Aman Verma

Hospital readmissions: prediction and inference

This thesis is submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy in Epidemiology.

Department of Epidemiology, Biostatistics and Occupational Health
McGill University, Montreal, Canada
December 2014

Abstract

Hospital readmission rates vary widely across regions and hospitals, suggesting that improvements are possible. To provide an incentive to improve quality of care, some jurisdictions have introduced legislation that financially penalizes hospitals with high readmission rates, but the models used to implement the legislation crudely adjusts for patient-mix. Additionally, hospitals have developed predictive models of readmission risk to better target enhanced transitional care. In this work, I examined how large healthcare administrative databases can help build better inferential and predictive models of hospital readmissions.

To target interventions at those patients with the highest readmission risk, hospitals can develop predictive models of readmission based on their own data (local models), they can pool their data with other hospitals (global models) or they can use sophisticated model combination techniques which avoid directly sharing patient data (combined models). In the first manuscript, I compared the accuracy of global, combined, and local models in predicting 30-day readmission risk, and found that the predictive accuracy of models developed with the three approaches were similar, suggesting that hospitals can use their own data to accurately predict hospital readmissions.

Although predictive models of hospital readmissions can be useful to guide resources to individual high-risk cases, inferential models can potentially lead to population-level interventions. In the second manuscript, I studied how the day-of-week of discharge affects readmission, and used both empiric (survival

model) and analytic (Markov model) approaches to study how this effect is confounded by the probability of admission on the weekend. I found that not only are Friday discharges more likely to be readmitted than Wednesday discharges, but also that the low probability of weekend admissions attenuates this effect if uncontrolled. Our results suggest that interventions that reduce the effect of Friday discharge on readmissions, such as increased weekend staffing, are likely to be more cost-effective than previous work has indicated.

In the third manuscript, I compare two techniques to measure the effect of twenty Montreal hospitals on readmissions: a standard regression approach that controls for the major, well-known confounders, and targeted maximum likelihood estimation (TMLE) where I could control for pre-admission diagnoses, procedures, and drug prescriptions using a machine learning technique (random forest). The standard model suggested that there was little difference between the hospitals, but the TMLE model showed that the confounders, particularly drug prescriptions, strongly confounded readmission risk, and revealed a wide variation in readmission risk between the hospitals.

My work suggests that: 1) predictive models of readmission are unlikely to be greatly improved by pooling hospital data or by using complex combination techniques, 2) inference on the causes of readmissions, particularly the day-of-week, can be confounded by the admission process, and 3) by using TMLE, the predictive power of machine learning techniques can be used to improve inference by reducing bias in our estimates of the effect of hospital care on readmissions.

Résumé

Les taux de réadmission à l'hôpital sont très variables selon les régions et les hôpitaux; ce qui suggère que des améliorations sont possibles. Pour les inciter à améliorer la qualité des soins, certains pays ont introduit une législation qui pénalise financièrement les hôpitaux ayant des taux de réadmission élevés, mais les techniques employées n'ajustement que grossièrement pour les caractéristiques des patients. En outre, les hôpitaux ont développé des modèles prédictifs du risque de réadmission afin d'améliorer la qualité des soins de transition entre l'hôpital et la communauté. Dans ce travail, j'ai examiné comment les grandes banques de données administratives en santé peuvent aider à construire de meilleurs modèles d'inférence et de prédictionn des réadmissions à l'hôpital.

Afin de mieux cibler les interventions chez les patients présentant le plus grand risque de réadmission, , les hôpitaux peuvent développer des modèles prédictifs de réadmission sur la base de leurs propres données (modèles locaux), ils peuvent aussi mettre en commun leurs données (modèles globaux) ou ils peuvent encore utiliser des techniques de combinaison faisant appel à des modèles statistiques sophistiqués qui évitent le partage des données entre les hôpitaux (modèles combinés). Dans le premier manuscrit, je comparé la précision des modèles locaux, globaux et combinés à prédire le risque de réadmission dans les 30 jours suivant le congé, et j'ai trouvé que la valeur prédictive de ces trois modèles était similaire; ce qui suggère que les hôpitaux peuvent utiliser leurs propres données pour prédire les réadmissions à l'hôpital sans significativement

réduire la précision de leurs estimés. Bien que les modèles prédictifs de réadmissions à l'hôpital peuvent être utiles pour orienter les ressources vers les cas présentant un risque élevé de réadmission, les modèles d'inférence peuvent potentiellement conduire à des interventions au niveau de la population. Dans le deuxième manuscrit, j' étudie comment le jour du congé affecte la réadmission, et j'utilise à la fois une approche empirique (modèle de survie) et une approche analytique (modèle de Markov) pour étudier comment cet effet est confondu par la probabilité d'admission durant le week-end. Je trouve que non seulement les départs survenant le vendredi sont plus susceptibles d'être réadmis que les départs survenant le mercredi, mais aussi que la faible probabilité d'admissions durant les week-ends atténue cet effet si elle n'est pas contrôlée.

Dans le troisième manuscrit, je compare deux techniques pour mesurer l'effet de vingt hôpitaux de Montréal sur les réadmissions: une approche standard de régression qui contrôle pour les principaux facteurs confondants, et l'estimation ciblée du maximum de vraisemblance (ECMV) où je pouvais contrôler pour les diagnostics, les procédures et les prescriptions de médicaments en utilisant une technique d'apprentissage machine (la forêt aléatoire). Le modèle standard a suggéré qu'il y avait peu de différence entre les hôpitaux, mais le modèle ECMV a montré que les facteurs de confusion, en particulier les prescriptions médicamenteuses, confondent fortement le risque de réadmission, et ont révélé une grande variation du risque de réadmission entre les hôpitaux.

En somme, mon travail suggère que : 1) des modèles prédictifs de réadmission sont peu susceptibles d'être grandement améliorés par la mise en commun des données hospitalières ou par l'usage de techniques sophistiquées de combinaison, 2) l'inférence sur les causes de réadmission, en particulier le jour de la semaine, peuvent être confondus par le processus d'admission, et 3) en utilisant l'ECMV, le pouvoir prédictif des techniques d'apprentissage machine peut être utilisée pour améliorer l'inférence quant à l'effet des soins hospitaliers sur les réadmissions.

Statement of originality

This work contains several original contributions to the analysis and understanding of hospital readmissions. Using epidemiologic analysis, I have precisely identified the reasons why the classification of hospital readmissions as preventable or not preventable is unnecessary, greatly simplifying the analysis of large healthcare databases. I also developed a unique technique to pool hospital data without breaching patient privacy, but found that pooling data in general may not significantly improve accuracy, and that in some cases, it may decrease accuracy. To my knowledge, this work is the first to analytically and empirically investigate how the differing probability of admission on the weekend confounds the effect of the day-of-the-week of discharge on readmission. This work is also the first to combine the predictive accuracy of machine learning methods with causal inference techniques (using targeted maximum likelihood estimation) to develop precise, less biased estimates of the effect of different hospitals on readmission. As a whole, this body of work combines the latest informatics and epidemiologic methods and applies them to vast data sources to develop original contributions to the study of hospital readmissions.

Although I have received guidance from my committee members on the substantive, methodological, and statistical aspects of this thesis, I declare that the conception, execution, and drafting of the work in this thesis were my own.

Contributions of authors

The three manuscripts in this thesis developed out of collaboration between my co-authors and myself. Under the guidance of my supervisor (David Buckeridge) and committee member (Christian Rochefort), I developed the research objectives, conducted literature reviews, performed all data management and statistical analyses, and wrote the first drafts of all manuscripts.

Manuscript 1: Verma AD, Izadi M, Reddy C, Rochefort CM, Hosseinzadeh A, Buckeridge DL. "Privacy-preserving predictions: a case study of hospital readmission"

Manuscript 2: Verma AD, Rochefort CM, Buckeridge DL. "Hospital readmissions and the day of the week"

Manuscript 3: Verma AD, Mamiya H, Schnitzer M, Rochefort CM, Buckeridge DL. "Hospital readmissions and targeted maximum likelihood estimation"

David Buckeridge is an Associate Professor in the Department of Epidemiology & Biostatistics at McGill University. Dr Buckeridge was extensively involved in all three manuscripts as my doctoral supervisor, providing support on the methodological and substantive aspects throughout the three manuscripts.

Arian Hosseinzadeh completed his Master of Computer Science degree while collaborating with me on the first manuscript in this work. Mr Hosseinzadeh contributed to the design and conception of the first manuscript.

Masoumeh Izadi is a Research Associate in the Department of Epidemiology &

Biostatistics at McGill University. Dr Izadi contributed to the conception, design, analysis and writing of the first manuscript.

Hiroshi Mamiya is pursuing a doctoral degree in the Department of Epidemiology & Biostatistics at McGill University. Mr Mamiya provided extensive feedback in the design and analysis of the third manuscript.

Chandan Reddy is an Associate Professor at Wayne State University in the Department of Computer Science. Dr Reddy helped in the design stage and provided comments on the first manuscript.

Christian Rochefort is an Assistant Professor in the Ingram School of Nursing at McGill University and an Associate Member of the Department of Epidemiology, Biostatistics and Occupational Health at McGill University. Dr Rochefort was extensively involved in all three manuscripts, providing both methodological and substantive support.

Mireille Schnitzer is an Assistant Professor of Biostatistics in the Faculty of Pharmacy at l'Université de Montréal. Dr Schnitzer provided extensive input and suggestions on the use of targeted maximum likelihood estimation in the third manuscript.

Abbreviations

AMI acute myocardial infarction. 22, 33, 38, 40, 41, 42, 70, 73, 75, 80, 81, 97

AUC area under the receiver-operating characteristic curve. 33, 40, 46

CCP Canadian Classification of Diagnostic, Therapeutic, and Surgical Procedures. 37, 76

CMA census metropolitan area. 24, 27, 53

CMS Centers for Medicare and Medicaid. 38

DRG diagnostic-related group. 6

FSA forward sortation area. 26

GAM generalized additive model. 25

ICD-9 International Classification of Diseases, 9th revision. 24, 37, 38, 53, 56, 75, 76

ILI influenza-like illness. 24, 25, 57, 113

ISQ *Institut de la statistiques de Québec*. 28

MDC major diagnostic category. 56

MEDÉCHO *Maintenance et exploitation des données pour l'étude de la clientèle hospitalière*. 24, 26, 28

MSSS *Ministère de santé et services sociaux*. 26

PPS prospective payment system. 6

RAMQ *Régie de l'assurance maladie du Québec*. 24, 26, 27, 28, 53, 75

ROC receiver-operating characteristic. 44

SD standard deviation. 42

TMLE targeted maximum likelihood estimation. 1, 22, 23, 72, 73, 74, 81, 85

Acknowledgements

I would like to thank my supervisor David Buckeridge for his mentorship, and for encouraging me to pursue a PhD in epidemiology. I would also like to thank my committee member, Christian Rochefort, for his excellent methodological and substantive reviews of my manuscripts. I am also grateful for the mentorship and financial assistance I received through the 4P (Prevention, Promotion and Public Policy) training program.

Over the past several years I have been lucky to have a very supportive environment in the Clinical and Health Informatics Research Lab, where I completed this dissertation, and made many lasting friendships. I have particularly enjoyed working with Luc, Anya, and Masoumeh, with whom I have enjoyed many academic and non-acadmic discussions. I would also like to thank those that completed and are completing their PhD along with me; Hiroshi, Kathryn, Kate, Alex, and Rolina have all made our lab an excellent place to work and study. I would also like to thank Tewodros (Teddy) for his friendship, patience and support throughout my PhD.

I would also like to thank my parents and my sister for their support. Finally, this thesis would not be possible without the unwavering support and encouragment of my partner Tatiana. I dedicate this thesis to her, and our daughter, who was born within hours of this writing.

Contents

Tables

Figures

Chapter 1

Introduction

> *"Is it too early to call this the Readmission Decade? Readmissions are on everybody's mind—identifying readmissions, preventing readmissions, considering the financial repercussions of having too many readmissions, lamenting the injustice of being held accountable for readmissions. Only time will tell whether this is a passing storm or here to stay, but for the moment the issue of readmissions is having its moment in the sun."*
>
> Karl E. Minges, Living in the Readmission Era, (February 2014) [1]

Recent legislation in both the US [2] and Canada [3] has mandated the public reporting of hospital readmission rates, and has introduced financial penalties for hospitals with high readmission rates. Although there is some evidence that public reporting and financial penalties for readmissions have caused hospitals to develop innovative programs to improve quality of care [4, 5], other authors [6–8] have suggested that readmissions cannot be meaningfully attributed to hospital quality of care, and that many other factors affect readmission rates. Hospitals with excellent quality of care may even be penalized for treating "sicker" patients [9], who are the most likely to be readmitted.

To improve quality of care and reduce readmissions, health systems can apply transitional care interventions, such as using dedicated transitional care

nurses for discharge planning, medication reconciliation, and telehealth monitoring, but these interventions are costly [10–13]. To help target costly transitional care interventions to those at highest risk of readmission, I investigated whether hospitals can develop predictive models of hospital readmission, and whether they can improve these models by combining data from different hospitals. Developing predictive models of readmission can help individual patients, but inferential models of readmission may help identify systemic health system changes that will improve quality of care and reduce readmissions. In this work, I explored how inference on one systemic-level factor, the effect of Friday discharge on readmissions [14], is complicated by the low probability of weekend admission. Recent hospital readmission legislation is designed to incent hospitals to improve quality of care and reduce readmissions by financially penalizing those with high risk-adjusted rates of readmissions, but the risk adjustment only accounts for a few well-known confounders. In this work, I applied targeted maximum likelihood estimation (TMLE) [15], a causal inference modeling technique, to take advantage of the rich confounder data available in hospital administrative databases and provide less biased estimates of hospital readmission risk.

In the following section, I first examine the history of the measurement of hospital readmissions, and trace its development from a cost measure to a measure of quality of care. I then discuss some of the causes of readmissions and review interventions that have been implemented to reduce readmissions, followed by an examination of how the concept of the preventability of readmissions has influenced how they are statistically modeled. I also examine how the inferential capacity of hospital readmission models has been measured by their predictive accuracy, and argue that while the inferential capacity and predictive accuracy are valuable, they are unrelated. Finally, I describe the motivation and objectives for three studies (Chapters 3-5), which explore ways to improve the predictive accuracy and inferential capacity of hospital readmission models.

