliper (a constant value). From this group of untreated patients, the patient who has the closest Mahalanobis distance to that of the treated patient is selected as the match. The third matching method produces the best balance for the covariate distributions between the treated and untreated groups and is considered to outperform the first two.

The two popular algorithms for creating PS matched sets are greedy matching and optimal matching . In greedy matching, a treated patient is first selected at random. The untreated patient whose PS is closest to this patient is chosen as a match for this patient. This process is repeated until all the untreated patients have been matched to all the treated patients or when all the treated patients have been matched. In contrast to the greedy matching, which finds the nearest untreated patient, optimal matching tries to minimize the total within-pair distance differences (83).

The most common implementation of PS matching is pair matching (one-to-one matching). With pair matching, only one untreated patient with a similar value of PS is matched to one treated patient to form matching pairs. Alternatives to pair matching include many-to-one (M:1) matching (66) and full matching (33, 37, 77). In many-to-one matching, a fixed or variable number of untreated patients are matched to each treated patients. The approach to match with a variable number of untreated patients was found

to have improved bias reduction. Full matching results in matched sets consisting of either one treated patient and multiple untreated patients or one untreated patient and at least one treated patient (77). Full matching removes bias better than pair matching and may achieve the closer match than M:1 matching (37). However, fullmatching method result in a wide range of ratio of treated to untreated patients (94).

In analyses using Cox proportional hazard models, PS pair-matching has been shown to have the smallest amount of bias when compared to other methods (26, 4). Martens showed the analysis from real data showed larger treatment effects from matching and subclassification compared to conventional Cox proportional hazard model (62).

1.3.2 Subclassification

Another method used to implement PS to control confounding is subclassification or stratification of PS into equal-sized strata, estimation of the treatment effect within these strata and then combining the stratum-specific effect estimates using a weighted-average approach. Both treated and untreated groups are grouped into equal-sized strata based on their PSs. In general, quintiles of PS are used to create the strata. Previous reports have shown that subclassification into five strata can remove approximately 90% of initial bias (81) despite the fact that it can also increase the variance of the estimator (104).

1.3.3 Propensity score covariate adjustment

Adding the PS to a multivariable outcome model is the least appealing method of PS implementation, since the validity of this approach depends upon a correct specification of both PS and outcome models. If the PS is included as a linear term in the outcome model, an assumption is made that there is a linear relationship between the PS and the outcome, which is likely to be violated in real world applications.

In situations where outcome models are linear models, this regression modeling method is the same as analysis of covariance (ANCOVA). A great amount of bias can be introduced when the covariances of the treated and untreated groups are unequal and the

variances of the treated and untreated groups differ (19). When this method is applied to studies using survival analysis to estimate effects, the fixed parameters of the PS variable suggest that the probability of receiving treatment given all measured covariates has a constant effect on the hazard ratio. Results from our simulation study showed that the efficiency of propensity score adjustment depended on the overlap area and the specification of variables included into propensity score model (35).

1.3.4 Weighting

Standardization with weights generated from the PS can also be used to control confounding. Unlike matching, this approach does not result in reduction of the original sample size. Individuals within the original sample are weighted based on their PS to create a pesudo-population where the covariates are well balanced in the two treatment groups, and no association exists between the confounders and treatment. The weighting method plays an importance role in the estimation of marginal treatment effects. There are two types of weighting commonly used: inverse probability of treatment weighted (IPTW), standardized mortality ratio weighted (SMRW).

IPTW is defined as the inverse of the estimated PS for treated patients and the inverse of one minus the estimated PS for untreated patients (76). These weights create a pseudo-pupulation where the weighted treated group and untreated group are representative of the patient characteristics of the entire population, resulting in an estimate that is generalizable to the entire population from which the observed study sample was drawn. The IPTW method has been shown to have good performance for estimating marginal treatment effects in Monte Carlo simulations (4). However, IPTW can be unstable in some situations so a stabilized weighting method may be preferred. The stabilized weight (STW) is calculated by multiplying the IPTW by the probability of being treated for the treated patients and the probability of not being treated for the untreated group. STW helps reduce the effects of extreme weights and produces a narrower confidence interval for the estimator compared to IPTW (17).

SMRW is set to one one for treated patients and defined as the ratio of the estimated PS to one minus the estimated PS for the untreated (86). The untreated patients are thus weighted to be representative of the treated population. The resulting effect estimate is generalizable to the treated population from which the observed treated sample was drawn. Unlike PS matching, which also often estimates average treatment effect in the treated, no treated patients are excluded from the analysis.

In all weighting analyses, statistical methods are needed to account for the relation between the replicated individuals created by the process of weighting. A robust sandwich variance estimate is required to calculate the variance. This approach results in confidence intervals that are conservative and wider than the nominal coverage.

Many studies have applied weighted methods to survival analysis. Cole and Hernan presented an approach to produce adjusted survival curves with inverse probability weights which offers direct interpretation of the data (16). Hernan et al (40) applied IPTW to study zidovedine (AZT) treatment effects on mortality and compared it to an unweighted analysis. The weighted analysis showed that AZT reduced risks of mortality but the unweighted analysis with a basic baseline adjustment model produced an adverse effect of AZT on mortality (which was contrary to reality).

Simulation studies showed survival analysis via Cox model with application of the IPTW method produced the smallest amount of bias for the estimator of interest (4, 26).

1.3.5 Variables selection for propensity score model

The identification of variables to be included in the PS model is an important process. Good subject-matter knowledge must be used to guide this process. Studies have shown that optimal PS models should include all variables that are related to the outcome of interest, regardless of whether they are associated with the treatment. Including variables that are related to treatment but which are either not related or only weakly related to outcome increase the variability of effect estimates and, in the presence of unmeasured confounding, increase bias (10). This idea of variable selection for the PS also exists for

the estimator of the average treatment effect for the treated group when applying the matching method and using a multinomial model for the outcome estimation (18, 110).

1.3.6 Assessment of the balance of the covariates

Once the PS model is fit, it is recommended to explicitly evaluate the performance of the estimated PS model by assessing the balance of covariates after PS implementation either through matching or weighting . Approaches to assess the covariate balance under PS methods have been developed (3).

Evaluation of the lack of fit of the PS model is performed using the logistic regression goodness-of-fit statistics and the c-statistic. the goodness of fit statistic summarizes the deviation between the observed and predicted outcomes. The c-statistics value indicates the capacity of a model to discriminate between treated subjects and untreated subjects (46). However, both methods fail to detect the balancing of background variables between treated and untreated groups (106).

The standardized mean difference can be applied to identify the importance of difference between the treated and untreated groups and to evaluate the covariate balance in PS matching and weighting methods. Values of 0.2, 0.5 and 0.8 represent a small, a moderate and a large difference, respectively; the sample size does not influence the standardized difference (105). For subclassification method, the covariate balance can be evaluated with a two-way analysis of variance (ANOVA). The treatment group and subclassification group are the factors of the ANOVA model for testing the balance of covariates in the treated and untreated groups (81). The weighted conditional standardized difference and quintile regression are the methods used to determine the similarity of the baseline covariates in the treated and untreated groups (2). Recently, Imai and Ratkovic (47) introduced a covariate balancing PS estimation method which optimizes the covariate balance while estimating PS (38). This method can improve the performance of PS weighting and matching methods.

1.4 Marginal and conditional models

A marginal treatment effect is the average treatment effect for the population, while a conditional treatment effect is the average treatment effect for the individual (3). In the absence of confounding, the difference in means and risk difference are collapsible, and the conditional and marginal effects are the same. In randomized studies, the covariates are balanced between the two treatment groups, therefore the crude difference in means and the adjusted difference in means will be equal. In non-experimental studies, the marginal and conditional estimates will coincide if there was no unmeasured confounding, the outcome was continuous, and the true outcome model was known (31). However, when the outcome is either binary or time-to-event, and the odds ratio or the hazard ratio is used as the effect measure, then the marginal and conditional effect will not coincide even in the absence of confounders (34, 25).

For PS methods, the estimators from the conventional model (adjusting for confounders in the outcome regression model), covariate adjustment (adding the PS to a multivariable outcome model) and matching method are estimating the conditional treatment effect. The PS based weighting methods yield estimates of the marginal treatment effect for the population.

1.5 Competing risks model

A survival analysis explores the time period from a certain point until the occurrence of the event of interest. A competing risks survival analysis is the special case of survival analysis where multiple events may occur and the occurrence of one event may preclude the occurrence of the other. Competing risk events threaten the validity of studies, even in randomized control trials. Individuals who are at higher risk of competing risk events may be less likely to experience the absolute benefit of treatment (64). Care is needed in statistical analysis to ensure that treatment effects are appropriately quantified.

Approaches have been developed to conduct competing risk survival analysis. The

likelihood of the occurrence of a particular competing risk can be summarized using the distribution of the observed data or using models representing the underlying mechanisms that generated the observed data (54). When modeling the observed data, so-called crude functions are utilized. The cause specific hazard and cumulative incidence functions described below are the most widely used quantities. A popular mechanistic model is the latent failure time model discussed below. The net function is the probability of the occurrence of the event of interest corresponding to the latent failure time in the hypothetical situation where the particular risk of event is the only risk present. Additional properties of the crude and net approaches are now discussed.

1.5.1 Notation

For this discussion, the following notations are used. Let Y and C be failure and censoring times, and $\varepsilon \in \{1, 2, ..J\}$ be the failure event. For each individual i, i=1,...,n, the observed failure time was T_i, $\mathbf{T}= (Y \wedge C)$, and the observed event be ε_i, $\Delta\varepsilon = I(Y \leq C)$ when I$(\cdot)$ is indicator function. Let Z be one for treated group and zero for untreated group.

1.5.2 Latent failure times

There are J mutually exclusive types of failure events, and the corresponding time until each failure type is $\widetilde{T}_{ij}$. The observed time is $min_j(\widetilde{T}_{ij})$, with each $\widetilde{T}_{ij}$ defined as if the other causes were not present, and with the observed cause of failure ε being the index of the observed latent failure time. The hazard function for latent failure times is a net hazard function defined as:

$$\widetilde{\lambda}_{ij} = \lim_{\Delta t \to 0} \frac{1}{\Delta t} P(t \leq \widetilde{T}_{ij} < t + \Delta t, \varepsilon = j | \widetilde{T}_{ij} \geq t), j = 1, ..., J.$$

The net function can be estimated when the latent failure times are independent. This assumption of independence cannot be verified, so this approach is not realistic for applications in data analyses.

1.5.3 Cause-specific hazard function

To model competing risk using a cause-specific hazard function, the type-specific or cause-specific hazard function is defined as:

$$h_j(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} P(t < T < t + \Delta t, \varepsilon = j | T \geq t), \text{ j=1,...m.}$$

The cause-specific hazard function is the instantaneous rate for a failure of type j at time t, in the presence of all other failure types (54). With this cause-specific hazard function approach, an event k (k≠j) is censored at time T if events other than type j are observed. The cumulative incidence function, $P(T \leq t, \varepsilon = j)$,which is the probability of occurrence of event j at time t is defined as:

$$F_j(t) = \int_0^t h_j(s) S(s) \mathrm{d}\, s \text{ , where S is the survival function for T.}$$

Defining $S_i(t)$ to be the survival function based on $h_i(t)$ where $S(t) = \prod_i S_j(t)$, one may show that the naive Kaplan-Meier estimator for $S_i(t)$ is a biased estimator for $F_i(t)$, generally underestimating this quantity (73).

From the definition of the cause-specific hazard function, the parameters of the cause-specific hazard model can be estimated using a Cox proportional hazard model. The treatment effects of event j can be obtained by maximizing the factor of the partial likelihood function involving event j when other event(s) are treated as censoring events. The semiparametric model of cause-specific hazard function is defined as:

$$h_j(t) = h_{0j}(t) exp\{\beta_j' Z\}$$

1.5.4 Subdistribution hazard function

The subdistribution hazard function is the other type of competing hazard function which is derived directly from cumulative incidence function (30),

$$\lambda_j = \{dF_j(t)\mathrm{d}t\}/\{1 - F_j(t)\}.$$

The subdistribution hazard function is probabilistically defined as

$$\lambda_j(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} P(t < T < t + \delta t, \varepsilon = j | T \geq t \cup (T \leq T \cap \varepsilon \neq j).$$

For the subdistribution function, the cumulative probability of occurrence of cause J remains less than one, the subdistribution satisfies the definition of an improper probability distribution. This occurs because an individual who had an event k no longer is at risk of failure from causes $j \neq k$.

Similarly to the cause specific hazard, a proportional hazards model may be specified for the subdistribution hazard function to describe the treatment effect on the risk of a particular cause of interest. The parameters of this model can be estimated using methods for the Fine-Gray model (21). The semiparametric model for subdistribution hazard function is:

$$\lambda_j(t) = \lambda_{0j}(t) exp\{\beta_j' \mathbf{Z}\}$$

1.6 The application using Real Data

For this example, we identified a cohort of Medicare beneficiaries who just had a hospitalization stay for acute MI (AMI) in 2008. The condition AMI was identified from Medicare inpatient claims files using relevant International Classification of Diseases, Ninth Revision, Clinical Medication (ICD-9-CM) codes (410.01, 410.11, 410.21, 410.31, 410.41, 410.51, 410.61, 410.71, 410.81, 410.91). Personal identifiers were removed from all analytical data files.

Eligible patients included in the cohort were at least 66 years old at the admission date, living in the United States and had been continuously enrolled in Medicare Part A at least one year. The exclusion criteria included the absence of specific ICD codes (410.x1) as the first or second discharged diagnosis, patients who died during hospital admission, patients who did not survive 30 days after discharge, AMI diagnosed only from skilled nurse facility (SNF) claims, patients who only had AMI admission prior

to 1/1/2008 and were discharged on or after 1/1/2008. Other exclusions were patients who had discharge code to hospice (40, 41, 42, 50, 51), transfered to other hospitals for inpatient care, patients who were discharged or transfered to SNF or long-term care facilities for inpatient care, or who did not have a prescription claim within 30 days after the index AMI discharge.

The exposure of interest was statin use after discharge. Patients initiating statin therapy were considered statin users while patients without statin prescriptions were considered nonusers. The follow up of these patients started at 31 days following the date of discharge and ended at the occurrence of the outcome or at the end of the study. The outcome of interest was the occurrence of MI or heart failure (HF) or stroke or all-cause mortality. The effect of interest in this example was the marginal and conditional treatment effect of statins on the cardiovascular outcomes or mortality in the presence of competing risks.

Potential confounding covariates were created, including demographic characteristics and clinical conditions based on claims occurring in the 12-month baseline period prior to admission were created. These covariates were identified a priori based on the literature, substantive knowledge, and the availability of covariates within the data. The variables included both demographic and medical record at baseline, during admission and follow up period. A list of these covariates appears in Table 1.1.

1.7 Objective and outline

The objective of this research study was to investigate the performance of estimators when using various PS implementation methods in competing risk survival analyses for estimating of marginal and condition treatment effect. The PSs were estimated using logistic regression. The implementation methods examined included subclassification, matching, PS adjusted into the model and weighting. The competing risks models used were the cause-specific hazard model and the subdistribution hazard model.

Chapter 2 describes the performance of PS methods incorporated in the competing risk survival analytic approach for estimation of marginal treatment effects. Chapter 3 presents the performance of PS methods incorporated in competing risk survival analyses for conditional treatment effects. The results presented in both chapters were from simulation studies as well as analyses of claims data evaluating the effect of statin treatment on the risk of cardiovascular end points and mortality. Chapter 4 presents an application of the proposed methods in an evaluation of the heterogeneous treatment effect of statins across different age groups and revascularization procedure groups.

Table 1.1: Variables included into propensity score model

General character	gender, age, race, income, medicare doughnut
Charlson Comorbidity index	AMI, Cerebrovascular Disease, Congentive HF, Periphral vascular disease, Renal disease, Chronic Obstructive Pulmonary disease, Diabetes,Peptic Ulcer disease, Cancer, Dementia, Connective Tissue disease, Rheumatic disease, Mild liver disease, Moderate to severe liver disease, Paralysis, Metastatic Carcinoma, AIDS/HIV, Diabetes with and without complication
Baseline disease	CABG, STENT, PTCA, unstable angina Ischemic heart disease, Hyperkalemia Atrial fibrillation, Hypertension, Hyperlipidemia, End-stage renal disease, Osteroporosis, Asthma, Hypotension,Rhabdomyolysis, Sinus bradycardia heart block, Angioedema&hyperkalemia, CCI total score
Baseline medication	statins, STENT/PTCA, Beta blocker, ACEI/ARB hospital admission in baseline, Number of admission Number of days stay in hospital
Admission procedure /diagnosis	Subendocardial infarction,Congestive HF Cardiogenic shock, Acute renal failure, Hypotension, Cardiac dysrhythmias, cardiac catheterization, CABG, PTCA, Angiocardiography, Platelet inhibitors Thrombolytics and platelet inhibitors Acute respiratory failure in AMI admission Septic shock in AMI admission, Days stay in ICU Days stay in coronary care unit, Total days in hospital
Medication record during Follow-up	Physician visit, Cardiologist visit, Revascularization procedure, Number of admission to short-term acute care hospital, Number of days to short term acute care hospital,
Co-medication	angiotensin-converting-enzyme inhibitor (ACEI)/ angiotensin receptor blocker (ARB), Beta blockers
Current comorbidity and assistance	Valvular disease or rheumatic heart disease Hypothyroidism, Other neurological disorders Obesity, Coagulation deficiency, Substance abuse Weight loss, Fluid/electrolyte disorder Blood loss, deficiency anemia Pulmonary Circ. Disorders, Parkinson's disease Osteoarthritis, Gastrointestinal bleed, Use of rehabilitation Use of screening, use of wheelchair Weakness, Vertigo, Fall/difficulty walking Bladder dysfunction, Decubitus, Use of Oxygen Use of hospital bed, Use of ambulance, Nail care Use of other assistive devices

Chapter 2

Marginal Treatment Effects of Completing Risks Model

2.1 Introduction

A marginal treatment effect is the average treatment effect at the population level, describing the difference in outcomes between two comparable populations where everyone in one population receives the treatment, while in the second population everyone receives an alternative treatment or is untreated (75). It can be estimated in both experimental and non-experimental settings. In the experimental setting, randomization balances the characteristics of the treatment and control groups. With the assumption that all patients could receive either of the treatments under comparison and independence of the treatment effects among patients, the balancing of covariates between treatment groups are expected, thus the crude analysis without any adjustment results in an unbiased estimate of the marginal treatment effect. Non-experimental studies based on administrative claims or clinical databases are increasingly used for post-marketing safety and effectiveness evaluation of treatments (39). These studies allow for estimation of treatment effects in settings where randomized trials are impractical or unethical. The treatment allocation in non-experimental studies is often influenced by subject characteristics, however, resulting in systematic differences between treatment groups possibly resulting in confounding or other bias. Thus, an unbiased estimate of the average treatment effect cannot be directly obtained by comparing outcomes between the two treatment groups.

PS methods have emerged as a useful approach, formalized by Rosenbaum and Rubin (80), to balance measured confounding between treatment groups and thus reduce or minimize the effect of confounding in the non-experimental setting when estimating treatment effects. The propensity score is the conditional probability of receiving treatment given measured baseline covariates. Among patients with the same propensity score, the treatment assignment is unrelated to measured covariates included in the propensity score model. The two treatment groups then have the same distribution of measured confounders, which enhances comparability between the groups. The propensity score, a continuous variable, can be implemented in various ways, including: matching, subclassification (stratification), inclusion in multivariate modeling, and inverse-probability weighting (80, 76). Of these, the matching and weighting methods play an important role in estimating the marginal treatment effect in the overall population. Under four assumptions of (1) no unmeasured confounders, (2) positivity, (3) no misspecification, and (4) consistency, weighting creates a pseudo-population in which exposure is independent of confounding (17). Under assumption of no unmeasured confounders, the exposure and outcome are independent within each level of a confounder variable, then the comparability of treated and untreated groups exists (65). The positivity assumption specifies that there are both treated and untreated subjects at every level of all confounder variables. The third assumption (no misspecification) is that all models, including both the PS models and the outcome model, are correctly specified (17). The consistency assumption is formally defined such that the potential outcome of individual for a specific exposure is the outcome that would be observed if he/she had received that specific exposure (15).

A competing risks survival analysis is the special case of survival analysis where multiple events may occur and the occurrence of one event (e.g., death) may preclude the occurrence of the event of interest. Competing events can threaten the validity of studies interested in marginal effects, even in randomized control trials (64). Competing risk models can be summarized using the distribution of the observed data or using underlying mechanistic models which generate the observed data. A popular mechanistic model

is the latent failure time model, but it cannot be used without strong and unverifiable assumptions making it unattractive for practical usage (72). When modeling the observed data, so-called crude functions are utilized. The cause-specific hazard and cumulative incidence functions, which are based on the observed data, are the most widely used. For non-experimental studies of a time-to-event outcome with competing events, statistical methods that account for both competing risks and the effect of confounding are needed to obtain a valid causal estimate of the marginal treatment effect.

IPTW, generated from the propensity score, can be used to control for confounding. The IPTW is defined as the inverse of the estimated likelihood of receiving the treatment actually received, which is the propensity score for the treated patients and one minus the propensity score for the untreated patients (76). The IPTW method creates a pseudo-population representative of the patient characteristics of the source population, thus producing an effect estimate generalizable to the population from which the sample was drawn. Assuming that the PS model was correctly specified, the measured covariates of the pseudo-population are balanced across the two treatment groups, and no association exists between measured confounders and treatment.

The IPTW method has been shown to have good performance in the estimation of marginal treatment effects in Cox proportional hazards models (98, 4, 26). However, IPTW can be unstable sometimes and a stabilized weighting method can be used to remedy the instability. The stabilized weight (STW) is calculated by multiplying the IPTW with the marginal probability of the treatment actually received (being treated for the treated group and one minus the probability of being treated for the untreated group) (76). STW methods reduce the influence of extreme weights and produce narrower confidence intervals for the estimator compared with the non-stabilized IPTW (17). In all weighting analyses, appropriate statistical methods are needed to account for the correlation between the replicated individuals created in the weighting process. A robust sandwich variance estimate is required to calculate the variance of the estimator. This approach results in confidence intervals that are conservative and wider leading to higher

than nominal coverage (59).

The objective of this study was to investigate the performance of the IPTW and the STW in estimating the marginal treatment effect using competing risks survival analysis model. Both simulation and application in a substantive data analysis were performed.

2.2 Methods

2.2.1 Monte Carlo simulation

A series of Monte Carlo simulations were conducted in order to examine the performance of inverse probability treatment weighting methods with two competing risk models: the subdistribution hazard model and the cause-specific hazard model.

Let Z be the binary exposure and X be a vector of 10 confounding variables. For each simulated dataset, the 10 covariates , X_l, l= (1,...,10) were generated from a Bernoulli distribution with parameter 0.5. The probability that the exposure Z was equal to one, which is the true propensity score, was estimated as a function of the covariates (α) using a logistic model. The exposure indicator Z was drawn from a Bernoulli distribution with probability set by

$$Pr(Z = 1|X_l) = \alpha_0 + \Sigma_l \alpha_l \mathbf{X}_l$$

In order to evaluate the performance of the weighting methods in a range of settings, four levels of confounding effect and three level of proportion of interesting event were considered. We generated the confounding effects by iteratively varying α_l untill the desired Komogorov-Smirnov (KS) distance of 0.10, 0.25, 0.40 and 0.55 (Figure 1) were reached. The larger KS distances represent the larger difference of covariates between treated and untreated groups found in a non-randomized study, KS distances closer to zero represent the differences found in a randomized study. The percentage of treated group in the simulation setting was roughly 40%. The observed events of this study were $\varepsilon \in \{1,2\}$, ε=1 indicated events of interest and ε=2 indicated competing event. Three

proportions of event of interest (ε =1) were generated, proportion (p) set in turn to 0.4, 0.6 and 0.8. The right censored times (C_i) were generated from a uniform distribution on [0, 6] to have 20% censoring. All failure time variables were generated using a competing risk model where treatment Z and covariates X were predictor variables. For individual i, where the effect of treatment, β_j,, for failure ε=1 (with either the cause-specific hazard or subdistribution hazard) was fixed at -0.5 and for failure ε=2 was fixed at -0.06, which is the conditional treatment effects.

The next two sections describe how the failure times were generated under the two different hazard functions under presence of competing risks. After which the detail of estimating true treatment effect for marginal model and PS estimated is described.

