Evaluating Chart Review Strategies Using Electronic Health Record Data in the Context of Risk Prediction Modeling

By

Elizabeth McNeer

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Biostatistics

August 10, 2018

Nashville, Tennessee

Approved:

Dandan Liu, Ph.D.

Qingxia Chen, Ph.D.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor, Dr. Dandan Liu, for her guidance and support as I have worked on this thesis. I would also like to thank Dr. Cindy Chen for being on my committee. I'm also thankful for the professors and my fellow students in the Department of Biostatistics. Finally, I would like to acknowledge and thank my family for always supporting and encouraging me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

Chapter 1

Introduction

Often in a clinical setting, it is of interest to know a patient's risk for developing a certain disease or other adverse outcomes to support clinical decision making. This can be accomplished through the use of a risk prediction model or risk score. Some examples of risk scores include the Framingham risk score (Hippisley-Cox et al., 2008), which estimates the 10-year risk of developing cardiovascular disease and is useful in outpatient or self-care settings and the Sepsis-related Organ Failure Assessment (SOFA) score (Vincent et al., 1996), which describes organ dysfunction/failure and is useful in an acute care setting. Being able to accurately predict a patient's risk for an adverse outcome can be beneficial at a population level by helping to control chronic diseases and at an individual level by helping the healthcare workers to deliver an effective intervention in a timely manner.

Risk prediction models are usually developed using research-based large cohort studies or clinical trials, which are often not representative of the local population of interest (Goldstein et al., 2016). Model validation using the local patient population has to be conducted before existing risk prediction models can be used in local clinical practice. Rapid deployment of electronic health record (EHR) systems provides unique opportunities and resources for such validation studies (Kolek et al., 2016; Riley et al., 2016).

Conducting research studies using EHR data presents several challenges as well. There could be systematically missing predictor values, meaning that a predictor is not measured for any individuals in one or more clusters. In this case, a method such as multiple imputation would have to be performed. Also of concern when using EHR data is the quality of the data. Potential concerns are issues such as missing data, non-standardized definitions of outcomes, and incomplete follow-up times and event dates (Riley et al., 2016). It is also possible that incorrect values for variables of interest could be present in EHR data. For example, in a validation study for the CHARGE-AF

risk prediction (Alonso et al., 2013), EHR data from a cohort of 33,494 patients is used (Kolek et al., 2016). For risk factor data extraction, treatment of hypertension and current smoking status were determined using previously validated algorithms with a sensitivity value of 88% for treatment of hypertension and positive predictive values of 93% for both. In other words, 88 out of 100 patients truly receiving treatment of hypertension were correctly identified and 93 out of 100 identified current smokers are true current smokers. Therefore, data errors could potentially have been present in those variables. To deal with these concerns, it is often necessary for physicians to perform chart reviews, which can be very time consuming and sometimes not feasible for large EHR datasets such as the CHARGE-AF validation study.

Another motivating study aims to evaluate the established Sepsis-related Organ Failure Assessment (SOFA) score in predicting 30-day mortality since ICU admission using EHR data. A SOFA score consists of six components corresponding to the following six organ systems: respiratory, cardiovascular, neurologic, hematological, renal, and hepatic (Vincent et al., 1996). Each component is assigned an integer value between 0 and 4. Sepsis is a syndrome of physiologic, pathologic, and bio-chemical abnormalities induced by infection and is the primary cause of death from infection, especially if not recognized and treated promptly. A consensus definition for sepsis has been developed and revised in the past two decades for early diagnosis of sepsis. Recently, the third international consensus definition for sepsis was released (Singer et al., 2016), where change in total SOFA score greater than two points was included as a major component of sepsis diagnosis criteria. Due to rapid deterioration of sepsis patients, it is critical that the SOFA score be calculated as quickly as possible for ICU patients, and thus development of computerized algorithms to automatically calculate the SOFA score by using natural language processing to extract data from the EHR is warranted. However, the information needed to calculate the SOFA score may not be accurately recorded in the EHR data, and it may require a check of the nurses' notes or other written sources of information, i.e. chart review. In this motivating study, it is of interest to compare an electronically derived SOFA score (eSOFA) with the gold standard manually derived SOFA score (mSOFA) based on chart review in predicting 30-day mortality since ICU admission.

The purpose of this paper is two-fold: (1) to evaluate the impact of data quality problems on risk prediction; (2) and to compare two chart review sampling strategies, case-cohort sampling and random sampling, to correct data errors and assess the impact of these corrections on risk prediction. The rest of the paper is organized as follows. Chapter 2 introduces chart reviews and proposes two sampling strategies for choosing the patients for which the chart reviews will be performed. Chapter 3 details simulation studies that were performed to assess the impact of data quality issues on the risk factor effects and risk prediction model performance under a Cox proportional hazards model. Various scenarios combining different risk factor distributions, event rates, and error rates are considered. The simulation also includes correction of the data errors using the chart review strategies introduced in Chapter 2 to examine how this impacts the models. In Chapter 4, the strategies introduced in Chapter 2 are applied to compare eSOFA and mSOFA in predicting 30-day mortality for ICU patients using EHR data from the Isotonic Solutions and Major Adverse Renal Events Trial (SMART) (Semler et al., 2018). Finally, the findings and future work are discussed in Chapter 5.

Chapter 2

Chart Review Strategies

Chart reviews are particularly important for retrospective studies. These types of studies use pre-recorded, patient focused data to address research questions that are not able to be answered using a prospective study, such as the effects of a harmful or beneficial exposure to which subjects cannot be randomized or the occurrence of a rare event after an exposure to which subjects cannot be randomized (Worster and Haines, 2004). With smaller datasets, it's relatively easy to review all the records. However, with the growing size of EHR data, this is more of a challenge, so a reasonable sample size of charts to be reviewed must be determined. Typically, the records to be reviewed are selected by a type of convenience sampling where all records within a given time frame are chosen or by a random sampling of records from the population of interest (Worster and Haines, 2004).

Chart reviews are also used for validating phenotyping algorithms where patients with certain traits, diseases, or responses to medication are identified through combined resources extracted from EHR data using various computation approaches such as natural language processing, as was discussed by Kirby et al. (2016) using the Electronic Medical Records and Genomics (eMERGE) network. Case-control sampling strategies are usually used to develop phenotyping algorithms, where the phenotyping algorithm performance is mainly evaluated using sensitivity and positive predictive values.

For studies to develop or evaluate risk prediction models using EHR data, a random sampled dataset might not include a sufficient number of events to provide reliable results, which is particularly concerning for rare diseases. Alternatively, we could consider a case-cohort sampling strategy for choosing the patients whose charts will be reviewed. The concept of the case-cohort sampling strategy comes from the case-cohort design proposed by Prentice (1986). In the case-cohort chart review sampling strategy, all cases and a random sample of non-cases will be chart

reviewed. Note that all the studies discussed so far only consider using the sampled chart reviewed records to draw conclusions while assuming results from chart reviewed data are applicable to the underlying population of interest.

The goal of this paper is to use the entire study population corroborated with chart reviewed records. The impact of chart review sampling strategies in the context of risk prediction models will be investigated. First, the risk prediction model will be fit without correcting any data errors. Next, chart reviews will be performed to correct any data errors in the sample of records that are reviewed. To choose the patients whose records are reviewed, a random sampling strategy will be used as well as the alternative strategy using case-cohort sampling. Then, the risk prediction model will be refit and the impact of the data correction will be assessed for each of the chart review sampling strategies.

Chapter 3

Simulation

3.1 Simulation Settings

In this chapter, extensive simulations are conducted. For each simulation, $N = 5,000$ subjects are generated from a Cox proportional hazards model

$$\lambda(t|\mathbf{Z}) = \lambda_0(t)\exp(\boldsymbol{\beta}^{\mathsf{T}}\mathbf{Z}),$$

where $\boldsymbol{\beta} = (0.3, 0.5, 0.3, 0.5, 0.3, 0.5)$ is a vector of coefficients and $\mathbf{Z} = (Z_1, Z_2, Z_3, Z_4, Z_5, Z_6)$ is a vector of covariates representative of the six types of risk factors that might be found in EHR data: Z_1 was continuous (e.g. age), Z_2 was binary (e.g. gender), Z_3 was continuous and dependent on Z_1 and Z_2 (e.g. systolic blood pressure), Z_4 was binary and dependent on Z_1 and Z_2 (e.g. medical history of cardiovascular disease), Z_5 was continuous and independent of all other covariates (e.g. BMI), and Z_6 was binary and independent of all other covariates (e.g. smoking status). The following distribution was used to generate $\mathbf{Z}$:

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \\ X_6 \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0.2 & 0.2 & 0 & 0 \\ 0 & 1 & 0.2 & 0.2 & 0 & 0 \\ 0.2 & 0.2 & 1 & 0 & 0 & 0 \\ 0.2 & 0.2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \right).$$

For the continuous covariates, $Z_1 = X_1$, $Z_3 = X_3$, and $Z_5 = X_5$. To generate the following distributions for binary covariates: $Z_2 \sim Bin(0.5)$, $Z_4 \sim Bin(0.3)$, and $Z_6 \sim Bin(0.3)$, let $Z_2 = I(X_2 > 0)$, $Z_4 = I(X_4 > \Phi^{-1}(0.7))$, and $Z_6 = I(X_6 > \Phi^{-1}(0.7))$, where I$(\cdot)$ is an indicator function and Φ is the

cumulative distribution function for the standard Normal distribution. Two scenarios for $\lambda_0(t)$ and censoring distributions are considered: (1) $\lambda_0(t) = 0.005$ and $C \sim exp(0.21)$; and (2) $\lambda_0(t) = 0.05$ and $C \sim exp(0.4)$, which results in event rate of approximately 5% and 20% respectively. The following was used to simulate the event times, T:

$$S(T) = e^{-\Lambda_0(T)e^{\boldsymbol{\beta}\mathbf{Z}}}$$

where $S(T) = U \sim U(0,1)$.