1.1 Background

1.1.1 History

Although the current study of readmissions is tied to improvements in quality of care, historically, interest in hospital readmissions has tended to wax and wane, peaking during proposed or realized changes to the reimbursement scheme for clinical services [16]. As governments assumed more of the cost of healthcare, they became interested in measuring rising healthcare utilization. Hospital readmissions are a particularly useful indicator because they tracked not only the increases in utilization (and therefore cost) but also quality (people come back less often with better care). Unfortunately, because governments often examined readmissions in the context of rising costs, the instrument itself, which is not a direct measure of health, is often looked at with mistrust.

The standardized collection of data and statistics in these hospitals began with Florence Nightingale's seminal work "Notes on Hospitals"[17]. Nightingale made specific recommendations to the Statistical Society at the time on the standardization of hospital statistics, but readmissions were not considered explicitly in that seminal publication. Given the nature of hospitals at the time, which were unlikely to provide anything more than palliative care, readmissions would have been rare [18]. However, the focus on Nightingale's work was to describe differences in mortality in hospitals, which resembles readmissions in the sense that an easy to measure, but crude proxy for the quality of care was used to directly compare hospitals.

In the mid to late nineteenth century, a new type of hospital emerged: the for-profit acute care hospital [18]. But not all those in need of care were able to make use of the new public hospitals, in particular, those with mental illnesses were "left behind" in what remained a place of charity. By the early 1900s, the governments in both Canada and the US had begun to subsidize the mental institutions, further entrenching the differences between hospitals and asy-

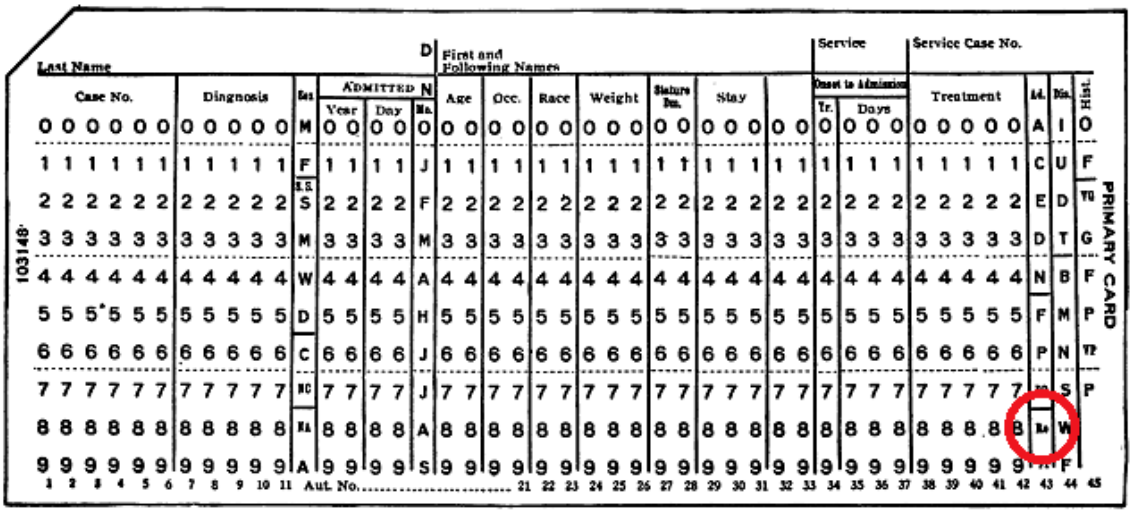

Fig. 4.—First or face card form for mechanical tabulation and indexing routine hospital statistics.

Figure 1.1: An early design for a punchcard to collect routine hospital statistics. This punchcard design, from Pearl[19], is one of the first direct references to hospital readmissions in the context of routine hospital statistics.

lums. The mental institution and the acute care hospital had opposing incentives for utilization. Because mental institutions were supported to some extent by the state, the state was interested in saving money by discharging patients to the population. For-profit acute care hospitals had a financial incentive to encourage people to use their services. Although they still had moral and ethical reasons to ensure that their patients received the best possible care, they were less motivated to explicitly measure what happened to their patients after discharge. Raymond Pearl first suggested the measurement of readmissions in 1921 for acute care hospitals (among other statistics), yet those were not implemented at the time[19]. The first careful reporting of readmissions appears to have been in a study by the New York State Department of Mental Hygiene, explicitly measuring the readmission rate of mental patients at the state institutions[20].

When socialized medicine was introduced in the middle of the century, the payment scheme for physicians changed, and the interest in hospital readmissions

was revived by both opponents and proponents of the new system. Following persecution by the Committee on Un-American Activities, Milton Roemer, a distinguished physician and strong advocate of socialized medicine, fled to Saskatchewan, Canada, where he helped develop the first hospital insurance program[21]. After the establishment of the hospital insurance program, Roemer published the first study of hospital readmissions[22]. The study was superficially a study of utilization, but in the context of the recent adoption of the hospital insurance program (which was vigorously opposed by many physicians), the study can be seen as a defense of the costs of the new program. This abridged excerpt of the discussion illustrates both viewpoints clearly:

> "... much attention is being given to methods of reducing so-called "excessive" hospital utilization. There is much talk of "abuse" of the hospital by patients and doctors, and a common remedy proposed is the imposition of "deterrent fees" ... [A] patient who is hospitalized... once in five years is [hardly] abusing the privilege. Why then should he be deterred by a non-insured fee? An undeterred admission, moreover, can hasten the early detection and prompt treatment of any illness. On the other hand, if he is a hospital-repeater, a frequent user of hospitals, he is likely to be a sufferer from chronic disease. If he is asked to pay a deterrent charge—such as a share of the cost of second or higher admissions in a year—then the chronically ill would bear a disproportionately large share of the over-all [sic] community costs of hospital service. This would vitiate the very purpose of hospitalization insurance, which is designed to spread the risks over the sick and the well alike."

Acheson and Barr [23] were the first to discuss hospital readmissions as a measure of *quality*. The primary concern of their study was to adjust mortality rates; instead of using simple discharges as the denominator for deaths soon after discharge, they would combine two discharges into a single discharge if they occurred soon after each other (a readmission). They note: "It would be attractive

if the readmission rate could be used as an index of the quality of medical care, but further work will be necessary to determine any possible use it may have for this purpose."

In the mid-sixties socialized medicine (Medicare) was introduced in the US and Canada, covering inpatient hospital costs for all people over 65. In the US, rapid rises in Medicare costs in the seventies led to the utilization review, in which Medicare would refuse to pay for hospital stays that were deemed "excessive". These controversial reviews were replaced with the Professional Standards Review Organization (PSRO) that nominally were about quality of care, but had essentially the same cost-control function as a utilization review. The PSROs collected information on every Medicare hospital stay and then developed indicators of performance, such as hospitalization rates, death rates, and readmission.

In 1983, the prospective payment system (PPS) was introduced in the US for hospitals[24]. Instead of being reimbursed for whatever they spent, hospitals would now receive a fixed sum based on the diagnostic-related group (DRG). This payment system meant that a single hospital stay would have a fixed reimbursement, regardless of the length of stay, making two short stays more lucrative than one long stay. Worried that the reimbursement structure of the PPS would incent early hospital discharge, in the late eighties, the US government mandated that the Health Care Financing Administration used early readmission as a quality indicator [25, 26].

Starting in the early 2000s, Quality Improvement Organizations (QIOs) (the former PROs), began publishing indicators for the Hospital Compare website [27], which makes readmission rates available for all hospitals in the US. In Canada, the Canadian Institutes for Health Information (CIHI) publishes the readmission rates, along with several other measures of quality of care. Formally, both organizations examine readmission rates as a means to measure quality, but readmissions are the only indicator that is directly related to healthcare utilization rather than health.

1.1.2 Readmissions in law

In 2010, the Ontario Excellent Care for All Act (ECFAA) received royal assent, which required all hospitals to submit a Quality Improvement Plan (QIP) to Health Quality Ontario (HQO)[28]. The legislation also requires that hospital executive compensation be linked to targets set out in the QIP. For the years 2012/2013, hospitals were required to include an "integrated" dimension in their QIP, which could include hospital readmissions. Twenty-seven hospitals chose the 30-day readmission indicator as a priority within their 2011/12 QIP, and eight more selected a readmission indicator other than the recommended core indicator [28]. Notably, the executive salary penalty associated with failing to reach an indicator was assigned by the executives themselves, with the recommended penalty being from 1 to 3%.

Similarly, also in 2010 the US president signed the Patient Protection and Affordable Care Act, amended by the Health Care and Education Reconciliation Act. The legislation reduces Medicare payments to hospitals with relatively high preventable readmission rates for three selected conditions: acute myocardial infarction, heart failure, and pneumonia. In 2013, the penalty was equal to 1% of all of the hospital's total Medicare billings, but has increased to 2% for 2014, and will increase to 3% in 2015. Additionally, beginning in 2015, two new conditions will be included in the hospital readmissions reduction program: acute exacerbation of chronic obstructive pulmonary disease, and patients readmitted after total hip arthroplasty or total knee arthroplasty [29]. Notably, while the CMS calculates a number of health-related indicators (surgical site infections, deep vein thrombosis, falls and trauma), 30-day readmissions are the only indicator used to reduce payment, and also the only indicator that is a measure of health utilization, rather than a direct measure of patient health.

In the United Kingdom, the Department of Health has introduced a system where the government will not pay for any emergency readmissions that occur within 30 days of discharge from an acute hospital, following an initial

planned stay. In Australia, the Health Innovation and Reform Council has recommended that the government consider financial consequences for high readmission rates[30].

1.1.3 Financial incentives

Financial penalties for poor readmission rates, especially as implemented in the US, can lead to perverse incentives that reduce the quality of care [31]. Following an emergency department visit, patients can be discharged or admitted, but it is also possible for the visit to be continued as an "observation service", a practice that has been increasing in the US in recent years[32]. Some authors have suggested that hospitals may be holding patients in an observation service until after 30-day readmission period has elapsed, which would allow the hospital to both reduce their readmission rate, but continue to charge for the patient admission[33].

Furthermore, some believe that, unless carefully designed, pay-for-performance measures will unfairly penalize those who treat underserved patients, increasing racial and ethnic disparities[34, 35]. Some argue that hospitals that do make improvements on specific measures (like readmissions) may do so at the expense of other untargeted or charity care [36, 37].

The financial penalty may not be big enough to incentivize better quality of care: if profit is the only consideration for readmitting patients, then the "loss" incurred by not readmitting patients after 30 days may exceed the penalty for readmitting them[38]. One alternative financial incentive would be a single-episode price (also called a bundled payment or a warranty system) that revokes payment for any hospital stay within 30 days of a previous admission[38, 39].

Implementing any financial penalty reinforces the notion that the reason why government agencies want to reduce readmissions is that they want to reduce cost. One of the original proposals was to both penalize hospitals with worse than average readmission rates and give bonuses to hospitals that had better

than average readmission rates[39] – a policy that would have made it clear that quality of care, rather than cost, was the government's motivation to reduce readmissions.

1.1.4 Summary

When considering the indicators of quality of care, hospital readmissions have a stronger connection to cost than most. Although readmissions may be a useful indicator of quality of care, they could easily be construed as a thinly veiled method to reduce utilization and therefore cost. In the literature, the arguments against the use of hospital readmissions typically conclude that they are a poor indicator of quality of care because other causes of readmission unrelated to hospital quality of care are more important drivers of readmission than quality of care. In the next section, I review some of the causes of readmissions, both related and unrelated to quality of care.

1.2 Causes of readmissions

Hospital readmission laws can incent hospitals to provide better quality of care and to reduce readmissions. However, since there are other causes of readmissions unrelated to quality of care, these laws can also unfairly penalize hospitals, and even create perverse incentives. Some reasons for hospital readmission can be connected to quality of care at hospitals, such as hospital discharge planning and adverse patient safety events, but other causes include patient-related factors (e.g., the natural deterioration of health), as well as system-related factors (e.g., the availability and the quality of outpatient care). To understand the relationship of hospital quality of care and readmissions in observational data, we must identify confounders of this relationship.

Hospital discharge planning includes the instructions that hospitals provide to

patients, caregivers, outpatient physicians, and other healthcare workers outside the hospital. Discharge planning is designed to help patients and caregivers manage their care outside of the hospital. In the US, Medicare regulations obligate hospitals to provide discharge planning, and note it in the medical record. However, despite these legal requirements, studies have found that discharge planning is often incomplete and untimely, and additionally outpatient providers rarely have access to these plans[40]. By paying greater attention to these important transitions in outpatient care, it may be possible to prevent readmissions to acute care.

Adverse patient safety events, which are caused by the medical management rather than the underlying condition of the patient, can result in a hospital readmission[41]. Adverse drug events (ADE) in the immediate post-discharge period is the most common reason for hospital readmission, and studies have found that medication reconciliation, and important component of the discharge planning process, could significantly reduce the rates of post-discharge ADEs, further highlighting the importance of high quality discharge planning [42].

From a system perspective, the variations in bed supply in long-term and acute care facilities can disrupt patient flow throughout the healthcare system. For example, if a hospital is overcrowded, it may begin to discharge patients slightly earlier than usual, which may lead to more readmissions. On the other hand, an overcrowded hospital might lead to fewer admissions, which would reduce readmissions. One study found that much of the variation in hospital readmissions could be attributed to variation in bed supply[43], and another recent study found that survey-reported admission rates were correlated with readmission rates[44]. Additionally, if long-term care services are at maximum capacity, then discharged patients may be more likely to be readmitted.

The natural deterioration of patients due to their health conditions will lead to hospital readmissions, but may not be attributable to suboptimal quality of care within hospitals, as some conditions cannot be treated with current medicine.

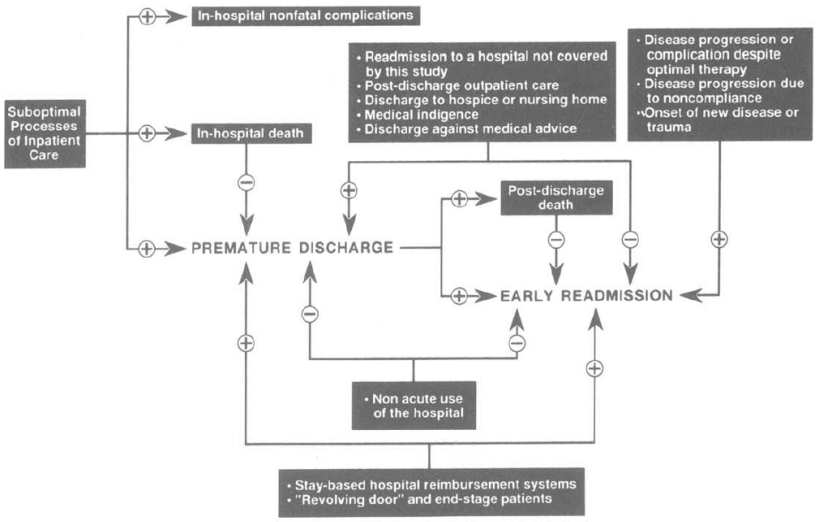

Fig. 1. Conceptual framework for the association between premature discharge and early readmission.