2.2.2 Simulation for subdistribution hazard function

The simulation method followed procedures described by Fine and Gray (21). The failure time for ε=1, was generated from

$$Pr(T_i \leq t, \varepsilon = 1|Z, X) = 1 - [1 - p\{1 - exp(-t)\}]^{exp(Z\beta_1 + X_l\gamma_l)}.$$

which is unit exponential with mass 1-p at infinity, Z, X =0 and p$\in \{0.4, 0.6, 0.8\}$. The subdistribution for $\varepsilon = 2$ was generated from $Pr(\varepsilon = 2|\mathbf{Z}) = 1 - Pr(\varepsilon = 1|Z)$ and the time until failure event conditionally on $\varepsilon = 2$ was generated from an exponential distribution with rate $\exp\{\beta_2 Z + \gamma_{l2} X_l\}$.

2.2.3 Simulation for cause-specific hazard function

Let $h_1(t|Z, X) = \theta_1 exp\{\beta_1 Z + \gamma_{l1} X_l\}$ and $h_2(t|Z, X) = \theta_2 exp\{\beta_2 Z + \gamma_{l2} X_l\}$ be the cause-specific hazard function, which is dependent on Z and X_l, of event $\varepsilon = 1$ and event ε = 2, respectively. The overall hazard rate was $h(t|Z, X) = h_1(t|Z, X) + h_2(t|Z, X)$. The failure time for each subject was generated from exponential distribution with hazard rate $h(t|Z, X)$. The event types, ε were generated from a Bernoulli experiment,

$$P(\varepsilon = 1|T < t) = \frac{h_1(t|Z,X)}{h(t|Z,X)}$$

where the parameter values were chosen such that $P(\varepsilon = 1|T < t)$ were 0.4, 0.6 and 0.8 for all t (9, 8).

2.2.4 True marginal treatment effect

As discussed earlier, the true marginal treatment effect can be estimated from a randomized trial. We simulated a dataset with the same parameters and settings described above to estimate the true marginal treatment effect. This approach is necessary because performing simulations using the models in 2.2.2 and 2.2.3 does not result in a marginal proportional hazards model, that is, the marginal treatment effect is not correctly specified via the proportional hazard models. If one employs the misspecified proportional hazard model in a data analysis, the resulting estimator estimates a parameter defined in large samples as the limiting value to which that estimator converges. This problem generally occurs in survival settings, when confounders are included in the true survival model and marginalizing over the confounders does not generally result in a proportional hazard model. The notion of convergence just described is standard with misspecified models, where parameter estimators do not converge to true values of parameters in the underlying models.

To determine the limiting value of the parameter estimator, a dataset with a size of 100,000 was generated. The same distributions for the confounders were used. The treatment variable Z was generated from a Bernoulli distribution with a parameter of 0.5, which imitates the randomization process. The true marginal treatment effect, measured as the difference in proportion of event of interest, was determined for both the subdistribution hazard model and the cause-specific hazard model, based on fitting these models using standard approaches. For the subdistribution hazard model, the approach of Fine and Gray was used without propensity scores, while for the cause-specific hazard, a standard partial likelihood analysis was utilized. Assuming consistency of the parameter estimator under the misspecified model, the estimator based on large sample size of 100,000 should be extremely close to the actual value being estimated.

2.2.5 Propensity-based estimation of the treatment effect in simulations

For the simulated dataset, four different types of marginal models were fitted with the two different competing risk analysis approaches: a crude model, a crude model based on region of common support (13), weighted regression using IPTW and weighted regression using STW. A conventional multivariable regression model, which is a conditional model, was also fitted to investigate the performance of estimating marginal treatment effects. For the two competing risk approaches, treatment effects from the cause-specific hazard function were estimated using a Cox proportional hazard model, while treatment effects from the subdistribution hazard function were obtained from the Fine-Gray model. The analyses were performed in R 3.0.1 (102) using the package *survival* for Cox models (103) and the package *cmprsk* for the Fine-Gray model (29).

For the weighted analyses, weights were calculated from the estimated PS. We estimated the PS, $\hat{p}$ for individual i, using a multivariable logistic regression model:

$$logit(\hat{p}) = \hat{\alpha_0} + \Sigma_l \hat{\alpha_l} \mathbf{X}_l$$

Weights for IPTW and STW were calculated using the following formulas:

$$IPTW = \frac{Z_i}{\hat{p_i}} + \frac{1 - Z_i}{1 - \hat{p_i}}$$
$$STW = Z\frac{\bar{Z}}{\hat{p_i}} + (1 - Z)\frac{1 - \bar{Z}}{1 - \hat{p_i}},$$

where $\bar{Z}$ was the proportion of treated patients in the study sample.

A weighted regression of outcome on treatment was fitted using these weights and the simulated data. Clustered analyses for both the Cox proportional hazard model and the Fine-Gray model were applied to adjust for the correlation introduced by the replication in the weighting process. Clustered analysis of Cox proportional hazard models is available in most of statistical packages. The clustered analysis of Fine-Gray model has been proposed (113)and was performed to estimate the marginal effect using the package *crrSC* in *R* (114).

The performance of the various estimation approaches were evaluated through several measures: the bias of the estimator from the true marginal treatment effect, the mean squared error (MSE) and the percent coverage (% coverage) of the nominal 0.95 confidence intervals. All performance measures were averaged over the 1,000 replications. Two different sample sizes, 500 and 2000, were used for each simulated sample.

2.2.6 Propensity-based estimation of the treatment effect in a case study

The five different models described above were also applied to the post-MI cohort to evaluate the marginal treatment effects of statins on three different cardiovascular outcomes (MI, heart failure, and stroke) and all-cause mortality in the presence of competing events. The PS was estimated using the covariates listed in Section 1.4.

2.3 Monte Carlo simulation results

2.3.1 True marginal treatment effects

The true marginal treatment effects for the three scenarios of competing risks are summarized in Table 2.1. The true marginal treatment effects decreased inversely to the proportion of the event of interest in both hazard models.

Table 2.1: True marginal treatment effect

	proportion of ε=1		
Model	0.4	0.6	0.8
subdistribution hazard	-0.420	-0.390	-0.373
cause-specific hazard	-0.426	-0.392	-0.384

2.3.2 Propensity-based estimation performance

The results obtained from the simulations for the four marginal models and one conditional model are summarized in Table 2.2 and Table 2.3. The PS distributions under the four different KS distances considered are shown in Figure 2.1.

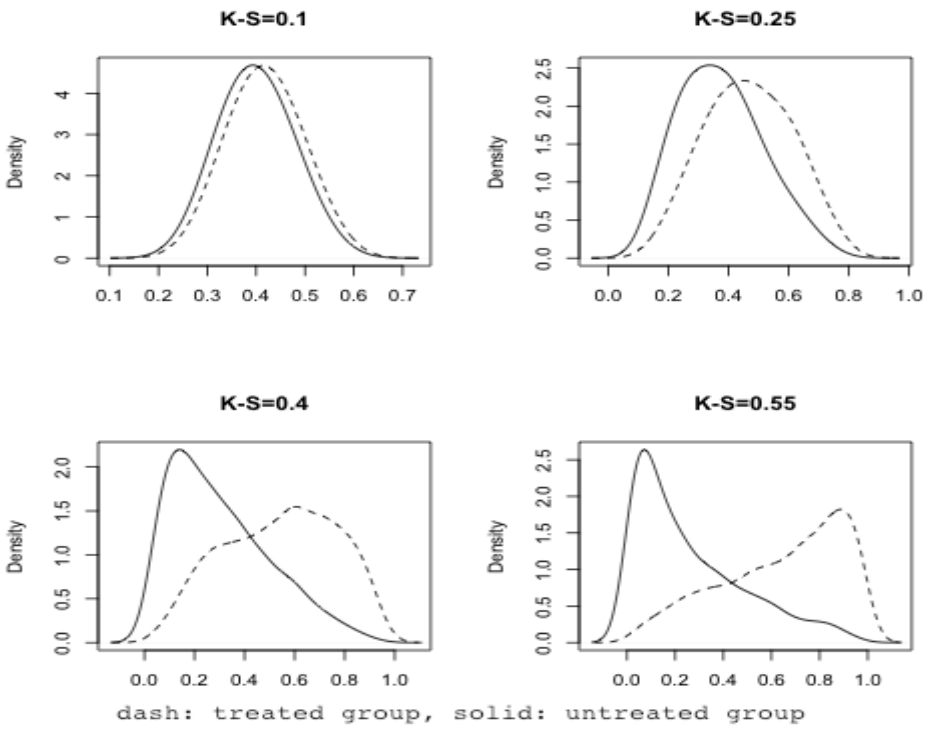

Figure 2.1: Density plots of treated and untreated group in four scenarios

Table 2.2: Simulation results of marginal treatment effect using competing risks (sample size =500)

		p=0.4						p=0.6						p=0.8					
		$\hat{\beta}$	Var($\hat{\beta}$)	E(V)	bias	MSE	%co.	$\hat{\beta}$	Var($\hat{\beta}$)	E(V)	bias	MSE	%co.	$\hat{\beta}$	Var($\hat{\beta}$)	E(V)	bias	MSE	%co.
K-S=0.10																			
SH	crude	-0.406	0.027	0.026	0.014	0.027	0.940	-0.379	0.019	0.018	0.011	0.019	0.940	-0.363	0.014	0.014	0.010	0.014	0.960
	cru. CS	-0.409	0.027	0.026	0.011	0.027	0.940	-0.381	0.018	0.019	0.009	0.018	0.950	-0.365	0.014	0.015	0.008	0.014	0.960
	IPTW	-0.439	0.023	0.027	-0.019	0.023	0.960	-0.410	0.015	0.020	-0.020	0.016	0.980	-0.391	0.010	0.015	-0.018	0.011	0.980
	STW	-0.414	0.025	0.026	0.005	0.025	0.960	-0.387	0.017	0.019	0.003	0.017	0.960	-0.371	0.013	0.015	0.002	0.013	0.970
	Conv.	-0.521	0.028	0.027	-0.101	0.039	0.910	-0.511	0.020	0.019	-0.122	0.035	0.870	-0.513	0.014	0.015	-0.140	0.034	0.810
CSH	crude	-0.395	0.030	0.028	0.031	0.031	0.940	-0.369	0.020	0.018	0.023	0.020	0.940	-0.352	0.015	0.013	0.032	0.016	0.930
	cru. CS	-0.398	0.030	0.029	0.028	0.031	0.950	-0.372	0.019	0.019	0.020	0.020	0.940	-0.355	0.014	0.014	0.029	0.015	0.940
	IPTW	-0.417	0.028	0.030	0.009	0.028	0.960	-0.392	0.017	0.020	-0.000	0.017	0.960	-0.378	0.012	0.014	0.006	0.012	0.970
	STW	-0.412	0.027	0.030	0.014	0.028	0.960	-0.387	0.017	0.019	0.005	0.017	0.960	-0.373	0.012	0.014	0.011	0.012	0.960
	Conv.	-0.511	0.034	0.031	-0.085	0.041	0.910	-0.511	0.022	0.020	-0.119	0.036	0.860	-0.514	0.017	0.014	-0.130	0.034	0.790
K-S=0.25																			
SH	crude	-0.262	0.026	0.025	0.158	0.051	0.810	-0.241	0.018	0.018	0.149	0.040	0.790	-0.233	0.014	0.014	0.140	0.033	0.790
	cru. CS	-0.277	0.025	0.026	0.143	0.046	0.840	-0.254	0.018	0.019	0.135	0.037	0.840	-0.246	0.013	0.015	0.127	0.030	0.830
	IPTW	-0.450	0.026	0.031	-0.030	0.027	0.960	-0.421	0.018	0.022	-0.031	0.019	0.980	-0.405	0.013	0.018	-0.032	0.014	0.980
	STW	-0.352	0.025	0.028	0.068	0.029	0.940	-0.327	0.017	0.021	0.062	0.021	0.940	-0.315	0.012	0.016	0.058	0.016	0.960
	Conv.	-0.510	0.031	0.029	-0.090	0.039	0.910	-0.502	0.023	0.021	-0.112	0.035	0.870	-0.507	0.016	0.016	-0.134	0.034	0.830
CSH	crude	-0.279	0.031	0.029	0.147	0.052	0.840	-0.256	0.019	0.019	0.136	0.037	0.830	-0.286	0.035	0.034	0.098	0.044	0.910
	cru. CS	-0.291	0.031	0.030	0.135	0.049	0.860	-0.268	0.019	0.019	0.124	0.034	0.850	-0.298	0.035	0.035	0.086	0.042	0.930
	IPTW	-0.433	0.033	0.035	-0.007	0.033	0.960	-0.407	0.019	0.023	-0.015	0.020	0.960	-0.444	0.038	0.041	-0.060	0.042	0.950
	STW	-0.413	0.032	0.035	0.013	0.032	0.970	-0.388	0.019	0.022	0.004	0.019	0.960	-0.425	0.037	0.040	-0.041	0.038	0.950
	Conv.	-0.512	0.038	0.034	-0.086	0.045	0.910	-0.509	0.022	0.021	-0.117	0.036	0.880	-0.513	0.042	0.039	-0.129	0.059	0.890
K-S=0.40																			
SH	crude	-0.640	0.027	0.028	-0.221	0.075	0.740	-0.595	0.020	0.019	-0.205	0.062	0.700	-0.566	0.016	0.015	-0.193	0.053	0.660
	cru. CS	-0.599	0.028	0.029	-0.179	0.060	0.830	-0.554	0.020	0.021	-0.165	0.047	0.800	-0.527	0.016	0.016	-0.154	0.040	0.790
	IPTW	-0.429	0.046	0.048	-0.009	0.046	0.960	-0.389	0.031	0.034	0.001	0.031	0.960	-0.375	0.022	0.026	-0.002	0.022	0.970
	STW	-0.474	0.042	0.044	-0.055	0.045	0.950	-0.433	0.028	0.031	-0.043	0.030	0.950	-0.415	0.020	0.024	-0.042	0.022	0.960
	Conv.	-0.524	0.038	0.036	-0.104	0.048	0.920	-0.511	0.028	0.025	-0.121	0.042	0.880	-0.509	0.021	0.020	-0.136	0.039	0.830
CSH	crude	-0.588	0.027	0.028	-0.162	0.053	0.850	-0.557	0.018	0.018	-0.165	0.045	0.780	-0.536	0.013	0.014	-0.152	0.037	0.770
	cru. CS	-0.555	0.028	0.030	-0.129	0.045	0.910	-0.524	0.019	0.020	-0.132	0.036	0.850	-0.503	0.014	0.015	-0.119	0.028	0.850
	IPTW	-0.422	0.043	0.048	0.004	0.043	0.960	-0.386	0.029	0.032	0.006	0.029	0.960	-0.368	0.021	0.023	0.016	0.021	0.960
	STW	-0.434	0.041	0.046	-0.008	0.041	0.960	-0.398	0.027	0.030	-0.006	0.027	0.970	-0.379	0.019	0.022	0.005	0.020	0.970
	Conv.	-0.525	0.040	0.038	-0.099	0.049	0.920	-0.516	0.027	0.025	-0.124	0.042	0.870	-0.515	0.020	0.018	-0.131	0.037	0.840
K-S=0.55																			
SH	crude	-0.744	0.030	0.028	-0.324	0.134	0.510	-0.691	0.020	0.019	-0.302	0.111	0.410	-0.658	0.014	0.015	-0.285	0.095	0.360
	cru. CS	-0.665	0.031	0.031	-0.245	0.091	0.710	-0.614	0.022	0.022	-0.224	0.072	0.690	-0.585	0.015	0.017	-0.212	0.060	0.650
	IPTW	-0.434	0.097	0.077	-0.014	0.097	0.920	-0.398	0.065	0.056	-0.008	0.065	0.940	-0.377	0.047	0.044	-0.004	0.047	0.960
	STW	-0.477	0.090	0.071	-0.057	0.094	0.920	-0.437	0.059	0.052	-0.047	0.062	0.940	-0.414	0.042	0.040	-0.040	0.044	0.960
	Conv.	-0.524	0.044	0.041	-0.104	0.055	0.910	-0.509	0.031	0.029	-0.119	0.045	0.880	-0.511	0.024	0.023	-0.138	0.043	0.840
CSH	crude	-0.675	0.031	0.028	-0.249	0.093	0.670	-0.647	0.019	0.018	-0.255	0.084	0.510	-0.631	0.013	0.014	-0.247	0.074	0.440
	cru. CS	-0.613	0.032	0.031	-0.187	0.067	0.820	-0.583	0.019	0.020	-0.191	0.055	0.720	-0.564	0.014	0.015	-0.180	0.047	0.690
	IPTW	-0.419	0.073	0.067	0.007	0.073	0.950	-0.405	0.048	0.046	-0.013	0.048	0.950	-0.382	0.035	0.037	0.002	0.035	0.970
	STW	-0.433	0.069	0.063	-0.007	0.069	0.950	-0.417	0.046	0.044	-0.025	0.046	0.950	-0.393	0.033	0.035	-0.009	0.033	0.970
	Conv.	-0.508	0.044	0.043	-0.082	0.051	0.930	-0.511	0.028	0.028	-0.119	0.042	0.900	-0.514	0.021	0.021	-0.130	0.038	0.860

p=proportion of ε=1, $\hat{\beta}$=estimator, Var($\hat{\beta}$)= empirical variance, E(V)=average variance, SH=subdistribution hazard function,
CSH=cause-specific hazard function, conv.=conventional model Crude = Crude model: $\lambda_1(t|Z) = \lambda_0(t)exp\{\beta Z\}$,Cru.-CS= crude model under common support ,
IPTW= weighted model by IPTW, STW=weighted model by stabilized weighted,Conv.=conventional model.

Table 2.3: Simulation results of marginal treatment effect using competing risks (sample size =2000)

		p=0.4						p=0.6						p=0.8					
		$\hat{\beta}$	Var($\hat{\beta}$)	E(V)	bias	MSE	%co.	$\hat{\beta}$	Var($\hat{\beta}$)	E(V)	bias	MSE	%co.	$\hat{\beta}$	Var($\hat{\beta}$)	E(V)	bias	MSE	%co.
K-S=0.10																			
SH	crude	-0.398	0.006	0.006	0.022	0.007	0.950	-0.376	0.004	0.005	0.013	0.004	0.960	-0.358	0.003	0.004	0.015	0.003	0.950
	cru. CS	-0.398	0.006	0.006	0.022	0.007	0.960	-0.377	0.004	0.005	0.013	0.004	0.960	-0.358	0.003	0.004	0.015	0.003	0.950
	IPTW	-0.430	0.006	0.007	-0.010	0.006	0.970	-0.406	0.004	0.005	-0.016	0.004	0.970	-0.386	0.003	0.004	-0.013	0.003	0.980
	STW	-0.401	0.006	0.006	0.019	0.006	0.960	-0.380	0.004	0.005	0.010	0.004	0.960	-0.361	0.003	0.004	0.012	0.003	0.960
	Conv.	-0.502	0.007	0.007	-0.082	0.013	0.830	-0.502	0.005	0.005	-0.113	0.017	0.620	-0.501	0.003	0.004	-0.128	0.020	0.450
CSH	crude	-0.395	0.007	0.007	0.031	0.008	0.930	-0.372	0.004	0.005	0.020	0.005	0.940	-0.349	0.003	0.003	0.035	0.005	0.910
	cru. CS	-0.395	0.007	0.007	0.031	0.007	0.940	-0.372	0.004	0.005	0.020	0.005	0.950	-0.349	0.003	0.003	0.035	0.004	0.910
	IPTW	-0.418	0.007	0.007	0.008	0.007	0.970	-0.396	0.004	0.005	-0.004	0.004	0.980	-0.374	0.003	0.004	0.010	0.003	0.980
	STW	-0.412	0.006	0.007	0.014	0.006	0.960	-0.391	0.004	0.005	0.001	0.004	0.980	-0.370	0.003	0.003	0.014	0.003	0.980
	Conv.	-0.502	0.008	0.007	-0.076	0.013	0.850	-0.504	0.005	0.005	-0.112	0.017	0.630	-0.502	0.004	0.004	-0.118	0.018	0.500
K-S=0.25																			
SH	crude	-0.263	0.007	0.006	0.157	0.031	0.480	-0.245	0.005	0.005	0.145	0.025	0.420	-0.231	0.004	0.004	0.142	0.024	0.340
	cru. CS	-0.267	0.007	0.006	0.153	0.030	0.500	-0.249	0.005	0.005	0.141	0.024	0.460	-0.235	0.003	0.004	0.138	0.022	0.360
	IPTW	-0.449	0.007	0.007	-0.029	0.007	0.950	-0.423	0.004	0.005	-0.033	0.005	0.960	-0.401	0.003	0.004	-0.028	0.004	0.970
	STW	-0.343	0.006	0.007	0.077	0.012	0.860	-0.323	0.004	0.005	0.067	0.009	0.860	-0.305	0.003	0.004	0.068	0.008	0.840
	Conv.	-0.505	0.007	0.007	-0.085	0.014	0.830	-0.502	0.005	0.005	-0.113	0.018	0.640	-0.500	0.004	0.004	-0.126	0.020	0.480
CSH	crude	-0.398	0.007	0.007	0.028	0.007	0.930	-0.377	0.005	0.004	0.015	0.005	0.940	-0.358	0.003	0.003	0.026	0.004	0.920
	cru. CS	-0.398	0.007	0.007	0.028	0.008	0.930	-0.376	0.005	0.004	0.016	0.005	0.940	-0.357	0.003	0.003	0.027	0.004	0.920
	IPTW	-0.413	0.007	0.008	0.013	0.007	0.960	-0.391	0.004	0.005	0.001	0.004	0.960	-0.370	0.003	0.004	0.014	0.003	0.960
	STW	-0.412	0.007	0.008	0.014	0.007	0.960	-0.390	0.004	0.005	0.002	0.004	0.960	-0.369	0.003	0.004	0.015	0.003	0.950
	Conv.	-0.501	0.008	0.008	-0.075	0.014	0.860	-0.503	0.006	0.005	-0.111	0.018	0.650	-0.500	0.004	0.004	-0.116	0.017	0.510
K-S=0.40																			
SH	crude	-0.628	0.007	0.007	-0.208	0.050	0.290	-0.594	0.005	0.005	-0.204	0.047	0.170	-0.566	0.004	0.004	-0.193	0.041	0.120
	cru. CS	-0.610	0.007	0.007	-0.190	0.043	0.370	-0.576	0.005	0.005	-0.187	0.040	0.240	-0.549	0.004	0.004	-0.176	0.035	0.190
	IPTW	-0.412	0.011	0.012	0.008	0.011	0.960	-0.386	0.007	0.008	0.004	0.007	0.970	-0.368	0.005	0.006	0.006	0.005	0.970
	STW	-0.465	0.010	0.011	-0.045	0.012	0.940	-0.436	0.006	0.008	-0.047	0.008	0.940	-0.416	0.005	0.006	-0.043	0.007	0.940
	Conv.	-0.501	0.009	0.009	-0.081	0.015	0.860	-0.499	0.006	0.006	-0.110	0.018	0.700	-0.500	0.005	0.005	-0.127	0.021	0.550
CSH	crude	-0.583	0.007	0.007	-0.157	0.032	0.560	-0.560	0.005	0.005	-0.168	0.033	0.290	-0.537	0.004	0.003	-0.153	0.027	0.260
	cru. CS	-0.569	0.007	0.007	-0.143	0.028	0.630	-0.545	0.005	0.005	-0.153	0.028	0.380	-0.522	0.004	0.003	-0.138	0.023	0.350
	IPTW	-0.398	0.011	0.012	0.028	0.012	0.950	-0.377	0.006	0.008	0.015	0.007	0.960	-0.356	0.004	0.006	0.028	0.005	0.960
	STW	-0.413	0.010	0.011	0.013	0.011	0.960	-0.391	0.006	0.007	0.001	0.006	0.970	-0.369	0.004	0.006	0.015	0.004	0.980
	Conv.	-0.503	0.009	0.009	-0.077	0.015	0.870	-0.507	0.006	0.006	-0.115	0.019	0.690	-0.503	0.004	0.004	-0.119	0.019	0.570
K-S=0.55																			
SH	crude	-0.732	0.007	0.007	-0.312	0.104	0.030	-0.693	0.005	0.005	-0.303	0.097	0.010	-0.663	0.004	0.004	-0.290	0.088	0.000
	cru. CS	-0.693	0.007	0.007	-0.273	0.082	0.090	-0.656	0.005	0.005	-0.266	0.076	0.040	-0.627	0.004	0.004	-0.254	0.068	0.020
	IPTW	-0.406	0.020	0.022	0.014	0.020	0.960	-0.388	0.015	0.016	0.002	0.015	0.960	-0.373	0.012	0.012	-0.000	0.012	0.960
	STW	-0.455	0.017	0.020	-0.035	0.019	0.950	-0.431	0.013	0.014	-0.041	0.015	0.940	-0.413	0.010	0.010	-0.040	0.012	0.930
	Conv.	-0.497	0.010	0.010	-0.078	0.016	0.870	-0.499	0.007	0.007	-0.109	0.019	0.760	-0.506	0.006	0.006	-0.133	0.024	0.580
CSH	crude	-0.672	0.007	0.007	-0.246	0.067	0.150	-0.641	0.005	0.005	-0.249	0.067	0.030	-0.622	0.003	0.003	-0.238	0.060	0.020
	cru. CS	-0.642	0.007	0.007	-0.216	0.054	0.290	-0.610	0.005	0.005	-0.218	0.052	0.100	-0.590	0.003	0.004	-0.206	0.046	0.070
	IPTW	-0.398	0.019	0.020	0.028	0.020	0.960	-0.374	0.012	0.013	0.018	0.012	0.960	-0.358	0.009	0.010	0.026	0.010	0.960
	STW	-0.414	0.018	0.019	0.012	0.018	0.960	-0.388	0.011	0.013	0.004	0.011	0.960	-0.371	0.008	0.010	0.013	0.009	0.970
	Conv.	-0.503	0.010	0.010	-0.077	0.016	0.880	-0.502	0.007	0.007	-0.110	0.019	0.730	-0.503	0.005	0.005	-0.119	0.019	0.620

p=proportion of ε=1, $\hat{\beta}$=estimator, Var($\hat{\beta}$)= empirical variance, E(V)=average variance, SH=subdistribution hazard function,
CSH=cause-specific hazard function, conv.=conventional model Crude = Crude model: $\lambda_1(t|Z) = \lambda_0(t)exp\{\beta Z\}$,Cru.-CS= crude model under common support ,
IPTW= weighted model by IPTW, STW=weighted model by stabilized weighted,Conv.=conventional model.