$$\begin{aligned} \implies \Lambda_0(T) &= -log(U)e^{-\boldsymbol{\beta}\mathbf{Z}} \\ T &= -\frac{T}{\Lambda_0(T)} log(U)e^{-\boldsymbol{\beta}\mathbf{Z}} \end{aligned}$$

3.2 Error Scenarios

Next, the observed covariates $\mathbf{Z}^* = (Z_1, Z_2, Z_3^*, Z_4^*, Z_5^*, Z_6^*)$ were created by adding errors to each of the four underlying true covariates Z_3, Z_4, Z_5, and Z_6 one at a time and with four different error rates, $ER = 5\%, 10\%, 20\%,$ and 30%, resulting in a total of 16 error scenarios. Specifically, for the first four scenarios, $Z_3^* = Z_3 + \varepsilon$ with $\varepsilon \sim N(0, 0.25)$ for $ER * N$ randomly selected observations (e.g. 250 observations for a 5% error rate), and the rest of the observations were the same as $\mathbf{Z}$. For the next four scenarios, $Z_5^* = Z_5 + \varepsilon$ with $\varepsilon \sim N(0, 0.25)$ for $ER * N$ randomly selected observations and the rest of the observations were the same as $\mathbf{Z}$. For the four scenarios involving Z_4, among all observations with $Z_4 = 1$, a proportion of ER observations were randomly selected such that $Z_4^* = 0$ and the rest of the observations were the same as $\mathbf{Z}$. Using an error rate of 20%, approximately $20\% * P(Z_4 = 1) = 6\%$ of observations had error introduced. Similar approaches were used for the last four scenarios involving Z_6.

3.3 Chart Review Sampling Scenarios

A sample size of 500 was chosen for the number of chart reviews to perform. The first sampling strategy used to choose which subjects had chart reviews performed was random sampling. In this strategy, 500 subjects were randomly chosen to be reviewed regardless of whether or not they experienced an event. Next, case-cohort sampling was used to choose which subjects had chart reviews. For the 5% event rate, all cases were reviewed and $500 - N_c$ non-cases were reviewed, where N_c was the number of cases. For the 20% event rate, 500 cases were reviewed. The following was performed for each of the four scenarios. For the 500 subjects chosen to have their charts reviewed, the true covariate values ($\mathbf{Z}$) were used instead of the observed covariate values ($\mathbf{Z}^*$). The same 16 Cox proportional hazards models discussed in section 3.2 were refit using the corrected data, and the summary measures discussed below were calculated for these models.

3.4 Summary Measures

Within each simulation, four models with different covariates were fitted for each of the 16 error scenarios. The first model used $\mathbf{Z}$ representing the true model. The second model used $\mathbf{Z}^*$ representing the error model. The third model used $\mathbf{Z}^*$ compensated by a case-cohort sampling chart review strategy. The fourth model used $\mathbf{Z}^*$ compensated by a random sampling chart review strategy. A total of 500 simulations were run, and the results for β estimators were summarized using percent bias, mean square error, and coverage probability. The C-index was used as a measure of discrimination of the models. Let β_i be the truth and let $\hat{\beta}_{ij}$ for i = 1,2,...,6 denote an estimate of β_i from the jth simulation. Let $\bar{\hat{\beta}}_i = \frac{\sum_{j=1}^{500} \hat{\beta}_{ij}}{500}$ denote the mean of $\hat{\beta}_i$ from all 500 simulations.

The bias was calculated using the formula Bias($\bar{\hat{\beta}}_i$,β_i) = $\bar{\hat{\beta}}_i - \beta_i$, and the percent bias was found using the formula $\frac{Bias(\bar{\hat{\beta}}_i,\beta_i)}{\beta_i} * 100$. A positive percent bias indicates overestimation of the true β_i, and a negative percent bias indicates underestimation of the true β_i.

The mean square error (MSE) is a measure that takes into account both variance and bias. It was found using the formula MSE($\bar{\hat{\beta}}_i$,β_i) = $\frac{1}{500}\sum_{j=1}^{500}(\hat{\beta}_{ij} - \beta_i)^2 = Var(\bar{\hat{\beta}}_i) + (Bias(\bar{\hat{\beta}}_i, \beta_i))^2$.

The coverage probability for $\hat{\beta}_i$ gives the proportion of the 500 $\hat{\beta}_{ij}$'s that fall within their associated 95% confidence intervals. The confidence interval for the jth simulation is found using the formula $\hat{\beta}_{ij} \pm 1.96 * SE(\hat{\beta}_{ij})$.

The C-index, or concordance probability, is a measure of discrimination for the Cox proportional hazards model. This measure is a generalization of the area under a receiver operating characteristic (ROC) curve. To calculate the C-index, the survival times of pairs of subjects are ordered. If both subjects are censored or if one subject has experienced an event and the follow-up time of the other subject is less than the event time of the first, then this pair of subjects' survival times cannot be ordered. The C-index is the proportion of all possible pairs of subjects such that the subject with the higher predicted survival is the one who survived longer. A C-index value of 0.5 indicates random predictions, and a value of 1 indicates that the model results in perfect predictions (Harrell, 2015).

3.5 Results

3.5.1 Continuous Covariates

The Cox proportional hazards model results from the simulations using the observed continuous covariates Z_3^* and Z_5^* and true continuous covariates Z_3 and Z_5 at all four error rate scenarios from both the 5% and 20% event rate scenarios can be found in Table 3.1. Using the true covariate Z_3, the percent bias for $\hat{\beta}_3$ was -1.39 at a 5% event rate and -1.38 at a 20% event rate. Using Z_3^*, the magnitude of the percent bias increased at both event rates, and it was more pronounced at higher error rates as expected. At the 5% event rate, the percent bias ranged from -2.75 at a 5% error rate to -8.88 at a 30% error rate, and at the 20% event rate, it ranged from -2.75 to -9.25. The C-index using the true covariates was 0.7 at a 5% event rate and 0.69 at a 20% event rate. The percent bias for the C-index when using Z_3^* was very small at all error rates, ranging from -0.03 to -0.14 at a 5% event rate and -0.03 to -0.16 at a 20% event rate.

The percent biases for $\hat{\beta}_1$ and $\hat{\beta}_2$ were also affected by the errors added to Z_3 because Z_3 was

generated to be dependent on Z_1 and Z_2. For β_1, the percent bias was higher when using Z_3^*, and it increased with increasing error rates at both event rates. The percent bias for $\hat{\beta}_1$ ranged from 1.58 at a 5% error rate to 2.81 at a 30% error rate compared to 1.32 with no errors at a 5% event rate. At a 20% event rate, this percent bias ranged from 1.31 to 2.48 compared to 1.06 without errors. The percent bias for $\hat{\beta}_2$ behaved similarly to what was described for $\hat{\beta}_1$, with a higher percent bias than $\hat{\beta}_2$ without errors and increasing percent bias as the error rate increased.

For the true independent covariate (Z_5), the percent bias with no errors at a 5% event rate was 1.05, and at a 20% event rate, it was -0.12. Similar patterns to what were described above for the true dependent covariate Z_3 were found here. When errors were introduced, the percent bias became more pronounced as the error rate increased. The magnitude of the percent bias for the observed independent covariate (Z_5^*) is slightly smaller than it is for the corresponding observed dependent covariate (Z_3^*) at all event rates and error rates. Also, the percent bias of the C-index when using Z_5^* is comparable to what was found for the scenarios using Z_3^*.

The results using the observed continuous covariates (Z_3^* and Z_5^*) after performing chart reviews using random sampling can be found in Table 3.2. Similarly, the results after performing chart reviews using case-cohort sampling are summarized in Table 3.3. The patterns for both chart review sampling strategies are similar to those that were found in Table 3.1. The percent bias for the C-index is very similar to what was found without performing chart reviews. However, the magnitude of the percent bias for β_3 and β_5 is slightly smaller after performing chart reviews. See Tables 3.2 and 3.3 for details. Also, there is about a 10% reduction in the percent bias after correcting errors using chart review. For example, the percent bias for $\hat{\beta}_5$ is reduced by an average of about 8.93% over all four error rates when using a 20% event rate with random sampling chart review, and it's reduced by an average of about 12.97% over all four error rates when using a 20% event rate with case-cohort sampling chart review.

Figure 3.1 displays the percent bias of the effect estimates for both the independent (Z_5) and dependent (Z_3) continuous covariates at each of the four error rates and at both the 5% and 20% event rates. The lines show the impact of increasing error rates on the percent bias, with a different

color for each of the chart review strategies. Overall, it appears that the case-cohort strategy tends to perform better in terms of percent bias than the random sampling strategy, and both chart review strategies perform slightly better than no chart review. Also, the difference is more distinct at higher error rates.