Figure 1.2: Conceptual framework for the association between premature discharge and early readmission. Ashton and Wray[45] schematically describe a plausible causal framework for the relationship between discharges and readmissions. A "+" sign indicates a positive relationship, while a "-" sign indicates an negative relationship.

The differences between hospital readmission rates may be attributable to the differences in patient-mix, particularly at academic hospitals, which treat more complicated patients who are at a higher risk of readmission.

Competing risks, especially mortality, may invert the relationship between hospital care and readmissions. A hospital with good quality of care might be preventing in-hospital deaths, but these high-risk patients might be at a greater risk of readmission[26]. Conversely, patients treated at a hospital with poor quality of care might be at more risk of dying in 30 days, reducing the probability of readmission[26].

Figure 1.2, reproduced from Ashton and Wray[45], provides a useful conceptual framework that describes several causes of readmissions, both related to and not related to quality of care. The Figure illustrates how some causes of read-

missions may also prevent readmissions. For example, a hospital with good quality of care can prevent adverse patient events, leading to fewer readmissions, but it can also prevent post-discharge deaths, leading to more readmissions. A crowded hospital may be more likely to discharge its patients prematurely (leading to readmission), but also may be less likely to admit patients from the emergency department (preventing readmission). Some of the causes could plausibly combine with each other to cause complex dynamics. For example, an overcrowded hospital may discharge its patient prematurely to a long-term care facility, crowding long-term care, leading to a longer length of stay in other hospitals, preventing premature discharge, leading to fewer readmissions in the other hospitals.

In the next section, I discuss ways of isolating those readmissions that are attributable to the hospital, that is, those readmissions that were preventable.

1.2.1 Preventability of admissions

The causes of readmission, as listed previously, can be classified as care-related (adverse patient safety events, effectiveness of patient discharge planning) and not care-related (natural deterioration of health, outpatient care). In the earliest work on readmissions [19, 20, 22], specific definitions of readmissions are omitted, suggesting that they assumed that a clinical expert could reliably classify any single readmission as related to the care of the initial visit (preventable) or not.

But recent work shows that clinicians cannot reliably classify the preventability of admissions[46]. In this study, after having a panel of 35 clinicians review readmissions, the study authors used latent class analysis to estimate the sensitivity and specificity of preventability classification in the absence of a reference standard, finding that half of clinicians had a sensitivity less than 50%, and a specificity of 88%, suggesting that clinicians cannot reliably measure the preventability of a readmission.

TABLE II—*Hypothetical comparison of adverse outcomes in two hospitals*

	Occurrence of avoidable adverse outcome				
	Yes	No	Total	Readmission rate (%)	Avoidable adverse outcome rate (%)
			Hospital A		
Readmission:					
Yes	50	50	100		
No	30	1870	1900	5·0 (100/2000)	4·0 (80/2000)
Total	80	1920	2000		
			Hospital B		
Readmission:					
Yes	40	80	120		
No	10	1870	1880	6·0 (120/2000)	2·5 (50/2000)
Total	50	1950	2000		

Figure 1.3: Avoidable readmissions in two hypothetical hospitals. Milne and Clarke [49] describe two theoretical hospitals A and B, where hospital B has a higher readmission rate, but a lower avoidable adverse outcome rate. Although hospital A has a larger proportion of preventable readmissions, it has a lower overall readmission rate.

Even if clinician review was a reliable instrument to measure avoidable readmissions, it would be too expensive to apply it to the massive administrative datasets in which it is now required. A compromise solution was to measure potentially avoidable readmissions. The 3M Corporation had clinical experts devise a list of admission/readmission diagnostic code pairs[39] where preventability was at least plausible, with other authors proposing similar lists[47, 48]. By assuming that potentially avoidable readmissions were a rough proxy for actually avoidable readmissions, we could easily measure readmission rates in large administrative databases.

Milne and Clarke[49] identify the critical implicit assumption with using po-

tentially preventable readmissions to compare hospitals: the proportion of preventable readmissions among the potentially preventable is the same across hospitals. Unfortunately, we have theoretical reasons to assume that the proportion of actually preventable readmissions varies greatly: 1) we know that the severity-of-illness varies across hospitals, especially when you compare academic to non-academic hospitals, so if you assume that unpreventable readmissions are largely a function of the severity-of-illness, then we know that the unpreventable proportion is likely to be different across hospitals. 2) The premise of comparing hospitals by using readmissions was that the quality of care differs across hospitals which results in different proportions of preventable readmissions. If this premise is not true, then hospital readmissions would not be able to discriminate the quality of care between hospitals. Empirical evidence also makes it clear that the proportion of preventable readmissions is not the same across hospitals: a meta-analysis of avoidable readmission proportions found that the actual proportion of avoidable readmissions among all readmissions ranged from 5% to 79% across study sites[46]. Although potentially preventable admissions were never intended to be a perfect proxy for preventable readmissions, the evidence suggests that they may be too biased for use as a financial penalty.

By precisely defining a preventable readmission in terms of counterfactuals, we do not need to measure the preventability of individual readmissions to draw inference on hospital rates of preventable readmissions. I define a preventable readmission as a readmission that would have occurred if a patient was treated at some hospital A, but would not have occurred if that patient was treated at hospital B. This definition of readmissions is unique in a few ways:

1) The definition is only valid with respect to a specific pair of hospitals. A readmission might be preventable when considering hospitals A and B, but not so when comparing hospitals A and C.

2) It is a binary outcome; you must specify a certain time frame x (for example, 30 days) for the definition to be complete.

3) As defined, it is impossible to measure whether any specific readmission is preventable, since one patient cannot be treated at both hospitals.

4) As defined, in a controlled trial (randomizing patients to two hospitals), we could estimate the difference in preventable readmissions in a population of patients. Similarly, in an observational setting, if we have no unadjusted confounding, we could estimate the difference in preventable readmissions.

We do not need to identify which individual readmissions were caused by a difference in treatment at the hospitals; we just wish to identify if there is a difference in the population rate of readmissions. This is the fundamental argument of causal inference in an observational setting – we can draw inference about population rates even when we cannot measure the effect on individuals. For example, in an observational trial, to understand if smoking causes lung cancer, we do not need to identify which individual smokers got lung cancer because of smoking, and we just need to measure the difference in rates of lung cancer between smokers and non-smokers. In the same way, we do not need to identify which individual readmissions were preventable to know that the rate of preventable admissions was different between hospitals. Given this definition of preventable readmissions, the main threat to validity when attempting to draw inference on how different hospitals change the rate of readmissions is confounding, particularly by severity-of-illness.

1.3 Predictive accuracy and inference

The techniques for estimating and adjusting hospital readmission rates in the US reflect the history of how they were implemented. Initially, the raw (unadjusted) rates of 30-day hospital readmissions per discharge were reported as a tool to measure quality of care. However, these rates reflected both the readmission rate and the severity-of-illness of the patients at the hospital. The rates needed to be adjusted for the patient-mix if they were to be used to infer quality

of care.

The patient-mix adjustment model now used in the US is essentially standardization technique. A logistic regression model is used to predict 30-day readmission risk based on a few well-known risk factors including age and sex, and then scaled the crude readmission rate at each hospital by the predicted readmission risk for all the hospital's patients. The technique is unusual for an inferential model, but allows the reporting of adjusted readmissions per discharge, rather than odds ratios (which were presumably considered too obscure for a wide audience). The Committee of Presidents of Statistical Societies released a report (commissioned by the CMS) which criticized the adjustment technique on several grounds, noting that further basic demographic variables should be included in the model, and more sophisticated techniques such as boosting and random forest should be considered to properly adjust for confounding [50].

Many analyses and meta-analyses use the predictive accuracy of readmission risk models to judge how well they can draw inference on hospital quality of care. In 2011, Kansagara[31] reviews 26 models of hospital readmission risk, concluding that these models should not be used to compare hospitals because the classification accuracy was "poor" (AUC below 0.7). There is no way to assess if there are unmeasured variables which confound the relationship between any exposure and outcome. The AUC is a measure of goodness-of-fit, and does not measure that validity of inference. For example, a model of smoking and lung cancer, even if perfectly adjusted, is unlikely to have a high AUC, because even among smokers, very few get lung cancer; smoking is not a good predictor of lung cancer, but we can infer that smoking causes lung cancer. The predictive accuracy and inferential capacity of a model are unrelated; one cannot be used to measure the other.

1.4 Interventions to reduce readmissions

Even if a particular hospital is associated with hospital readmissions, it may not be immediately obvious how hospitals could improve patient's health and reduce readmissions. The Institute for Healthcare Improvement (IHI) categorizes readmission interventions into four broad categories: 1) enhanced transition care 2) patient education 3) multidisciplinary team management and 4) end-of-life care planning. Several of these interventions, in a variety of forms, have been adopted in the US[51], and in Canada.

1.4.1 Transitional care models

Transitional care refers to the care designed to 1) coordinate care between hospital and post-hospital providers, 2) temporarily monitor the patient just before and after the discharge, and 3) educate the patient and caregivers on how to manage changing needs (pharmaceutical, nutritional, social, et cetera) after discharge. Several hospital networks have implemented transitional care strategies that have been shown to reduce readmissions.

At the University of Colorado, the Care Transitions Intervention was developed as a four-week post-discharge program which provides an advanced practice nurse who acted as a "transition coach" to patients discharged from the hospital. The transition coach assists patients with medication self-management, creates a patient centered health record to facilitate information transfer throughout different sites of care, ensures timely follow-up with primary or specialty care, and watches for a series of "red flag" conditions which required immediate medical attention. In a randomized controlled trial, those patients receiving the care intervention had significantly lower 30-day and 90-day readmission rates[11].

In the Transitional Care Model, used at the University of Pennsylvania, a multidisciplinary team, led by a dedicated transitional care nurse, treats, moni-

tors and guides patients just before, during and just after hospital discharge. Specifically, they focus on patient and caregiver education, medication reconciliation, and facilitating access to primary care providers. In a randomized, controlled trial of older adults hospitalized with heart failure, those randomized to transitional care had significantly fewer 52-week readmissions than the control group, and that the time-to-readmission was also significantly longer in the intervention group[52].

The Re-Engineered Discharge program, at the Boston University Medical Center uses nurses (discharge advocates) to improve transitional care[10]. The discharge advocate nurse works with patients to ensure medication reconciliation (including contacting the pharmacist 2 to 4 days after follow-up), arranges follow-up appointments with primary care, and provides individualized instruction booklets to help patients manage their treatment outside the hospital. A randomized controlled trial found that intervention group patients had a lower rate of hospital utilization than the control group[10].

In several states in the US, Quality Improvement Organizations (QIOs) also provide individual interventions similar to those above, but also attempt broader structural improvements to improve care transitions[53]. Although these interventions vary widely, and take local needs into consideration, they can broadly be categorized into 1) interventions that improve the transition processes, which can include new protocols for transfer from hospitals to long-term care and improvement of information technology, and 2) interventions that address access to services, which can include the provision of new services to palliative care, and providing better access to nutritious meals.

In Ontario, several interventions are being piloted to reduce readmission. St. Michael's Hospital partnered with the Toronto Central Community Care Access Centre (CCAC) and several hospitals in the region to create the Virtual Ward, which uses a hospital-like approach to treat patients in their own home. Patients receive care from a multidisciplinary team that meets daily and share a common set of notes, all coordinated through dedicated staff at the CCAC.

The South West Local Health Integration Network (LHIN) implemented a small pilot project to improve transitions from the hospital in which a nurse practitioner visited the home shortly after discharge to enhance education, reconcile medications, and develop a focused plan to prevent readmission[54], finding that readmissions could be reduced. Finally, the "Home at last" program, currently running in several LHINs, provides a comprehensive transition service from hospital to home, which may include driving the patient home, picking up medications and groceries, preparing a small meal, providing personal care and homemaking services, follow-up phone calls, and referral services to other community support services.

1.4.2 Hospices, long-term and palliative care

Kaiser Permanente (a major healthcare provider in the United States) has created the TriCentral Palliative Care Program Toolkit to help organizations create palliative care programs. These palliative care programs would have interdisciplinary teams help patients to manage both physical pain and other symptoms, as well as provide emotional support for family and other caregivers. A randomized, controlled trial found that the intervention group had significantly less hospital days than the control group[55]. Other studies have found that inpatient palliative care consultation services reduce the likelihood of ICU readmission [56–58], but others found that this was true for only those discharged to a hospice[59].

1.4.3 Telehealth

One randomized control trial found that three months of telemonitoring (post-discharge video conferencing with a nurse, and daily transmission of weight, blood pressure and electrocardiogram) significantly reduced hospital readmissions[60]. Another randomized controlled compared home telemonitoring,

nurse telephone support, and usual care for high risk heart failure patients, and found that among both the telemonitoring and nurse telephone support groups, mortality was decreased, but rehospitalizations were increased[13].

While prior studies have provided some evidence that transitional care interventions can significantly reduce hospital readmissions, these interventions, which are people and resource intensive, remain very costly. In the current era of cost containment and resource scarcity, there is a need for better methods to identify patients at high risk of hospital readmission so that preventive interventions can be targeted at these individuals. There is also a need to better understand the potentially modifiable structural drivers to hospital readmission to assist decision-makers in designing more efficient systems and discharge policies that contribute to reduce readmission rates. This purpose of this doctoral thesis is to address these issues.

1.4.4 Summary

Although there is some evidence that transitional care interventions can reduce hospital readmissions, these interventions are costly. In the current era of shrinking hospital budgets, interventions need to applied to those who need them the most; predictive readmission risk models can help to target these interventions to those at highest readmission risk. Additionally, inferential models of readmission could help us to estimate how effective broader structural level interventions could reduce readmission risk. The purpose of this thesis was to explore how improvements to predictive and inferential models of readmission risk could help to develop cost-effective interventions to improve quality of care and prevent readmissions.

1.5 Objectives of this work

An important requirement to target transitional care interventions to highest risk patients is to have accurate predictive models of hospital readmissions. Three broad approaches can be used for that purpose. First, hospitals can develop predictive models of readmission based on their own data (local models), but the accuracy of these models has been questioned. Alternatively, hospitals could pool data from several geographically distributed sites and take advantage of the larger variability of the information in this 'big dataset' to develop potentially more accurate models of hospital readmission (global models). However, in practice legal (privacy-related), technical, and administrative issues makes data pooling very difficult. Lastly, recent research has shown that accurate models can be developed across hospitals by sharing information about model fit to proposed parameters rather than directly sharing patient data (combined models).