As shown in Table 2.2 and Table 2.3, the model adjusted by weighed methods, either IPTW or STW, had the highest percent of true estimator coverage and the least amount of bias under all scenarios considered. The percent of true value coverage of the weighted model was higher than 95% for all scenarios. The bias-variance trade-off can be observed in weighted models, however, the percent coverage of both weighted methods were close. In the weighted model, the bigger amount of bias showed with the larger KS distance. The two weighted regressions also performed better than the other models when sample sizes were smaller. In situations with a low occurrence of the event of interest and a large KS distance, the estimators of the weighted methods had variances remarkably different from the other models. In the IPTW model and KS =0.55, the variance estimator from the model based approach was reduced by 80% when using the bootstrapping method. In the STW, variance from bootstrapping method was reduced to 60% with the same KS.

For the crude model, which is unadjusted, the marginal treatment effects can be estimated from a randomized study. However, the same model, used in an observational study where confounding variables are present, may result in a biased estimator. As the level of confounding increases, the bias increases. In this study, the crude models were more sensitive than the weighted models to the level of confounding as measured by KS distance. When the KS distance increased from 0.1 to 0.55, the amount of bias increased less than three times for the weighted estimators. In comparison to the crude model, the crude model under common support showed improved efficiency in terms of bias, MSE and percent coverage.

The conventional model, which yielded estimators of the conditional treatment effects by including all confounding in the model, was shown to be a poor estimator of the marginal treatment effects. The estimators from the conventional model were more stable than the other models when the confounding level increased. However, the estimator from conventional model may not be the true causal estimator.

2.4 Case study results

In the study of statin treatment effects in the elderly who experienced myocardial infarction, randomization was unethical and the mortality event highly competed other cardiovascular disease outcomes. The cardiovascular outcomes of interest were MI, HF and stroke while all-cause mortality was a competing event. A total of 71,030 patients from the Medicare claims data were included in the analysis, and 64.5% of them received statins after discharge. As presented in Table 2.4, the unadjusted percentage of patients that had the occurrence of MI, mortality, stroke and heart failure were lower in the statin users group. All cause mortality constituted the highest percentage among observed events in both statin user and non-user groups.

Table 2.4: Number and percent of observed outcomes

		Survived	Mortality	MI	stroke	heart failure
Non-statins	number	14529	5749	2812	506	1643
	percent	57.6	22.8	11.1	2.0	6.5
Statins	number	31509	6318	4415	831	2718
	percent	68.8	13.8	9.6	1.8	5.9

For the PS estimated using a logistic regression, the strong predictors for the probability of receiving statin treatment were baseline statins ($\hat{\alpha}$= 1.66), admission for CABG ($\hat{\alpha}$=0.68), admission for stent/PTCA ($\hat{\alpha}$=0.66), beta blocker users ($\hat{\alpha}$=0.99) and moderate to severe liver disease ($\hat{\alpha}$=-0.58). Those variables were consistent with the indications for treatment with statins (56, 52, 101) The KS distance between the PS distribution of statin users and non-statin user was 0.418, and the c statistics was 0.767.

As shown in Table 2.5, the unadjusted crude model showed large marginal treatment effects for all outcomes, indicating bias in non-experimental studies. The results from crude model showed that statins significantly reduced the risk of recurrent MI, stroke, heart failure and all cause mortality. The use of the IPTW resulted in a large amount of weighting, the maximum weighting was 46. The weighted models, which were expected

to show a good efficiency for the estimator, did not show a significant effect for statins for statins for reducing the risk of stroke or recurrent MI, but did show that statins significantly reduced the risk of HF and all-cause mortality. IPTW and STW methods showed an inconsistency when the event was recurrent MI.

The estimated statin treatment effects for stroke from STW presented different results for the Fine-Gray model and the Cox proportional hazard model. The estimators from the cause-specific hazard function were larger than subdistribution hazard function for all outcomes. In addition, the estimator from the cause-specific hazard function can be overestimated when summarized from the competing risks survival model with the cumulative incidence function. The estimators from applying STW to subdistribution hazard function seem to be reasonable to use for the marginal treatment effects for this data.

To account for the large variance resulting from the low occurrence of event of interest and the great separation in the propensity score distributions, a bootstrapping method was used to obtain variance estimates. With this approach, the variance of the estimators were decreased to 40% and 20% of the model-based sandwich variance estimator for IPTW and STW methods, respectively (Table 2.6). As a result, the effect estimate for MI and heart failure from the IPTW model changed from non-significant to significant ($\alpha = 0.05$). STW also showed a different result, in terms of significance of effect, for the stroke endpoint. This finding of the large reduction in variances observed here has not been described previously. This may be due, in part, to the large number of confounders being adjusted for in the propensity score. In general, greater reductions in variance can be expected from larger propensity score models.

Table 2.5: The treatment effects of statins for elderly who were hospitalized for an AMI

		subdistribution model			cause-specific model		
Failure		$\hat{\beta}$	SE	p-value	$\hat{\beta}$	SE	p-value
MI	Crude	-0.153	0.024	0.000	-0.222	0.024	0.000
	Crude-CS	-0.153	0.024	0.000	-0.222	0.024	0.000
	IPTW*	0.041	0.016	0.011	0.039	0.016	0.017
	STW*	0.003	0.024	0.905	-0.010	0.024	0.676
	Conv.	-0.003	0.027	0.900	-0.021	0.027	0.440
All-cause	Crude	-0.553	0.018	0.000	-0.581	0.018	0.000
mortality	Crude-CS	-0.553	0.018	0.000	-0.581	0.018	0.000
	IPTW*	-0.045	0.013	0.001	-0.042	0.014	0.002
	STW*	-0.126	0.018	0.000	-0.129	0.018	0.000
	Conv.	-0.132	0.021	0.000	-0.130	0.021	0.000
Stroke	Crude	-0.101	0.056	0.075	-0.196	0.056	0.001
	Crude-CS	-0.101	0.056	0.074	-0.196	0.056	0.001
	IPTW*	-0.056	0.036	0.125	-0.057	0.036	0.116
	STW*	-0.097	0.053	0.068	-0.112	0.053	0.034
	Conv.	-0.095	0.063	0.130	-0.112	0.063	0.075
HF	Crude	-0.095	0.031	0.002	-0.176	0.031	0.000
	Crude-CS	-0.095	0.031	0.002	-0.176	0.031	0.000
	IPTW*	-0.067	0.022	0.003	-0.065	0.022	0.003
	STW*	-0.077	0.032	0.016	-0.088	0.032	0.006
	Conv.	-0.048	0.035	0.170	-0.066	0.035	0.059

*Standard error estimated from bootstrapping method

Table 2.6: Comparison of the variance of the model-base sandwich and bootstrapping method

			subdistribution hazard				cause-specific hazard			
			$\hat{\beta}$	SE	p-value	% reduc.	$\hat{\beta}$	SE	p-value	% reduc.
IPTW	MI	model	0.041	0.031	0.186		0.039	0.031	0.208	
		boots.	0.040	0.016	0.012	73.4	0.038	0.016	0.018	73.4
	Death	model	-0.045	0.023	0.050		-0.042	0.023	0.068	
		boots.	-0.044	0.013	0.001	68.1	-0.058	0.014	0.000	62.9
	Stroke	model	-0.056	0.070	0.424		-0.057	0.070	0.415	
		boots.	-0.057	0.036	0.113	73.6	-0.058	0.036	0.107	73.6
	HF	model	-0.067	0.040	0.094		-0.065	0.041	0.113	
		boots.	-0.068	0.022	0.002	69.8	-0.067	0.022	0.002	71.2
STW	MI	model	0.003	0.030	0.920		-0.010	0.030	0.739	
		boots.	0.005	0.024	0.835	36.0	-0.008	0.024	0.739	36.0
	Death	model	-0.126	0.023	0.000		-0.129	0.023	0.000	
		boots.	-0.125	0.018	0.000	38.8	-0.127	0.018	0.000	38.8
	Stroke	model	-0.097	0.068	0.154		-0.112	0.068	0.100	
		boots.	-0.098	0.053	0.064	39.3	-0.113	0.053	0.033	39.3
	HF	model	-0.077	0.039	0.048		-0.088	0.039	0.024	
		boots.	-0.077	0.032	0.016	32.7	-0.088	0.032	0.006	32.7

$\hat{\beta}$=treatment effect estimator,

%reduc. = percent of variance reduction from model-base variance

model=model-base sandwich variance estimator, boots=bootstrapping variance estimator

2.5 Discussion

In this study we examined the performance of propensity score-based methods in the presence of competing risks . Two hazard functions for competing risk methods were considered: the cause-specific hazard function and the subdistribution hazard function. Two weighting methods, IPTW and STW were shown to have the smallest amount of bias and the highest percent coverage compared to other models. The advantage of the STW method had over the IPTW methods shown in Xu et al (111) for conventional linear regression models was not clearly exhibited for the competing risks model in this study. Suarez et al (99) examined the differences between marginal structural models (MSMs) and conventional models and found that the estimators from MSMs had a 20% difference in effect size from a conventional model, but MSMs estimators had a 19% larger standard error. This study also saw inflated variances in the weighted models in comparison with the conventional model and the crude models. This inflation was more pronounced when the occurrence of the event of interest was low and the separation between propensity score distributions of the two treatment groups was high. Hahn (36) demonstrated that PS does not decrease the asymptotic variance bound when estimating average treatment effect. In this situation, the bootstrapping method is recommended to obtain the variance.

The estimates for the true marginal treatment effect calculated from the randomization trial simulation were not equal to the true parameters specified in the setup of simulations. This was due to the non-collapsibility characteristic of estimators from nonlinear models. The conventional conditional model yielded greater biases in the effect estimator due to the same non-collapsibility of estimators from nonlinear models. The marginal treatment effects from nonlinear models, such as survival analysis, are usually closer to the null compared to a conditional model (71). It is also possible that there was an imbalance in the data. Gail et al (25) showed that the estimator from unadjusted analyses of randomized trials could be biased when important variables were not well

controlled. It is similar to the situation where the assumption of no model misspecification was violated in MSMs. In order to obtain an unbiased estimator, one needs an excellent knowledge of confounders and risk factors for the question of interest. Our analyses were limited to data available in Medicare claims, which may not capture all potential risk factors.

The KS distance was used to characterize the differences between density functions of the propensity scores of the two treatment groups. The KS distance increases when the confounding variables are highly predictive of the probability of receiving treatment. In simulations presented in this study, efficiency of estimators decreased as the KS distance increased, which also produced an increase in bias and a decrease in percent coverage. We chose KS distance as a measure of confounding for several reasons. In studies using propensity score methods, c-statistics are widely used to show how well the propensity score model can discriminate between treatment groups (46). Westreich (107) demonstrated that higher values of the c-statistics may be associated with less overlap of propensity score distributions between the treated and untreated groups. Austin et al (5) found that c-statistics has no association with the ability of propensity scores to balance confounding factors between the treated and untreated groups when matching methods were used. In comparison, the KS distance is based on the empirical distribution function, and the supremum of absolute differences between the empirical distribution of treated and untreated is reported (63). We thus believe the KS distance is a better measure for this study.

Our findings should be interpreted in light of some limitations. This study considered a relatively small number of dichotomous confounding factors in the simulations. In practice, the number of confounding factors in observational studies can be large, the confounding factors may be continuous (such as age), and continuous confounding factors increase the chance of model misspecification. Further investigations are needed to evaluate the performances of these models when continuous confounding variables are present.

In conclusion, the weighted methods, IPTW and STW, could be used to control for confounding when conducting competing risk analyses to estimate the marginal treatment effects. Appropriate variance estimation approaches are needed when the frequency of the event of interest is low and the distance between propensity score distributions of the two treatment groups are high. In this situation, a bootstrapping method can be used to obtain variance estimates which appropriately take advantage of the estimation of the propensity score.

Chapter 3

Conditional Treatment Effects of Competing Risks Models

3.1 Introduction

A conditional treatment effect is the average effect of treatment on the individual. Non-experimental studies are increasingly used to evaluate the effect of treatments on outcomes, especially for medications that have been released on the market (11). In these studies, the treatment choice is often influenced by the characteristics of patients, and this may lead to confounding due to characteristics present in patients before the treatment choice is made that may influence both the choice and future events (96). Therefore, in order to obtain a valid estimate of the treatment effect on outcomes, one must account for these systematic differences.

Multivariable regression adjustment has been widely used to control for confounding by including potentially confounding baseline covariates in the outcome model. Recently, methods based on PS have been increasingly used to control for confounding in estimation of the treatment effects in non-experimental studies. Such approaches may be used to obtain unbiased estimates of adjusted treatment effects from regression models (which include confounding variables also present in the PS models) as well as the marginal treatment effects discussed in Chapter 2.

The performances of PS methods have been investigated in both Monte Carlo simulations and systematic reviews. When applied to linear regression, the PS methods showed similar or better results as a multivariate regression (88, 89, 70). Using PS methods

with logistic regression leads to better control of imbalances between treatment groups than conventional logistic regression provides when the event to confounder ratio is less than 7:1 (14). In time-to-event analyses where PS methods were applied to Cox proportional hazard models, the PS pair-matching method had the smallest amount of bias compared to other methods (26, 4). PS matching and subclassification yielded a larger treatment effect compared to conventional Cox proportional hazard models (62). Little has been done to evaluate the performance of PS methods when used with competing risks analyses.

Competing risk methods extend the applicability of survival analysis to situations where there are multiple outcome events but only the first occurring event for each subject is observable. These methods are being increasingly used in studies where competing events are likely to occur. However, competing risk models conventionally use multivariable adjustment, including all the covariates in the outcome model, which is not suitable in studies where the outcome of interest is rare and a large number of confounders are present (11). It was natural to consider applying PS methods to reduce or diminish the effect of confounding in competing risks analysis. This study aimed to investigate the performance of PS methods to estimate conditional treatment effects when applied to competing risks analyses where a large number of confounders are present.

3.2 Methods

3.2.1 Monte Carlo simulation

We conducted a series of Monte Carlo simulations in order to examine the performance of several PS methods with two competing risk models for estimating conditional treatment effects: the subdistribution hazard model and the cause-specific hazard model.

Let $Z_{n\times 1}$ be the matrix of binary treatment, $Z \in \{0, 1\}$, where z=1 if treated and z=0 if untreated. Let $X_{n\times l}$ T be the matrix of confounding variables with sample size $n \in \{500,\ 1000,\ 2000\}$ and l numbers of confounding variables $l \in \{5,\ 15,\ 50,\ 100\}$.

The set of confounders consisted of both binary and continuous variables. The binary confounding were generated from a Bernoulli distribution with parameter $p = 0.5$. The continuous confounding variables were generated from a normal distribution with zero mean and unit variance. The total number of confounding variables dictated the number of continuous variables that were generated: for a total of 5 or 15, one continuous variable was generated; for a total of 50 confounders, 3 continuous variables were generated; for a total of 100, 5 variables were generated.

The probability that the exposure was equal to one, which is the true PS, was generated as a function of the confounding. The vector Z was drawn from a Bernoulli distribution with probability set by

$$logit(\theta) = \alpha_0 + \Sigma_l \alpha_l X_l$$

where θ is the probability to receive treatment, and α_l are parameters for covariates included in the PS model. The values of α were fixed at levels such that the distance between PS distributions of the treated and untreated group measured by KS distance was equal to 0.2.

Two types of hazard model used in competing risks analysis were considered: cause-specific hazard function (51) and subdistribtion hazard function (30). Failure times were generated using the two hazard functions given the presence of competing events using methods similar to those presented in Chapter 2. Two types of observing event and the proportion of the event of interest was 0.6. The true treatment effects (β) on the outcome considered were 0 for the null model and -0.5 for the alternative model. The censored data (C_i) were generated from a uniform distribution to generate 20% of censoring.

3.2.2 Propensity-based estimation of the treatment effect in simulations

For each simulated dataset, five different models were fitted with the two different competing risk analysis approaches: PS matching, PS matching under common support area, subclassification, PS adjustment in the model.

3.2.2.1 PS matching

PS matching methods attempt to choose a single or multiple patients from the untreated group with the same values of the PS variable for each patient in the treated group. There are three ways to implement PS matching: pair matching (one-to-one matching), many-to-one matching and full matching. Pair matching method is the most common method for PS matching. Full matching includes all subjects and groups them to their best matching sets. This method reduces bias more effectively than pair matching and result in closer matches than many-to-one matching (37). Full matching results in matched sets consisting at least one treated and at least one untreated, thus potentially creating a wide range of treated to untreated ratio (94).

In this simulation, we used full matching. The distance matrix was constructed from rank based Mahalanobis distance and the caliper was 0.2×S.D of PS (79). The sample was matched using two different methods, the entire sample approach and the common support approach. The common support approach included only subjects who fell under the overlap area between PS distribution of the treated and untreated.

A stratified Cox model was applied to estimate the conditional treatment effect with the cause-specific hazard function. For the full matching method, the number of strata increases as the sample size increases, which could impact the power of Cox model (103). The cause-specific hazard function for the k^{th} stratum is

$$h_k(t, Z, X) = h_{0k}(t)exp(\beta Z + \gamma X)$$

Stratification for the Fine-Gray model was developed with two stratification regimes , highly stratified and regular stratified (115). The highly stratification method applies when the size of stratum is finite as n $\to \infty$. The regular stratification method is used when the number of strata is finite, a larger total sample size can produce larger sample size in the stratum.

The subdistribution hazard function for the k^{th} stratum (group of matching PS) is

$$\lambda_k(t, Z) = \lambda_{0k}(t)exp(\beta Z)$$

whereλ_{0k} is the baseline hazard function for stratum k.

3.2.2.2 Subclassification

Subclassification is a method used to adjust for confounding factors by classifying the treated and untreated into strata based on percentiles of PS. We used quintiles of PS, as is generally done to divide the subjects into strata, to conduct subclassification. The quintiles of PS can remove approximately 90% of the initial bias due to confounding (81).

For the cause-specific hazard function analysis, a stratified Cox model was employed to estimate the treatment effect. With this model, the baseline hazard function varies by each subclass or stratum. The number of strata from subclassification is smaller than the full PS matching used in Section 3.2.2.1, and it may have less impact on the sample sizes and power of the Cox model.

The regular stratification method of the Fine-Gray model was applied to the subdistribution hazard function data analysis (115).

3.2.2.3 Adjustment for PS in the model

Including PS adjustment in an outcome regression model is another standard PS implementation method. This approach assumes a specific function form for the relationship of the PS to the outcome. Violation of this assumption can introduce a large amount of bias when the covariance of the treated and untreated group are unequal (19).

For this study, a multivariable Cox model was used to estimate the treatment effect for the cause-specific hazard function and a multivariable Fine-Gray model was used to estimate the treatment effect for the subdistribution hazard function.

All analyses in this study were performed in R version 3.0.1 (102). The package *survival* (103) was employed for the cause-specific hazard function. The packages *cmprsk* (29) and *crrSC* (114) were used to analyze the data from the subdistribution hazard

function. The package *optmatch* (38) were used for PS matching. The performance of the various estimation approaches were evaluated through several measures: the bias of the estimator from the true treatment effect, the mean squared error (MSE), power and the percent coverage of the nominal 0.95 confidence intervals. All performance measures were averaged over the 1,000 replications. Three different sample sizes, 500, 100 and 2000 were used for each simulated sample. The null model and alternative model were tested, the true treatment effect for null model was fixed at 0 while true treatment effect for alternative model was fixed at -0.5.

3.2.3 Propensity-based estimation of the conditional treatment effect in a case study

The five different models described above were also used with the dataset extracted from the Medicare claims data of the elderly post-AMI cohort to evaluate the conditional treatment effects of statins on three different cardiovascular outcomes (MI, HF, and stroke) and all-cause mortality. The patients who died within 30 days after discharge were excluded from the analysis. The patients were followed up from 31 days after discharge until the occurrence of the first event: MI, stroke, HF or death.

The PS was estimated in a logistic regression model using a list of 134 covariates identified a priori, including baseline demographics, comorbidities, medications, procedures, current cormorbidities, medication and diagnoses during admission, interaction terms between age and group of total CCI also included in the PS model. Details of the variables included in the model and estimators appear in Appendix 1.

The competing risks models, one using the subdistribution hazard function and the other using the cause-specific hazard function, were applied to estimate treatment effects of statins on the risk of MI, stroke, HF and mortality. The R package was used to analyze the data.

3.3 Simulation results

The estimators for conditional treatment effects under all considered scenarios are summarized in Table 3.1- Table 3.4. The performance of the estimators was different for each of the following: different numbers of confounding variables, the type of model (null, alternative), and the hazard function (subdistribution, cause-specific). For smaller numbers of confounding variables, the PS-based estimator's performance was similar to that from the conventional model.

PS methods yielded good efficiency for the estimator in the model with heavy confounding. In the scenario of small sample size and heavy confounding, the large variance of the estimator from the conventional model was noticeable. In the scenario of heavy confounding, the power of the estimator from the subdistribution hazard model was larger than that from the cause-specific hazard model. The estimator from PS adjusted into multivariate model presented a high percent coverage and small bias in the null model. Subclassification methods showed the smallest MSE for alternative model with heavy confounding.

3.3.1 Under the null model

The objective of simulations under a null model was to study the false negative results. Estimators of covariate adjustment method using the PS performed the best and presented the smallest bias, MSE and the highest percent coverage ($> 95\%$) for both the cause-specific hazard function and the subdistribution hazard function under all scenarios considered.

For the PS subclassification method, high percent coverage and small MSE were present under the cause-specific hazard function but not under the subdistribution hazard function. Under the subdistribution hazard function, the percent coverage of the estimators from PS subclassification method noticeably decreased with heavier confounding.

The PS matching estimators showed low percent coverage and a large MSE under all

settings. Matching under the common support approach yielded estimators with minimal improved efficiency for the treatment effect for both types of hazard functions. Among all methods, the PS matching estimators had the largest variance.

3.3.2 Under the alternative model

In the scenario of heavy confounding (50 and 100 confounder variables), estimator from subclassification hazard model had small MSE and high power of the test. However, those results did not show in the estimators from the cause-specific hazard model. For the cause-specific hazard model, the conventional model presented a small bias, a small MSE and a high percent coverage. For the non-heavy confounding (5 and 15 confounder variables) the conventional model outperformed the other methods. The conventional model estimators had a smaller bias, a smaller MSE and greater percent coverage for both the subdistribution hazard function and the cause-specific hazard function.