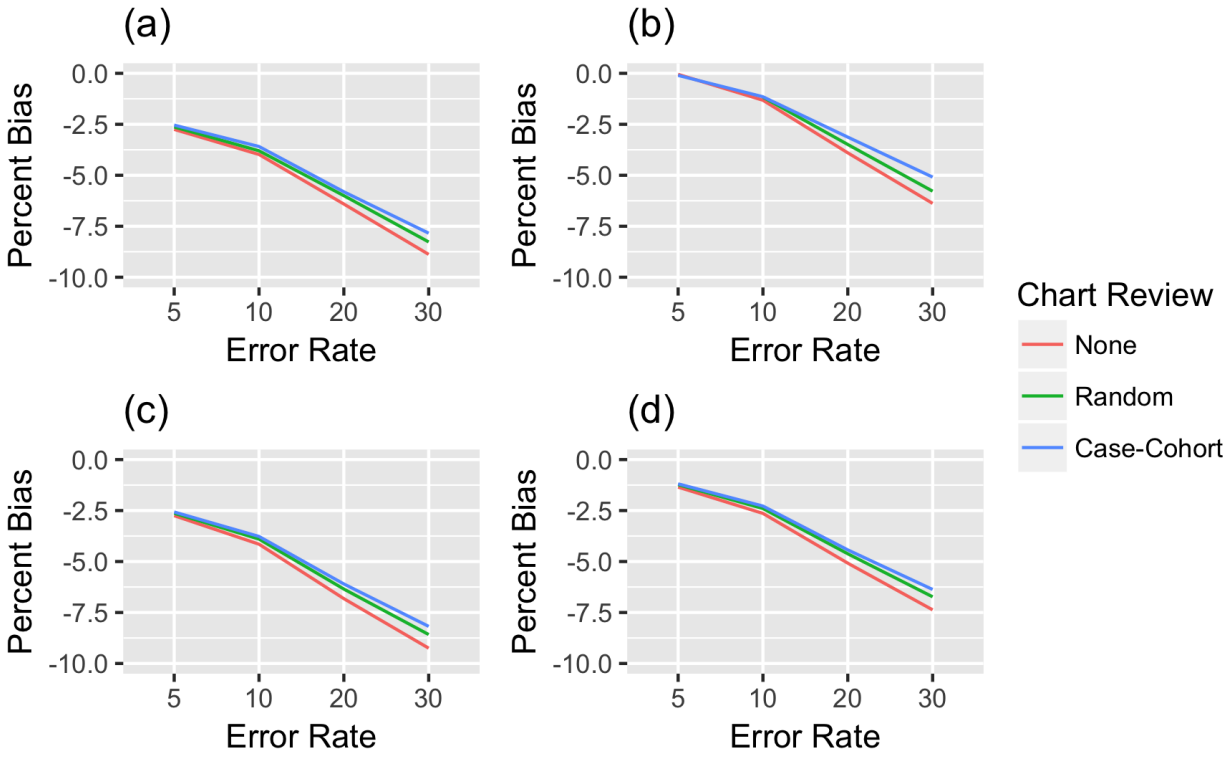

Figure 3.1: Percent Bias for Continuous Covariates
(a) Z_3, 5% event rate (b) Z_5, 5% event rate
(c) Z_3, 20% event rate (d) Z_5, 20% event rate

3.5.2 Binary Covariates

Table 3.4 displays the Cox proportional hazard model results from the simulations using the observed binary covariates Z_4^* and Z_6^* and true binary covariates Z_4 and Z_6 at all four error rates from both the 5% and 20% event rate scenarios. For the dependent covariate (Z_4) at a 5% event rate, the percent bias for the effect estimate ($\hat{\beta}_4$) was -0.11 when using the true covariate Z_4. When using the observed covariate Z_4^*, the magnitude of the percent bias increased with increasing error rates. Also, the percent bias became more pronounced as the error rate increased. At a 5% error

rate, the percent bias for $\hat{\beta}_4$ was -3.50, and at a 30% error rate, the percent bias was -17.04. A similar pattern was found for the dependent covariate at a 20% event rate. The percent bias in this scenario was -0.01 with no errors, -2.87 at a 5% error rate, and -15.66 at a 20% error rate. At a 5% event rate, the percent bias of the C-index when using Z_4^* ranged from -0.09 at a 5% error rate to -0.54 at a 30% error rate, and at a 20% event rate, the it ranged from -0.11 to -0.58.

The percent biases for $\hat{\beta}_1$ and $\hat{\beta}_2$ were also affected by the errors added to Z_4 because Z_4 was dependent on Z_1 and Z_2. The percent bias for $\hat{\beta}_1$ when using the true Z_4 was 1.32 at a 5% event rate and 1.06 at a 20% event rate. For $\hat{\beta}_2$, the percent bias was -1.22 at a 5% event rate and 0.53 at a 20% event rate when using Z_4 with no errors. The percent bias was higher for both effect estimates when using Z_4^*, and it increased with increasing error rates at both event rates. See Table 3.4 for details.

The percent bias for the effect estimate for the independent covariate (Z_6) when using Z_6 with no errors was 1.68 at a 5% event rate and 1.05 at a 20% event rate. A similar pattern was found for the observed independent covariate Z_6^* as was described above for the observed dependent covariate Z_4^*. The magnitude of the percent bias increased with increasing error rates, and the percent bias became more pronounced as the error rate increased. Also, the percent bias for the C-index when using Z_6^* was very similar to what was found for the C-index when using Z_4^*. See table 3.4 for details.

The results from fitting the models after performing chart reviews with random sampling are shown in Table 3.5. The behavior resembles what was found from the modesl without performing chart reviews. The percent bias is negative in all scenarios and the magnitude increases as the event rate increases. Also, the percent bias of the dependent covariate is more pronounced than the percent bias of the independent covariate for both event rates and all error rates. The percent biases for the C-index were also very similar at all error rates and both event rates to what was found without performing chart reviews. See Table 3.5 for details. Also, there is about a 10% reduction in the percent bias after correcting errors using random sampling chart review. For example, the percent bias for $\hat{\beta}_6$ is reduced by an average of about 11.64% over all four error rates when using

a 20% event rate with random sampling chart review.

In Table 3.6, the results from fitting the models with the observed binary covariates (Z_4^* and Z_6^*) after performing chart reviews using case-cohort sampling are shown. The pattern is the same as was described for the binary covariates without chart reviews performed; the magnitude of the percent bias increased as the error rate increased. However, the percent biases after performing chart reviews using case-cohort sampling are all positive and a much greater magnitude than those without using chart review. At a 30% error rate and a 5% event rate, the effect estimate for the dependent covariate ($\hat{\beta}_4$) after performing case-cohort sampling chart review had a percent bias of 84.49 compared to -17.04 without chart review. During the chart review process, incorrect 0's were changed to correct 1's for all cases and a portion of non-cases at the 5% event rate and for 500 cases and no non-cases at the 20% event rate. This resulted in different Z_4 and Z_6 distributions for the cases and the non-cases, which explains the large positive percent bias seen here. Also, the percent bias for the C-index is positive at all error rates, and it is more pronounced than the percent bias for the C-index without performing chart reviews.

Figure 3.2 displays the percent bias for the effect estimates of the observed dependent (Z_4^*) and independent (Z_6^*) binary covariates at both the 5% event rate and the 20% event rate. As was discussed above, performing chart reviews with case-cohort sampling results in a large positive percent bias, so this scenario was not included in this plot. Random sampling chart review has slightly less pronounced bias than no chart review with a larger difference at higher error rates.

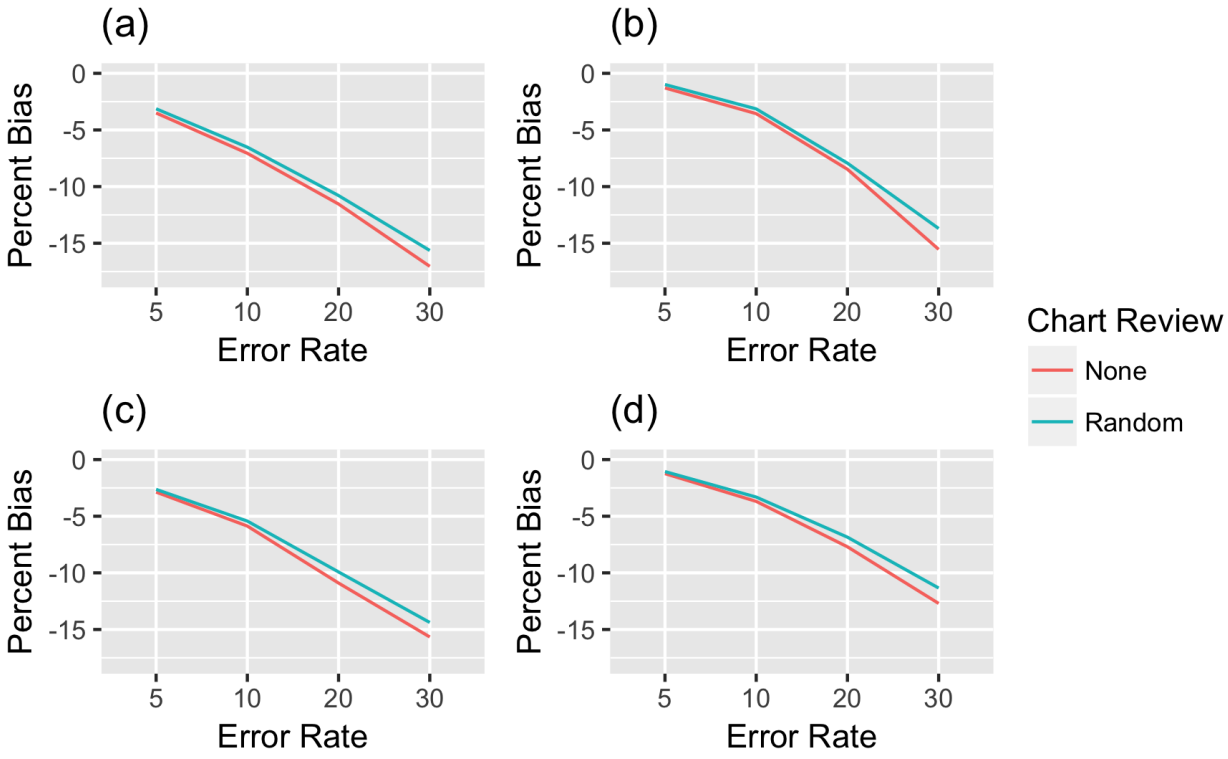

Figure 3.2: Percent Bias for Binary Covariates
(a) Z_4, 5% event rate (b) Z_6, 5% event rate
(c) Z_4, 20% event rate (d) Z_6, 20% event rate