We may be able to prevent hospital readmissions through transitional care interventions, but the cost of these interventions makes it infeasible to apply them to all discharges. To target these interventions at those patients with the highest readmission risk, hospitals can develop predictive models of readmission based on their own data (local models). Ideally, hospitals would pool their patient data to develop more accurate readmission risk models (global models), but legal (privacy-related) technical, and administrative issues make pooling infeasible. Recent research has shown that accurate models can be developed across hospitals by sharing information about model fit to proposed parameters rather than directly sharing patient data (combined models) [61–65]. Interestingly, no prior research work has directly compared the relative accuracy of the three approaches (local, global, combined) to model development. In my first manuscript, I compared the accuracy of the following models to predict 30-day readmission risk: 1) an ideal "global" model, fit to data pooled across hospitals, 2) "local" models, fit to individual hospital data, and 3) "combined" models,

fit to data from an individual hospital, but additionally using coefficients from local models from other hospitals.

Predictive models of hospital readmissions can be useful guides for resource allocation to individual high-risk cases, but inferential models can potentially lead to population-level interventions. One known risk factor for hospital readmissions is that patients discharged from hospitals on a Friday (Friday discharges) are readmitted sooner than Wednesday discharges. Some authors [66–68] have suggested that readmissions related to Friday discharges can be decreased by increasing weekend staff or allowing admissions into long-term care on the weekend. However, the effect of differing admission probability by day-of-week has not been investigated, meaning the estimated effect of interventions may be biased. In my second manuscript, I used analytic (an absorbing Markov model) and empiric (a time-varying Cox proportional hazards model) approaches to investigate how admission probability confounds the effect of Friday discharges on readmission.

The US and other jurisdictions financially penalize hospitals with poor (confounder-adjusted) 30-day readmission rates. Although hospital administrative data are information-rich, confounder adjustment tends to be crude. Non-parametric machine learning techniques can take advantage of these rich data to predict readmission, but cannot isolate the independent effect of hospitals on readmission risk. In my third manuscript, I estimate the effect of care at different hospitals on 30-day readmission risk, using targeted maximum likelihood estimation TMLE, which allowed the use of a non-parametric machine learning technique (random forest) to take advantage of the rich confounder data. I developed three models to estimate the marginal readmission risk at each of the hospitals after hospitalization for heart failure, acute myocardial infarction (AMI), and pneumonia. I controlled for hundreds of confounders including pre-admission drug prescriptions, medical procedures, and diagnoses. We compared the TMLE-estimated risk to a logistic regression model that only adjusted for well-known confounders. Using TMLE, I could use the predictive

power of machine learning techniques to take advantage of rich confounder data, and draw inference on the wide differences in quality of care between hospitals.

In this work, I examine how large healthcare administrative databases can help build better inferential and predictive models of hospital readmissions. In the first manuscript (Chapter 3), I study how pooling hospital data improves predictive readmission models. In the second manuscript (Chapter 4), I study how the day-of-week of discharge affects readmission, and use both empiric and analytic approaches to study how this effect is confounded by the probability of admission on the weekend. In the third manuscript (Chapter 5), I investigate how TMLE can be used to improve inference on the effect of hospital quality of care on readmissions by using machine learning techniques on high-dimensional confounder data. In all three manuscripts, I analyzed hospital discharges in a cohort of people 65 years of age or older between 1996 and 2006 (inclusive) from the Quebec hospital administrative database which included both the in-hospital procedures and diagnoses, and outpatient diagnoses and dispensed drug prescriptions (described in detail in Chapter 2).

Chapter 2

Data

In all three studies presented, I used a cohort of patients extracted from the *Régie de l'assurance maladie du Québec* (RAMQ) to study hospital readmissions. In this section, I describe these data in detail.

2.1 Cohort selection

This extract includes all Quebec hospitalizations and outpatient visits for those who ever reported an influenza-like illness (ILI) between the years 1996 and 2006 (inclusive), while living in the census metropolitan area (CMA) of Montreal, as defined by the 2006 Canadian census. Having an ILI was defined as any of the International Classification of Diseases, 9th revision (ICD-9) codes listed in the Table 6.1, a relatively broad set of codes, among which there are several extremely common codes such as cough and fever. This was a dynamic cohort; a person "entered" the cohort on the day that they met both of these criteria: 1) at least one ILI code was recorded in either the RAMQ or *Maintenance et exploitation des données pour l'étude de la clientèle hospitalière* (MEDÉCHO) data after January 1, 1996 while they were living in the Montreal CMA, and 2) they were at least 65 years of age. A person who changed residence from the Mon-

treal CMA to another address inside Quebec would remain in the cohort, but a person who changed address to outside of Quebec would be removed from the cohort.

Importantly, the majority of admissions in the following three studies were not related to influenza. ILI was only used as one of the criteria for cohort entry; after entry, all hospitalizations for a patient were included into the study.

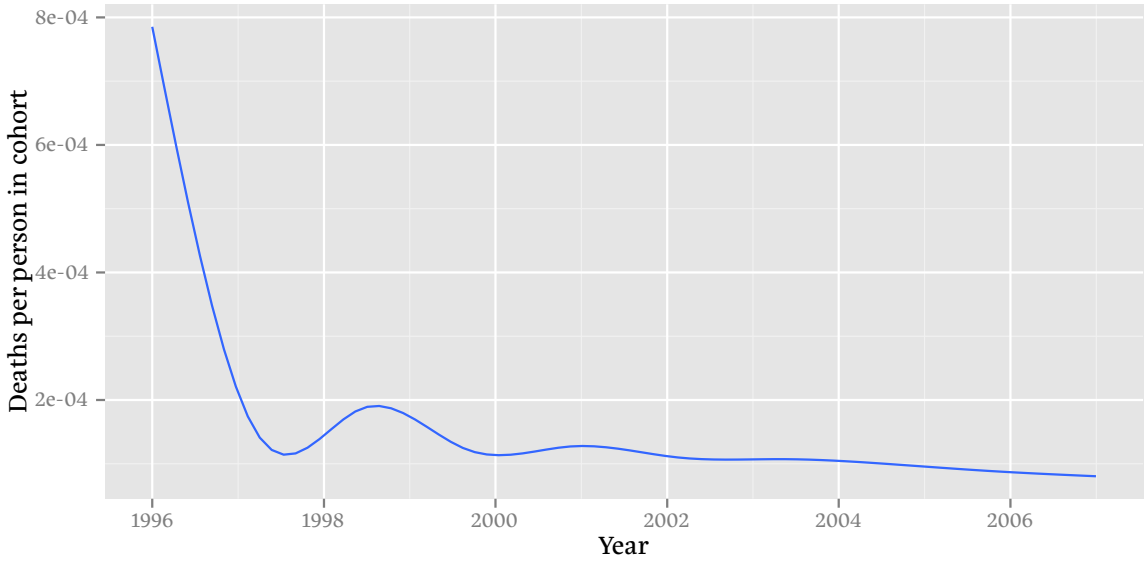

Figure 2.1: Number of deaths per person in the cohort by time. A GAM smoother has been applied to the rate to emphasize long-term trends.

The use of ILI codes as a selection criterion is not ideal for this work. This data source was extracted for a different study; we used it because it included enough discharges to allow us to measure small effects, was unlikely to have a selection bias with respect to hospital readmissions, and was readily available. However, in this cohort, due to the selection by presence of a diagnostic code, a selection bias has induced a relationship between time and outcomes associated with severity-of-illness such as death. People with the highest frequency of ILI codes will also be the most probable to enter the cohort early, meaning that early in the study, we have a very high proportion of those who have very high ILI code

frequency. Those with the highest frequency of diagnostic codes are also likely to be the most severely ill, and those who are severely ill are most likely to die. This leads to a relationship between the year of discharge and death, as shown in Figure 2.1. In light of this bias, I cannot draw inference on the relationship between the year of discharge and hospitals readmissions.

2.1.1 Hospital identification

The hospital's identification numbers were anonymized in these data. However, if a hospital license number changed (which happened roughly one time per hospital over 1996 to 2006), the anonymized identification number changed, which was unfortunate because I had no way of identifying pairs of anonymized hospital license numbers that represented the same hospital. Although I did not need to deanonymize the hospital identification numbers, because I wanted to compare hospital's effect on readmissions, I had to identify which identification numbers represented the same hospital.

To correct this problem, I identified the times at which each hospital identification number started and stopped having visits. Unfortunately, many hospital identification numbers stopped or started on the same day, indicating that many hospitals changed license numbers on the same day (which I confirmed by consulting the website of the *Ministère de santé et services sociaux* (MSSS)). However, in the RAMQ data, hospitals changed the license number on exactly the date indicated by the MSSS, but in the MEDÉCHO data, the hospitals changed license number at the end of the financial year (March 31). This created a "gap" which could be exploited; since doctors bill in RAMQ data for hospitalizations, I could find which hospital number they billed in for a patient that appeared in MEDÉCHO. For each major discharging hospital that changed numbers, I was able to clearly identify one clear candidate as the likely number it changed to. Finally, to verify that my guess was correct, I plotted a choropleth map of the addresses of the visitors to each hospital identification number by

forward sortation area (FSA). In all of the following work, I have used my "estimate" (which I believe is exactly accurate) of which hospital license number paired with which other hospital number.

2.1.2 Hospital selection

Although cohort members must have lived in Montreal at least once, they remained inside the cohort even if they moved to other Quebec locations. This meant that the database (likely) contained a significant number of hospital discharges from outside the Montreal CMA. To ensure that I selected only hospitals in the Montreal CMA, I selected the twenty hospitals that discharged the most 65 year old patients. Although I only had selected patients from the top twenty hospitals, I counted a person who was readmitted to any hospital in Quebec as readmitted.

2.2 Basic description

2.2.1 Addresses

The RAMQ database does not consistently record whether a person has changed their residence to a place outside of Quebec. The address of a person is recorded every year, so a missing address may indicate that a person was no longer in Quebec. However, the address data can also be missing because RAMQ did not record that data, and the quality of the address data is known to be poor (often missing) for the period of 1996-1998.

In this analysis, I took a relatively specific, less sensitive measure of leaving Quebec. If a person had a consecutive missing addresses from any year all the way to the final year (2006), and they had no healthcare utilization in all of those years, then that person was considered to have left Quebec, and was censored.

Because I had to censor on a specific day, I picked a random day between the last day of healthcare utilization and the end of year.

2.2.2 Transfers

If people were transferred between hospitals, it was considered one stay (the transfer is not counted as a discharge). Transfers are recorded in the database, but are of dubious quality (it is clear that many transfers go unrecorded). Some hospitals record transfers *within* their own hospital. If a person was readmitted on the same day as discharge, or the next day, then it was considered a transfer. This means that a person cannot be readmitted in exactly one day. Exactly 4010 discharges (0.6%) were considered a transfer because of a hospital admission the *next* day.

If a patient was transferred between hospitals, then only the last hospital that the patient stayed at was considered the discharging hospital. Some bias could have resulted from this choice: if a hospital was treating patients particularly poorly, but then transferred the patient to more specialized care, it may cause the specialized hospital to be "assigned" more readmissions. Among hospital stays that resulted in a discharge of someone 65 years of age, 32 121 stays (5.1%) included at least one transfer. 75% of hospital stays that included a transfer only included one transfer, while 21% contained two transfers. The maximum number of transfers was 10.

2.2.3 Death

In these data, we had three sources of information on the date of death. The MEDÉCHO data included the day of death for those who died inside a hospital. The RAMQ data included the year-month (for example, March of the year 2000) of death. Finally, we had the *Institut de la statistiques de Québec* (ISQ) data which also included the year-month of death, as well as some cause-of-death

information. Theoretically, all three sets of death data should be exactly concordant, but practically, there were inconsistencies. When the dates of death differed in the three datasets, I used the MEDÉCHO data when possible, and favoured the RAMQ data over the ISQ data in the small number of cases (374) when they were inconsistent.

For the survival analyses used in the first and second manuscripts, the time unit used was the day, and a person was censored if they died outside of the hospital. Since only the year-month of death was recorded for those who died outside of the hospital, it was necessary to estimate their day of death. In cases where a hospital visit or outpatient was recorded in the data during that month, it was possible to estimate the earliest date of death during the month. I used the date halfway between the earliest date of death and the end of the month, rounded down.

2.3 Cohort entry

Figure 2.3 shows the number of people alive in the cohort by time. The rate of cohort entry slightly accelerates during the winters, because diagnoses of influenza-like illnesses are more likely in the winter.

Figure 2.3 displays the absolute number of people entering the cohort by time, smoothed with a Gaussian kernel. Every winter season, the rate of entry increases, because diagnoses of influenza-like illnesses are more likely in the winter. The rate of cohort entry also declines over the long-term. We believe that the selection criteria was common enough that the majority of 65-year olds in Montreal were entered into the cohort in the initial few years, and following this, the rate flattens because our rate of entry is limited by the number of people turning 65 in Montreal every year. RAMQ changed their diagnostic code system from ICD-9 to ICD-10 in March 2006, but not all hospitals and outpatient clinics adopted this uniformly. The drop in cohort entry after March 2006

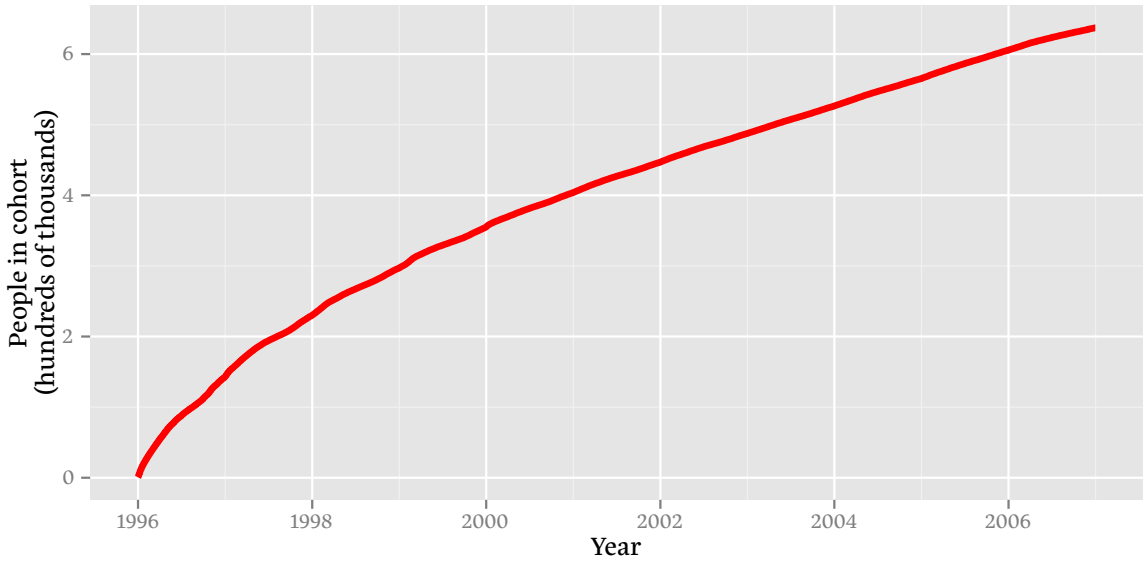

Figure 2.2: People alive in cohort by time. Calculated at a daily resolution.