For the estimators from the PS covariate adjustment method, although they performed the best under the null model, they were shown to be the worst estimator for the alternative models. The estimators from PS matching and matching under common support showed the largest variance for all scenarios. In the scenario of small sample size and heavy confounding (100 confounding variables), the performance of PS matching improved, the estimator had the smallest bias and the highest percent coverage.

Table 3.1: Simulation result of conditional model of 5 confounders

No. confounders=5		Subdistribution model								Cause-specific model						
Sample size		β	$\hat{\beta}$	V($\hat{\beta}$)	E(V)	bias	MSE	%cov.	power	$\hat{\beta}$	V($\hat{\beta}$)	E(V)	bias	MSE	%cov.	power
500	CONV.	0.000	0.000	0.019	0.019	0.000	0.019	0.948	0.052	0.004	0.018	0.018	0.004	0.018	0.948	0.052
	PS Adj.	0.000	0.000	0.017	0.019	0.000	0.017	0.962	0.038	0.003	0.016	0.018	0.003	0.016	0.962	0.038
	Subclass	0.000	0.079	0.017	0.017	0.079	0.023	0.908	0.092	0.006	0.016	0.018	0.006	0.016	0.965	0.035
	Matching	0.000	0.078	0.029	0.030	0.078	0.036	0.926	0.074	0.089	0.036	0.035	0.089	0.044	0.931	0.069
	Match.-CS	0.000	0.079	0.031	0.031	0.079	0.037	0.923	0.077	0.072	0.037	0.037	0.072	0.043	0.939	0.061
	CONV.	-0.500	-0.505	0.024	0.023	-0.005	0.024	0.945	0.918	-0.498	0.021	0.021	0.002	0.021	0.950	0.942
	PS Adj.	-0.500	-0.467	0.022	0.023	0.033	0.023	0.952	0.887	-0.447	0.018	0.021	0.053	0.021	0.955	0.896
	Subclass	-0.500	-0.387	0.021	0.021	0.113	0.034	0.877	0.767	-0.443	0.019	0.021	0.057	0.022	0.947	0.888
	Matching	-0.500	-0.390	0.036	0.035	0.110	0.048	0.893	0.560	-0.375	0.039	0.041	0.125	0.055	0.910	0.475
	Match.-CS	-0.500	-0.388	0.039	0.036	0.112	0.051	0.887	0.531	-0.394	0.044	0.043	0.106	0.055	0.909	0.486
1000	CONV.	0.000	0.002	0.009	0.009	0.002	0.009	0.946	0.054	-0.001	0.009	0.009	-0.001	0.009	0.947	0.053
	PS Adj.	0.000	0.001	0.008	0.009	0.001	0.008	0.963	0.037	-0.000	0.008	0.009	-0.000	0.008	0.968	0.032
	Subclass	0.000	0.080	0.008	0.008	0.080	0.015	0.849	0.151	0.004	0.008	0.009	0.004	0.008	0.964	0.036
	Matching	0.000	0.075	0.014	0.015	0.075	0.020	0.909	0.091	0.071	0.017	0.017	0.071	0.022	0.917	0.083
	Match.-CS	0.000	0.079	0.014	0.015	0.079	0.020	0.910	0.090	0.071	0.017	0.018	0.071	0.022	0.923	0.077
	CONV.	-0.500	-0.497	0.011	0.011	0.003	0.011	0.955		-0.505	0.011	0.010	-0.005	0.011	0.949	0.999
	PS Adj.	-0.500	-0.461	0.009	0.011	0.039	0.011	0.954	0.998	-0.454	0.009	0.010	0.046	0.011	0.936	0.997
	Subclass	-0.500	-0.381	0.009	0.010	0.119	0.024	0.790	0.969	-0.450	0.009	0.010	0.050	0.012	0.928	0.996
	Matching	-0.500	-0.378	0.017	0.017	0.122	0.032	0.840	0.823	-0.398	0.020	0.020	0.102	0.030	0.892	0.819
	Match.-CS	-0.500	-0.381	0.017	0.018	0.119	0.031	0.858	0.820	-0.396	0.021	0.021	0.104	0.031	0.878	0.802
2000	CONV.	0.000	-0.003	0.005	0.005	-0.003	0.005	0.940	0.060	-0.001	0.009	0.009	-0.001	0.009	0.947	0.053
	PS Adj.	0.000	-0.003	0.004	0.005	-0.003	0.004	0.948	0.052	-0.000	0.008	0.009	-0.000	0.008	0.968	0.032
	Subclass	0.000	0.072	0.005	0.004	0.072	0.010	0.803	0.197	0.004	0.008	0.009	0.004	0.008	0.964	0.036
	Matching	0.000	0.065	0.008	0.008	0.065	0.012	0.884	0.116	0.071	0.017	0.017	0.071	0.022	0.917	0.083
	Match.-CS	0.000	0.073	0.008	0.008	0.073	0.014	0.855	0.145	0.071	0.017	0.018	0.071	0.022	0.923	0.077
	CONV.	-0.500	-0.506	0.005	0.006	-0.006	0.005	0.940	1	-0.505	0.011	0.010	-0.005	0.011	0.949	0.999
	PS Adj.	-0.500	-0.469	0.005	0.006	0.031	0.006	0.947	1	-0.454	0.009	0.010	0.046	0.011	0.936	0.997
	Subclass	-0.500	-0.392	0.005	0.005	0.108	0.017	0.682	1	-0.450	0.009	0.010	0.050	0.012	0.928	0.996
	Matching	-0.500	-0.394	0.009	0.009	0.106	0.020	0.797	0.991	-0.398	0.020	0.020	0.102	0.030	0.892	0.819
	Match.-CS	-0.500	-0.386	0.009	0.009	0.114	0.022	0.758	0.976	-0.396	0.021	0.021	0.104	0.031	0.878	0.802

β=treatment effect, Var($\hat{\beta}$)= empirical variance, E(V)=average variance,
CONV.=conventional model, PS adj.= Propensity score adjusted into the model, Subclass=subclassification,
Match.-CS= Matching under common support

Table 3.2: Simulation result of conditional model of 15 confounders

No. confounders=15		Subdistribution model								Cause-specific model						
Sample size		β	$\hat{\beta}$	V($\hat{\beta}$)	E(V)	bias	MSE	%cov.	power	$\hat{\beta}$	V($\hat{\beta}$)	E(V)	bias	MSE	%cov.	power
500	CONV.	0.000	0.003	0.018	0.017	0.003	0.018	0.944	0.056	-0.000	0.018	0.018	-0.000	0.018	0.943	0.057
	PS Adj.	0.000	0.005	0.013	0.017	0.005	0.013	0.975	0.025	0.005	0.014	0.018	0.005	0.014	0.976	0.024
	Subclass	0.000	0.039	0.015	0.015	0.039	0.016	0.936	0.064	0.007	0.014	0.018	0.007	0.014	0.973	0.027
	Matching	0.000	0.035	0.028	0.030	0.035	0.030	0.952	0.048	0.039	0.038	0.038	0.039	0.039	0.948	0.052
	Match.-CS	0.000	0.039	0.032	0.031	0.039	0.034	0.944	0.056	0.045	0.038	0.040	0.045	0.040	0.959	0.041
	CONV.	-0.500	-0.514	0.021	0.020	-0.014	0.021	0.944	0.950	-0.516	0.023	0.021	-0.016	0.024	0.941	0.941
	PS Adj.	-0.500	-0.409	0.015	0.020	0.091	0.024	0.921	0.867	-0.409	0.017	0.020	0.091	0.026	0.922	0.849
	Subclass	-0.500	-0.373	0.017	0.017	0.127	0.033	0.835	0.829	-0.408	0.018	0.021	0.092	0.026	0.922	0.833
	Matching	-0.500	-0.380	0.035	0.034	0.120	0.049	0.889	0.553	-0.391	0.041	0.044	0.109	0.053	0.929	0.450
	Match.-CS	-0.500	-0.373	0.035	0.035	0.127	0.051	0.892	0.522	-0.384	0.045	0.046	0.116	0.059	0.908	0.421
1000	CONV.	0.000	-0.001	0.009	0.008	-0.001	0.009	0.939	0.061	0.006	0.009	0.009	0.006	0.009	0.941	0.059
	PS Adj.	0.000	0.000	0.006	0.008	0.000	0.006	0.974	0.026	0.009	0.007	0.009	0.009	0.007	0.976	0.024
	Subclass	0.000	0.036	0.007	0.007	0.036	0.009	0.926	0.074	0.011	0.007	0.009	0.011	0.007	0.976	0.024
	Matching	0.000	0.036	0.015	0.015	0.036	0.016	0.944	0.056	0.057	0.019	0.020	0.057	0.022	0.936	0.064
	Match.-CS	0.000	0.042	0.015	0.016	0.042	0.017	0.945	0.055	0.056	0.020	0.020	0.056	0.023	0.929	0.071
	CONV.	-0.500	-0.510	0.010	0.010	-0.010	0.010	0.949	0.998	-0.503	0.011	0.010	-0.003	0.011	0.924	0.996
	PS Adj.	-0.500	-0.412	0.007	0.010	0.088	0.015	0.891	0.991	-0.405	0.009	0.010	0.095	0.018	0.863	0.986
	Subclass	-0.500	-0.375	0.009	0.009	0.125	0.024	0.735	0.979	-0.404	0.009	0.010	0.096	0.018	0.867	0.988
	Matching	-0.500	-0.382	0.016	0.017	0.118	0.030	0.861	0.841	-0.375	0.023	0.023	0.125	0.039	0.858	0.710
	Match.-CS	-0.500	-0.379	0.017	0.018	0.121	0.032	0.851	0.835	-0.379	0.022	0.023	0.121	0.037	0.871	0.729
2000	CONV.	0.000	-0.000	0.004	0.004	-0.000	0.004	0.939	0.061	0.002	0.004	0.004	0.002	0.004	0.949	0.051
	PS Adj.	0.000	-0.000	0.003	0.004	-0.000	0.003	0.972	0.028	0.006	0.003	0.004	0.006	0.003	0.968	0.032
	Subclass	0.000	0.035	0.004	0.004	0.035	0.005	0.911	0.089	0.007	0.003	0.004	0.007	0.003	0.965	0.035
	Matching	0.000	0.035	0.008	0.008	0.035	0.009	0.935	0.065	0.049	0.009	0.010	0.049	0.012	0.926	0.074
	Match.-CS	0.000	0.037	0.008	0.008	0.037	0.009	0.936	0.064	0.049	0.010	0.010	0.049	0.012	0.918	0.082
	CONV.	-0.500	-0.503	0.005	0.005	-0.003	0.005	0.949	1	-0.500	0.004	0.005	-0.000	0.004	0.963	1
	PS Adj.	-0.500	-0.409	0.004	0.005	0.091	0.012	0.757	1	-0.405	0.004	0.005	0.095	0.013	0.741	1
	Subclass	-0.500	-0.372	0.004	0.004	0.128	0.020	0.514	1	-0.404	0.004	0.005	0.096	0.013	0.745	1
	Matching	-0.500	-0.384	0.009	0.009	0.116	0.022	0.777	0.989	-0.382	0.011	0.011	0.118	0.025	0.804	0.953
	Match.-CS	-0.500	-0.378	0.009	0.009	0.122	0.024	0.749	0.985	-0.379	0.011	0.011	0.121	0.026	0.794	0.953

β=treatment effect, Var($\hat{\beta}$)= empirical variance, E(V)=average variance,
CONV.=conventional model, PS adj.= Propensity score adjusted into the model, Subclass=subclassification,
Match.-CS= Matching under common support

Table 3.3: Simulation result of conditional model of 50 confounders

No. confounders=50		Subdistribution model								Cause-specific model						
Sample size		β	$\hat{\beta}$	V($\hat{\beta}$)	E(V)	bias	MSE	%cov.	power	$\hat{\beta}$	V($\hat{\beta}$)	E(V)	bias	MSE	%cov.	power
500	CONV.	0.000	0.003	0.026	0.020	0.003	0.026	0.919	0.081	0.009	0.026	0.020	0.009	0.026	0.922	0.078
	PS Adj.	0.000	-0.001	0.013	0.018	-0.001	0.013	0.983	0.017	0.017	0.014	0.018	0.017	0.014	0.979	0.021
	Subclass	0.000	-0.163	0.017	0.016	-0.163	0.043	0.732	0.268	0.014	0.014	0.019	0.014	0.014	0.978	0.022
	Matching	0.000	-0.166	0.033	0.031	-0.166	0.061	0.835	0.165	-0.011	0.037	0.037	-0.011	0.037	0.947	0.053
	Match.-CS	0.000	-0.159	0.033	0.033	-0.159	0.059	0.845	0.155	-0.012	0.039	0.040	-0.012	0.039	0.955	0.045
	CONV.	-0.500	-0.551	0.029	0.024	-0.051	0.031	0.915	0.928	-0.548	0.031	0.024	-0.048	0.033	0.906	0.923
	PS Adj.	-0.500	-0.362	0.015	0.021	0.138	0.034	0.875	0.749	-0.361	0.017	0.021	0.139	0.036	0.874	0.726
	Subclass	-0.500	-0.524	0.019	0.018	-0.024	0.020	0.942	0.971	-0.365	0.017	0.021	0.135	0.036	0.878	0.715
	Matching	-0.500	-0.532	0.040	0.036	-0.032	0.041	0.947	0.813	-0.407	0.044	0.043	0.093	0.053	0.932	0.498
	Match.-CS	-0.500	-0.531	0.037	0.038	-0.031	0.038	0.948	0.792	-0.411	0.049	0.046	0.089	0.057	0.923	0.472
1000	CONV.	0.000	-0.008	0.010	0.009	-0.008	0.010	0.934	0.066	0.002	0.010	0.009	0.002	0.010	0.930	0.070
	PS Adj.	0.000	-0.005	0.006	0.009	-0.005	0.006	0.981	0.019	0.009	0.006	0.009	0.009	0.006	0.978	0.022
	Subclass	0.000	-0.166	0.008	0.008	-0.166	0.035	0.544	0.456	0.007	0.006	0.009	0.007	0.006	0.979	0.021
	Matching	0.000	-0.168	0.017	0.017	-0.168	0.045	0.757	0.243	-0.019	0.019	0.020	-0.019	0.020	0.959	0.041
	Match.-CS	0.000	-0.172	0.017	0.018	-0.172	0.046	0.762	0.238	-0.014	0.021	0.021	-0.014	0.022	0.958	0.042
	CONV.	-0.500	-0.527	0.011	0.010	-0.027	0.012	0.935	1	-0.524	0.012	0.010	-0.024	0.013	0.917	1
	PS Adj.	-0.500	-0.365	0.007	0.010	0.135	0.025	0.756	0.984	-0.367	0.008	0.010	0.133	0.025	0.739	0.987
	Subclass	-0.500	-0.523	0.009	0.009	-0.023	0.010	0.944	1	-0.369	0.008	0.010	0.131	0.025	0.757	0.987
	Matching	-0.500	-0.542	0.020	0.019	-0.042	0.021	0.933	0.988	-0.415	0.024	0.024	0.085	0.031	0.908	0.774
	Match.-CS	-0.500	-0.535	0.021	0.020	-0.035	0.022	0.942	0.976	-0.414	0.025	0.024	0.086	0.033	0.899	0.760
2000	CONV.	0.000	0.001	0.005	0.004	0.001	0.005	0.939	0.061	0.001	0.004	0.004	0.001	0.004	0.941	0.059
	PS Adj.	0.000	0.003	0.003	0.004	0.003	0.003	0.980	0.020	0.008	0.003	0.004	0.008	0.003	0.982	0.018
	Subclass	0.000	-0.160	0.004	0.004	-0.160	0.030	0.285	0.715	0.006	0.003	0.004	0.006	0.003	0.983	0.017
	Matching	0.000	-0.169	0.009	0.009	-0.169	0.038	0.578	0.422	-0.017	0.011	0.011	-0.017	0.011	0.954	0.046
	Match.-CS	0.000	-0.159	0.009	0.009	-0.159	0.035	0.584	0.416	-0.018	0.010	0.011	-0.018	0.011	0.952	0.048
	CONV.	-0.500	-0.511	0.005	0.005	-0.011	0.005	0.936	1	-0.510	0.005	0.005	-0.010	0.005	0.941	1
	PS Adj.	-0.500	-0.365	0.003	0.005	0.135	0.022	0.506	1	-0.366	0.004	0.005	0.134	0.022	0.493	1
	Subclass	-0.500	-0.521	0.004	0.004	-0.021	0.005	0.952	1	-0.368	0.004	0.005	0.132	0.021	0.511	1
	Matching	-0.500	-0.539	0.010	0.010	-0.039	0.012	0.930	0.999	-0.415	0.013	0.012	0.085	0.020	0.868	0.966
	Match.-CS	-0.500	-0.545	0.009	0.010	-0.045	0.011	0.947	1	-0.413	0.013	0.013	0.087	0.020	0.869	0.969

β=treatment effect, Var($\hat{\beta}$)= empirical variance, E(V)=average variance,
CONV.=conventional model, PS adj.= Propensity score adjusted into the model, Subclass=subclassification,
Match.-CS= Matching under common support

Table 3.4: Simulation result of conditional model of 100 confounders

No. confounders=100		Subdistribution model								Cause-specific model						
Sample size		β	$\hat{\beta}$	V($\hat{\beta}$)	E(V)	bias	MSE	%cov.	power	$\hat{\beta}$	V($\hat{\beta}$)	E(V)	bias	MSE	%cov.	power
500	CONV.	0.000	0.003	0.040	0.028	0.003	0.040	0.891	0.109	0.001	0.036	0.030	0.001	0.036	0.926	0.074
	PS Adj.	0.000	-0.001	0.012	0.020	-0.001	0.012	0.983	0.017	0.012	0.013	0.021	0.012	0.013	0.988	0.012
	Subclass	0.000	-0.216	0.015	0.015	-0.216	0.061	0.578	0.422	0.011	0.013	0.022	0.011	0.013	0.991	0.009
	Matching	0.000	-0.223	0.026	0.025	-0.223	0.075	0.696	0.304	-0.020	0.029	0.030	-0.020	0.030	0.955	0.045
	Match.-CS	0.000	-0.185	0.027	0.028	-0.185	0.061	0.802	0.198	-0.009	0.033	0.034	-0.009	0.033	0.951	0.049
	CONV.	-0.500	-0.623	0.046	0.034	-0.123	0.061	0.865	0.906	-0.621	0.046	0.037	-0.121	0.060	0.881	0.885
	PS Adj.	-0.500	-0.308	0.014	0.023	0.192	0.051	0.800	0.535	-0.335	0.017	0.024	0.165	0.044	0.855	0.590
	Subclass	-0.500	-0.531	0.017	0.017	-0.031	0.018	0.947	0.989	-0.337	0.017	0.025	0.163	0.044	0.856	0.587
	Matching	-0.500	-0.544	0.029	0.028	-0.044	0.031	0.937	0.908	-0.374	0.035	0.034	0.126	0.051	0.896	0.540
	Match.-CS	-0.500	-0.515	0.029	0.032	-0.015	0.029	0.965	0.828	-0.371	0.038	0.039	0.129	0.055	0.889	0.472
1000	CONV.	0.000	0.000	0.012	0.010	0.000	0.012	0.917	0.083	-0.000	0.014	0.010	-0.000	0.014	0.908	0.092
	PS Adj.	0.000	-0.001	0.005	0.009	-0.001	0.005	0.987	0.013	0.009	0.007	0.009	0.009	0.007	0.975	0.025
	Subclass	0.000	-0.220	0.008	0.007	-0.220	0.056	0.268	0.732	0.006	0.006	0.009	0.006	0.007	0.978	0.022
	Matching	0.000	-0.234	0.016	0.014	-0.234	0.070	0.492	0.508	-0.018	0.016	0.017	-0.018	0.016	0.957	0.043
	Match.-CS	0.000	-0.218	0.016	0.015	-0.218	0.064	0.553	0.447	-0.013	0.018	0.018	-0.013	0.018	0.937	0.063
	CONV.	-0.500	-0.546	0.015	0.011	-0.046	0.017	0.886	0.999	-0.543	0.015	0.012	-0.043	0.016	0.907	0.995
	PS Adj.	-0.500	-0.306	0.006	0.010	0.194	0.044	0.515	0.916	-0.328	0.007	0.011	0.172	0.037	0.635	0.940
	Subclass	-0.500	-0.523	0.008	0.008	-0.023	0.009	0.942	1	-0.332	0.007	0.011	0.168	0.035	0.653	0.945
	Matching	-0.500	-0.544	0.016	0.016	-0.044	0.018	0.939	0.990	-0.374	0.017	0.020	0.126	0.033	0.868	0.780
	Match.-CS	-0.500	-0.542	0.016	0.017	-0.042	0.018	0.933	0.987	-0.365	0.020	0.020	0.135	0.038	0.844	0.731
2000	CONV.	0.000	-0.000	0.005	0.004	-0.000	0.005	0.932	0.068	0.007	0.005	0.005	0.007	0.005	0.924	0.076
	PS Adj.	0.000	0.000	0.002	0.004	0.000	0.002	0.989	0.011	0.012	0.003	0.004	0.012	0.003	0.980	0.020
	Subclass	0.000	-0.218	0.004	0.004	-0.218	0.051	0.048	0.952	0.010	0.003	0.004	0.010	0.003	0.978	0.022
	Matching	0.000	-0.230	0.008	0.008	-0.230	0.061	0.249	0.751	-0.011	0.010	0.009	-0.011	0.010	0.951	0.049
	Match.-CS	0.000	-0.238	0.008	0.008	-0.238	0.064	0.236	0.764	-0.014	0.009	0.009	-0.014	0.009	0.956	0.044
	CONV.	-0.500	-0.525	0.005	0.005	-0.025	0.006	0.923	1	-0.519	0.006	0.005	-0.019	0.007	0.926	1
	PS Adj.	-0.500	-0.312	0.003	0.005	0.188	0.038	0.151	1	-0.334	0.004	0.005	0.166	0.031	0.319	0.999
	Subclass	-0.500	-0.524	0.004	0.004	-0.024	0.004	0.942	1	-0.337	0.004	0.005	0.163	0.030	0.326	0.999
	Matching	-0.500	-0.544	0.008	0.009	-0.044	0.010	0.940	1	-0.376	0.011	0.011	0.124	0.026	0.756	0.959
	Match.-CS	-0.500	-0.568	0.009	0.009	-0.068	0.013	0.887	1	-0.378	0.011	0.011	0.122	0.026	0.774	0.952

β=treatment effect, Var($\hat{\beta}$)= empirical variance, E(V)=average variance,
CONV.=conventional model, PS adj.= Propensity score adjusted into the model, Subclass=subclassification,
Match.-CS= Matching under common support

3.4 Case study results

A total of 71,030 patients were included in the analysis, and 64.5% of them received statins after discharge. At the end of two years of follow up, the percentage of patients who had an MI, stroke or HF was 10.2%, 1.8%, 6.1% respectively, the percent who died was 17.0% (mortality was a competing risk for the CVD events).

We summarized the results from three different confounding control methods under the cause-specific hazard function and the subdistribution hazard function in Table 3.5. We were not able to apply the PS matching methods due to technical difficulties resulting from the large sample size. The results from both competing risk models suggested that statins treatment had no effect on the risks of MI, stroke and HF in this cohort. The reduction in the risk of all-cause mortality was substantial among the treated group. This may represent healthy user bias, namely, that a healthier or stronger patient is more likely to get treatment.

In summary, the conditional models suggest that statin use resulted in a large and statistically significant reduction in mortality, smaller marginally significant effects on HF and stroke, and a small and not statistically significant effect on MI. These results can be interpreted as the effect at the individual level conditioning on all the confounders.