Table 3.1: Simulation Results for Continuous Covariates

			5% Error Rate			10% Error Rate			20% Error Rate			30% Error Rate			No Error		
		Truth	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP
5% Event Rate	β_1	0.3	1.58	0.41	0.95	1.83	0.41	0.95	2.32	0.41	0.96	2.81	0.41	0.96	1.32	0.41	0.95
	β_2	0.5	-0.96	1.98	0.94	-0.70	1.97	0.94	-0.24	1.96	0.94	0.24	1.96	0.94	-1.22	1.99	0.94
	$\boldsymbol{\beta_3}$	**0.3**	**-2.75**	**0.42**	**0.95**	**-3.98**	**0.43**	**0.94**	**-6.41**	**0.45**	**0.95**	**-8.88**	**0.45**	**0.95**	**-1.39**	**0.43**	**0.95**
	β_4	0.5	-0.19	1.82	0.95	-0.25	1.82	0.95	-0.44	1.82	0.95	-0.63	1.83	0.95	-0.11	1.83	0.95
	β_5	0.3	1.04	0.38	0.96	1.05	0.38	0.96	1.01	0.38	0.96	0.98	0.39	0.96	1.05	0.38	0.96
	β_6	0.5	1.68	1.79	0.94	1.68	1.80	0.94	1.65	1.80	0.94	1.62	1.79	0.94	1.68	1.80	0.94
	C-Index	0.7	-0.03			-0.05			-0.09			-0.14					
	β_1	0.3	1.34	0.41	0.95	1.31	0.41	0.95	1.30	0.41	0.96	1.26	0.41	0.95	1.32	0.41	0.95
	β_2	0.5	-1.23	1.99	0.94	-1.27	1.98	0.93	-1.25	1.99	0.93	-1.30	1.99	0.94	-1.22	1.99	0.94
	β_3	0.3	-1.4	0.43	0.95	-1.42	0.43	0.95	-1.41	0.43	0.95	-1.39	0.43	0.95	-1.39	0.43	0.95
	β_4	0.5	-0.13	1.83	0.95	-0.14	1.82	0.95	-0.15	1.83	0.95	-0.13	1.83	0.95	-0.11	1.83	0.95
	$\boldsymbol{\beta_5}$	**0.3**	**-0.05**	**0.38**	**0.96**	**-1.32**	**0.37**	**0.97**	**-3.90**	**0.38**	**0.96**	**-6.38**	**0.40**	**0.94**	**1.05**	**0.38**	**0.96**
	β_6	0.5	1.68	1.80	0.93	1.67	1.80	0.93	1.69	1.80	0.93	1.69	1.80	0.94	1.68	1.80	0.94
	C-Index	0.7	-0.02			-0.04			-0.10			-0.16					
20% Event Rate	β_1	0.3	1.31	0.12	0.95	1.58	0.12	0.95	2.05	0.12	0.95	2.48	0.12	0.95	1.06	0.12	0.95
	β_2	0.5	0.77	0.46	0.94	1.03	0.46	0.94	1.48	0.46	0.94	1.92	0.47	0.94	0.53	0.46	0.94
	$\boldsymbol{\beta_3}$	**0.3**	**-2.75**	**0.12**	**0.95**	**-4.15**	**0.12**	**0.93**	**-6.82**	**0.15**	**0.90**	**-9.25**	**0.18**	**0.85**	**-1.38**	**0.11**	**0.95**
	β_4	0.5	-0.11	0.45	0.95	-0.22	0.45	0.95	-0.45	0.45	0.94	-0.65	0.45	0.95	-0.01	0.45	0.95
	β_5	0.3	-0.14	0.10	0.95	-0.16	0.10	0.95	-0.24	0.10	0.95	-0.28	0.10	0.95	-0.12	0.10	0.96
	β_6	0.5	1.03	0.44	0.95	0.99	0.44	0.95	0.94	0.44	0.95	0.89	0.44	0.95	1.05	0.44	0.95
	C-Index	0.69	-0.03			-0.06			-0.11			-0.16					
	β_1	0.3	1.05	0.12	0.95	1.02	0.12	0.95	0.96	0.12	0.95	0.89	0.12	0.95	1.06	0.12	0.95
	β_2	0.5	0.5	0.46	0.94	0.44	0.46	0.94	0.45	0.46	0.94	0.35	0.45	0.94	0.53	0.46	0.94
	β_3	0.3	-1.41	0.11	0.95	-1.43	0.11	0.95	-1.46	0.11	0.95	-1.48	0.11	0.95	-1.38	0.11	0.95
	β_4	0.5	-0.04	0.45	0.95	-0.07	0.45	0.94	-0.13	0.45	0.95	-0.10	0.45	0.95	-0.01	0.45	0.95
	$\boldsymbol{\beta_5}$	**0.3**	**-1.35**	**0.10**	**0.96**	**-2.64**	**0.10**	**0.95**	**-5.08**	**0.12**	**0.92**	**-7.37**	**0.14**	**0.88**	**-0.12**	**0.10**	**0.96**
	β_6	0.5	1.04	0.44	0.95	1.00	0.44	0.95	0.97	0.44	0.95	0.93	0.44	0.95	1.05	0.44	0.95
	C-Index	0.69	-0.03			-0.05			-0.11			-0.16					

Table 3.2: Simulation Results for Continuous Covariates with Random Sampling Chart Review

			5% Error Rate			10% Error Rate			20% Error Rate			30% Error Rate			No Error		
		Truth	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP
5% Event Rate	β_1	0.3	1.55	0.41	0.95	1.80	0.41	0.95	2.23	0.41	0.96	2.69	0.41	0.96	1.32	0.41	0.95
	β_2	0.5	-0.98	1.98	0.94	-0.74	1.97	0.94	-0.33	1.97	0.94	0.10	1.96	0.94	-1.22	1.99	0.94
	$\boldsymbol{\beta_3}$	**0.3**	**-2.63**	**0.43**	**0.95**	**-3.80**	**0.43**	**0.95**	**-6.00**	**0.44**	**0.95**	**-8.27**	**0.45**	**0.95**	**-1.39**	**0.43**	**0.95**
	β_4	0.5	-0.18	1.82	0.95	-0.24	1.82	0.95	-0.40	1.82	0.95	-0.57	1.83	0.95	-0.11	1.83	0.95
	β_5	0.3	1.05	0.38	0.96	1.05	0.38	0.96	1.01	0.38	0.96	0.98	0.39	0.96	1.05	0.38	0.96
	β_6	0.5	1.68	1.80	0.93	1.67	1.80	0.94	1.66	1.80	0.94	1.64	1.79	0.94	1.68	1.80	0.94
	C-Index	0.7	-0.03			-0.04			-0.08			-0.13					
	β_1	0.3	1.35	0.41	0.96	1.31	0.41	0.95	1.30	0.41	0.96	1.28	0.41	0.95	1.32	0.41	0.95
	β_2	0.5	-1.22	1.99	0.94	-1.26	1.98	0.94	-1.25	1.99	0.93	-1.30	1.99	0.94	-1.22	1.99	0.94
	β_3	0.3	-1.40	0.43	0.95	-1.42	0.43	0.96	-1.41	0.43	0.95	-1.41	0.43	0.95	-1.39	0.43	0.95
	β_4	0.5	-0.13	1.83	0.95	-0.13	1.82	0.95	-0.16	1.83	0.95	-0.14	1.83	0.95	-0.11	1.83	0.95
	$\boldsymbol{\beta_5}$	**0.3**	**0.08**	**0.38**	**0.96**	**-1.17**	**0.37**	**0.97**	**-3.49**	**0.38**	**0.96**	**-5.78**	**0.40**	**0.95**	**1.05**	**0.38**	**0.96**
	β_6	0.5	1.69	1.80	0.93	1.66	1.80	0.93	1.70	1.80	0.93	1.68	1.80	0.94	1.68	1.80	0.94
	C-Index	0.7	-0.02			-0.04			-0.09			-0.14					
20% Event Rate	β_1	0.3	1.29	0.12	0.95	1.53	0.12	0.95	1.96	0.12	0.95	2.34	0.12	0.95	1.06	0.12	0.95
	β_2	0.5	0.75	0.46	0.94	0.99	0.46	0.94	1.40	0.46	0.94	1.80	0.46	0.94	0.53	0.46	0.94
	$\boldsymbol{\beta_3}$	**0.3**	**-2.63**	**0.12**	**0.94**	**-3.90**	**0.12**	**0.94**	**-6.35**	**0.15**	**0.91**	**-8.58**	**0.17**	**0.86**	**-1.38**	**0.11**	**0.95**
	β_4	0.5	-0.10	0.45	0.95	-0.19	0.45	0.95	-0.42	0.45	0.94	-0.60	0.45	0.95	-0.01	0.45	0.95
	β_5	0.3	-0.14	0.10	0.95	-0.16	0.10	0.95	-0.23	0.10	0.95	-0.28	0.10	0.95	-0.12	0.10	0.96
	β_6	0.5	1.03	0.44	0.95	0.99	0.44	0.95	0.94	0.44	0.95	0.89	0.43	0.95	1.05	0.44	0.95
	C-Index	0.69	-0.03			-0.05			-0.10			-0.15					
	β_1	0.3	1.05	0.12	0.95	1.02	0.12	0.95	0.97	0.12	0.95	0.91	0.12	0.95	1.06	0.12	0.95
	β_2	0.5	0.50	0.46	0.94	0.45	0.46	0.94	0.45	0.46	0.94	0.37	0.45	0.94	0.53	0.46	0.94
	β_3	0.3	-1.41	0.11	0.95	-1.43	0.11	0.95	-1.45	0.11	0.95	-1.47	0.11	0.95	-1.38	0.11	0.95
	β_4	0.5	-0.04	0.45	0.95	-0.07	0.45	0.94	-0.11	0.44	0.95	-0.09	0.45	0.95	-0.01	0.45	0.95
	$\boldsymbol{\beta_5}$	**0.3**	**-1.23**	**0.10**	**0.95**	**-2.40**	**0.10**	**0.95**	**-4.62**	**0.11**	**0.93**	**-6.73**	**0.14**	**0.89**	**-0.12**	**0.10**	**0.96**
	β_6	0.5	1.05	0.44	0.95	1.01	0.44	0.94	0.98	0.44	0.95	0.93	0.44	0.95	1.05	0.44	0.95
	C-Index	0.69	-0.02			-0.05			-0.10			-0.15					