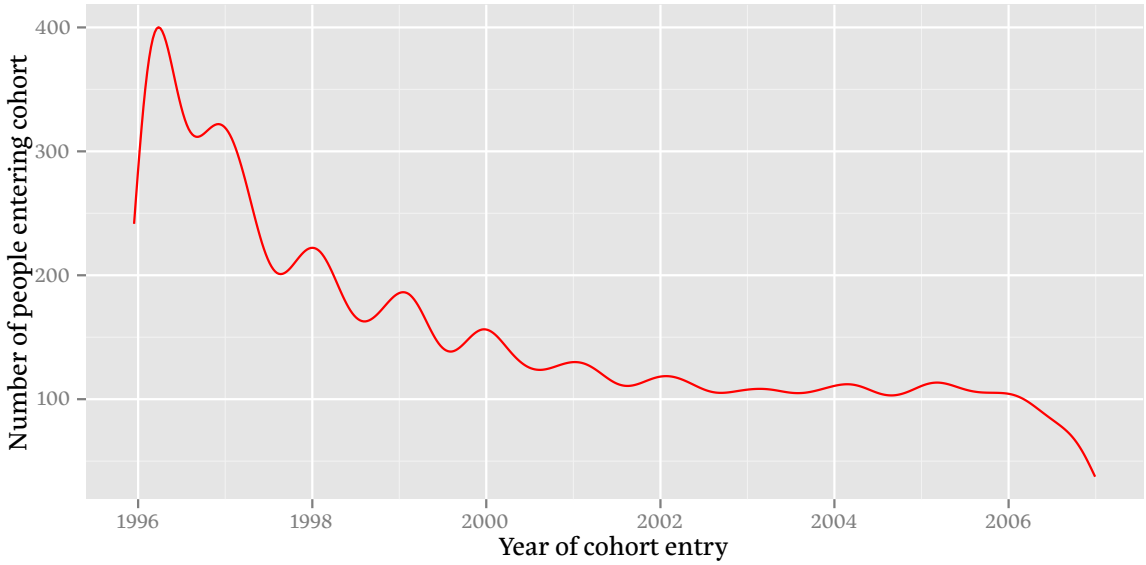

Figure 2.3: Number of people entering the cohort by time. The number of people entering each day have been smoothed with a Gaussian kernel with an automated bandwidth selector.

reflects the diagnostic code system change, as our selection criteria was based largely on ICD-9 codes.

Chapter 3

Privacy-preserving predictions: a case study of hospital readmissions

3.1 Preamble

Hospitals can address readmissions through system-level interventions, like increasing staff or improving adherence to clinical guidelines, or they can use individual-level interventions, like identifying patients with high readmission risk and improving their individual transitional care. Since these transitional care interventions can be costly, we can improve their cost-effectiveness by applying them only to patients at high risk of readmission. To identify patients at high readmission risk, we can use predictive models of readmission, based on the rich data available within administrative databases.

A hospital can use their own data to develop predictive readmission risk models, but for patients with relatively rare illnesses, their statistical power may not be sufficient to estimate risk accurately. To improve the accuracy of readmission risk models, hospitals can pool their data and fit models to that pool. However, typically, hospitals cannot share patient data due to confidentiality agreements, and technical challenges. Recently, privacy-preserving model combi-

nation techniques have been developed, which allow hospitals to fit models on pooled data, without directly sharing patient data.

In the following manuscript, I focused on how much we can improve the use of individual-level transitional care interventions by improving the accuracy of our predictive readmission risk models by pooling data or combining models.

3.2 Abstract

Introduction: Hospital readmissions may be preventable through transitional care interventions, but these interventions are costly. To target these interventions at high risk patients, hospitals can develop readmission risk models. Ideally, hospitals would pool patient data to develop risk models, but typically, private patient data cannot be shared. Privacy-preserving model combination techniques exist, but require technical expertise, and there is little evidence that models estimated from pooled data are more accurate than those estimated from data at a single hospital.

Objectives: To compare the accuracy of the following three models in predicting 30-day readmission risk: 1) an ideal "global" model, fit to data pooled across hospitals, 2) "local" models, fit to individual hospital data, and 3) "combined" models, fit to data from an individual hospital, but additionally using coefficients from local models from other hospitals.

Research Design: We used 11 years of data from 20 hospitals to build the predictive models. We compared the accuracy (area under the receiver-operating characteristic curve (AUC)) of local, global, and combined models for predicting readmission following hospitalization for heart failure, AMI, and pneumonia.

Results: Within 30 days of discharge, there were 7 355 (22%) heart failure readmissions, 4 127 (15%) pneumonia readmissions, and 3 414 (16%) AMI readmissions. For global models, the AUC was 0.65, for local models the AUC was 0.63, and for the combined models the AUC was between 0.63 and 0.64.

Conclusion: Hospitals can use their own data to predict hospital readmissions and achieve similar performance to models that pool data.

3.3 Introduction

Hospital readmissions are both common and costly: in the United States, 13.3% of discharges end in a potentially preventable readmission within 30 days, costing an estimated $12 billion per year [39]. Not only does readmission suggest suboptimal transitional care, but hospitalization itself can pose a health risk, particularly for older adults who can experience serious functional decline in hospital that is unrelated to the original admission [69, 70]. Research has identified substantial variation in readmission rates across regions and hospitals, suggesting that improvement is possible [39, 71]. Furthermore, some jurisdictions in United States[2] and Canada[3] have recently introduced legislation that financially penalizes hospitals with high readmission rates, providing another incentive to prevent readmissions.

To prevent readmissions, transitional and post-discharge care interventions exist, but they are too costly to be applied to all discharged patients [10–12]. To target transitional care interventions to patients at-risk of readmission, hospitals can develop statistical models of readmissions based on their own data (local models). A hospital with low patient volume may not be able to accurately estimate the effect of rare patient characteristics on readmission, reducing the performance of these local models. Ideally, hospitals would pool their patient data [72] and develop a "global" readmission risk model, accumulating enough data to estimate the effect of rare conditions and drug prescriptions on readmission risk. However, there are many practical barriers to data sharing involving data ownership, privacy and security.

Recently, much interest has developed in privacy-preserving model learning techniques, which build models across hospitals without directly sharing any

patient data (combined models). These systems [61, 62] build models by iteratively sharing information about model fit (for example, the residuals) to proposed parameters rather than directly sharing patient data. Although the performance of privacy-preserving model fitting techniques have been shown to be similar to global models [61–65], the techniques all require significant network communication infrastructure, hospital coordination, and technical expertise, making them costly for hospitals to implement.

Global readmission risk models, which need data pooled from multiple hospitals, may be more accurate than local models, but they require data sharing agreements and pose a risk to patient privacy. Combined readmission risk models can greatly reduce the risk to patient privacy, but require technical expertise. Local readmission risk models, which are fit only to local hospital data, are less costly but may also be less accurate than global or combined models. The main objective of this study was to compare the performance of local, global, and combined models of hospital readmissions. To our knowledge, this study will be the first to directly compare these three approaches to model estimation. Understanding the relative accuracy of these approaches should help hospitals to assess the potential benefit of investing in different approaches to risk modeling.

3.4 Methods

3.4.1 Study Data

We used a cohort extracted from a Canadian provincial (Quebec) administrative database of hospitalizations, obtained from the Régie de l'assurance maladie du Québec (RAMQ). We enrolled patients into this cohort on the month that two conditions were satisfied: 1) they had at least one diagnosis of a respiratory illness (the exact list of respiratory International Classification of Diseases, 9th Revision [ICD-9] codes is given in the Appendix) between January 1st, 1996 and

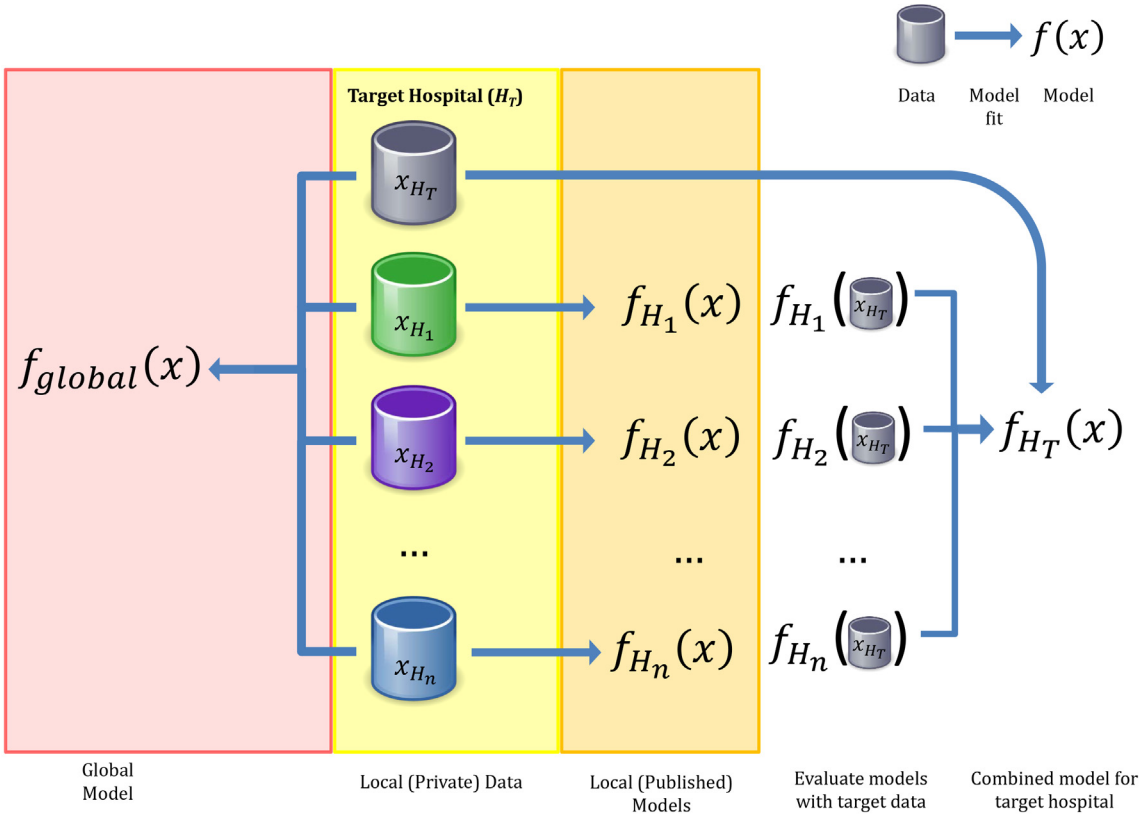

Figure 3.1: Schematic of local, global, and combined model fit.

March 31, 2006, while living in the 2006 census metropolitan area of Montreal, and 2) were at least 65 years of age. We used this cohort because it represents the majority of 65-year olds who were hospitalized in the region during the study period. We restricted our data to only the discharges from the twenty hospitals with the most discharges of patients 65 years or older within the study period; the twenty hospitals accounted for 75% of all such discharges.

3.4.2 Hospital readmissions

The unit of analysis in all models was the hospital discharge; a single patient could be discharged multiple times. A hospital readmission was defined as an emergency hospital admission to any Quebec hospital in the 30 days following a discharge. A person who died or had a non-emergency readmission in the 30 days following discharge was considered not readmitted (right censored).

3.4.3 Discharge data

For each hospital discharge, we used the following variables to predict readmission: a) admission type (emergency, semi-emergency, or non-emergency), b) number of hospital transfers during stay, c) discharge location (home or nursing home), which may influence the type of post-discharge care and follow-up, d) length of stay (days), which may act as a proxy for severity-of-illness, e) demographic characteristics, including birth year-month, sex, and age at discharge (years), f) the number of previous readmissions (within the study period), which has been shown to be associated with readmission [73] and, g) admission diagnosis (major diagnostic category – one of 24 groups of ICD-9 codes). We also included the day of week of discharge, which has been previously shown to have an association with readmissions [14], and the month of discharge, because we hypothesized that readmission risk would vary by seasons in Montreal.

Additionally, for each discharge, we used the recorded diagnoses, procedures, and dispensed drugs at the time of admission to predict hospital readmissions. The procedures performed were recorded in the Canadian Classification of Diagnostic, Therapeutic, and Surgical Procedures (CCP) system. Hospital diagnostic codes were coded using the ICD-9 system. Finally, drugs which were prescribed and dispensed outside the hospital, and were being taken on the day of admission were also recorded for each patient in the *code commune* system,

which records the chemical compound being taken. To ease computation, before fitting any model, we removed any diagnosis, procedure or drug that occurred less than 30 times among all discharges. We chose 30 because it appeared to be a natural breakpoint; if the number of variables included is a function f of the threshold, then the first derivative of f dropped at 30 for all three disease categories.

3.4.4 Disease types

We selected three high-volume admission diagnoses with high rates of hospital readmissions: pneumonia, AMI, and heart failure, the three initial conditions selected by the Centers for Medicare and Medicaid (CMS) to implement the Hospital Readmissions Reduction Program mandated by the Affordable Care Act. We identified each of the admission diagnoses using ICD-9 codes; for pneumonia we used codes ranging from 480-487, for heart failure we used all 428 codes, and for AMI we used all 410 codes. We predicted 30-day hospital readmission in all three disease subsets.

3.4.5 Regularized Cox model

All of the data were fit using a Cox proportional hazards model, with the hospital discharge as the unit of analysis, and the date of hospital discharge as time zero. Since Cox's model uses time-to-event as an outcome, we did not classify each discharge as readmitted or not within 30 days. Instead, we measured the time-to-readmission of each discharge: if the discharge resulted in an emergency admission the discharge was considered to have the “event”, and if the discharge resulted in a non-emergency admission, death, or the study ended before either, the discharge was considered “right-censored”. We did not account for the variance reduction induced by the correlation between repeated discharges of the same patient, because we did not estimate the variance of the

model parameters.