Table 3.5: The estimated treatment effects of statins from conditional models

		subdistribution model			cause-specific model		
		$\hat{\beta}$	SE	p-value	$\hat{\beta}$	SE	p-value
MI	Conv.	-0.003	0.027	0.900	-0.021	0.027	0.440
	PS adj	0.012	0.028	0.660	-0.001	0.028	0.970
	subclass	-0.011	0.027	0.680	-0.031	0.027	0.260
all-cause	Conv.	-0.132	0.021	0.000	-0.130	0.021	0.000
mortality	PS adj	-0.115	0.021	0.000	-0.115	0.021	0.000
	subclass	-0.163	0.021	0.000	-0.166	0.021	0.000
Stroke	Conv.	-0.095	0.063	0.130	-0.112	0.063	0.075
	PS adj	-0.105	0.064	0.100	-0.120	0.064	0.060
	subclass	-0.112	0.063	0.076	-0.135	0.063	0.032
HF	Conv.	-0.048	0.035	0.170	-0.066	0.035	0.059
	PS adj	-0.052	0.036	0.140	-0.064	0.035	0.069
	subclass	-0.065	0.035	0.065	-0.084	0.035	0.016

β=treatment effect estimator, MI=myocardial infarction,

CONV.=conventional model,

PS adj.= Propensity score adjusted into the model,

Subclass=subclassification,

Match.-CS= Matching under common support

3.5 Discussion

In this study, we examined the performance of several PS methods in the estimation of the conditional treatment effect in competing risk analysis. The simulation studies showed that the covariate adjusted PS method performed the best under the null, and subclassification performed best under the non-null models, particularly for the subdistribution hazard function with heavy confounding.

Gayat et al (26) and Austin et al (6) examined performance of the PS used in the standard Cox model via Monte Carlo simulation. The number of confounders of both studies was small with only 9 variables. They also did not present the extent of overlap in PS distributions between the treated and untreated groups. Gayat et al. showed that the matched-robust-adjusted method gave an unbiased estimator, however, this approach used matched samples, and included confounders into the model. The estimator from this model was the treatment effects for the treated group. The estimators from the matched model without adjustment were poor. Austin et al. presented similar results but also a larger relative bias.

The conventional model was sensitive to sample size. When the sample size was small and there was a large number of confounding variables, the estimator was biased and had a large variance. In the real world applications of these methods to studies using administrative claims data, where the sample size is large and there are many confounding variables, we recommend that the PS approach be used to reduce the large number of confounding variables dimensions to one dimension.

The PS matching methods may not be applicable to big datasets. The optimal matching method needs to create a distance matrix with dimensions of treatment by the number of untreated patients and use the caliper to find the best matching group (79). This method requires a lot of time and computing power. In addition, the matching method has been mostly applied to studies with more untreated patients than treated patients. However, in studies where treatment is common, the treatment group may be

larger than the untreated group and in this case PS matching may result in a reduction of the sample size.

In conclusion, the PS method is a popular method to balance the different baseline characteristics between treatment groups in observational studies. We have shown through simulations that the PS methods can be incorporated into competing risk analyses. Recommendations for appropriate PS methods in specific situations have been provided.

Chapter 4

Application of Propensity Score Methods for Competing Risks Under Heterogeneity

4.1 Introduction

Cardiovascular diseases (CVD) are a major leading cause of mortality in the elderly worldwide (108). High levels of low-density lipoprotein cholesterol (LDL-c) are associated with an increased risk for CVD events and statins, a class of 3-hydroxy-3-methlglutaryl coenzyme A reductase inhibitors, are used to lower LDL-c levels, thus helping to prevent CVD events (100). Clinical trials have demonstrated that statins are highly effective at lowering LDL-c, which in turn yields great benefit in reducing the number of cardiovascular events (32).

The incidence of CVD rises steeply with age. In 2012, the rate of CVD in men ages 85-94 (7.4%) was more than 20 times that seen in 35-44 year of men (0.3%). Women followed the same pattern, but 10 years later in life than the men. Adults who were free from CVD had lifetime cumulative risks of developing CVD of 51.7% (men) and 39.0% (women). The factors related to recurrence of CVD included age, CVD burden and whether any prevention procedures were performed (109). People surviving an initial acute myocardial infarction (AMI) are 1.5 to 15 times more likely to suffer a repeat AMI or death, depending on the comorbidities present and their health status (27). This strong effect

The incidence of CVD rises steeply with age, in 2012, CVD rate increased from 0.3% at age 35-44 to 7.4% at age 85-94 in men, and schedule 10 years later for women (92).

The adults who were free from CVD illustrated the life time risks for developing CVD after 50 year of age with 51.7% in men and 39.2% in women. The factors related to recurrence of CVD included age, CVD burden and prevention procedure (109). The people who survived after acute myocardial infarction (AMI) are more likely to suffer and death 1.5 to 15 times more depending on clinical conditions (27). This strong effect of age on cardiovascular related events led to research on how to treat older patients and whether the treatment effects of statins seen in clinical trials in fact exist for the older population.

The effects of statins treatments in the elderly have been investigated, but most results come from small subgroups, thus are less statistically reliable (28). A meta-analysis of clinical studies showed the benefit of statins in reducing the risks of all-cause mortality, MI and stroke in an elderly population, but the most study included the patients who were free from CVD (87, 1). Also, little evidence is available on the treatment effects of statins for different (combinations of) condition. Due to the potential complexity of existing comorbidities in the elderly, investigation of treatment effects in this group poses several challenges including risk of competing events and confounding.

Decreasing rates of coronary heart disease (CHD) during the last two decades come from improvements in treatment approaches, which Coronary Artery Bypass Grafting (CABG) or stent or Percutaneous transluminal coronary angioplasty (PTCA) contributed 11% the decrease. Improvements in surgical techniques (including CABG or stent or PTCA) have improved the survival of MI patients compared with medical therapy alone (50). The patients who underwent CABG/stent/PTCA possibly lower risks of mortality and CHD events.

In the previous two chapters, we have shown that propensity score (PS) methods can be incorporated into competing risk models to evaluate marginal and conditional treatment effects in non-experimental studies. The approaches have advantages over conventional models in studies where a large number of confounders exist. However, the results assume that the treatment effects of statins are homogeneous. In this chapter,

where we are interested in estimating heterogeneous treatment effects, the PS methods based on competing risk model are deemed more appropriate.

Heterogeneity treatment effects are the variation in treatment effects among different populations (55). The inclusion criteria of clinical research reduces the heterogeneity treatment effects and the randomized treatment assignment results in an unbiased estimator. For the non-randomized study, an increase in sample size can reduce sampling variability but it may not lower heterogeneity (78). A method which can eliminate major sources of heterogeneity is important to a non-randomized study. The homogeneity analyzing by subgroup is the common approach to verify heterogeneity treatment effects.

Age and CABG/stent/PTCA during admission may cause the heterogeneity of statins's effects in elderly who AMI. The increasing frailty of old age is likely related to the risks and the benefits of statins (61). CABG/stent/PTCA procedures are the preferable medical methods to restore the blood flow to the heart by removing or bypassing the atherosclerosis that causes the blockages in coronary arteries. These procedures can reduce the risk of overall mortality and recurring MI (53) which may be related to the interaction of treatment effects between those procedures and statins. The baseline risks of patients who underwent CABG/stent/PTCA may be different from those who did not have those procedures done; the subgroup analysis resulted in a reduction of those differences.

The objective of this study is to investigate the treatment effect of statins in an elderly population which recently experienced an acute MI. The treatment effect across several age groups and CABG/stent/PTCA procedures were examined to glean information of the treatment effects, which hopefully can inform the management and prevention of cardiovascular events in the elderly population.

4.2 Method

4.2.1 Source of data and population

We identified a cohort of Medicare beneficiaries who just had a hospitalization stay for AMI in 2008. AMI was identified from Medicare inpatient claims files using relevant International Classification of Diseases, Ninth Revision, Clinical Medication (ICD-9-CM) codes (410.01, 410.11, 410.21, 410.31, 410.41, 410.51, 410.61, 410.71, 410.81, 410.91). Eligible patients included in the cohort were at least 66 years of age at the index date, living in the United States and had been continuously enrolled in Medicare Part A, B and D for at least one year prior. The exposure of interest was statin use after discharge.

4.2.2 Outcome and covariates

These patients were followed up from the 31st day after discharge until the occurrence of the first cardiovascular event (recurrent MI, stroke, heart failure) or death. We created a number of covariates for demographic characteristics and clinical conditions based on claims occurring in the 12-month baseline period prior to statin initiation. These covariates were identified a priori based on the literature, substantive knowledge, and the availability of covariates within the data. The variables included demographic characteristics age and gender, Charlson comorbidity index (CCI), baseline and current comorbidities, baseline and current medications, diagnoses and procedures for admission, and medical procedures during hospital stay (details in Section 1.4).

4.2.3 Statistical analysis

To determine if the treatment effect of statin is heterogeneous across age groups and patients with or without cardiovascular related procedures during hospital stay for AMI, we conducted the aforementioned analyses for subgroups. For age, the groups were 66 to 74, 75 to 84, and 85 an older. For the procedure subgroup analysis, patients were grouped together if they had CABG, stent, or PTCA during the hospital stay. The comparison

group consisted of patients who did not have any of these procedures.

To estimate the marginal and conditional effect of statin treatment on the risk of cardiovascular events (recurrent MI, heart failure, and stroke) as a risk in competition with mortality, we conducted a competing risk analysis with appropriate PS methods to control for confounding. The competing risk models based on subdistribution hazard function and cause-specific hazard function were used. For the marginal treatment effect, we used a crude model, weighted regression with inverse probability treatment weighting (IPTW) (76) and stabilized treatment weighting (STW) (17). For the conditional treatment effect, we used a conventional model, covariate adjustment using PS and a subclassification model. The R statistical computing software was used for all analyses. The R library crr , crrc and crrs were used for subdistribution hazard function based Fine-Gray model (21). The R library coxph was used to estimate the treatment effect from cause-specific hazard function based Cox model (103).

4.3 Results

The age distribution of the patients was: 35.1% ages 66-74 years, 40.3% ages 75-84 years, and 24.6% age 85 years and above. Older age groups showed a greater percentage of patients - the group aged 85 years and older accounted for 40% of total mortality. Non-statin users ages 66-74 years had twice the risk of dying as did statin users in the same age group, and 25% of non-statin users aged 85 and older died.

The patients who established hospitalized AMI received CABG/stent/PTCA 42.9%. 75.7% of the patients receiving CABG/stent/PTCA during hospitalization were prescribed statins, but only 56% of patients who did not undergo a revascularization procedure were prescribed statins. Approximately one-fifth of patients not undergoing CABG/stent/PTCA who were not prescribed statins died, compared with only 5% of patients receiving both revascularization and statins. Nearly 14% of patients not undergoing revascularization had heart failure, whether they were prescribed statins or not.

Table 4.1: Number and percent of observed events

	Survived	Death	MI	Stroke	HF
	Number(%)	Number(%)	Number(%)	Number(%)	Number(%)
statins					
age 66-74	14609 (73.51)	1545 (7.77)	1481 (7.45)	447 (2.25)	1791 (9.01)
age 75-84	9932 (67.25)	1549 (10.49)	1267 (8.58)	401 (2.72)	1619 (10.96)
age > 84	3985 (53.84)	1246 (16.83)	911 (12.31)	225 (3.04)	1035 (13.98)
non-Vasc.	12293 (56.36)	3341 (15.32)	2508 (11.50)	616 (2.82)	3055 (14.01)
Vasc.	17561 (79.58)	1152 (5.22)	1274 (5.77)	499 (2.26)	1580 (7.16)
Non-statins					
age 66-74	5680 (61.83)	1236 (13.45)	864 (9.40)	275 (2.99)	1132 (12.32)
age 75-84	4638 (57.25)	1347 (16.63)	766 (9.45)	266 (3.28)	1085 (13.39)
age > 84	2610 (43.61)	1598 (26.70)	689 (11.51)	165 (2.76)	923 (15.42)
non-Vasc.	8369 (49.04)	3801 (22.27)	1831 (10.73)	535 (3.14)	2529 (14.82)
Vasc.	5083 (71.83)	515 (7.28)	563 (7.96)	196 (2.77)	719 (10.16)

4.3.1 Propensity Score model

A multiple logistic regression model was used to estimate propensity score for each age group and each procedure group. The distribution of PS for treated and untreated groups were different for age groups and revascularization group (Figure 4.1 and Figure 4.2). The discrimination of distribution between statins users and non-statins users increased by age groups. The shape of the density function of statins and non-statin group differed among patients who underwent CABG/stent/PTCA during admission and patients who did not undergo those procedures.

The important variables that predicted the probability to be prescribed statins for all age groups were baseline statins, beta blocker users and admission for CABG or stent or PTCA. The patients who had CCI for moderate to severe liver disease reduced the probability of statin in elderly less than 85 years old. Patients ages 75 to 84 years who had hyperkalemia at baseline were more likely to be prescribed statins, as were patients

with rhabdomyolysis in the 85+ age group.

The important variables that predicted the probability of receiving statins for the patients who did not have CABG/stent/PTCA at admission were statin or beta blocker users at baseline, and the CCI for moderate to severe liver disease. For patients who received CABG/stent/PTCA, the important variables that predicted the probability of being prescribed statins were the CCI for diabetes, the CCI for uncomplicated diabetes, a baseline CCI score greater than 9, ACEI, baseline statins and beta blocker users.

4.3.2 Treatment effects of statins for different age groups

The marginal model presented in the previous chapter showed that statins reduced the risk of mortality and heart failure in the elderly patient who recently experience AMI. The treatment effect of statins from the conditional model was associated only with a reduction of mortality hazard. The treatment effect of statins varied by age group - for patients aged 66 to 74 years, statins were associated with a reduction of the risk of heart failure, but this was not the case for the other age groups. The treatment effect of statins with respect to the risk of heart failure was the lowest for those age 85+, whereas the treatment effect of statins with respect to the risk of mortality was the least in the younger age groups.

The estimator from marginal model for mortality and heart failure outcome differed in the conditional model for the 66-74 age group. The weighted model from both subdistribution hazard model and cause-specific model showed a non-significant effect for statins on reduce the risks of mortality but patients who prescribed statins showed significantly lower risks of mortality in the the conventional model and subclassification model. In contrast, the average treatment effects of statins significantly reduced the risk of heart failure but conditional model showed the opposite results.

Then, the average treatment effects of statins reduced the risk of heart failure in patients age 66-74 with a recent AMI but this results did not apply at the individual level. The individual who was prescribed statins has a similar risk of a cardiovascular

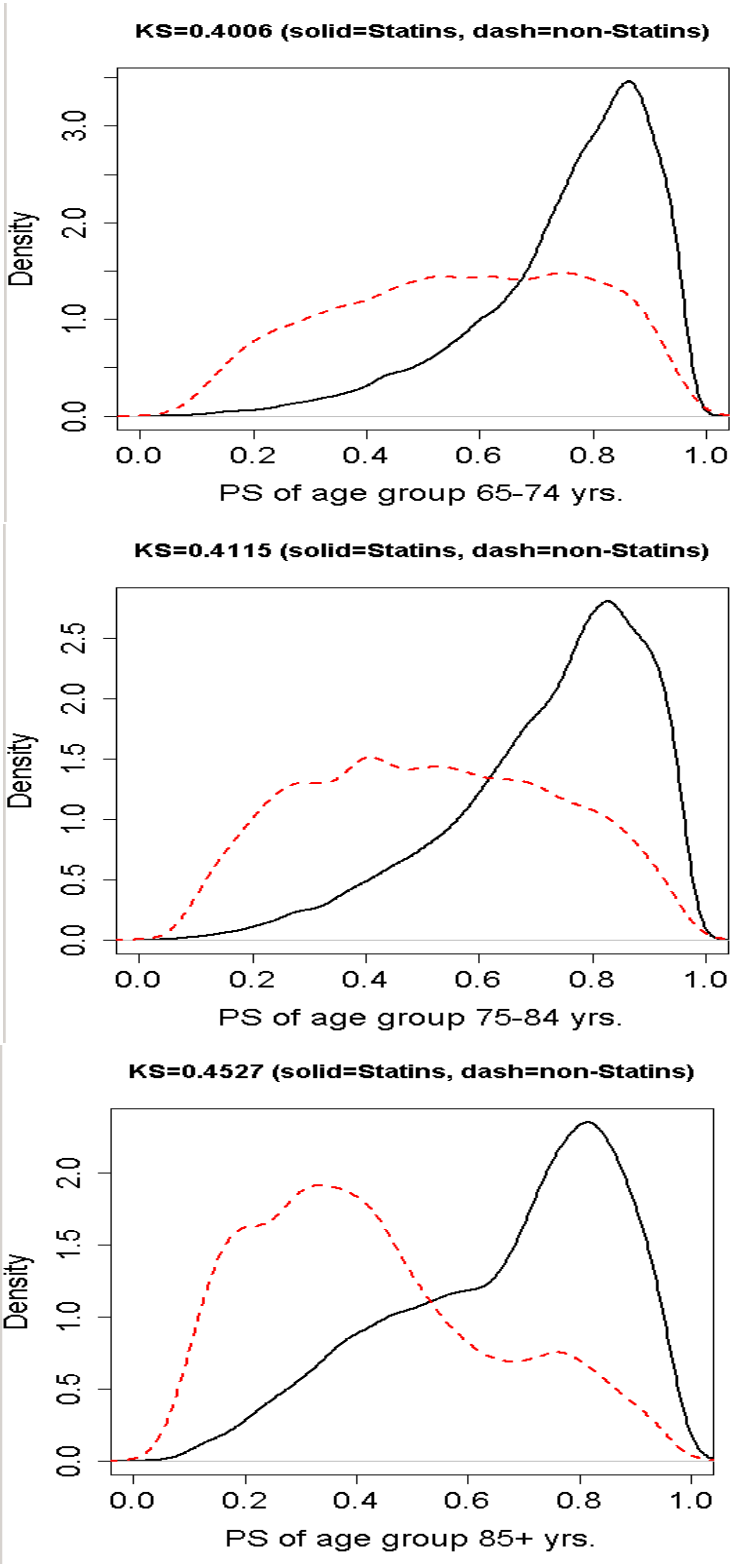

Figure 4.1: Distribution of propensity score of statins users and non-users by age group– 66-74 yrs. (above), 75-84 yrs. (middle), at least 85 (below)

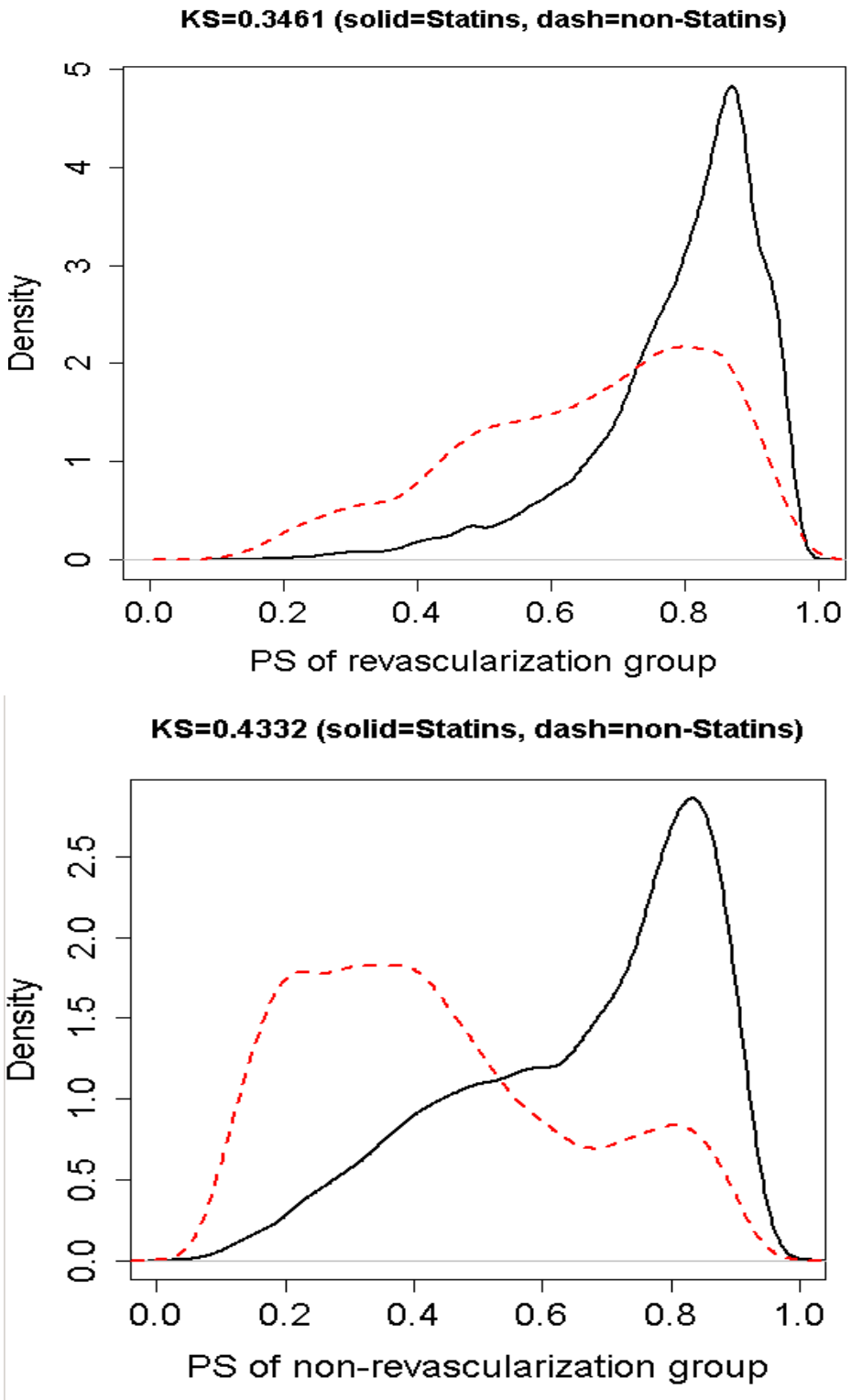

Figure 4.2: Distribution of propensity score of statins user and non-user by revasucularization status

event as the individual not prescribed statins and was as likely to be prescribed statins.

In summary, statins showed less of an effect on the risk of CVD, MI, stroke or heart failure in the older age groups. Healthier older patients may have had a greater chance of being prescribed statins, thus the statins group showed a lower risk of mortality, especially for the older patients.