Table 3.3: Simulation Results for Continuous Covariates with Case-Cohort Sampling Chart Review

			5% Error Rate			10% Error Rate			20% Error Rate			30% Error Rate			No Error		
		Truth	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP
5% Event Rate	β_1	0.3	1.54	0.41	0.95	1.76	0.41	0.95	2.21	0.41	0.96	2.62	0.41	0.96	1.32	0.41	0.95
	β_2	0.5	-0.99	1.99	0.93	-0.77	1.98	0.93	-0.34	1.98	0.94	0.06	1.98	0.94	-1.22	1.99	0.94
	$\boldsymbol{\beta_3}$	**0.3**	**-2.54**	**0.42**	**0.95**	**-3.59**	**0.42**	**0.94**	**-5.81**	**0.42**	**0.95**	**-7.84**	**0.43**	**0.95**	**-1.39**	**0.43**	**0.95**
	β_4	0.5	-0.18	1.83	0.95	-0.23	1.83	0.95	-0.41	1.83	0.95	-0.53	1.83	0.95	-0.11	1.83	0.95
	β_5	0.3	1.04	0.38	0.96	1.05	0.38	0.96	1.01	0.38	0.96	0.99	0.39	0.96	1.05	0.38	0.96
	β_6	0.5	1.69	1.79	0.94	1.68	1.80	0.94	1.66	1.80	0.94	1.64	1.80	0.94	1.68	1.80	0.94
	C-Index	0.7	-0.01			-0.02			-0.05			-0.07					
	β_1	0.3	1.33	0.41	0.96	1.30	0.41	0.95	1.31	0.41	0.95	1.30	0.41	0.95	1.32	0.41	0.95
	β_2	0.5	-1.23	1.99	0.93	-1.27	1.99	0.94	-1.24	1.99	0.93	-1.28	1.99	0.94	-1.22	1.99	0.94
	β_3	0.3	-1.40	0.43	0.95	-1.40	0.43	0.95	-1.39	0.43	0.95	-1.37	0.43	0.95	-1.39	0.43	0.95
	β_4	0.5	-0.13	1.83	0.95	-0.12	1.82	0.95	-0.14	1.83	0.95	-0.09	1.83	0.95	-0.11	1.83	0.95
	$\boldsymbol{\beta_5}$	**0.3**	**-0.10**	**0.38**	**0.97**	**-1.15**	**0.37**	**0.97**	**-3.13**	**0.36**	**0.96**	**-5.09**	**0.36**	**0.97**	**1.05**	**0.38**	**0.96**
	β_6	0.5	1.70	1.80	0.93	1.69	1.80	0.93	1.70	1.80	0.93	1.72	1.80	0.94	1.68	1.80	0.94
	C-Index	0.7	-0.01			-0.02			-0.05			-0.07					
20% Event Rate	β_1	0.3	1.29	0.12	0.95	1.53	0.12	0.95	1.95	0.12	0.95	2.32	0.12	0.95	1.06	0.12	0.95
	β_2	0.5	0.74	0.46	0.94	0.98	0.46	0.94	1.39	0.46	0.94	1.78	0.46	0.94	0.53	0.46	0.94
	$\boldsymbol{\beta_3}$	**0.3**	**-2.57**	**0.12**	**0.95**	**-3.77**	**0.12**	**0.94**	**-6.10**	**0.14**	**0.92**	**-8.18**	**0.17**	**0.87**	**-1.38**	**0.11**	**0.95**
	β_4	0.5	-0.10	0.45	0.95	-0.19	0.45	0.95	-0.38	0.45	0.95	-0.54	0.45	0.95	-0.01	0.45	0.95
	β_5	0.3	-0.13	0.10	0.95	-0.15	0.10	0.95	-0.21	0.10	0.95	-0.23	0.10	0.95	-0.12	0.10	0.96
	β_6	0.5	1.03	0.44	0.95	1.01	0.44	0.95	0.96	0.44	0.95	0.93	0.44	0.95	1.05	0.44	0.95
	C-Index	0.69	-0.02			-0.04			-0.08			-0.12					
	β_1	0.3	1.05	0.12	0.95	1.04	0.12	0.95	0.98	0.12	0.95	0.95	0.12	0.95	1.06	0.12	0.95
	β_2	0.5	0.51	0.46	0.94	0.45	0.46	0.94	0.47	0.46	0.94	0.40	0.45	0.94	0.53	0.46	0.94
	β_3	0.3	-1.41	0.11	0.95	-1.41	0.11	0.95	-1.43	0.11	0.95	-1.44	0.11	0.95	-1.38	0.11	0.95
	β_4	0.5	-0.02	0.45	0.95	-0.05	0.45	0.94	-0.10	0.45	0.95	-0.05	0.45	0.95	-0.01	0.45	0.95
	$\boldsymbol{\beta_5}$	**0.3**	**-1.19**	**0.10**	**0.96**	**-2.28**	**0.10**	**0.95**	**-4.43**	**0.11**	**0.93**	**-6.37**	**0.13**	**0.90**	**-0.12**	**0.10**	**0.96**
	β_6	0.5	1.04	0.44	0.95	1.02	0.44	0.95	1.00	0.44	0.95	0.96	0.44	0.95	1.05	0.44	0.95
	C-Index	0.69	-0.02			-0.04			-0.08			-0.12					

Table 3.4: Simulation Results for Binary Covariates

			5% Error Rate			10% Error Rate			20% Error Rate			30% Error Rate			No Error		
		Truth	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP
5% Event Rate	β_1	0.3	2.34	0.42	0.96	3.29	0.43	0.95	4.95	0.44	0.95	6.52	0.46	0.94	1.32	0.41	0.95
	β_2	0.5	-0.28	1.99	0.94	0.70	1.99	0.94	2.29	2.00	0.94	3.82	2.04	0.94	-1.22	1.99	0.94
	β_3	0.3	-1.75	0.43	0.95	-2.21	0.43	0.95	-2.85	0.43	0.95	-3.31	0.44	0.94	-1.39	0.43	0.95
	$\boldsymbol{\beta_4}$	**0.5**	**-3.50**	**2.00**	**0.94**	**-7.06**	**1.91**	**0.93**	**-11.54**	**2.24**	**0.92**	**-17.04**	**2.85**	**0.90**	**-0.11**	**1.83**	**0.95**
	β_5	0.3	1.11	0.39	0.96	0.99	0.39	0.96	1.00	0.38	0.96	0.92	0.39	0.96	1.05	0.38	0.96
	β_6	0.5	1.72	1.80	0.94	1.62	1.80	0.94	1.59	1.81	0.94	1.53	1.82	0.93	1.68	1.80	0.94
	C-Index	0.7	-0.09			-0.21			-0.36			-0.54					
	β_1	0.3	1.30	0.41	0.96	1.27	0.41	0.96	1.20	0.41	0.95	1.19	0.41	0.96	1.32	0.41	0.95
	β_2	0.5	-1.19	1.99	0.93	-1.29	1.99	0.94	-1.35	2.00	0.93	-1.35	2.00	0.93	-1.22	1.99	0.94
	β_3	0.3	-1.41	0.43	0.95	-1.47	0.43	0.95	-1.55	0.43	0.95	-1.53	0.43	0.95	-1.39	0.43	0.95
	β_4	0.5	-0.14	1.83	0.95	-0.22	1.83	0.95	-0.12	1.82	0.95	-0.23	1.83	0.95	-0.11	1.83	0.95
	β_5	0.3	1.01	0.38	0.96	0.95	0.38	0.97	0.90	0.39	0.96	0.90	0.38	0.96	1.05	0.38	0.96
	$\boldsymbol{\beta_6}$	**0.5**	**-1.29**	**1.85**	**0.93**	**-3.56**	**1.87**	**0.95**	**-8.47**	**2.10**	**0.94**	**-15.54**	**2.90**	**0.90**	**1.68**	**1.80**	**0.94**
	C-Index	0.7	-0.11			-0.21			-0.38			-0.63					
20% Event Rate	β_1	0.3	1.94	0.12	0.95	2.76	0.12	0.94	4.29	0.13	0.92	5.73	0.15	0.91	1.06	0.12	0.95
	β_2	0.5	1.33	0.46	0.94	2.14	0.47	0.93	3.66	0.48	0.93	5.05	0.52	0.92	0.53	0.46	0.94
	β_3	0.3	-1.78	0.11	0.95	-2.25	0.11	0.95	-2.98	0.12	0.94	-3.64	0.12	0.94	-1.38	0.11	0.95
	$\boldsymbol{\beta_4}$	**0.5**	**-2.87**	**0.49**	**0.94**	**-5.87**	**0.54**	**0.93**	**-10.90**	**0.77**	**0.88**	**-15.66**	**1.11**	**0.81**	**-0.01**	**0.45**	**0.95**
	β_5	0.3	-0.11	0.10	0.96	-0.30	0.10	0.96	-0.41	0.10	0.96	-0.62	0.10	0.96	-0.12	0.10	0.96
	β_6	0.5	1.01	0.43	0.95	0.84	0.44	0.95	0.66	0.44	0.94	0.47	0.44	0.94	1.05	0.44	0.95
	C-Index	0.69	-0.11			-0.21			-0.40			-0.58					
	β_1	0.3	0.99	0.12	0.95	0.91	0.12	0.95	0.73	0.11	0.95	0.62	0.12	0.94	1.06	0.12	0.95
	β_2	0.5	0.47	0.46	0.94	0.38	0.46	0.94	0.22	0.46	0.94	0.07	0.46	0.94	0.53	0.46	0.94
	β_3	0.3	-1.46	0.11	0.95	-1.56	0.11	0.95	-1.75	0.11	0.94	-1.89	0.11	0.95	-1.38	0.11	0.95
	β_4	0.5	-0.09	0.45	0.95	-0.26	0.45	0.94	-0.29	0.45	0.94	-0.56	0.45	0.94	-0.01	0.45	0.95
	β_5	0.3	-0.21	0.10	0.96	-0.31	0.10	0.96	-0.46	0.10	0.95	-0.62	0.10	0.96	-0.12	0.10	0.96
	$\boldsymbol{\beta_6}$	**0.5**	**-1.24**	**0.44**	**0.95**	**-3.69**	**0.49**	**0.94**	**-7.70**	**0.66**	**0.90**	**-12.70**	**0.93**	**0.86**	**1.05**	**0.44**	**0.95**
	C-Index	0.69	-0.11			-0.21			-0.39			-0.60					