For all analyses, we used a regularized version of the Cox model with the lasso penalty [74] to both efficiently fit a model to hundreds of covariates (including one indicator variable for each drug, procedure, and diagnosis) and to produce sparse (parsimonious) models. The scale of the penalty was determined by the parameter λ. For each model fit, we optimized the selection of the penalty scale for the best partial likelihood using a nested a 10-fold cross-validation. Within each fold we assessed 100 λ-values spaced evenly between $max(\lambda)\times10^{-4}$ and $max(\lambda)$, where $max(\lambda)$ was the smallest λ-value that would result in a model with no non-zero coefficients.

3.4.6 Local and global data prediction

For each of the 20 hospitals, we fit a "local" readmission model to only the hospital's data, using 10 fold cross-validation to measure performance. We then fit a "global" model for each hospital: instead of using just the local hospital's data, we fit a model to the data from all hospitals. To measure performance for the global models, we used 10 fold cross-validation, where each fold excluded $^1/_{10}$ of the local hospital data. Since we had 20 hospitals, and 10 folds each, we fit 200 global models, and 200 local models. We repeated this process for all three disease types.

3.4.7 Combination model

We then applied a simple, completely privacy-preserving, meta-analysis technique, which uses only model coefficients to combine the data from all of the hospitals, described schematically in Figure 3.1, and algorithmically in Algorithm 1. For each of the 20 hospitals, we fit a local model to all of the data within the hospital, resulting in a set of coefficients for each hospital. This step simulates the usual approach in health care data analysis, where each hospital an-

alyzes its own data, and publishes summary statistics of the resulting model. We then had each hospital fit another model to only its own patients, but we added 19 new variables: the predicted outcome using the other 19 models (sets of coefficients) fit to the 19 other hospital's data. This step simulates each hospital gathering the model coefficients from other hospitals, and applying it to their own data. This kind of model only requires access to the final coefficients of other hospital models that are typically reported in other research articles.

3.4.8 Performance measure

We measured the performance of each model for predicting 30-day readmission using the AUC. A Cox model estimates the hazard ratio and not a probability of 30-day readmission. To calculate the AUC, we used the hazard ratio as the discrimination threshold. For each model, 10 AUCs could be calculated from each of the 10 folds. We calculated a single AUC on the pooled hazard ratios from all 10 folds [75].

3.4.9 Software

The models were fit in R using the coxnet [76] function in the glmnet [77] package.

3.5 Results

The twenty sampled hospitals varied widely in size and scope: teaching hospitals, large regional hospitals, and smaller urban hospitals were all included. The number of beds ranged from 182 to 571, with a mean of 336 beds. Within these twenty hospitals, there were 33 696 heart failure discharges among 21 363

Algorithm 1 Regularized Cox Combination (Hospital k)

for all $i \in \{1, ..., n\}$ **do**

P_i=set of parameters selected through nested cross validation with Lasso penalty in a Cox model for hospitals i

end for

let $X_r^k = (x_1^k, ..., x_m^k)$ be a set of training data for hospital k

let $Y_r^k = (y_1^k, ..., y_m^k)$ be a real value outcome in the range $(0, \infty)$ indicating time to event with the Cox model for hospital X^k

let $C = \{\beta_{ij}\}$; $i \neq k \in \{1, 2, ..., n\}$ and $j \in \{1, ..., f_i\}$ be a set of coefficients from a set of Cox models H_i of all other hospitals

Expand X_r^k to have $\cup_{i=1}^n f_i$ features by inserting zeros for missing features

Run each model H_i on X_r^k to create hazard ratios $R_i(X_r^k)$

F =A new Cox model fit to X_r^k and R_i

	Pneumonia	AMI	Heart failure
Dispensed drug (code commune)	308	258	304
Medical procedure (CCP)	66	86	84
Diagnosis (ICD-9)	616	486	649

Table 3.1: Number of variables present in more than 30 discharges by disease type and variable category.

patients, 28 121 pneumonia discharges among 22 910 patients, and 21 468 AMI discharges among 18 876 patients.

In Table 3.1, we describe the number of included variables after excluding diagnoses, procedures or drugs that occurred fewer than 30 times. Table 3.5 presents descriptive information about the cohort and compares the average characteristics of the patients who were eventually readmitted and not readmitted. Across all three disease categories, previous readmissions were correlated weakly with future readmissions (Pearson's correlation coefficient was 0.12, 0.15, and 0.14

Hospital	Pneumonia discharges	30-day readmission	Local Model	Combined Model	Global Model
1	1 356	0.15	0.65	0.65	0.67
2	247	0.15	0.63	0.62	0.64
3	1 215	0.16	0.63	0.63	0.64
4	1 072	0.14	0.63	0.65	0.68
5	1 110	0.16	0.62	0.63	0.64
6	945	0.13	0.59	0.61	0.64
7	2 618	0.16	0.62	0.62	0.65
8	1 705	0.14	0.60	0.64	0.66
9	1 266	0.17	0.66	0.69	0.68
10	1 451	0.15	0.63	0.63	0.65
11	1 066	0.14	0.62	0.64	0.63
12	1 294	0.11	0.59	0.60	0.61
13	2 287	0.18	0.62	0.63	0.65
14	2 217	0.15	0.66	0.66	0.67
15	1 624	0.13	0.59	0.60	0.63
16	737	0.09	0.63	0.67	0.69
17	1 575	0.16	0.63	0.63	0.64
18	1 544	0.16	0.65	0.65	0.68
19	1 355	0.13	0.64	0.64	0.61
20	1 437	0.12	0.60	0.59	0.63
All	28 121	0.15	0.63	0.64	0.65

Table 3.2: Pneumonia discharges

for AMI, heart failure, and pneumonia respectively). The number of discharges increased from Sunday to Friday for all three disease categories, suggesting that considerations other than health status influenced the decision to discharge.

Tables 3.2–3.4 show the number of discharges, readmission rate, and performance of the local, global, and combined models for each of the 20 hospitals, and a summary of their performance. The hospitals had widely varying number of discharges: among the pneumonia discharges (Table 3.2) there was a mean of 1 406 discharges, with a standard deviation (SD) of 537 discharges, similar to

Hospital	AMI discharges	30-day readmission	Local Model	Combined Model	Global Model
1	887	0.16	0.65	0.66	0.68
2	1 717	0.13	0.65	0.64	0.65
3	686	0.17	0.64	0.63	0.64
4	1 166	0.15	0.59	0.61	0.63
5	855	0.16	0.65	0.66	0.66
6	997	0.15	0.57	0.59	0.60
7	1 734	0.18	0.67	0.68	0.68
8	1 401	0.14	0.65	0.65	0.67
9	904	0.20	0.63	0.63	0.64
10	884	0.19	0.55	0.57	0.58
11	820	0.16	0.59	0.61	0.63
12	1 356	0.16	0.61	0.62	0.63
13	950	0.16	0.61	0.62	0.65
14	1 582	0.17	0.68	0.68	0.68
15	1 454	0.13	0.59	0.59	0.60
16	487	0.13	0.61	0.64	0.68
17	941	0.19	0.63	0.65	0.66
18	1 007	0.15	0.66	0.67	0.68
19	552	0.14	0.58	0.61	0.65
20	1 088	0.17	0.60	0.61	0.62
All	21 468	0.16	0.63	0.63	0.65

Table 3.3: Acute myocardial infarction (AMI) discharges.

the heart failure discharges, which had a mean of 1 685 discharges, and a SD of 584 discharges; the AMI discharges had a mean 1 073 discharges and SD of 360 discharges.

The performance of the global models was typically higher than the combined and local models, but there were a few exceptions. In hospital 19, in Table 3.4, the global model (AUC: 0.61) underperforms the local model (AUC: 0.64), suggesting a heterogeneity of the predictors on readmission in this hospital compared to others. When pooling the performance across all twenty hospitals, the

Hospital	Heart failure discharges	30-day readmission	Local Model	Combined Model	Global Model
1	1 496	0.23	0.62	0.64	0.65
2	2 352	0.22	0.64	0.65	0.65
3	1 452	0.25	0.64	0.64	0.65
4	1 313	0.21	0.62	0.63	0.64
5	1 982	0.22	0.61	0.62	0.64
6	1 025	0.18	0.59	0.61	0.63
7	3 297	0.26	0.62	0.63	0.63
8	1 725	0.18	0.63	0.64	0.66
9	1 028	0.22	0.66	0.67	0.68
10	1 604	0.24	0.63	0.64	0.66
11	1 684	0.22	0.65	0.64	0.66
12	1 523	0.16	0.62	0.62	0.61
13	1 437	0.22	0.62	0.63	0.64
14	2 502	0.22	0.64	0.65	0.65
15	1 722	0.17	0.61	0.62	0.63
16	842	0.18	0.65	0.65	0.66
17	1 729	0.23	0.61	0.61	0.62
18	2 372	0.24	0.67	0.67	0.68
19	1 196	0.21	0.62	0.63	0.65
20	1 415	0.20	0.61	0.62	0.65
All	33 696	0.22	0.63	0.64	0.65

Table 3.4: Heart failure discharges.

AUC for heart failure, AMI, and pneumonia was (0.65, 0.65, 0.65) respectively for the global models, (0.63, 0.63, 0.63) for the local models, and (0.64, 0.63, 0.64) for the combined models.

The receiver-operating characteristic (ROC) is a summary of the trade-off between sensitivity and specificity; to understand how many more discharges would have been detected at a discharge level, we must choose a particular sensitivity and specificity pair. If we fix the specificity to the set {0.25, 0.50, 0.75, 0.90}, the sensitivities for the global model are {0.89, 0.71, 0.45, 0.13} respec-

tively, and for the local model the sensitivities are {0.87, 0.67, 0.41, 0.11}. On an absolute scale, since there were 7 355 total readmitted over 11 years, the global model would have detected {18, 28, 31, 10} more 30-day readmissions per year after heart failure (across all 20 hospitals).

3.6 Discussion

In this study, we compared the performance of three 30-day hospital readmission risk models: 1) an ideal "global" model, fit to data pooled across hospitals, 2) "local" models, fit to individual hospital data, and 3) "combined" models, fit to data from an individual hospital, but additionally using coefficients from local models from other hospitals. We fit the models to 11 years of administrative data from 20 hospitals. We found that the local models were nearly as accurate in predicting readmissions as the global and combined models. Although model combination or pooling all the data may provide some increase in performance, our results suggest that the benefit is small and may not be worth the cost.

Other model combination studies [61–65] focused on improving effect estimates of a few, common exposures (features) on rare outcomes, while this study focused on improving predictive accuracy for a relatively common outcome (readmissions) for many exposures. Both Wang and El Emam developed model combination techniques that compared well with global models, but did not provide a comparison to local models. In our study, we found that global and combined models had similar accuracy, but not meaningfully more accurate than local models. Rassen used propensity scores to combine models, and compared local and combined models. He found that local models were underpowered to precisely detect an effect, but that combined models could detect the expected effect (global model results were not provided). Wiens [78] compared global and local models in predicting infections, and found a slight improvement in accuracy in the global models, similar to our own study. Our results

suggest that for a common outcome, the improvement in accuracy would have a very small benefit in the number of newly detected cases; for a rarer outcome, the benefit would be even smaller.

In our study, the outcome was relatively common, but we had many sparse exposures; we had originally hypothesized that local models would have insufficient statistical power to precisely estimate the effects of the sparse exposure on the outcome. However, our results suggest that even with many, sparse exposures, local models could achieve similar accuracy to the global model.

One strength of our study was that we used data from hospitals which serve patient populations with heterogeneous relationships between predictors and readmission. Model combination studies using real or simulated data tend to horizontally partition the data into uniform samples [61–63], assuming that each hospital's patient population is a random sample from the global population of patients. Implicitly, this assumes that the relationship between the exposures and outcome is homogenous across all hospitals. All meta-analyses rely on the homogeneity of the relationship between exposures and outcome. However, pooling data across sites with heterogeneous relationships between exposures and readmissions may lead to a decrease in accuracy. Because we used real data, we could assess the difference in accuracy between local and pooled models in the presence of heterogeneity in the relationship between exposures and the outcome.

One limitation of our study is that we assessed model combination within a dataset where each site had exactly the same features, recorded in the same way, in the same city, making the data perhaps more homogeneous than might be expected in other settings. We expect that this homogeneity would have improved the accuracy of combined and global models.

We used the AUC to measure model accuracy, but we did not report variance because it was irrelevant for our conclusions. Although the AUC concisely summarizes the tradeoff between sensitivity and specificity (making it suitable for tabulation of the accuracy of a large number of models), it does not convey abso-

lute differences in correct predictions of readmission. From a clinical perspective, the absolute difference in the number of correctly identified patients for different models is more meaningful. Because the point estimates did not result in a meaningful clinical difference in correctly identified patients, we did not estimate the variance of these point estimates; whether these estimates were statistically significant would not have changed our conclusions.

The performance of all models, as measured by the AUC, was generally low, but very similar to other hospital readmission models [31]. The purpose of hospital readmission models is to focus transitional care interventions on those who would most likely benefit from them. If readmissions were often inevitable, they would likely be easier to predict, but transitional care interventions would not prevent them. Future work should focus on developing models that predict not just which patients will be readmitted, but specifically those who can benefit from a transitional care intervention.

3.6.1 Conclusion

When compared to the performance of locally built models, the improvement in performance from models built from pooled data was negligible, suggesting that for hospital readmissions, locally built models may suffice for practical use.

3.6.2 Acknowledgements

This work was supported in part by the U.S. National Science Foundation grants IIS-1231742 and IIS-1242304.