Table 4.2: Results of treatment effects from the marginal, conditional risks and competing risks models, by age group

	age 66-75 yrs.						age75-84 yrs.						age $\geq$ 85 yrs.					
	SH			CSH			SH			CSH			SH			CSH		
	$\hat{\beta}$	SE	p	$\hat{\beta}$	SE	p	$\hat{\beta}$	SE	p	$\hat{\beta}$	SE	p	$\hat{\beta}$	SE	p	$\hat{\beta}$	SE	p
Mortality																		
Crude	-0.537	0.039	0.000	-0.563	0.039	0.000	-0.503	0.029	0.000	-0.528	0.029	0.000	-0.406	0.029	0.000	-0.427	0.029	0.000
Crude-CS	-0.536	0.039	0.000	-0.563	0.039	0.000	-0.502	0.029	0.000	-0.528	0.029	0.000	-0.040	0.029	0.000	-0.425	0.029	0.000
IPTW*	0.003	0.028	0.909	0.008	0.028	0.786	-0.060	0.021	0.004	-0.060	0.021	0.004	0.058	0.022	0.009	-0.060	0.023	0.008
STW*	-0.029	0.043	0.505	0.027	0.043	0.526	-0.124	0.029	0.000	-0.126	0.029	0.000	-0.184	0.034	0.000	-0.193	0.027	0.000
Conv.	-0.120	0.044	0.006	-0.096	0.044	0.030	-0.140	0.033	0.000	-0.154	0.033	0.000	-0.124	0.034	0.000	-0.118	0.034	0.000
PS-Adj.	-0.078	0.045	0.084	-0.070	0.045	0.120	-0.122	0.034	0.000	-0.121	0.034	0.000	-0.115	0.034	0.001	-0.121	0.033	0.000
Subclass.	-0.152	0.043	0.000	-0.152	0.044	0.000	-0.173	0.033	0.000	-0.176	0.033	0.000	-0.145	0.033	0.000	-0.155	0.033	0.000
Non-Fetal MI																		
Crude	-0.256	0.046	0.000	-0.302	0.046	0.000	-0.102	0.040	0.010	-0.166	0.040	0.000	0.013	0.041	0.750	-0.062	0.041	0.130
Crude-CS	-0.256	0.046	0.000	-0.302	0.046	0.000	-0.103	0.040	0.009	-0.166	0.040	0.000	0.015	0.041	0.710	-0.060	0.041	0.150
IPTW*	-0.016	0.032	0.616	-0.011	0.032	0.736	0.025	0.027	0.359	0.019	0.027	0.476	0.091	0.029	0.001	0.076	0.029	0.008
STW*	-0.053	0.047	0.259	-0.051	0.047	0.276	0.009	0.040	0.822	-0.003	0.040	0.941	0.060	0.049	0.222	0.024	0.049	0.624
Conv.	-0.057	0.051	0.265	-0.061	0.052	0.240	-0.008	0.045	0.866	-0.033	0.044	0.457	0.050	0.047	0.293	0.031	0.047	0.513
PS-Adj.	-0.036	0.052	0.490	-0.036	0.052	0.490	0.010	0.045	0.820	-0.003	0.045	0.950	0.053	0.047	0.260	0.028	0.047	0.550
Subclass.	-0.081	0.051	0.110	-0.090	0.051	0.077	-0.010	0.044	0.820	-0.030	0.044	0.500	0.038	0.047	0.430	0.006	0.047	0.890
Stroke																		
Crude	-0.212	0.095	0.026	-0.279	0.095	0.003	-0.105	0.087	0.220	-0.188	0.087	0.030	0.001	0.120	0.990	-0.097	0.120	0.420
Crude-CS	-0.212	0.095	0.026	-0.279	0.095	0.003	-0.106	0.087	0.220	-0.188	0.087	0.030	-0.005	0.120	0.970	-0.103	0.120	0.390
IPTW*	-0.102	0.067	0.129	-0.097	0.067	0.150	-0.079	0.064	0.216	-0.083	0.063	0.191	0.007	0.083	0.936	-0.004	0.084	0.966
STW*	-0.121	0.092	0.186	-0.122	0.092	0.184	-0.119	0.082	0.144	-0.132	0.082	0.107	0.028	0.144	0.847	-0.014	0.108	0.893
Conv.	-0.128	0.106	0.225	-0.136	0.106	0.200	-0.126	0.097	0.194	-0.142	0.097	0.142	-0.048	0.137	0.723	-0.075	0.136	0.582
PS-Adj.	-0.138	0.107	0.200	-0.142	0.108	0.190	-0.129	0.098	0.180	-0.144	0.097	0.140	-0.054	0.138	0.700	-0.079	0.137	0.560
Subclass.	-0.143	0.106	0.180	-0.158	0.106	0.140	-0.156	0.096	0.110	-0.178	0.096	0.065	-0.044	0.137	0.750	-0.082	0.136	0.550
Heart Failure																		
Crude	-0.175	0.057	0.002	-0.229	0.057	0.000	-0.093	0.048	0.052	-0.164	0.048	0.001	0.025	0.061	0.680	-0.061	0.061	0.310
Crude-CS	-0.176	0.057	0.002	-0.230	0.057	0.000	-0.091	0.048	0.056	-0.163	0.048	0.001	0.027	0.061	0.660	-0.583	0.061	0.340
IPTW*	-0.093	0.036	0.010	-0.086	0.036	0.018	-0.036	0.033	0.276	-0.039	0.033	0.234	-0.087	0.044	0.051	-0.095	0.044	0.032
STW*	-0.108	0.052	0.039	-0.105	0.052	0.045	-0.066	0.045	0.140	-0.076	0.045	0.089	-0.023	0.075	0.759	-0.057	0.052	0.277
Conv.	-0.064	0.063	0.312	-0.070	0.064	0.272	-0.033	0.053	0.534	-0.057	0.053	0.282	-0.075	0.069	0.277	-0.098	0.069	0.156
PS-Adj.	-0.059	0.064	0.360	-0.061	0.065	0.350	-0.035	0.054	0.510	-0.048	0.054	0.380	-0.084	0.069	0.220	-0.104	0.069	0.130
Subclass.	-0.082	0.063	0.200	-0.092	0.063	0.140	-0.044	0.053	0.410	-0.064	0.053	0.230	-0.079	0.070	0.260	-0.109	0.069	0.120

*SE estimated from bootstrap method, SH=subdistribution model, CSH=cause-specific model

4.3.3 Treatment effects of statins for patients who had a revascularization procedure

The effects of statins for patients who underwent CABG/stent/ptca procedures at admission were different from those of patients who did not undergo CABG/stent/PTCA. The treatment effects of statins of patients with those procedure showed significant reductions in the risks of recurring MI, heart failure and all cause mortality. However, prescriptions of statins to patients of this group did not significantly reduce the risks of stroke. The estimator from the cause-specific model were larger than the estimator from the subdistribution hazard function.

Prescribing statins to non-CABG/stent/ptca patients did not reduce the risks of MI, stroke and heart failure but it significantly reduced the risks of all-cause mortality. However, the estimator of marginal model from the subdistribution hazard model showed a significant increasing the risk of MI for patients from non-CABG/stent/PTCA group. The results of significant level between marginal and conditional were the same direction except the estimator of the non-CABG/stent/PTCA group for the risk of MI. The results from marginal model Fine-Gray model and Cox PH model were different when the outcome was MI in the non-CABG/stent/PTCA group.

In conjunction with CABG/stent/PTCA, prescribing statins improved the effects of medical treatment for recurrent MI and heart failure, compared with only prescribing statins. The association of statins with a reduced risk of death was demonstrated in both groups. These results may have healthy patient bias effects. The Cox PH model showed larger treatment effects than did the Fine-Gray model.

4.4 Discussion

Statins did not reduce the risks of CVD for those aged 75 years or more but do provide a reduction in the risk of all-cause mortality. A Canadian observational study of patients ages 66-85 years with heart failure showed an effect of statins on mortality and stroke,

Table 4.3: Treatment effects of statins by status of revascularization procedure

	CABG/stent/PTCA						non-CABG/stent/PTCA					
	SH			CSH			SH			CSH		
	$\hat{\beta}$	SE	p	$\hat{\beta}$	SE	p	$\hat{\beta}$	SE	p	$\hat{\beta}$	SE	p
Mortality												
Crude	-0.395	0.043	0.000	-0.427	0.043	0.000	-0.338	0.021	0.000	-0.343	0.021	0.000
Crude-CS	-0.394	0.043	0.000	-0.425	0.043	0.000	-0.338	0.021	0.000	-0.343	0.021	0.000
IPTW*	-0.133	0.032	0.000	-0.141	0.032	0.000	-0.060	0.014	0.000	-0.056	0.014	0.000
STW*	-0.095	0.050	0.060	-0.101	0.050	0.045	-0.175	0.019	0.000	-0.171	0.019	0.000
Conv.	-0.186	0.046	0.000	-0.210	0.046	0.000	-0.107	0.023	0.000	-0.099	0.023	0.000
PS-Adj.	-0.170	0.047	0.000	-0.181	0.047	0.000	-0.100	0.023	0.000	-0.095	0.023	0.000
Subclass.	-0.205	0.046	0.000	-0.222	0.046	0.000	-0.127	0.023	0.000	-0.124	0.023	0.000
Non-Fetal MI												
Crude	-0.323	0.047	0.000	-0.357	0.047	0.000	0.088	0.028	0.002	0.035	0.028	0.220
Crude-CS	-0.322	0.047	0.000	-0.356	0.047	0.000	0.088	0.028	0.002	0.035	0.028	0.220
IPTW*	-0.093	0.033	0.005	-0.102	0.033	0.002	0.044	0.020	0.030	0.037	0.020	0.063
STW*	-0.098	0.050	0.048	-0.105	0.050	0.036	0.067	0.027	0.013	0.044	0.027	0.102
Conv.	-0.144	0.051	0.005	-0.174	0.051	0.001	0.059	0.032	0.069	0.047	0.032	0.150
PS-Adj.	-0.128	0.051	0.013	-0.141	0.051	0.006	0.061	0.032	0.060	0.048	0.032	0.140
Subclass.	-0.158	0.050	0.002	-0.175	0.050	0.001	0.060	0.032	0.063	0.042	0.032	0.190
Stroke												
Crude	-0.161	0.089	0.073	-0.214	0.089	0.017	-0.099	0.075	0.190	-0.158	0.075	0.035
Crude-CS	-0.155	0.090	0.084	-0.208	0.090	0.020	-0.099	0.075	0.190	-0.158	0.075	0.035
IPTW*	-0.060	0.061	0.322	-0.074	0.061	0.224	-0.106	0.053	0.045	-0.111	0.053	0.035
STW*	-0.081	0.096	0.397	-0.092	0.096	0.338	-0.088	0.065	0.178	-0.111	0.065	0.090
Conv.	-0.088	0.097	0.362	-0.115	0.096	0.232	-0.116	0.085	0.174	-0.127	0.085	0.135
PS-Adj.	-0.082	0.097	0.400	-0.104	0.097	0.290	-0.125	0.086	0.150	-0.135	0.085	0.110
Subclass.	-0.100	0.096	0.300	-0.128	0.096	0.180	-0.129	0.085	0.130	-0.147	0.085	0.084
Heart Failure												
Crude	-0.251	0.053	0.000	-0.290	0.053	0.000	0.053	0.039	0.180	-0.002	0.039	0.950
Crude-CS	-0.251	0.053	0.000	-0.290	0.053	0.000	0.054	0.039	0.170	-0.002	0.039	0.970
IPTW*	-0.132	0.036	0.000	-0.141	0.036	0.000	-0.044	0.028	0.114	-0.047	0.028	0.092
STW*	-0.117	0.057	0.039	-0.123	0.057	0.030	0.017	0.034	0.609	-0.004	0.033	0.904
Conv.	-0.125	0.057	0.028	-0.155	0.057	0.006	-0.023	0.045	0.610	-0.033	0.045	0.457
PS-Adj.	-0.123	0.058	0.033	-0.138	0.058	0.017	-0.023	0.045	0.600	-0.033	0.045	0.460
Subclass.	-0.151	0.057	0.008	-0.171	0.057	0.003	-0.037	0.045	0.400	-0.053	0.044	0.240

*SE estimated from bootstrap method, SH=subdistribution model, CSH=cause-specific model

but not for MI (12). Younger age groups (less than 85 years old) showed a greater effect of statins than did those over 84 years of age. A meta-analysis in patients (maximum age 82) who did not have cardiovascular disease but had cardiovascular risk factors found that statins reduce the risk of death, major coronary events and major cerebrovascular events (12).

Statins were associated with a reduced risk of mortality in all age groups but older elderly showed a greater reduction in mortality risk than did the younger elderly. The effect of statins on mortality has been documented by many studies but with contradictory results. Observational studies of patients with heart failure found that statins can reduce risks of mortality (69, 24, 45, 22). In contrast, the meta-analysis of randomized controlled trials of subjects without CVD at baseline did not show the same effects of statins on all cause mortality (74). A long-term follow-up study of CHD patients or high risks in CHD patients found that statins significantly reduced mortality from cardiovascular disease but it did not reduce non-cardiovascular mortality (60). The difference in medical history of patients affected the treatment effects of statins on mortality.

Statins are associated with a reduced risk of death in the oldest age group, but show non-significant treatment effects on the risk of MI, stroke and heart failure. Moreover, the treatment effects are reversed for MI. For the oldest age group, statins appear to operate using a different prediction model. The oldest elderly were vulnerable and frail. Optimal medication use in the elderly must take into account both the benefit and risks, especially for patients with a short life expectancy (44). In addition, the elderly often have several comorbidities and polymedication, therefore the standard guidelines may not be suitable for all elderly patient (93). Although statins are associated with the prevention of CVD in the elderly, they also may cause myopathy (41). Lee and colleagues (58) found that intermediate acute coronary syndrome patients aged 81+ years were less likely to be prescribed secondary prevention medication such as ARB, ACE or statins. Whether statins are prescribed depends on the patient's medical condition and the decision of their physician. However, statins are recommended for the elderly, particularly the healthy

elderly with more years of life left who may benefit more from statin' effects (28).

Although the effects of statins in patients who underwent CABG/stent/PTCA procedure have been shown, adherence to statins after a procedure was low compared to revascularization procedures alone (49). Prescribing statins tend to be performed more in younger age groups (almost 80% of those 75 years of age or older are not routinely treated), demonstrating a healthy bias effect.

The heterogeneity of statins treatment effects can be seen in this study, since age and CABG/stent/PTCA procedures modify the effect of statins. Overall treatment effects under heterogeneity need the weighted methods. The standardized method including IPTW allows the summary estimation of treatment effects in certain populations (97).

Limitation

Including both new and recurring MI into the cohort causes the different probabilities of recurrence of MI (68). Clinical data, such as the size and the area of infarction, were not included in the model. Underlying cause of death cannot be reliably linked to CHD. The dataset constructed for this analysis did not include the initiation date nor the length of use of statins. Discontinuing use of statins after discharge showed strong harmful effects in patients, which were not seen in patients not prescribed statins (20).

Chapter 5

Conclusion

This study investigated the efficiency of estimators when propensity score methods were used to reduce or eliminate the effect of confounding in competing risks survival analysis. We investigated the treatment effects estimator from both marginal and conditional models. The PS methods for estimating marginal treatment effects included a crude model and weighted model. The subclassification model, matching and PS were included in a multivariate model; the matching model used the PS methods to estimate conditional treatment effects. This study also applied a conventional model which included all confounding into the model to compare efficiency of the conventional method to PS methods. The competing risks models of this study included a cause-specific hazard model and a subdistribution hazard model.

The PS methods showed the good performance for estimate marginal treatment effect when applied to competing risks analysis. The weighted model showed a small amount of bias in the MSE, and a high percent of coverage. However, we found the inflated variance of estimator in the scenario of low occurrence of the interesting event and the heavy confounder. The bootstrapping methods applied to obtain the variance estimator appropriately take advantage of the estimation of variance for weighed models. A future research topic for the estimator from the marginal model will be the investigation of the behavior of the inflated variance from the weighted model in different scenarios. Another topic for future research is the development of a better method to estimate the variance of the weighted model.

The estimator from PS methods also showed the good performance for conditional model especially given the large amount of confounding present. PS adjusted into multivariate competing risk model showed good performance for null model, which illustrated high percent coverage, small bias and MSE for all scenarios. For alternative model, when the PS model included heavy confounders, the subclassification model showed the smallest MSE of subdistribution hazard model. However, subclassification showed a larger bias compared to conventional model. For the cause-specific hazard model, the conventional model showed a small bias, small MSE and high percent coverage. Under Monte Carlo simulation, the confounder variables were independent, but may have been correlated. The performance of PS methods under certain level of correlation between confounder variables should to be investigated in the future.

We presented the case study of applying PS methods to estimate treatment effects with a competing risks model. The Medicare claims data included a cohort of patients who were recently hospitalized for AMI in 2008. The treatment effect of statins in this population was investigated. The outcomes of interest were MI, stroke, heart failure and all cause mortality. We believe that the estimators from weighted model showed the best marginal estimator. The estimators from PS, added to the multivariate model, produced a good estimator for null model, while the good estimators for the alternative model of subdistribution hazard model and the cause-specific model were the subclassification model and the conventional model respectively. The marginal estimator from the weighted model demonstrated the benefit of statins in reducing hazard of heart failure and all cause mortality in this population. However, the conditional model showed the effect of statins of all-cause mortality.

The dataset for this hospitalized elderly contains the variation contributed by individuals and their medical condition. The heterogeneity treatment effects may show in the results for the different age and medical procedure group. The heterogeneity treatment effect represents the interaction between treatment effects and individual patient characteristics. However, the PS methods cannot include the interaction between effect

and confounding factors needed to investigate heterogeneity of treatment. The subgroup analysis is the considering tool to estimate treatment effect when the heterogeneity treatment effect possibly present.

The results showed a heterogeneity of treatment effects by age group and procedure of CABG/stent/PTCA. The younger elderly showed a greater effect of statins on the secondary prevention of heart failure than was seen in older elderly. The patients who underwent CABG/stent/PTCA showed more of an effect of statins on MI and heart failure than those patients who were only prescribed statins. However, the results of heterogeneity treatment effect may be related to the healthy bias effect. The younger elderly and the patients who had CABG/stent/PTCA were stronger and healthier than other groups.

The appropriate PS model showed good performance in controlling confounder in competing risks survival analysis. PS methods produced good estimators for both subdistribution and cause-specific hazard models. When the heterogeneity treatment effects may be present, subgroup analysis should be performed to investigate the treatment effects for each subgroup.

APPENDIX: Propensity Score model

Table 5.1: Propensity score model

Variables	Estimation	SE	P-value
intercept	-1.094	0.092	0.000
Age ×comobidity	0.027	0.016	0.099
gender	0.013	0.020	0.516
Age 65-74 yrs.			
Age 75-84 yrs.	-0.045	0.022	0.045
Age 85 yrs. and more	-0.118	0.029	0.000
White			
Black	0.042	0.034	0.207
Asian	0.072	0.056	0.201
Hispanic	0.205	0.071	0.004
Other race	0.216	0.071	0.002
Income less than $30,001			
Income $30,001-$60,000	0.105	0.073	0.141
Income $60,001-$100,000	0.168	0.074	0.024
Income $100,001-$150,000	0.201	0.083	0.015
Income more than $150,000	0.211	0.106	0.046
Charlson Comorbidity index			
Acute Myocaridal Infarction	0.108	0.045	0.016
Cerebrovascular Disease	-0.012	0.031	0.687
Congentive heart failure	0.011	0.031	0.758
Periphral vascular disease	0.032	0.030	0.291
Diabetes	-0.036	0.111	0.747
Renal disease	0.147	0.039	0.000
Chronic Obstructive Pulmonary disease	0.095	0.036	0.009
Peptic Ulcer disease	0.064	0.072	0.371
Cancer	-0.007	0.039	0.860
Dementia	0.003	0.055	0.956
Connective Tissue disease Rheumatic disease	0.008	0.051	0.877
Mild liver disease	-0.160	0.071	0.024
Moderate to severe liver disease	-0.578	0.213	0.007
Paralysis	0.079	0.097	0.414
Metastatic Carsinoma	-0.024	0.103	0.813
AIDS/HIV	-0.033	0.374	0.931
Diabetes without complication	0.000	0.109	0.999
Diabetes with complications	0.036	0.042	0.384
Baseline disease			
Stent	-0.002	0.162	0.510
PTCA	-0.101	0.112	0.365
Unstable angina	0.034	4.020	0.391
Ischemic heart disease	-0.334	0.023	0.000
Atrial fibrillation	-0.065	0.030	0.031

Variables	Estimation	SE	P-value
Hypertension	0.006	0.025	0.794
Hyperlipidemia	-0.221	0.022	0.000
End-stage renal disease	0.027	0.088	0.760
Osteroporosis	0.011	0.036	0.756
Asthma	0.098	0.042	0.019
Angiodedma&hyperkalemia	0.084	0.198	0.672
Hypotension	-0.015	0.043	0.730
Sinus bradycardia&heart block	-0.012	0.026	0.647
Rhabdomyolysis	-0.234	0.138	0.090
Hyperkalemia	0.021	0.204	0.917
Baseline CCI=0	-0.080	0.035	0.023
Baseline CCI=1-2	-0.171	0.038	0.000
Baseline CCI=3-5	-0.233	0.063	0.000
Baseline CCI=6-8	-0.305	0.100	0.002
Baseline CCI more than 9	-0.387	0.146	0.008
Baseline medication			
Baseline Beta blocker	-0.230	0.020	0.000
Baseline ACEI/ARB	-0.157	0.021	0.000
Baseline STATINS	1.659	0.022	0.000
BaselineCoronary Artery bypass grafting	0.076	0.116	0.510
Baseline STENT/PTCA	-0.127	0.190	0.503
Admission procedure/diagnosis			
Subendocardial infarction	-0.104	0.022	0.000
Congestive heart failure	-0.094	0.021	0.000
Cardiogenic shock	0.097	0.064	0.129
Acute renal failure	0.029	0.029	0.314
Hypotension	0.047	0.041	0.248
Cardiac dysrhythmias	-0.062	0.020	0.002
Cardiac catheterization	0.118	0.036	0.001
CABG	0.682	0.043	0.000
PTCA	-0.142	0.518	0.784
STENT/PTCA	0.655	0.518	0.206
Angiocardiography	0.162	0.035	0.000
Thromblytics and patelet inhibitors	-0.029	0.121	0.811
Platelet inhibitors	0.177	0.046	0.000
Number of days in ICU=0			
Number of days in ICU=1-3	0.044	0.022	0.042
Number of days in ICU=4-10	0.015	0.027	0.579
Number of days in ICU ¿10	-0.016	0.064	0.807
Number of coronary care unit=0			
Number of coronary care unit= 1-3	0.129	0.024	0.000
Number of coronary care unit= 4-10	0.113	0.030	0.000
Number of coronary care unit ¿10	0.057	0.077	0.457
Total number of days in hospital			

Variables	Estimation	SE	P-value
1 day in hospital	-0.091	0.053	0.087
2-5 days in hospital	0.048	0.032	0.130
6-10 days in hospital	-0.045	0.031	0.141
More than 10 days in hospital	-0.102	0.049	0.039
Acute respiratory failure/	-0.061	0.032	0.056
Mechventilation in AMI admission			
Septic shock in AMI admission	-0.161	0.129	0.212
Physician visit during follow-up	0.143	0.023	0.000
Cardiologist visit during follow-up	0.120	0.021	0.000
Revascularization procedure during follow-up	0.127	0.037	0.001
Number of hospital admission in baseline	-0.027	0.018	0.143
Number of days in hospital in baseline	0.001	0.001	0.656
Indicator for hospital admission in baseline	0.044	0.033	0.184
Number of admission to short-term	-0.004	0.029	0.893
Acute care hospital during follow up			
Number of days to short term	-0.009	0.004	0.033
Acute care hospital during follow up			
Medicare DOUGNUT	-0.012	0.030	0.691
Low comorbidity level			
Moderate comorbidity level	-0.122	0.046	0.008
High comorbidity level	-0.095	0.076	0.211
Current Beta blockers	0.987	0.022	0.000
Current ACEI/ARB	0.319	0.081	0.000
CurrentACEI	0.342	0.081	0.000
CurrentARB	0.186	0.076	0.014
Current comorbidity			
Valvular disease or rheumatic heart disease	-0.010	0.023	0.667
Other neurological disorders	-0.030	0.029	0.304
Obesity	-0.013	0.041	0.746
Coagulation deficiency	-0.028	0.041	0.486
Weight loss	-0.021	0.036	0.560
Fluid/electrolyte disorder	-0.054	0.025	0.033
Substance abuse	-0.025	0.072	0.731
Blood loss&deficiency anemia	-0.087	0.022	0.000
Hypothyroidism	-0.041	0.022	0.063
Pulmonary Circ. Disorders	-0.034	0.039	0.379
Osteoarthritis	-0.026	0.020	0.210
GI bleed	-0.031	0.037	0.396
Parkinson's disease	-0.045	0.071	0.522
Weakness	0.041	0.046	0.371
Vertigo	0.010	0.026	0.708
Fall/difficulty walking	0.036	0.028	0.193
Bladder dysfunction	0.006	0.030	0.848
Decubitus	-0.032	0.039	0.412

Variables	Estimation	SE	P-value
Use of Oxygen	-0.031	0.033	0.348
Use of hospital bed	0.055	0.057	0.336
Use of ambulance	0.019	0.023	0.398
Nail care	0.025	0.027	0.345
Use of other assistive devices	-0.011	0.045	0.815
Use of screening	0.029	0.023	0.196
Use of wheelchair	0.053	0.042	0.207
Use of rehabilitation	-0.009	0.026	0.721

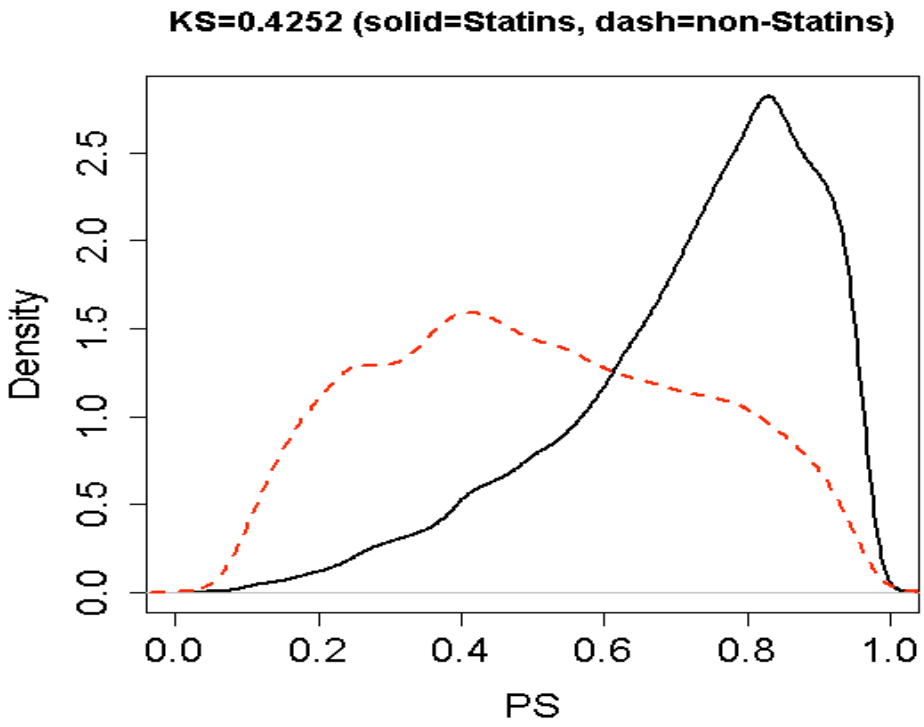

Figure 5.1: Distribution of the propensity score of statins users and non-users by revascularization procedure group