Table 3.5: Simulation Results for Binary Covariates with Random Sampling Chart Review

			5% Error Rate			10% Error Rate			20% Error Rate			30% Error Rate			No Error		
		Truth	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP
5% Event Rate	β_1	0.3	2.24	0.42	0.96	3.11	0.42	0.95	4.61	0.44	0.95	6.07	0.45	0.94	1.32	0.41	0.95
	β_2	0.5	-0.38	1.99	0.94	0.51	1.99	0.94	1.97	1.99	0.94	3.42	2.03	0.94	-1.22	1.99	0.94
	β_3	0.3	-1.72	0.43	0.95	-2.13	0.43	0.95	-2.73	0.43	0.95	-3.14	0.44	0.94	-1.39	0.43	0.95
	$\boldsymbol{\beta_4}$	**0.5**	**-3.13**	**1.97**	**0.94**	**-6.52**	**1.89**	**0.93**	**-10.79**	**2.17**	**0.93**	**-15.64**	**2.65**	**0.91**	**-0.11**	**1.83**	**0.95**
	β_5	0.3	1.11	0.39	0.96	1.00	0.39	0.96	0.99	0.38	0.96	0.92	0.39	0.96	1.05	0.38	0.96
	β_6	0.5	1.72	1.80	0.94	1.60	1.79	0.94	1.60	1.80	0.94	1.55	1.81	0.94	1.68	1.80	0.94
	C-Index	0.7	-0.08			-0.20			-0.34			-0.50					
	β_1	0.3	1.30	0.41	0.95	1.29	0.41	0.96	1.21	0.41	0.95	1.22	0.41	0.96	1.32	0.41	0.95
	β_2	0.5	-1.22	1.99	0.93	-1.28	1.99	0.94	-1.36	1.99	0.94	-1.37	2.00	0.93	-1.22	1.99	0.94
	β_3	0.3	-1.41	0.43	0.95	-1.45	0.43	0.95	-1.56	0.43	0.95	-1.52	0.43	0.95	-1.39	0.43	0.95
	β_4	0.5	-0.13	1.83	0.95	-0.20	1.83	0.95	-0.11	1.81	0.95	-0.24	1.83	0.95	-0.11	1.83	0.95
	β_5	0.3	1.02	0.38	0.96	0.96	0.38	0.96	0.90	0.39	0.96	0.91	0.39	0.96	1.05	0.38	0.96
	$\boldsymbol{\beta_6}$	**0.5**	**-0.99**	**1.83**	**0.94**	**-3.14**	**1.85**	**0.95**	**-7.93**	**2.09**	**0.93**	**-13.70**	**2.69**	**0.91**	**1.68**	**1.80**	**0.94**
	C-Index	0.7	-0.10			-0.19			-0.35			-0.58					
20% Event Rate	β_1	0.3	1.86	0.12	0.95	2.61	0.12	0.94	4.00	0.13	0.93	5.30	0.14	0.92	1.06	0.12	0.95
	β_2	0.5	1.25	0.46	0.94	1.99	0.47	0.93	3.37	0.48	0.93	4.64	0.51	0.93	0.53	0.46	0.94
	β_3	0.3	-1.75	0.11	0.95	-2.17	0.11	0.95	-2.85	0.12	0.95	-3.44	0.12	0.94	-1.38	0.11	0.95
	$\boldsymbol{\beta_4}$	**0.5**	**-2.63**	**0.49**	**0.94**	**-5.43**	**0.54**	**0.93**	**-9.92**	**0.72**	**0.88**	**-14.38**	**1.02**	**0.83**	**-0.01**	**0.45**	**0.95**
	β_5	0.3	-0.11	0.10	0.96	-0.29	0.10	0.96	-0.40	0.10	0.95	-0.58	0.10	0.96	-0.12	0.10	0.96
	β_6	0.5	1.00	0.44	0.95	0.85	0.44	0.95	0.69	0.44	0.95	0.53	0.44	0.94	1.05	0.44	0.95
	C-Index	0.69	-0.10			-0.19			-0.36			-0.53					
	β_1	0.3	0.99	0.12	0.95	0.93	0.12	0.95	0.77	0.11	0.95	0.68	0.12	0.94	1.06	0.12	0.95
	β_2	0.5	0.46	0.46	0.94	0.38	0.46	0.94	0.24	0.46	0.94	0.10	0.45	0.94	0.53	0.46	0.94
	β_3	0.3	-1.46	0.11	0.95	-1.54	0.11	0.95	-1.72	0.11	0.94	-1.84	0.11	0.95	-1.38	0.11	0.95
	β_4	0.5	-0.08	0.45	0.94	-0.24	0.45	0.94	-0.26	0.45	0.94	-0.51	0.45	0.94	-0.01	0.45	0.95
	β_5	0.3	-0.19	0.10	0.96	-0.29	0.10	0.96	-0.43	0.10	0.96	-0.57	0.10	0.96	-0.12	0.10	0.96
	$\boldsymbol{\beta_6}$	**0.5**	**-1.06**	**0.44**	**0.95**	**-3.31**	**0.47**	**0.95**	**-6.85**	**0.61**	**0.92**	**-11.34**	**0.82**	**0.87**	**1.05**	**0.44**	**0.95**
	C-Index	0.69	-0.10			-0.19			-0.36			-0.54					

Table 3.6: Simulation Results for Binary Covariates with Case-Cohort Sampling Chart Review