		Pneumonia		AMI		Heart failure	
		N	30-day readmission	N	30-day readmission	N	30-day readmission
Sex	F	14 183	1 918 (13.5)	9 703	1 660 (17.1)	17 526	3 754 (21.4)
	M	13 938	2 209 (15.8)	11 765	1 754 (14.9)	16 170	3 601 (22.3)
Age (years)	(65,75]	10 102	1 389 (13.7)	9 851	1342 (13.6)	12 390	2 660 (21.5)
	(75,85]	12 244	1 872 (15.3)	8 536	1461 (17.1)	14 737	3 325 (22.6)
	(75,85]	12 244	1 872 (15.3)	8 536	1 461 (17.1)	14 737	3 325 (22.6)
	(85,95]	5 375	800 (14.9)	2925	580 (19.8)	6 194	1 303 (21.0)
	>95	400	66 (16.5)	156	31 (19.9)	375	67 (17.9)
Length of stay (days)	(1,5]	8 532	1 053 (12.3)	6667	843 (12.6)	10 009	2 076 (20.7)
	(5,8]	6 330	966 (15.3)	4847	761 (15.7)	7 600	1 586 (20.9)
	(8,14]	6 544	1 035 (15.8)	4884	876 (17.9)	6 831	1 557 (22.8)
	>14	6 398	1 036 (16.2)	4997	928 (18.6)	8 406	1 980 (23.6)
Transfers	0	346	67 (19.4)	2 365	333 (14.1)	903	234 (25.9)
	1	80	17 (21.2)	960	134 (14.0)	260	58 (22.3)
	2	6	2 (33.3)	150	26 (17.3)	47	9 (19.1)
	≥3	4	0 (0.0)	36	6 (16.7)	11	1 (9.1)
Previous readmissions	1	5 280	685 (13.0)	4 387	680 (15.5)	6 163	1 080 (17.5)
	2	3 609	569 (15.8)	2 532	457 (18.0)	4 807	957 (19.9)
	3	2 550	413 (16.2)	1 521	286 (18.8)	3 638	803 (22.1)
	4	1 824	324 (17.8)	931	211 (22.7)	2 777	687 (24.7)
	5	1 312	253 (19.3)	635	146 (23.0)	2 085	548 (26.3)
	6	909	203 (22.3)	403	98 (24.3)	1 551	465 (30.0)
	≥7	2 760	727 (26.3)	1 013	313 (30.9)	5 082	1 707 (33.6)
Nursing home?	No	26 429	3 904 (14.8)	20 833	3 320 (15.9)	32 393	7 141 (22.0)
	Yes	1 692	223 (13.2)	635	94 (14.8)	1 303	214 (16.4)
Day of week of discharge	Sun	2 051	297 (14.5)	1 499	234 (15.6)	2 431	521 (21.4)
	Mon	3 744	538 (14.4)	2 480	404 (16.3)	4 291	977 (22.8)
	Tue	4 506	679 (15.1)	3 261	508 (15.6)	5 000	1110 (22.2)
	Wed	4 760	693 (14.6)	3 663	598 (16.3)	5 500	1159 (21.1)
	Thu	4 597	648 (14.1)	3 685	575 (15.6)	5 699	1243 (21.8)
	Fri	5 839	907 (15.5)	4 487	730 (16.3)	7 056	1525 (21.6)
	Sat	2 624	365 (13.9)	2 393	365 (15.3)	3 719	820 (22.0)
Year of discharge	1996	1 978	272 (13.8)	1 215	227 (18.7)	2 582	636 (24.6)
	1997	2 558	321 (12.5)	1 533	262 (17.1)	3 329	778 (23.4)
	1998	2 983	450 (15.1)	1 807	311 (17.2)	3 369	757 (22.5)
	1999	2 857	410 (14.4)	1 824	327 (17.9)	3 147	730 (23.2)
	2000	2 716	393 (14.5)	2 106	340 (16.1)	3 317	704 (21.2)
	2001	2 548	389 (15.3)	2 133	322 (15.1)	3 223	692 (21.5)
	2002	2 367	349 (14.7)	2 122	328 (15.5)	3 052	641 (21.0)
	2003	2 109	314 (14.9)	2 227	326 (14.6)	2 979	618 (20.7)
	2004	2 574	431 (16.7)	2 269	343 (15.1)	2 831	602 (21.3)
	2005	3 013	470 (15.6)	2 265	343 (15.1)	2 939	641 (21.8)
	2006	2 418	328 (13.6)	1 967	285 (14.5)	2 928	556 (19.0)

Table 3.5: Selected variables and their distribution by disease and 30-day readmission.

Chapter 4

Hospital readmissions and the day-of-the-week

4.1 Preamble

Individual-level interventions, like transitional care interventions, can reduce readmissions, but even small system-level interventions have the potential to reduce readmission significantly. For example, a system-level intervention like increased weekend staffing may reduce the rate of adverse patient safety events slightly, but since the intervention affects all patients, the total effect may be large. But system-level interventions can have complex dynamics that make inference difficult. For example, increased weekend staffing may reduce adverse patient safety events, but it may also have unexpected effects, like increased weekend admission.

Previous studies have shown that patients discharged on a Friday tend to be readmitted sooner than Wednesday discharges. Although the mechanism of this effect is not clear, several system-level interventions have been proposed, such as increasing weekend staff or allowing admissions into long-term care on the weekend. Since over 20% of discharges occur on Friday any systemic change

to the time-to-readmission of Friday discharges can have large population effects. However, the probability of admission on the weekend is also much lower than on weekdays. The effect of discharge day-of-week on readmission may be complicated by the effect of the day-of-week on admission.

In the following manuscript, I examine how the association between Friday discharge and hospital readmission is modified by the low probability of weekend admission.

4.2 Abstract

Background: Patients discharged from hospitals on a Friday (Friday discharges) tend to be readmitted sooner than Wednesday discharges. Some authors have suggested that readmissions related to Friday discharges can be decreased by increasing weekend staff or allowing admissions into long-term care on the weekend. However, because the effect of differing *admission* probability by day-of-week has not been investigated, the estimated effect of interventions may be biased.

Objective: To examine how differing healthcare-seeking behaviour and admission practices by day-of-week influences the effect of discharge day on readmission.

Methods: We extracted discharges of people 65 years of age or older between 1996 and 2006 (inclusive) from the Quebec hospital administrative database. We used a Markov model to determine the effect of admission on readmission if Friday discharges were not at increased risk of readmission. We then used a Cox proportional hazards model to fit the time-to-emergency-readmission to any Quebec hospital as a function of the day of week of discharge and readmission. We then fit another Cox model with an additional time-varying covariate for the current day of week, to model differing admission probabilities by day-of-week.

Results: We identified 619 274 hospital discharges during the study period; 22% were Friday discharges, and 17% were Wednesday discharges. Our Markov model showed we should expect Friday discharges to be admitter later if admission probability is lower on the weekends. Using only the discharge and admission day of week, we found that Friday discharges were readmitted slightly earlier than Wednesday discharges [HR: 1.03 95% CI: (1.02, 1.05)]. After adding a time-varying covariate for the current day of week, a Friday discharge was still readmitted sooner than a Wednesday discharge [HR: 1.04 95% CI: (1.01, 1.07)].

Conclusions: Lower admission probabilities on the weekend *increase* the time-to-readmission for Friday discharges, because Fridays occur soon before the weekend. Not controlling for low weekend admission causes an underestimate of the effect of Friday discharge on readmission.

4.3 Introduction

Patients discharged on Friday (Friday discharges) have been found to be readmitted sooner than Wednesday discharges [14, 79, 80]. Although the relationship between Friday discharges and readmissions is weak, the exposure is common, so any intervention that ameliorates the effect may prevent many patient-days in the hospital. However, developing effective interventions requires a clear understanding of the mechanism through which discharges on different days of the week affect readmission.

Several mechanisms have been proposed to explain the relationship between Friday discharges and readmission, including factors that affect patient and physician preference, and the availability of social and health services on the weekends. Because hospitals are typically run with fewer, more inexperienced staff on the weekends, weekday physicians may prefer to discharge patients before the weekend, while the quality of discharge preparation is under their control. Furthermore, there may be subtle social pressures to “clean up” the wards

(by discharging patients -- perhaps before they are completely medically stable) before a new physician starts their shift on Monday. These factors may explain why Friday is the most common discharge day. However, the multiple discharges on Fridays may result in rushed and incomplete discharge preparation, leading to quicker readmission. Also, if a patient is discharged on a Friday to their home, then community health and social services may not be available for the first few critical days after discharge, when readmission risk is highest. Indeed, because long-term care services typically do not admit patients on the weekend, physicians may discharge on Friday to ensure a direct transition, rather than wait till the following Monday. Finally, patients themselves may desire to be discharged before the weekend (and before they are completely medically stable), to spend time with family and friends who are typically more available on the weekend.

Depending on the true mechanism of the effect of Fridays on readmissions, increasing staff during the weekends in both hospitals and other long-term facilities may help prevent readmissions [66, 67]. Some authors have suggested that increasing weekend services may also help to reduce congestion in the emergency department [68].

Although patient, physician, and weekend staff levels may underlie the mechanism of Friday discharges affecting readmission, differing probabilities of *admission* by day-of-week may also influence the probability of *readmission* by day-of-discharge. For certain illnesses and conditions, admission is certain, but for others, the responsible physician has a strong influence in deciding if the patient is admitted [81–84]. Epstein [44] found that patient readmission rates and survey-reported hospital admission rates were correlated. If the day-of-week affects admission probabilities, then the increased time-to-readmission of Friday discharges may simply be an artifact of the admission probabilities.

If patient behaviours, physician practices, or the availability of weekend services lead to quicker readmission of Friday discharges, then we may be able to identify a health-improving, cost-saving intervention. However, before im-

plementing such interventions, we need to be sure that they are cost-effective, meaning that we need to precisely measure the effect of Friday discharge on readmission. In particular, we need to control for the effect of the differing probability of admission on different days of the week to estimate the effect of Friday discharges on readmission. In this study, we used analytic and empirical methods to study how the probability of admission influences the effect of discharging on a Friday on readmission.

4.4 Methods

To study the effect of day-of-discharge on hospital readmissions, we extracted a large cohort from the Quebec administrative database of hospitalization information, obtained from the RAMQ. This extract includes all hospitalizations and outpatient visits on all 3.6 million people who ever reported a respiratory illness between the years 1996 and 2006 (inclusive) and lived in the CMA of Montreal, as defined by the 2006 census. Respiratory illness was defined as any of the ICD-9 codes listed in Table 6.1, a relatively broad set of codes, among which there are several extremely common codes. This was a dynamic cohort; a patient entered the cohort after their first respiratory illness. This data source was extracted for a different study; we used it because it includes enough discharges to allow us to measure small effects, is unlikely to have a selection bias with respect to hospital readmissions, and is readily available.

We extracted all discharges of people who were 65 years of age or older at the time of discharge. We then calculated the number of admissions by hospital, and only included the top 20 hospitals.

4.4.1 Definition of readmission

Notably, any readmission to another hospital on the same day as discharge, or the next day, was considered a transfer, rather than a readmission. We applied this rule because we suspect that many transfers are not correctly coded in the administrative data as a transfer, but simply as a readmission.

An "event" occurred if the discharged patient had an emergency readmission to *any* Quebec hospital, including those that were not part of the 20 hospitals described in the selection criteria. The time-to-readmission was measured in the number of days. If the discharged patient died, had a non-emergency hospital readmission, moved out of Quebec, or if the study ended (December 31st, 2006) they were considered censored (they did not have the outcome event).

4.4.2 Descriptive analysis

We first developed descriptive summary statistics and visualizations to determine on which day people tended to be admitted and discharged. We plotted the Kaplan-Meier survival curves for Friday discharges and Wednesday discharges. To emphasize failure times by day, we also plotted the probability of readmission on each day after discharge.

For each year in the cohort, we plotted a heat map of the discharge rate by day, and arranged the pixels like a calendar. All *jours fériés* (Quebec statutory holidays) were identified on our heat map, namely: New Year's Day, the day after New Year's Day, Easter Monday, *Journée nationale des Patriotes*, *Fête nationale du Québec*, Canada Day, Labour Day, Thanksgiving Day, Christmas Eve, Christmas Day, Boxing Day, and New Year's Eve. These heat maps allowed us to visually inspect how the days of the week affected discharges, and how that effect was modified by the presence of holidays.

4.4.3 Analytic model

We suspected that the effect of the day of discharge on time-to-readmission was modified by the probability of *admission* on different days of the week. We used an absorbing Markov model to predict the time-to-readmission for each discharge day, assuming that there was no effect of the day of week on discharge independent of the varying admission probability on each day. For example, the Markov model could compare the difference in time-to-readmission for Friday discharges and Wednesday discharges if the only factor driving readmission was that weekend admissions were less likely than weekday admissions. This model was not fit to any data. The purpose of the model was to understand how weekend admission probabilities change the relationship between discharge day and readmission time.

Our absorbing Markov model consisted of eight states, seven for the days of the week, and one "absorbing" state (with zero probability of transitioning into another state) which represented readmission. Each day-of-week state could transition into two states: the readmission state, and the next day of the week. A person's expected time-to-readmission could be calculated by finding the expected number of transitions before reaching the readmission state, if the starting state was the state associated with the discharge day. For example, a person discharged on a Friday would begin in the Friday state, and the expected time-to-readmission (in days) would be the expected number of transitions before reaching the readmission state.

The expected number of state transitions for any beginning state for an absorbing Markov model can be derived in closed form (see the 4.5 section for details). Informed by our descriptive analysis, we assumed that the probability of admission on a weekday was some fixed probability $1 - a$, and the probability of admission on a weekend was some fixed probability $1 - b$. We then calculated the expected time-to-readmission for Friday and Wednesday discharges as functions of a and b.

4.4.4 Statistical analysis

We then fit a Cox proportional hazards model to estimate the effect of discharge day on readmission, only controlling for time-fixed confounders. The exposure, the day of discharge, was represented in the model using an indicator variable, except Wednesday, which was the reference category. The outcome was the time-to-readmission in days, with "time zero" being the day of discharge. The unit of analysis was the discharge; a single person could have several discharges.

We controlled for several time-fixed confounders, the admission day-of-week, whether the discharge or admission day was a holiday, whether the discharge or admission day was the day after a holiday, age, sex, the number of previous drug prescriptions within the past year, the number of previous discharges within the past year, and the classification of the admission ICD-9 code into one of the 24 major diagnostic category (MDC). We did not expect strong confounding by age, sex, previous drug prescriptions, admission class and MDC because we expected a weak correlation between these variables and our exposure, the day of discharge. We included these factors mainly as a diagnostic tool to ensure that our models could reproduce expected effects, such as severe illnesses and age having strong effects of readmission.

We then fit another Cox proportional hazards model that added a single, time-varying covariate for the day-of-week *after* discharge. For example, a single discharge has several time-fixed covariates: they were discharged on a Friday, admitted on a Monday, and had two previous admissions, etc. But on the next day after discharge, the day-of-week changes, so the time-varying variable changes to Saturday, while all of the other time-fixed variables remain the same. This allows the model to account for a varying hazard of admission by day-of-week.

4.4.5 Software

The data were prepared for statistical analysis using the Postgres relational database (version 9.3.6). We implemented our models using the R statistical package (version 3.1.1), [85] using the "survival" package (version 2.38.1) to fit the Cox proportional hazards models [86]. We used YACAS (Yet Another Computer Algebra System) 1.3.3 to conduct the linear algebra computations used in the Markov models [87]. We plotted our figures using the "ggplot2" package (version 1.0.1) [88].