Table 5.2: Propensity score model by age groups

	age 66-74 yrs.			age 75-84 yrs.			age 85 and older		
Variables	Estimation	SE	P-value	Estimation	SE	P-value	Estimation	SE	P-value
Intercept	-0.490	0.479	0.307	-0.901	0.427	0.035	-0.029	0.501	0.954
Gender	0.000	0.033	0.999	0.024	0.031	0.430	-0.002	0.041	0.969
Age	-0.005	0.006	0.396	-0.323	0.005	0.522	-0.015	0.005	0.003
White									
Black	0.083	0.055	0.132	0.004	0.054	0.939	0.035	0.071	0.620
Asian	-0.044	0.112	0.693	0.111	0.082	0.179	0.092	0.106	0.386
Hispanic	0.066	0.132	0.616	0.244	0.109	0.025	0.286	0.134	0.032
Other race	0.133	0.112	0.234	0.285	0.118	0.016	0.280	0.149	0.060
Income less than $30,001									
Income $30,001-$60,000	0.151	0.120	0.211	0.103	0.118	0.382	0.058	0.149	0.696
Income $60,001-$100,000	0.186	0.123	0.131	0.189	0.120	0.113	0.112	0.151	0.458
Income $100,001-$150,000	0.137	0.142	0.337	0.231	0.133	0.082	0.203	0.164	0.215
Income more than $150,000	0.088	0.198	0.657	0.323	0.169	0.056	0.166	0.198	0.402
Charlson Comorbidity index									
Acute Myocaridal Infarction	0.170	0.079	0.031	0.070	0.072	0.326	0.029	0.087	0.737
Cerebrovascular Disease	0.042	0.056	0.453	-0.036	0.047	0.437	-0.051	0.062	0.404
Congentive heart failure	0.005	0.066	0.943	0.066	0.058	0.257	-0.073	0.074	0.318
Periphral vascular disease	-0.012	0.054	0.828	0.028	0.047	0.548	0.084	0.061	0.166
Diabetes	0.101	0.190	0.596	-0.385	0.176	0.029	0.288	0.215	0.181
Renal disease	0.188	0.070	0.007	0.094	0.061	0.121	0.155	0.078	0.047
Chronic Obstructive Pulmonary disease	0.116	0.063	0.065	0.120	0.056	0.032	0.005	0.077	0.944
Peptic Ulcer disease	-0.116	0.128	0.363	0.184	0.111	0.099	0.086	0.142	0.546
Cancer	-0.052	0.070	0.458	0.003	0.059	0.964	0.004	0.081	0.965
Dementia	-0.024	0.149	0.869	0.120	0.088	0.170	-0.052	0.088	0.550
Connective Tissue disease Rheumatic disease	-0.022	0.087	0.800	-0.070	0.077	0.363	0.194	0.109	0.076
Mild liver disease	-0.167	0.108	0.122	-0.142	0.114	0.211	-0.251	0.171	0.143
Moderate to severe liver disease	-0.692	0.319	0.030	-0.617	0.332	0.063	-0.271	0.593	0.647
Paralysis	0.197	0.162	0.234	0.038	0.154	0.805	-0.199	0.203	0.325
Metastatic Carsinoma	0.199	0.173	0.250	-0.121	0.156	0.439	-0.247	0.232	0.287
AIDS/HIV	0.049	0.421	0.908	-0.211	0.884	0.812			
Diabetes without complication	-0.142	0.187	0.449	0.364	0.174	0.036	-0.337	0.212	0.112
Diabetes with complications	-0.028	0.070	0.686	0.135	0.065	0.038	-0.046	0.091	0.611
Baseline disease									
BaselineCoronary Artery bypass grafting	0.185	0.161	0.251	0.072	0.179	0.688	0.006	0.519	0.990
Stent	0.022	0.263	0.932	0.002	0.236	0.993	0.060	0.421	0.888
PTCA	-0.006	0.160	0.970	-0.213	0.176	0.227	-0.171	0.349	0.623
Unstable angina	-0.039	0.067	0.557	0.142	0.063	0.025	-0.052	0.085	0.539
Ischemic heart disease	-0.370	0.041	0.000	-0.380	0.036	0.000	-0.212	0.044	0.000
Atrial fibrillation	-0.014	0.062	0.821	-0.090	0.046	0.054	-0.056	0.054	0.297
Hypertension	-0.054	0.043	0.209	0.047	0.039	0.228	0.087	0.048	0.069
Hyperlipidemia	-0.313	0.040	0.000	-0.194	0.035	0.000	-0.168	0.045	0.000
End-stage renal disease	0.007	0.135	0.960	0.003	0.135	0.980	0.138	0.222	0.533
Osteroporosis	-0.031	0.077	0.683	0.054	0.056	0.328	-0.003	0.060	0.959

	age 66-74 yrs.			age 75-84 yrs.			age 85 and older		
Variables	Estimation	SE	P-value	Estimation	SE	P-value	Estimation	SE	P-value
Asthma	0.134	0.069	0.052	0.136	0.066	0.038	0.003	0.090	0.976
Angiodedma&hyperkalemia	0.314	0.336	0.351	-0.270	0.303	0.373	0.300	0.426	0.481
Hypotension	-0.017	0.078	0.823	0.083	0.067	0.216	-0.155	0.081	0.056
Sinus bradycardia&heart block	-0.082	0.051	0.107	0.041	0.041	0.316	-0.040	0.048	0.400
Rhabdomyolysis	-0.081	0.255	0.751	-0.155	0.228	0.496	-0.497	0.246	0.043
Hyperkalemia	-0.273	0.349	0.434	0.503	0.313	0.108	-0.313	0.437	0.474
Baseline CCI=0	-0.079	0.060	0.188	-0.055	0.056	0.327	-0.059	0.068	0.387
Baseline CCI=1-2	-0.247	0.066	0.000	-0.111	0.059	0.061	-0.137	0.076	0.070
Baseline CCI=3-5	-0.265	0.109	0.016	-0.201	0.097	0.038	-0.202	0.129	0.119
Baseline CCI=6-8	-0.337	0.174	0.053	-0.298	0.154	0.054	-0.233	0.209	0.265
Baseline CCI more than 9	-0.412	0.251	0.101	-0.408	0.225	0.070	-0.265	0.311	0.393
Baseline medication									
Baseline Beta blocker	-0.268	0.036	0.000	-0.240	0.032	0.000	-0.157	0.039	0.000
Baseline ACEI/ARB	-0.119	0.037	0.001	-0.197	0.032	0.000	-0.119	0.040	0.003
Baseline STATINS	1.483	0.037	0.000	1.652	0.034	0.000	1.901	0.044	0.000
Baseline STENT/PTCA	-0.218	0.298	0.464	-0.029	0.284	0.919	0.032	0.530	0.952
Admission procedure/diagnosis									
Subendocardial infarction	-0.167	0.037	0.000	-0.057	0.035	0.100	-0.100	0.045	0.025
Congestive heart failure	0.010	0.038	0.804	-0.100	0.032	0.002	-0.167	0.038	0.000
Cardiogenic shock	0.002	0.096	0.985	0.122	0.102	0.232	0.250	0.158	0.114
Acute renal failure	0.003	0.052	0.954	0.067	0.044	0.132	0.015	0.055	0.790
Hypotension	0.045	0.072	0.534	-0.016	0.064	0.800	0.162	0.082	0.047
Cardiac dysrhythmias	-0.059	0.037	0.113	-0.050	0.032	0.118	-0.080	0.039	0.043
Cardiac catheterization	0.181	0.056	0.001	0.021	0.055	0.704	0.127	0.095	0.181
CABG	0.674	0.064	0.000	0.693	0.067	0.000	0.611	0.160	0.000
PTCA	-0.153	0.830	0.854	-0.214	0.670	0.749	-8.110	119.468	0.946
Angiocardiography	0.036	0.054	0.503	0.223	0.053	0.000	0.247	0.094	0.009
Thromblytics and patelet inhibitors	0.078	0.189	0.680	-0.180	0.184	0.327	0.134	0.311	0.665
Platelet inhibitors	0.040	0.068	0.551	0.243	0.072	0.001	0.443	0.134	0.001
Number of days in ICU=0									
Number of days in ICU=1-3	-0.028	0.037	0.461	0.070	0.034	0.040	0.081	0.043	0.061
Number of days in ICU=4-10	-0.032	0.049	0.520	0.006	0.042	0.880	0.053	0.052	0.038
Number of days in ICU ¿10	0.017	0.107	0.875	-0.021	0.096	0.825	-0.151	0.151	0.317
Number of coronary care unit=0									
Number of coronary care unit= 1-3	0.094	0.040	0.018	0.123	0.038	0.001	0.168	0.050	0.001
Number of coronary care unit= 4-10	0.018	0.053	0.734	0.175	0.046	0.000	0.098	0.059	0.096
Number of coronary care unit ¿10	-0.044	0.125	0.728	0.070	0.115	0.543	0.184	0.188	0.329
STENT/PTCA	0.634	0.830	0.445	0.723	0.670	0.280			
Total number of days in hospital									
1 day in hospital	-0.072	0.089	0.421	-0.083	0.085	0.329	-0.139	0.111	0.210
2-5 days in hospital	0.027	0.060	0.654	0.029	0.051	0.561	0.106	0.058	0.069
6-10 days in hospital	-0.094	0.057	0.097	-0.031	0.049	0.529	-0.022	0.057	0.699
More than 10 days in hospital	-0.203	0.089	0.022	-0.057	0.077	0.460	-0.096	0.099	0.334
Acute respiratory failure/									

	age 66-74 yrs.			age 75-84 yrs.			age 85 and older		
Variables	Estimation	SE	P-value	Estimation	SE	P-value	Estimation	SE	P-value
Mechventilation in AMI admission	-0.011	0.052	0.836	-0.116	0.049	0.018	-0.097	0.070	0.164
Septic shock in AMI admission	0.170	0.214	0.429	-0.356	0.193	0.064	-0.506	0.324	0.118
Physician visit during follow-up	0.191	0.042	0.000	0.123	0.038	0.001	0.119	0.044	0.006
Cardiologist visit during follow-up	0.091	0.036	0.012	0.129	0.033	0.000	0.127	0.044	0.004
Revascularization procudure during follow-up	0.055	0.056	0.325	0.135	0.058	0.020	0.264	0.099	0.007
Number of hospital admission in baseline	-0.024	0.031	0.447	-0.043	0.028	0.124	-0.011	0.040	0.779
Number of days in hospital in baseline	0.001	0.003	0.765	0.004	0.002	0.064	-0.004	0.003	0.148
Indicator for hospital admission in baseline	0.116	0.058	0.045	0.007	0.052	0.897	0.035	0.070	0.623
Number of admission to short-term acute care hospital during follow up	-0.037	0.051	0.471	-0.019	0.045	0.676	0.048	0.058	0.412
Number of days to short term acute care hospital during follow up	-0.003	0.008	0.747	-0.012	0.007	0.080	-0.014	0.009	0.099
DOUGNUT	-0.029	0.051	0.572	-0.002	0.047	0.961	0.002	0.063	0.976
Low comorbidity level									
Moderate comorbidity level	-0.067	0.060	0.265	-0.047	0.052	0.361	-0.078	0.062	0.211
High comorbidity level	-0.047	0.072	0.515	-0.026	0.062	0.672	0.133	0.081	0.100
Current Beta blockers	1.113	0.037	0.000	0.921	0.034	0.000	0.952	0.043	0.000
Current ACEI/ARB	0.324	0.133	0.015	0.291	0.135	0.031	0.389	0.163	0.017
CurrentACEI	0.312	0.128	0.015	0.436	0.131	0.001	0.210	0.158	0.185
CurrentARB	0.199	0.121	0.102	0.284	0.127	0.025	0.010	0.153	0.947
Current comorbidity									
Valvular disease or rheumatic heart disease	0.028	0.043	0.510	-0.050	0.036	0.160	-0.008	0.043	0.853
Other neurological disorders	-0.016	0.056	0.769	-0.025	0.045	0.576	-0.043	0.053	0.414
Obesity	0.064	0.057	0.259	-0.106	0.066	0.111	-0.055	0.134	0.683
Coagulation deficiency	-0.055	0.075	0.468	-0.015	0.062	0.810	-0.043	0.079	0.587
Weight loss	-0.023	0.072	0.755	-0.073	0.056	0.191	0.021	0.065	0.744
Fluid/electrolyte disorder	-0.113	0.047	0.016	-0.047	0.040	0.233	-0.002	0.048	0.975
Substance abuse	-0.018	0.099	0.855	-0.127	0.124	0.307	-0.093	0.213	0.661
Blood loss&deficiency anemia	-0.065	0.042	0.122	-0.077	0.035	0.027	-0.116	0.041	0.005
Hypothyroidism	-0.051	0.042	0.221	0.011	0.035	0.751	-0.109	0.041	0.008
Pulmonary Circ. Disorders	-0.123	0.071	0.082	-0.047	0.060	0.439	0.056	0.074	0.444
Osteoarthritis	-0.069	0.037	0.062	0.004	0.032	0.895	-0.019	0.039	0.627
GI bleed	0.008	0.069	0.912	-0.086	0.056	0.124	0.004	0.070	0.954
Use of screening	0.050	0.041	0.227	-0.024	0.037	0.511	0.044	0.042	0.301
Use of wheelchair	-0.028	0.082	0.733	0.049	0.069	0.480	0.121	0.072	0.090
Parkinson's disease	-0.094	0.150	0.529	-0.046	0.105	0.662	-0.052	0.130	0.692
Use of rehabilitation	-0.072	0.047	0.125	0.075	0.041	0.067	-0.072	0.051	0.162
Weakness	0.200	0.095	0.035	-0.020	0.072	0.785	0.027	0.080	0.737
Vertigo	-0.078	0.048	0.103	0.045	0.040	0.259	0.035	0.048	0.464
Fall/difficulty walking	0.082	0.055	0.135	0.008	0.043	0.857	0.047	0.048	0.332
Bladder dysfunction	-0.040	0.059	0.501	-0.010	0.046	0.833	0.075	0.056	0.178
Decubitus	0.010	0.074	0.897	-0.053	0.062	0.400	-0.031	0.070	0.659
Use of Oxygen	0.013	0.057	0.815	-0.053	0.051	0.298	-0.077	0.068	0.256
Use of hospital bed	0.126	0.111	0.259	0.146	0.093	0.114	-0.052	0.098	0.597
Use of ambulance	0.090	0.041	0.030	-0.027	0.035	0.444	-0.015	0.043	0.723

	age 66-74 yrs.			age 75-84 yrs.			age 85 and older		
Variables	Estimation	SE	P-value	Estimation	SE	P-value	Estimation	SE	P-value
Nail care	0.035	0.057	0.545	-0.015	0.042	0.714	0.052	0.046	0.255
Use of other assistive devices	-0.014	0.090	0.878	-0.128	0.071	0.051	0.125	0.081	0.121

Table 5.3: Propensity score model by revascularization procedure group

	CABG/stent/PTCA			No CABG/stent/PTCA		
Variables	Estimator	SE	p-value	Estimator	SE	p-value
Intercept	0.139	0.205	0.498	-1.368	0.116	0.000
Gender	-0.043	0.032	0.170	0.034	0.026	0.185
Age 66-74 yrs.						
Age 75-84 yrs.	-0.047	0.032	0.138	-0.017	0.030	0.556
Age 85+ yrs.	-0.038	0.049	0.435	-0.052	0.034	0.119
White						
Black	-0.037	0.062	0.555	0.100	0.041	0.013
Asian	0.086	0.098	0.383	0.091	0.069	0.191
Hispanic	0.322	0.130	0.013	0.169	0.086	0.051
Other race	0.362	0.129	0.005	0.160	0.087	0.066
Income less than $30,001						
Income $30,001-$60,000	0.007	0.125	0.952	0.168	0.092	0.068
Income $60,001-$100,000	0.078	0.127	0.538	0.229	0.094	0.014
Income $100,001-$150,000	0.094	0.141	0.503	0.269	0.104	0.010
Income more than $150,000	0.237	0.176	0.177	0.199	0.137	0.145
Charlson Comorbidity index						
Acute Myocaridal Infarction	0.104	0.087	0.234	0.059	0.053	0.262
Cerebrovascular Disease	-0.031	0.056	0.573	-0.011	0.037	0.777
Congentive heart failure	0.049	0.071	0.485	-0.043	0.044	0.329
Periphral vascular disease	0.053	0.055	0.340	0.017	0.037	0.648
Diabetes	-0.619	0.181	0.001	0.268	0.138	0.052
Renal disease	0.123	0.075	0.101	0.131	0.047	0.005
Chronic Obstructive Pulmonary disease	0.178	0.066	0.007	0.055	0.044	0.219
Peptic Ulcer disease	0.156	0.137	0.253	0.007	0.086	0.939
Cancer	0.077	0.070	0.270	-0.061	0.049	0.212
Dementia	0.207	0.140	0.139	-0.036	0.062	0.563
Connective Tissue disease Rheumatic disease	0.040	0.086	0.638	-0.034	0.065	0.602
Mild liver disease	-0.213	0.121	0.079	-0.132	0.089	0.135
Moderate to severe liver disease	-0.272	0.461	0.556	-0.679	0.247	0.006
Paralysis	0.021	0.206	0.919	0.071	0.111	0.523
Metastatic Carsinoma	0.421	0.207	0.043	-0.149	0.122	0.222
AIDS/HIV	0.354	0.651	0.587	-0.213	0.462	0.644
Diabetes without complication	0.570	0.179	0.001	-0.301	0.136	0.027
Diabetes with complications	0.080	0.076	0.293	0.001	0.051	0.987
Baseline disease						
BaselineCoronary Artery bypass grafting	-0.055	0.176	0.756	0.232	0.154	0.133
Stent	-0.025	0.228	0.914	0.069	0.226	0.760
PTCA	-0.214	0.136	0.115	0.206	0.212	0.332
Unstable angina	0.065	0.066	0.322	-0.006	0.051	0.903
Ischemic heart disease	-0.495	0.038	0.000	-0.234	0.029	0.000
Atrial fibrillation	0.086	0.059	0.144	-0.108	0.036	0.003
Hypertension	0.049	0.039	0.210	-0.003	0.032	0.932
Hyperlipidemia	-0.425	0.037	0.000	-0.111	0.029	0.000
End-stage renal disease	-0.061	0.167	0.714	0.064	0.103	0.537
Osteroporosis	-0.010	0.065	0.877	0.022	0.043	0.616
Baseline CCI=0	-0.037	0.052	0.475	-0.082	0.049	0.090
Baseline CCI=1-2	-0.163	0.063	0.009	-0.128	0.049	0.010
Baseline CCI=3-5	-0.295	0.113	0.009	-0.144	0.078	0.063
Baseline CCI=6-8	-0.468	0.188	0.013	-0.194	0.121	0.110
Baseline CCI more than 9	-0.635	0.279	0.023	-0.246	0.176	0.161
Baseline Beta blocker	-0.246	0.033	0.000	-0.204	0.026	0.000
Baseline ACEI/ARB	-0.089	0.034	0.008	-0.185	0.027	0.000
Baseline STATINS	1.211	0.035	0.000	1.896	0.027	0.000
Baseline STENT/PTCA	0.003	0.256	0.990	-0.323	0.302	0.285
Asthma	0.073	0.076	0.337	0.093	0.051	0.068
Baseline medication						
Angiodedma&hyperkalemia	-0.313	0.330	0.343	0.262	0.247	0.289
Hypotension	0.017	0.084	0.836	-0.048	0.050	0.338
Sinus bradycardia&heart block	-0.065	0.048	0.176	0.009	0.032	0.783
Rhabdomyolysis	0.106	0.287	0.712	-0.362	0.164	0.027
Hyperkalemia	0.270	0.346	0.436	-0.106	0.254	0.678
Admission procedure/diagnosis						
CABG	-0.013	0.119	0.914			
PTCA	-0.143	0.517	0.782			
Stent/PTCA	-0.078	0.531	0.883			

	CABG/stent/PTCA			No CABG/stent/PTCA		
Variables	Estimator	SE	p-value	Estimator	SE	p-value
Subendocardial infarction	-0.155	0.032	0.000	-0.029	0.031	0.346
Congestive heart failure	-0.103	0.037	0.006	-0.088	0.025	0.000
Cardiogenic shock	0.081	0.080	0.309	-0.012	0.113	0.914
Acute renal failure	-0.006	0.053	0.910	0.055	0.034	0.107
Hypotension	0.046	0.064	0.477	0.037	0.054	0.493
Cardiac dysrhythmias	-0.029	0.034	0.382	-0.087	0.026	0.001
Cardiac catheterization	0.007	0.046	0.883	0.035	0.091	0.703
Angiocardiography	0.044	0.043	0.307	0.259	0.091	0.004
Thromblytics and patelet inhibitors	-0.164	0.161	0.307	0.088	0.184	0.633
Platelet inhibitors	0.035	0.055	0.525	0.425	0.092	0.000
Number of days in ICU=0						
Number of days in ICU=1-3	-0.030	0.035	0.387	0.069	0.028	0.015
Number of days in ICU=4-10	-0.002	0.048	0.968	0.000	0.033	0.993
Number of days in ICU ¿10	-0.086	0.097	0.379	-0.003	0.088	0.976
Number of coronary care unit=0						
Number of coronary care unit= 1-3	0.076	0.036	0.034	0.137	0.033	0.000
Number of coronary care unit= 4-10	0.080	0.049	0.100	0.105	0.038	0.006
Number of coronary care unit ¿10	-0.106	0.106	0.319	0.249	0.116	0.032
Total number of days in hospital						
1 day in hospital	0.068	0.090	0.451	-0.197	0.068	0.004
2-5 days in hospital	0.044	0.060	0.459	0.045	0.039	0.245
6-10 days in hospital	-0.080	0.056	0.155	-0.038	0.038	0.317
More than 10 days in hospital	-0.090	0.087	0.299	-0.117	0.062	0.054
Acute respiratory failure/						
Mechventilation in AMI admission	-0.024	0.054	0.659	-0.076	0.040	0.056
Septic shock in AMI admission	-0.081	0.290	0.781	-0.168	0.146	0.250
Physician visit during follow-up	0.245	0.041	0.000	0.096	0.029	0.001
Cardiologist visit during follow-up	0.034	0.034	0.321	0.147	0.028	0.000
Revascularization procudure						
during follow-up	-0.002	0.044	0.960	0.412	0.071	0.000
Number of hospital admission in baseline	0.028	0.041	0.491	-0.050	0.021	0.017
Number of days in hospital in baseline	0.002	0.003	0.507	0.001	0.001	0.595
Indicator for hospital admission in baseline	0.058	0.063	0.353	0.041	0.041	0.319
Number of admission to short-term						
acute care hospital during follow up	0.005	0.047	0.913	-0.028	0.037	0.449
Number of days to short term acute						
care hospital during follow up	-0.024	0.007	0.001	-0.003	0.006	0.600
DOUGNUT	-0.076	0.053	0.158	0.016	0.037	0.663
Low comorbidity level						
Moderate comorbidity level	-0.076	0.053	0.158	-0.072	0.040	0.071
HIgh comorbidity level	-0.092	0.077	0.231	0.022	0.048	0.650
Current Beta blockers	1.128	0.036	0.000	0.896	0.027	0.000
Current ACEI/ARB	0.065	0.143	0.648	0.476	0.103	0.000
CurrentACEI	0.600	0.139	0.000	0.185	0.099	0.063
CurrentARB	0.482	0.134	0.000	0.004	0.096	0.964
Current comorbidity						
Valvular disease or						
Rheumatic heart disease	-0.021	0.041	0.620	-0.008	0.028	0.765
Other neurological disorders	-0.037	0.056	0.510	-0.015	0.034	0.654
Obesity	-0.056	0.065	0.387	0.028	0.053	0.599
Coagulation deficiency	0.029	0.080	0.715	-0.048	0.048	0.315
Weight loss	0.010	0.073	0.895	-0.018	0.043	0.677
Fluid/electrolyte disorder	0.059	0.048	0.213	-0.096	0.031	0.002
Substance abuse	-0.218	0.134	0.105	0.084	0.086	0.330
Blood loss&deficiency anemia	-0.052	0.040	0.193	-0.094	0.027	0.001
Hypothyroidism	-0.047	0.038	0.220	-0.038	0.028	0.175
Pulmonary Circ. Disorders	-0.099	0.078	0.200	-0.020	0.045	0.659
Osteoarthritis	-0.048	0.034	0.159	-0.014	0.026	0.598
GI bleed	-0.021	0.071	0.768	-0.044	0.044	0.316
Use of screening	0.045	0.039	0.246	0.004	0.029	0.901
Use of wheelchair	0.108	0.097	0.268	0.044	0.048	0.353
Parkinson's disease	-0.069	0.137	0.618	-0.038	0.084	0.648
Use of rehabilitation	-0.017	0.045	0.695	-0.004	0.033	0.898
Weakness	0.124	0.096	0.198	0.026	0.053	0.630
Vertigo	-0.002	0.044	0.973	0.017	0.032	0.592
Fall/difficulty walking	0.003	0.052	0.954	0.056	0.033	0.091
Bladder dysfunction	-0.035	0.053	0.512	0.037	0.037	0.326

	CABG/stent/PTCA			No CABG/stent/PTCA		
Variables	Estimator	SE	p-value	Estimator	SE	p-value
Decubitus	-0.121	0.077	0.114	-0.005	0.046	0.913
Use of Oxygen	-0.016	0.065	0.809	-0.021	0.039	0.581
Use of hospital bed	0.145	0.146	0.321	0.052	0.063	0.411
Use of ambulance	0.060	0.040	0.130	0.001	0.028	0.975
Nail care	0.067	0.052	0.197	0.131	0.032	0.683
Use of other assistive devices	-0.156	0.092	0.091	0.029	0.053	0.588

BIBLIOGRAPHY

[1] J. Afilalo, G. Duque, R. Steele, J. W. Jukema, A. J. de Craen, and M. J. Eisenberg. Statins for secondary prevention in elderly patients: A hierarchical bayesian meta-analysis. *Journal of the American College of Cardiology*, 51(1):37–45, 2008.