			5% Error Rate			10% Error Rate			20% Error Rate			30% Error Rate			No Error		
		Truth	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP	% Bias	$\frac{MSE}{0.01}$	CP
5% Event Rate	β_1	0.3	-0.02	0.41	0.95	-1.36	0.41	0.95	-3.92	0.42	0.95	-6.47	0.45	0.94	1.32	0.41	0.95
	β_2	0.5	-2.54	2.00	0.94	-3.85	2.03	0.93	-6.25	2.08	0.93	-8.63	2.17	0.93	-1.22	1.99	0.94
	β_3	0.3	-0.94	0.43	0.95	-0.66	0.43	0.95	-0.05	0.42	0.95	0.55	0.43	0.95	-1.39	0.43	0.95
	$\boldsymbol{\beta_4}$	**0.5**	**13.14**	**2.27**	**0.92**	**26.27**	**3.54**	**0.83**	**54.58**	**9.23**	**0.45**	**84.49**	**19.64**	**0.12**	**-0.11**	**1.83**	**0.95**
	β_5	0.3	1.11	0.39	0.96	0.91	0.39	0.97	0.78	0.39	0.96	0.41	0.39	0.96	1.05	0.38	0.96
	β_6	0.5	1.72	1.80	0.93	1.55	1.79	0.95	1.37	1.80	0.94	0.98	1.81	0.94	1.68	1.80	0.94
	C-Index	0.7	0.35			0.75			1.68			2.77					
	β_1	0.3	1.31	0.41	0.95	1.29	0.41	0.95	1.07	0.41	0.95	0.99	0.41	0.95	1.32	0.41	0.95
	β_2	0.5	-1.23	2.00	0.93	-1.36	1.99	0.93	-1.59	2.00	0.93	-1.87	2.01	0.94	-1.22	1.99	0.94
	β_3	0.3	-1.43	0.43	0.95	-1.47	0.43	0.95	-1.70	0.43	0.95	-1.92	0.43	0.95	-1.39	0.43	0.95
	β_4	0.5	-0.16	1.83	0.95	-0.30	1.83	0.95	-0.34	1.82	0.95	-0.67	1.84	0.95	-0.11	1.83	0.95
	β_5	0.3	1.01	0.38	0.97	0.96	0.38	0.96	0.87	0.39	0.96	0.67	0.39	0.96	1.05	0.38	0.96
	$\boldsymbol{\beta_6}$	**0.5**	**13.78**	**2.26**	**0.92**	**26.42**	**3.54**	**0.82**	**52.79**	**8.77**	**0.46**	**80.95**	**18.19**	**0.14**	**1.68**	**1.80**	**0.94**
	C-Index	0.7	0.35			0.74			1.63			2.68					
20% Event Rate	β_1	0.3	1.00	0.12	0.95	0.89	0.12	0.95	0.68	0.12	0.95	0.46	0.12	0.94	1.06	0.12	0.95
	β_2	0.5	0.43	0.46	0.94	0.35	0.46	0.94	0.21	0.45	0.94	-0.02	0.46	0.93	0.53	0.46	0.94
	β_3	0.3	-1.46	0.11	0.94	-1.59	0.11	0.95	-1.78	0.11	0.95	-2.05	0.11	0.94	-1.38	0.11	0.95
	$\boldsymbol{\beta_4}$	**0.5**	**3.52**	**0.50**	**0.93**	**7.16**	**0.60**	**0.91**	**15.54**	**1.07**	**0.76**	**25.09**	**2.04**	**0.55**	**-0.01**	**0.45**	**0.95**
	β_5	0.3	-0.12	0.10	0.96	-0.28	0.10	0.96	-0.42	0.10	0.96	-0.77	0.10	0.96	-0.12	0.10	0.96
	β_6	0.5	1.00	0.43	0.95	0.87	0.44	0.95	0.63	0.44	0.94	0.34	0.44	0.95	1.05	0.44	0.95
	C-Index	0.69	0.09			0.18			0.38			0.61					
	β_1	0.3	0.99	0.12	0.95	0.89	0.12	0.95	0.72	0.11	0.95	0.51	0.12	0.95	1.06	0.12	0.95
	β_2	0.5	0.46	0.46	0.94	0.37	0.46	0.94	0.15	0.46	0.94	-0.10	0.46	0.94	0.53	0.46	0.94
	β_3	0.3	-1.43	0.11	0.94	-1.52	0.11	0.95	-1.73	0.11	0.94	-1.93	0.11	0.95	-1.38	0.11	0.95
	β_4	0.5	-0.08	0.45	0.94	-0.22	0.45	0.94	-0.29	0.45	0.95	-0.65	0.45	0.94	-0.01	0.45	0.95
	β_5	0.3	-0.19	0.10	0.96	-0.29	0.10	0.95	-0.38	0.10	0.95	-0.63	0.10	0.96	-0.12	0.10	0.96
	$\boldsymbol{\beta_6}$	**0.5**	**4.58**	**0.49**	**0.93**	**8.43**	**0.62**	**0.90**	**17.23**	**1.22**	**0.73**	**26.27**	**2.19**	**0.51**	**1.05**	**0.44**	**0.95**
	C-Index	0.69	0.08			0.18			0.40			0.60					

Chapter 4

Application to Evaluating the SOFA Score Using EHR Data

4.1 SMART Dataset

Data has been obtained from the Isotonic Solutions and Major Adverse Renal Events Trial (SMART), which consists of 15,802 adult patients from five intensive care units at Vanderbilt University Medical Center. This was a pragmatic, unblinded, cluster-randomized, multiple-crossover trial comparing clinical outcomes after the use of balanced crystalloids or saline for intravenous fluid administration in critically ill adults (Semler et al., 2018). For this paper, a random sample of 300 observations was used, which includes demographic information as well as the eSOFA score and mSOFA score for the six components of the SOFA score. There were 39 cases in this sample for an event rate of 39/300 = 13%. Descripive statistics for this dataset can be found in Table 4.1. The cases were older, and there were a higher proportion of males in the cases than the non-cases. Also, the overall SOFA scores, both eSOFA and mSOfA, were higher for the cases. The neurologic component had the largest difference between the cases and non-cases.

4.2 Comparison of eSOFA Score and mSOFA Score

An eSOFA score has been developed that can extract data from EHR using natural language processing. This score has been calculated for all patients in the SMART trial. A benefit of using the eSOFA score is that it can be calculated much faster than the mSOFA score. However, there will be errors present when the eSOFA score doesn't match the mSOFA score. The mSOFA score has been calculated for a random sample of 300 patients from the SMART trial.

The bias between the mSOFA score and the eSOFA score was found by subtracting the mSOFA scores from the eSOFA scores for each patient in the sample. The bias was calculated for each component of the SOFA score. The results can be found in Figure 4.1. Note that this figure was produced with jitter to more easily see how many patients were at each point. The respiratory

Table 4.1: Descriptive Statistics for SMART Data

	Non-Cases	Cases	Combined
	$N = 261$	$N = 39$	$N = 300$
Age	44.81 58.71 71.12	52.52 64.71 69.44	45.00 59.14 71.12
Race			
Asian	0.01 (2)	0.00 (0)	0.01 (2)
Black	0.10 (25)	0.10 (4)	0.10 (29)
Unknown	0.07 (17)	0.18 (7)	0.08 (24)
White	0.83 (217)	0.72 (28)	0.82 (245)
Gender			
Male	0.56 (145)	0.67 (26)	0.57 (171)
Unit			
CVICU	0.18 (46)	0.23 (9)	0.18 (55)
MICU	0.34 (90)	0.46 (18)	0.36 (108)
NEICU	0.18 (48)	0.21 (8)	0.19 (56)
SICU	0.09 (24)	0.03 (1)	0.08 (25)
TICU	0.20 (53)	0.08 (3)	0.19 (56)
Source of Admission			
Emergency department	0.54 (140)	0.46 (18)	0.53 (158)
Hospital ward	0.07 (18)	0.15 (6)	0.08 (24)
Operating room	0.19 (50)	0.00 (0)	0.17 (50)
Other	0.07 (17)	0.03 (1)	0.06 (18)
Transfer from another hospital	0.14 (36)	0.36 (14)	0.17 (50)
eSOFA	3 4 7	6 8 12	3 5 8
mSOFA	2 4 7	5.5 8 12	2 5 8
eSOFA (Cardiovascular)	0 1 1	0 1 3	0 1 1
mSOFA (Cardiovascular)	0 1 1	0 1 3	0 1 1
eSOFA (Respiratory)	1 1 2	1 2 3	1 2 2
mSOFA (Respiratory)	1 2 2	1 2 3	1 2 2
eSOFA (Renal)	0 0 3	0 1 3	0 0 3
mSOFA (Renal)	0 0 1	0 1 3	0 0 1
eSOFA (Hepatic)	0 0 0	0 0 1.5	0 0 0
mSOFA (Hepatic)	0 0 0	0 0 1.5	0 0 0
eSOFA (Neurologic)	0 1 2	0.5 4 4	0 1 3
mSOFA (Neurologic)	0 1 2	0.5 3 4	0 1 3
eSOFA (Hematological)	0 0 0	0 0 1	0 0 0
mSOFA (Hematological)	0 0 0	0 0 1	0 0 0

$_a$ b $_c$ represent the lower quartile a, the median b, and the upper quartile c for continuous variables. Numbers after proportions are frequencies.

and renal components have the most bias, while the hepatic and neurologic components have the least. Table 4.2 shows a summary of the distributions of the bias for each of the components. The respiratory component had the most patients (n=79, 26%) with disagreement between the mSOFA and eSOFA scores, while the hepatic and neurologic components had the least (n=1, 0.33%). For the cardiovascular, respiratory, and renal components, there was bias in both the positive and negative directions, whereas the four instances of bias in the hepatic, neurologic, and hematological components were all in the positive direction, meaning that the eSOFA score was lower than the mSOFA score.