[2] P. C. Austin. Goodness-of-fit diagnostics for the propensity score model when estimating treatment effects using covariate adjustment with the propensity score. *Pharmacoepidemiology and drug safety*, 17(12):1202–1217, 2008.

[3] P. C. Austin. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate behavioral research*, 46(3):399–424, 2011.

[4] P. C. Austin. The performance of different propensity score methods for estimating marginal hazard ratios. *Statistics in medicine*, 2012.

[5] P. C. Austin, P. Grootendorst, and G. M. Anderson. A comparison of the ability of different propensity score models to balance measured variables between treated and untreated subjects: a monte carlo study. *Statistics in medicine*, 26(4):734–753, 2007.

[6] P. C. Austin, P. Grootendorst, S.-L. T. Normand, and G. M. Anderson. Conditioning on the propensity score can result in biased estimation of common measures of treatment effect: a monte carlo study. *Statistics in medicine*, 26(4):754–768, 2007.

[7] R. J. BELLOTO. On statins, strokes, meta-analyses, competing risks, and the onward march of science. *The Annals of pharmacotherapy*, 41(12):2055–2057, 2007.

[8] R. Bender, T. Augustin, and M. Blettner. Generating survival times to simulate cox proportional hazards models. *Statistics in medicine*, 24(11):1713–1723, 2005.

[9] J. Beyersmann, A. Latouche, A. Buchholz, and M. Schumacher. Simulating competing risks data in survival analysis. *Statistics in Medicine*, 28(6):956–971, 2009.

[10] M. A. Brookhart, S. Schneeweiss, K. J. Rothman, R. J. Glynn, J. Avorn, and T. Stürmer. Variable selection for propensity score models. *American journal of epidemiology*, 163(12):1149–1156, 2006.

[11] M. A. Brookhart, R. Wyss, J. B. Layton, and T. Stürmer. Propensity score methods for confounding control in nonexperimental research. *Circulation: Cardiovascular Quality and Outcomes*, 6(5):604–611, 2013.

[12] J. Brugts, T. Yetgin, S. Hoeks, A. Gotto, J. Shepherd, R. Westendorp, A. de Craen, R. Knopp, H. Nakamura, P. Ridker, et al. The benefits of statins in people without established cardiovascular disease but with cardiovascular risk factors: meta-analysis of randomised controlled trials. *BMJ: British Medical Journal*, 338, 2009.

[13] M. Caliendo and S. Kopeinig. Some practical guidance for the implementation of propensity score matching. 22(1):31–72, 2008.

[14] M. S. Cepeda, R. Boston, J. T. Farrar, and B. L. Strom. Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. *American Journal of Epidemiology*, 158(3):280–287, 2003.

[15] S. R. Cole and C. E. Frangakis. The consistency statement in causal inference: a definition or an assumption? *Epidemiology*, 20(1):3–5, 2009.

[16] S. R. Cole and M. A. Hern ' an. Adjusted survival curves with inverse probability weights. *Computer methods and programs in biomedicine*, 75(1):45–49, 2004.

[17] S. R. Cole and M. A. Hernán. Constructing inverse probability weights for marginal structural models. *American journal of epidemiology*, 168(6):656–664, 2008.

[18] N. V. Cuong. Which covariates should be controlled in propensity score matching? evidence from a simulation study. *Statistica Neerlandica*, 2011.

[19] R. B. d'Agostino. Tutorial in biostatistics: propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med*, 17(19):2265–2281, 1998.

[20] S. S. Daskalopoulou, J. A. Delaney, K. B. Filion, J. M. Brophy, N. E. Mayo, and S. Suissa. Discontinuation of statin therapy following an acute myocardial infarction: a population-based study. *European heart journal*, 29(17):2083–2091, 2008.

[21] J. P. Fine and R. J. Gray. A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*, 94(446):496–509, 1999.

[22] R. J. Folkeringa, D. J. Van Kraaij, R. G. Tieleman, F. H. Nieman, Y. M. Pinto, and H. J. Crijns. Statins associated with reduced mortality in patients admitted for congestive heart failure. *Journal of cardiac failure*, 12(2):134–138, 2006.

[23] G. C. Fonarow. Statin therapy after acute myocardial infarction: Are we adequately treating high-risk patients? 4(2):99–106, 2002.

[24] H. Fukuta, D. C. Sane, S. Brucks, and W. C. Little. Statin therapy may be associated with lower mortality in patients with diastolic heart failure a preliminary report. *Circulation*, 112(3):357–363, 2005.

[25] M. H. Gail, S. Wieand, and S. Piantadosi. Biased estimates of treatment effect in randomized experiments with nonlinear regressions and omitted covariates. *Biometrika*, 71(3):431–444, 1984.

[26] E. Gayat, M. Resche-Rigon, J.-Y. Mary, and R. Porcher. Propensity score applied to survival data analysis through proportional hazards models: a monte carlo study. *Pharmaceutical Statistics*, 11(3):222–229, 2012.

[27] A. S. Go, D. Mozaffarian, V. L. Roger, E. J. Benjamin, J. D. Berry, M. J. Blaha, S. Dai, E. S. Ford, C. S. Fox, S. Franco, et al. Heart disease and stroke statistics—2014 update a report from the american heart association. *Circulation*, 129(3):e28–e292, 2014.

[28] A. M. Gotto. Statin therapy and the elderly sage advice? *Circulation*, 115(6):681–683, 2007.

[29] B. Gray. *cmprsk: Subdistribution Analysis of Competing Risks*, 2013. R package version 2.2-6.

[30] R. J. Gray. A class of k-sample tests for comparing the cumulative incidence of a competing risk. *The Annals of statistics*, pages 1141–1154, 1988.

[31] S. Greenland, J. M. Robins, J. Pearl, et al. Confounding and collapsibility in causal inference. *Statistical Science*, 14(1):29–46, 1999.

[32] L. S. Group et al. The lipid research clinics coronary primary prevention trial results. *JAMA*, 251:351–364, 1984.

[33] X. S. Gu and P. R. Rosenbaum. Comparison of multivariate matching methods: Structures, distances, and algorithms. *Journal of Computational and Graphical Statistics*, 2(4):405–420, 1993.

[34] J. Guo and Z. Geng. Collapsibility of logistic regression coefficients. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 263–267, 1995.

[35] E. M. Hade and B. Lu. Bias associated with using the estimated propensity score as a regression covariate. *Statistics in medicine*, 33(1):74–87, 2014.

[36] J. Hahn. On the role of the propensity score in efficient semiparametric estimation of average treatment effects. *Econometrica*, pages 315–331, 1998.

[37] B. B. Hansen. Full matching in an observational study of coaching for the sat. *Journal of the American Statistical Association*, 99(467):609–618, 2004.

[38] B. B. Hansen and S. O. Klopfer. Optimal full matching and related designs via network flows. *Journal of Computational and Graphical Statistics*, 15(3):609–627, 2006.

[39] S. Hennessy. Use of health care databases in pharmacoepidemiology. *Basic clinical pharmacology and toxicology*, 98(3):311–313, 2006.

[40] M. Á. Hernán, B. Brumback, and J. M. Robins. Marginal structural models to estimate the causal effect of zidovudine on the survival of hiv-positive men. *Epidemiology*, 11(5):561–570, 2000.

[41] S. Hilmer and D. Gnjidic. Statins in older adults. *Australian Prescriber*, 36(3), 2013.

[42] K. Hirano, G. W. Imbens, and G. Ridder. Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, 71(4):1161–1189, 2003.

[43] P. W. Holland. Statistics and causal inference. *Journal of the American statistical Association*, 81(396):945–960, 1986.

[44] H. Holmes. Rational prescribing for patients with a reduced life expectancy. *Clinical pharmacology & Therapeutics*, 85(1):103–107, 2008.

[45] T. B. Horwich, W. R. MacLellan, and G. C. Fonarow. Statin therapy is associated with improved survival in ischemic and non-ischemic heart failure. *Journal of the American College of Cardiology*, 43(4):642–648, 2004.

[46] D. W. Hosmer and S. Lemesbow. Goodness of fit tests for the multiple logistic regression model. *Communications in Statistics-Theory and Methods*, 9(10):1043–1069, 1980.

[47] K. Imai and M. Ratkovic. Covariate balancing propensity score. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(1):243–263, 2014.

[48] K. Imai and D. A. Van Dyk. Causal inference with general treatment regimes. *Journal of the American Statistical Association*, 99(467), 2004.

[49] C. A. Jackevicius, M. Mamdani, and J. V. Tu. Adherence with statin therapy in elderly patients with and without acute coronary syndromes. *Jama*, 288(4):462–467, 2002.

[50] A. Jeremias, S. Kaul, T. K. Rosengart, L. Gruberg, and D. L. Brown. The impact of revascularization on mortality in patients with nonacute coronary artery disease. *The American Journal of Medicine*, 122(2):152 – 161, 2009.

[51] J. D. Kalbfleisch and R. L. Prentice. *The statistical analysis of failure time data*, volume 360. John Wiley & Sons, 2011.

[52] H. Kamishirado, T. Inoue, M. Sakuma, T. Tsuda, T. Hayashi, K. Takayanagi, and K. Node. Effects of statins on restenosis after coronary stent implantation. *Angiology*, 58(1):55–60, 2007.

[53] E. C. Keeley, J. A. Boura, and C. L. Grines. Primary angioplasty versus intravenous thrombolytic therapy for acute myocardial infarction: a quantitative review of 23 randomised trials. *The Lancet*, 361(9351):13–20, 2003.

[54] J. Klein and M. Moeschberger. Survival analysis: statistical methods for censored and truncated data. *Springer-Verlag, New York, NY*, 2003.

[55] R. L. Kravitz, N. Duan, and J. Braslow. Evidence-based medicine, heterogeneity of treatment effects, and the trouble with averages. *Milbank Quarterly*, 82(4):661–687, 2004.

[56] A. Kulik, M. A. Brookhart, R. Levin, M. Ruel, D. H. Solomon, and N. K. Choudhry. Impact of statin use on outcomes after coronary artery bypass graft surgery. *Circulation*, 118(18):1785–1792, 2008.

[57] S. T. Lanza, J. E. Moore, and N. M. Butera. Drawing causal inferences using propensity scores: A practical guide for community psychologists. *American journal of community psychology*, 52(3-4):380–392, 2013.

[58] H. Y. Lee, C. E. Cooke, and T. A. Robertson. Use of secondary prevention drug therapy in patients with acute coronary syndrome after hospital discharge. *Journal of Managed Care Pharmacy*, 14(3):271, 2008.

[59] D. Y. Lin and L.-J. Wei. The robust inference for the cox proportional hazards model. *Journal of the American Statistical Association*, 84(408):1074–1078, 1989.

[60] S. M. Lloyd, D. J. Stott, A. J. de Craen, P. M. Kearney, N. Sattar, I. Perry, C. J. Packard, A. Briggs, L. Marchbank, H. Comber, et al. Long-term effects of statin treatment in elderly people: Extended follow-up of the prospective study of pravastatin in the elderly at risk (prosper). *PloS one*, 8(9):e72642, 2013.

[61] A. A. Mangoni and S. H. Jackson. Age-related changes in pharmacokinetics and pharmacodynamics: basic principles and practical applications. *British journal of clinical pharmacology*, 57(1):6–14, 2004.

[62] E. P. Martens, A. de Boer, W. R. Pestman, S. V. Belitser, B. H. C. Stricker, and O. H. Klungel. Comparing treatment effects after adjustment with multivariable cox proportional hazards regression and propensity score methods. *pharmacoepidemiology and drug safety*, 17(1):1–8, 2008.

[63] F. J. Massey Jr. The kolmogorov-smirnov test for goodness of fit. *Journal of American Statistical Association*, 46(253):68–78, 1951.

[64] L. K. Mell and J.-H. Jeong. Pitfalls of using composite primary end points in the presence of competing risks. *Journal of Clinical Oncology*, 28(28):4297–4299, 2010.

[65] J. M. R. Miguel A. Hernan. Causal inference. May 2012.

[66] K. Ming and P. R. Rosenbaum. Substantial gains in bias reduction from matching with a variable number of controls. *Biometrics*, 56(1):118–124, 2000.

[67] P. L. Morgan, M. L. Frisco, G. Farkas, and J. Hibel. A propensity score matching analysis of the effects of special education services. *The Journal of special education*, 2008.

[68] A. A. Motivala, U. Tamhane, V. S. Ramanath, F. Saab, D. G. Montgomery, J. Fang, E. Kline-Rogers, N. May, G. Ng, J. Froehlich, et al. A prior myocardial infarction: how does it affect management and outcomes in recurrent acute coronary syndromes? *Clinical cardiology*, 31(12):590–596, 2008.

[69] D. Mozaffarian, R. Nye, and W. C. Levy. Statin therapy is associated with lower mortality among patients with severe heart failure. *The American journal of cardiology*, 93(9):1124–1129, 2004.

[70] A. S. Nowacki, B. J. Wells, C. Yu, and M. W. Kattan. Adding propensity scores to pure prediction models fails to improve predictive performance. *PeerJ*, 1:e123, 2013.

[71] P. Peduzzi, W. Henderson, P. Hartigan, and P. Lavori. Analysis of randomized controlled trials. *Epidemiologic reviews*, 24(1):26–38, 2002.

[72] R. L. Prentice, J. D. Kalbfleisch, A. V. Peterson Jr, N. Flournoy, V. Farewell, and N. Breslow. The analysis of failure times in the presence of competing risks. *Biometrics*, pages 541–554, 1978.

[73] H. Putter, M. Fiocco, and R. Geskus. Tutorial in biostatistics: competing risks and multi-state models. *Statistics in medicine*, 26(11):2389–2430, 2007.

[74] K. Ray, S. Kondapally, and S. Erqou. Statins and all-cause mortality in high-risk primary prevention: A meta-analysis of 11 randomized controlled trials involving 65,229 participants. *Journal of Vascular Surgery*, 53(3):878, 2011.

[75] J. M. Robins. Association, causation, and marginal structural models. *Synthese*, 121(1):151–179, 1999.

[76] J. M. Robins, M. Á. Hernán, and B. Brumback. Marginal structural models and causal inference in epidemiology. *Epidemiology*, 11(5):550–560, 2000.

[77] P. R. Rosenbaum. A characterization of optimal designs for observational studies. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 597–610, 1991.

[78] P. R. Rosenbaum. Heterogeneity and causality. *The American Statistician*, 59(2), 2005.

[79] P. R. Rosenbaum. *Design of observational studies*. Springer, 2010.

[80] P. R. Rosenbaum and D. B. Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 1983.

[81] P. R. Rosenbaum and D. B. Rubin. Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association*, 79(387):516–524, 1984.

[82] P. R. Rosenbaum and D. B. Rubin. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, 39(1):33–38, 1985.

[83] P. R. Rosenbuam. *Observational studies*. Springer, 2nd edition, 2002.

[84] D. B. Rubin. Bias reduction using mahalanobis metric matching. 36(2):293–298, 2002.

[85] F. M. Sacks, M. A. Pfeffer, L. A. Moye, J. L. Rouleau, J. D. Rutherford, T. G. Cole, L. Brown, J. W. Warnica, J. M. O. Arnold, C.-C. Wun, et al. The effect of pravastatin on coronary events after myocardial infarction in patients with average cholesterol levels. *New England Journal of Medicine*, 335(14):1001–1009, 1996.

[86] T. Sato and Y. Matsuyama. Marginal structural models as a tool for standardization. *Epidemiology*, 14(6):680–686, 2003.

[87] G. Savarese, A. M. Gotto Jr, S. Paolillo, C. D'Amore, T. Losco, F. Musella, O. Scala, C. Marciano, D. Ruggiero, F. Marsico, et al. Benefits of statins in elderly subjects without established cardiovascular disease: a meta-analysis. *Journal of the American College of Cardiology*, 62(22):2090–2099, 2013.

[88] S. Senn, E. Graf, and A. Caputo. Stratification for the propensity score compared with linear regression techniques to assess the effect of treatment or exposure. *Statistics in Medicine*, 26(30):5529–5544, 2007.

[89] B. R. Shah, A. Laupacis, J. E. Hux, and P. C. Austin. Propensity score methods gave similar results to traditional regression modeling in observational studies: a systematic review. *Journal of clinical epidemiology*, 58(6):550–559, 2005.

[90] J. Shepherd, G. J. Blauw, M. B. Murphy, E. L. Bollen, B. M. Buckley, S. M. Cobbe, I. Ford, A. Gaw, M. Hyland, J. W. Jukema, et al. Pravastatin in elderly individuals at risk of vascular disease (prosper): a randomised controlled trial. *The Lancet*, 360(9346):1623–1630, 2002.

[91] M. Sheppard, A. Begg, and S. Heart. Statins in older adults. *Australian Prescriber*, 36(6), 2013.

[92] S. Sidney, W. D. Rosamond, V. J. Howard, and R. V. Luepker. The heart disease and stroke statistics-2013 update and the need for a national cardiovascular surveillance system. *Circulation*, 127(1):21–23, 2013.

[93] A. Spinewine, K. E. Schmader, N. Barber, C. Hughes, K. L. Lapane, C. Swine, and J. T. Hanlon. Appropriate prescribing in elderly people: how well can it be measured and optimised? *The Lancet*, 370(9582):173–184, 2007.

[94] E. A. Stuart and K. M. Green. Using full matching to estimate causal effects in nonexperimental studies: examining the relationship between adolescent marijuana use and adult outcomes. *Developmental Psychology*, 44(2):395, 2008.

[95] T. A. Stukel, E. S. Fisher, D. E. Wennberg, D. A. Alter, D. J. Gottlieb, and M. J. Vermeulen. Analysis of observational studies in the presence of treatment selection bias: effects of invasive cardiac management on ami survival using propensity score and instrumental variable methods. *Jama*, 297(3):278–285, 2007.

[96] T. Stümer, M. Jshi, R. J. Glynn, J. Avorn, K. J. Rothman, and S. Schneeweiss. A review of the application of propensity score methods yielded increasing use, advantages in specific setting, but not substantially different estimates compared with conventional multivariable methods. *Journal of clinical epidemiology*, 59(5).

[97] T. Stürmer, K. J. Rothman, and R. J. Glynn. Insights into different results from different causal contrasts in the presence of effect-measure modification. *Pharmacoepidemiology and drug safety*, 15(10):698–709, 2006.

[98] T. Stürmer, S. Schneeweiss, M. A. Brookhart, K. J. Rothman, J. Avorn, and R. J. Glynn. Analytic strategies to adjust confounding using exposure propensity scores and disease risk scores: nonsteroidal antiinflammatory drugs and short-term mortality in the elderly. *American Journal of Epidemiology*, 161(9):891–898, 2005.

[99] D. Suarez, R. Borras, and X. Basagaña. Differences between marginal structural models and conventional models in their exposure effect estimates: a systematic review. *Epidemiology*, 22(4):586–588, 2011.

[100] D. T. Sugerman. Statins. *JAMA*, 309(13):1419, 2013.

[101] S. Tandra and R. Vuppalanchi. Use of statins in patients with liver disease. *Current treatment options in cardiovascular medicine*, 11(4):272–278, 2009.

[102] R. C. Team et al. R: A language and environment for statistical computing, 2005.

[103] T. M. Therneau. *Modeling survival data: extending the Cox model*. Springer, 2000.

[104] E. Tipton. Improving generalizations from experiments using propensity score subclassification assumptions, properties, and contexts. *Journal of Educational and Behavioral Statistics*, 38(3):239–266, 2013.

[105] D. Tritchler. Interpreting the standardized difference. *Biometrics*, 51(1):351–353, 1995.

[106] S. Weitzen, K. L. Lapane, A. Y. Toledano, A. L. Hume, and V. Mor. Weaknesses of goodness-of-fit tests for evaluating propensity score models: the case of the omitted confounder. *Pharmacoepidemiology and drug safety*, 14(4):227–238, 2005.

[107] D. Westreich, S. R. Cole, M. J. Funk, M. A. Brookhart, and T. Stürmer. The role of the c-statistic in variable selection for propensity score models. *Pharmacoepidemiology and drug safety*, 20(3):317–320, 2011.

[108] WHO. The top 10 causes of death, July 2013.

[109] P. W. Wilson, R. D'Agostino Sr, D. L. Bhatt, K. Eagle, M. J. Pencina, S. C. Smith, M. J. Alberts, J. Dallongeville, S. Goto, A. T. Hirsch, et al. An international model to predict recurrent cardiovascular disease. *The American journal of medicine*, 125(7):695–703, 2012.

[110] R. Wyss, C. J. Girman, R. J. LoCasale, M. Alan Brookhart, and T. Stürmer. Variable selection for propensity score models when estimating treatment effects on multiple outcomes: a simulation study. *Pharmacoepidemiology and drug safety*, 22(1):77–85, 2013.

[111] S. Xu, C. Ross, M. A. Raebel, S. Shetterly, C. Blanchette, and D. Smith. Use of stabilized inverse propensity scores as weights to directly estimate relative risk and its confidence intervals. *Value in Health*, 13(2):273–277, 2010.

[112] S. YUSUF. Mrc/bhf heart protection study of cholesterol lowering with simvastatin in 20536 high-risk individuals: a randomised placebo-controlled trial. commentary. *Lancet*, 360(9326):7–22, 2002.

[113] B. Zhou, J. Fine, A. Latouche, and M. Labopin. Competing risks regression for clustered data. *Biostatistics*, 13(3):371–383, 2012.

[114] B. Zhou and A. Latouche. *crrSC: Competing risks regression for Stratified and Clustered data*, 2013. R package version 1.1.

[115] B. Zhou, A. Latouche, V. Rocha, and J. Fine. Competing risks regression for stratified data. *Biometrics*, 67(2):661–670, 2011.