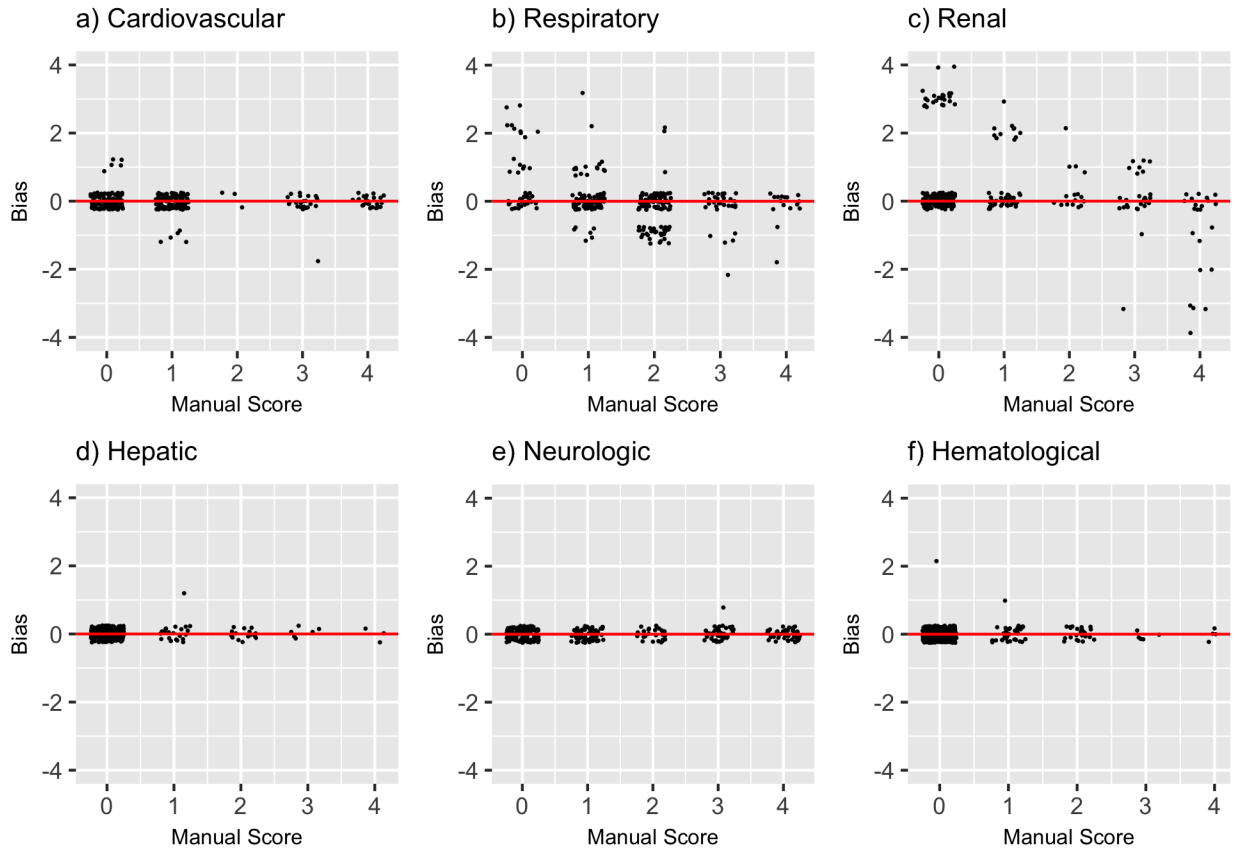

Figure 4.1: Bias Between mSOFA and eSOFA Scores

Table 4.2: Summary of Errors for SOFA Score Components

	# with error (%)	Mean	S.D.	Min.	Max.
Cardiovascular	12 (4%)	-0.02	0.32	-4	1
Respiratory	79 (26%)	0.01	0.68	-2	3
Renal	63 (21%)	0.33	1.17	-4	4
Hepatic	1 (0.33%)	0.00	0.06	0	1
Neurologic	1 (0.33%)	0.00	0.06	0	1
Hematological	(0.67%) 2	0.01	0.13	0	2

4.3 Application of Chart Review Strategies

A Cox proportional hazards model was fit using the mSOFA components of the SOFA score to obtain the gold standard coefficients. These can be found in the "Truth" column of Table 4.3. The model was also fit using the eSOFA score components without correcting any errors; the percent bias for these estimates can be found in the "No Correction" column of Table 4.3. Next, the chart review sampling strategies were used to correct errors by replacing the eSOFA score with the mSOFA score for the sampled patients. Two different sample sizes were examined for the random sampling (N=100 and N=150). The case-cohort sampling strategy corrected errors for the 39 cases and a random sample of 100-39 = 61 non-cases. The model was fit 100 times using each of these strategies, and the percent bias for the effect estimate corresponding to each of the SOFA score components as well as the percent bias for the C-index was obtained in each of the repetitions. The average of these percent biases after correcting errors using each of the chart review strategies can be found in Table 4.3.

Table 4.3: Cox PH Model Results

		% Bias			
	Truth	No Correction	Random Sampling (N=100)	Random Sampling (N=150)	Case-Cohort Sampling (N=100)
β_1 (Cardiovascular)	-0.09	-5.11	-8.91	-11.29	-58.79
β_2 (Respiratory)	0.16	-30.49	-19.21	-12.30	-1.33
β_3 (Renal)	0.21	-77.17	-54.92	-47.35	-73.19
β_4 (Hepatic)	0.68	4.88	2.93	2.28	1.07
β_5 (Neurologic)	0.53	6.58	3.35	2.18	-2.08
β_6 (Hematological)	-0.36	-11.23	-10.40	-8.98	-10.01
C-Index	0.79	-0.40	-0.34	-0.25	-0.90

The random sampling with 150 patients seemed to perform the best out of the three chart review strategies examined in terms of percent bias. This method resulted in a lower magnitude percent bias compared to the model fit using the eSOFA score without correcting errors for the coefficients β_2 through β_6. However, for β_1, the percent bias after performing chart reviews on a random sample of 150 patients was -11.29 compared to a percent bias of -5.11 when fitting the model without correcting errors and a percent bias of -8.91 when fitting the model after correcting errors for a random sample of 100 patients. Similarly, the percent bias for the estimates from the model after correcting for errors using chart reviews with case-cohort sampling compared to no correction had a smaller magnitude for all of the coefficients β_2 through β_6, but a larger magnitude for β_1 (-58.79 vs. -5.11). The C-index from fitting the model with the mSOFA components was 0.79, and there was less than 1% bias in all scenarios.

Chapter 5

Discussion

This paper evaluated the impact of data quality on risk factor assessment and risk prediction performance and investigated different chart review strategies in reducing bias in the context of a Cox proportional hazards model. Unlike traditional utilities of chart review where only chart reviewed data are used in the analysis, this work aims to maximize data resources using both chart reviewed and non-chart reviewed records.

The simulation studies presented in this paper revealed that data quality problems result in bias of risk factor effects under a Cox proportional hazards model framework as expected. Specifically, a greater proportion of errors results in more bias of the effects regardless of the event rate or whether the covariate of interest is continuous or binary, correlated or uncorrelated with other covariates. For continuous covariates, performing chart review reduces the bias of the effect estimates, with case-cohort sampling performing slightly better than random sampling. For the binary variables, the random sampling strategy resulted in improved estimates in terms of percent bias. The case-cohort sampling strategy for the binary covariates resulted in larger bias due to the over-representative cases where the value of 0 (representing absence of a risk factor) was corrected to 1 (representing presence of a risk factor). As expected, the simulation studies also revealed that the proportion of bias reduction after performing chart reviews may be linked to the proportion of observations being reviewed. For example, in the simulation studies where 10% of the observations were reviewed, the bias improved by about 5% to 15% after correcting errors using chart reviews.

For the application using the SOFA score, we were only able to obtain a random sample of 300 patients from the 15,802 patients enrolled in the SMART study. Performing chart review on a random sample of 150 patients resulted in the smallest percent bias of the effect estimates. Applying these methods to the entire dataset is a potential future analysis that would be more representative of how these chart review sampling strategies would behave in real EHR data.

This paper is our first attempt in the broad area of EHR data quality problems in a risk prediction model framework. With better understanding of the bias in model estimation and performance due to data errors and potential utilities of chart review in bias reduction, we identified opportunities to develop statistical methods to address data quality problems through chart review. Future statistical methods development will focus on areas such as accounting for chart review sampling probability, combining data error modeling with non-chart reviewed records to further reduce bias, and identifying optimal sampling strategies for chart review.

BIBLIOGRAPHY

Alonso, A., Krijthe, B. P., Aspelund, T., Stepas, K. A., Pencina, M. J., Moser, C. B., Sinner, M. F., Sotoodehnia, N., Fontes, J. D., Janssens, A. C. J. et al. (2013), Simple risk model predicts incidence of atrial fibrillation in a racially and geographically diverse population: the charge-af consortium, *Journal of the American Heart Association* **2**(2), e000102.

Goldstein, B. A., Navar, A. M. and Pencina, M. J. (2016), Risk prediction with electronic health records: the importance of model validation and clinical context, *JAMA Cardiology* **1**(9), 976–977.

Harrell, F. (2015), *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*, 2nd edn, Springer.

Hippisley-Cox, J., Coupland, C., Vinogradova, Y., Robson, J., Minhas, R., Sheikh, A. and Brindle, P. (2008), Predicting cardiovascular risk in england and wales: prospective derivation and validation of qrisk2, *BMJ* **336**(7659), 1475–1482.

Kirby, J. C., Speltz, P., Rasmussen, L. V., Basford, M., Gottesman, O., Peissig, P. L., Pacheco, J. A., Tromp, G., Pathak, J., Carrell, D. S. et al. (2016), Phekb: a catalog and workflow for creating electronic phenotype algorithms for transportability, *Journal of the American Medical Informatics Association* **23**(6), 1046–1052.

Kolek, M. J., Graves, A. J., Xu, M., Bian, A., Teixeira, P. L., Shoemaker, M. B., Parvez, B., Xu, H., Heckbert, S. R., Ellinor, P. T. et al. (2016), Evaluation of a prediction model for the development of atrial fibrillation in a repository of electronic medical records, *JAMA Cardiology* **1**(9), 1007–1013.

Prentice, R. L. (1986), A case-cohort design for epidemiologic cohort studies and disease prevention trials, *Biometrika* **73**(1), 1–11.

Riley, R. D., Ensor, J., Snell, K. I., Debray, T. P., Altman, D. G., Moons, K. G. and Collins, G. S. (2016), External validation of clinical prediction models using big datasets from e-health records or ipd meta-analysis: opportunities and challenges, *BMJ* **353**, i3140.

Semler, M. W., Self, W. H., Wanderer, J. P., Ehrenfeld, J. M., Wang, L., Byrne, D. W., Stollings, J. L., Kumar, A. B., Hughes, C. G., Hernandez, A. et al. (2018), Balanced crystalloids versus saline in critically ill adults, *New England Journal of Medicine* **378**(9), 829–839.

Singer, M., Deutschman, C. S., Seymour, C. W., Shankar-Hari, M., Annane, D., Bauer, M., Bellomo, R., Bernard, G. R., Chiche, J.-D., Coopersmith, C. M. et al. (2016), The third international consensus definitions for sepsis and septic shock (sepsis-3), *JAMA* **315**(8), 801–810.

Vincent, J.-L., Moreno, R., Takala, J., Willatts, S., De Mendonça, A., Bruining, H., Reinhart, C., Suter, P. and Thijs, L. (1996), The sofa (sepsis-related organ failure assessment) score to describe organ dysfunction/failure, *Intensive Care Medicine* **22**(7), 707–710.

Worster, A. and Haines, T. (2004), Advanced statistics: understanding medical record review (mrr) studies, *Academic Emergency Medicine* **11**(2), 187–192.