4.5 Results

4.5.1 Descriptive analysis

In Figure 4.1, the relative proportions indicate that there are both fewer emergency discharges and admissions on the weekend, as compared to weekdays. During the weekdays, the admissions peak on Monday (16%), and then decline slowly. However, the Friday was the most common discharge day (22%), while only 10% were discharged on a Saturday, and only 6% were discharged on Sunday.

The calendar plot (Figure 4.2), which plots the discharge rate per day, indicates that discharge rates are lower on weekends than weekdays, and that Fridays have a particularly high discharge rate, reflecting the trends in Figure 4.1. Additionally, the rate of discharge declines over the study period. Because one selection criterion for cohort entry was an ILI diagnosis, patients with higher frequency of diagnoses (the severely ill) tended to enter the cohort earlier in the study period. Since a greater proportion of the cohort was severely ill in the early part of the study period, we expected that the discharge rate would be higher in the early part of the cohort.

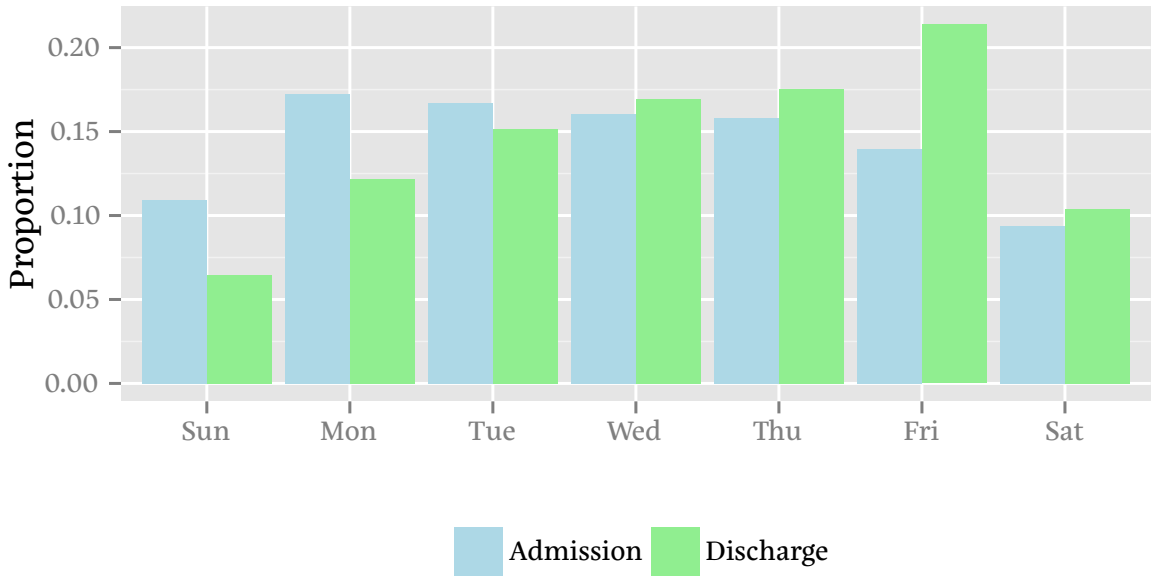

Figure 4.1: Emergency admissions and discharges by day-of-week. The denominator in each proportion is the number of emergency admissions that ended in live discharges, meaning that those that died during the hospital stay or were admitted for non-emergency reasons are not counted.

Additionally, holidays that occur on a weekday tend to result in lower discharge rates. The holiday that occurs on one of the Mondays of the 14th to 18th week of each year is Easter Monday. In Quebec, some employers give employees the option to take the Friday preceding Easter Monday as a holiday instead of Easter Monday. In the plot, it appears that the Friday preceding Easter Monday acts as a holiday in the sense that it reduces the discharge rate. Additionally, however, it appears that a holiday on a Friday "displaces" the discharge effect; in the week preceding Easter Monday, *Thursday* has the highest discharge rate.

The Kaplan-Meier plot of survival (the probability of not being readmitted) is shown for Friday and Wednesday discharges in Figure 4.3. The figure suggests a very slight increase in risk for Friday discharges. The decrease is very smooth; there is no particular day with a sudden decrease in risk.

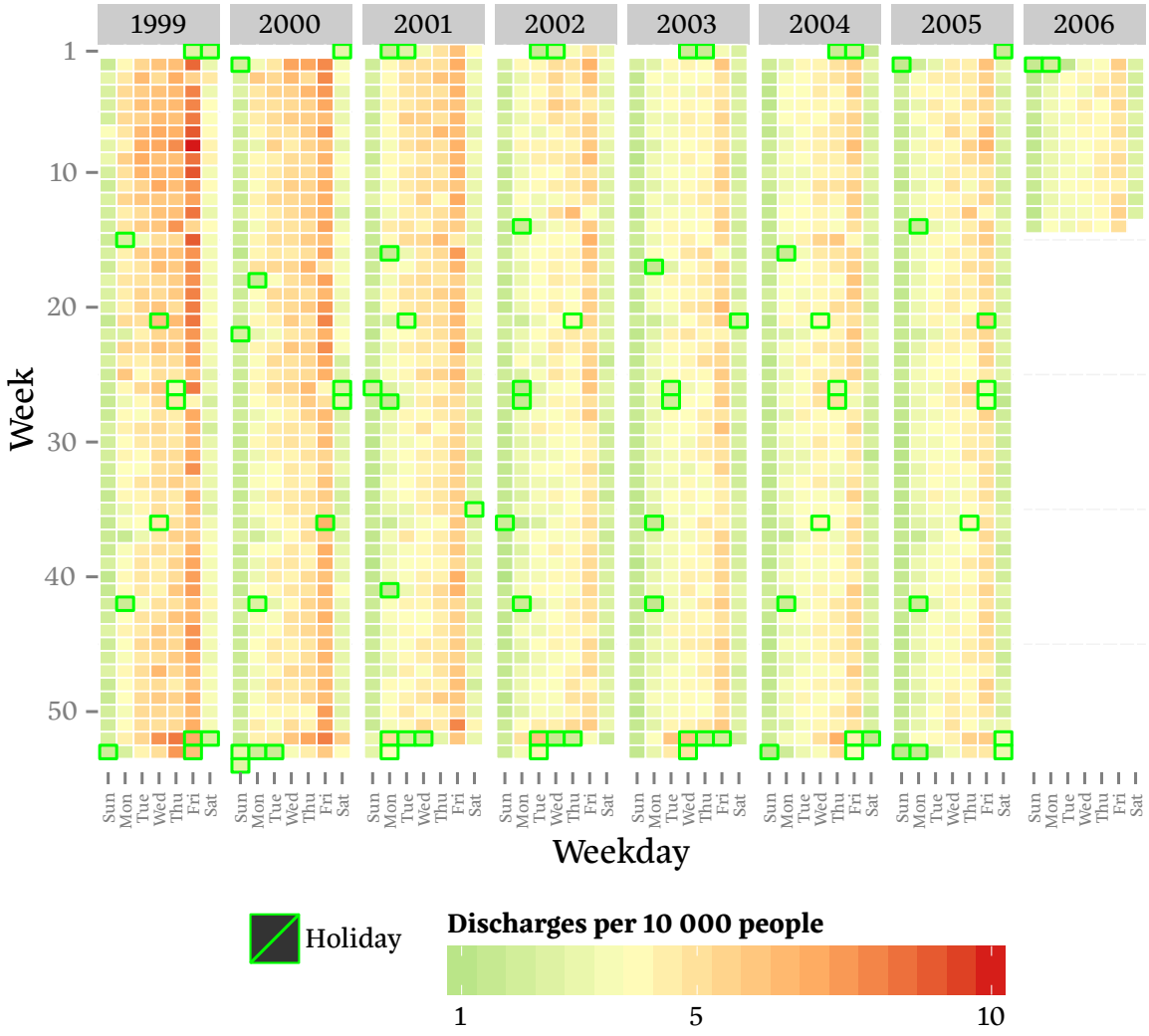

Figure 4.2: Discharge rate by day. Each pixel represents a single day in the study period. Not all years in the study period are displayed, for aesthetic reasons. The pixels are arranged to resemble calendars; within each year, each row is a week and each column is a weekday. The colour of each pixel represents the proportion of the cohort that was discharged on that day. Holidays have a green border.

In Figure 4.4, we show, for Friday and Wednesday discharges, the readmission probability by days after discharge, given survival to that day. For example, if

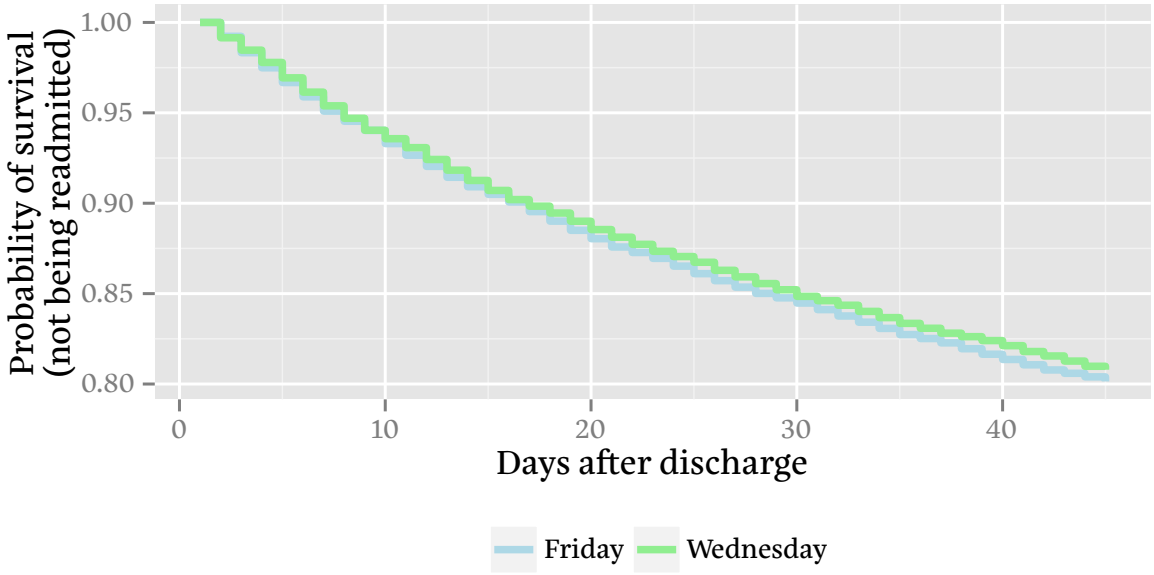

Figure 4.3: Probability of not being readmitted by day-of-week. This figure is a Kaplan-Meier plot of probability of readmission for discharges on a Friday, and discharges on a Wedensday. Notably, the vertical axis does not display the full range of probabilities. Only the first 45 days after discharge are displayed, but patients were not automatically censored after this time.

a patient discharged on a Friday hasn't been readmitted for six days after discharge (a Thursday), the probability of readmission on the seventh day (which will be a Friday) is 0.8%. The figure shows a marked difference in readmission probability depending on the day of the week after discharge: regardless of the day you were discharged, you are much less likely to be readmitted on Saturdays and Sundays.

4.5.2 Markov models

Informed by our descriptive analysis, we assumed that the probability of admission on a weekday was some fixed probability a, and the probability of admission on a weekend was some fixed probability b. We used a Markov model

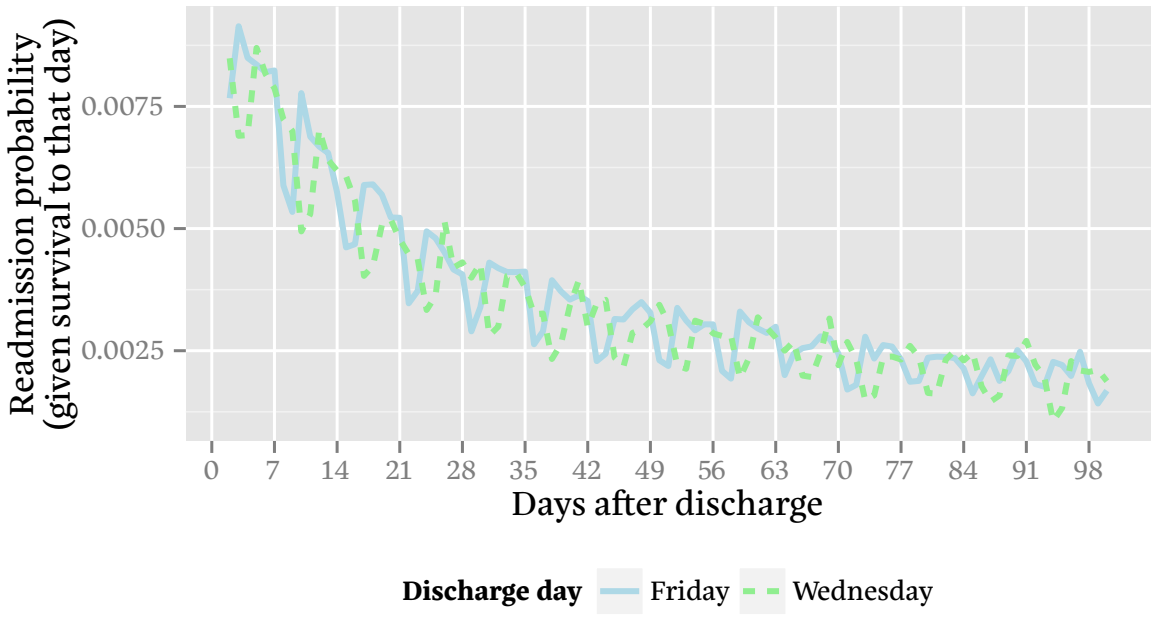

Figure 4.4: Readmission probability by day of discharge. For both Friday and Wednesday discharges, this figures describes the probability of readmission given survival for the number of days in the horizontal axis. The horizontal axis is marked every seven days to emphasize day-of-week trends; for Friday discharges the seventh day after discharge is a Friday.

to calculate the time-to-readmission for Friday discharges and Wednesday discharges, if the probability of admission was a function of whether the current day-of-week was a weekend or weekday. The graph G, described schematically in Figure 4.5 represents such a model of readmission in which the probability of admission on the weekdays is $1 - a$, and the probability of admission on weekends is $1 - b$.

The graph G can also be described as a matrix of transition probabilities, where the probability of transition from state s_1 to state s_2 is in the row for s_1 and the column for s_2 in the matrix. The matrix M is such a matrix describing G:

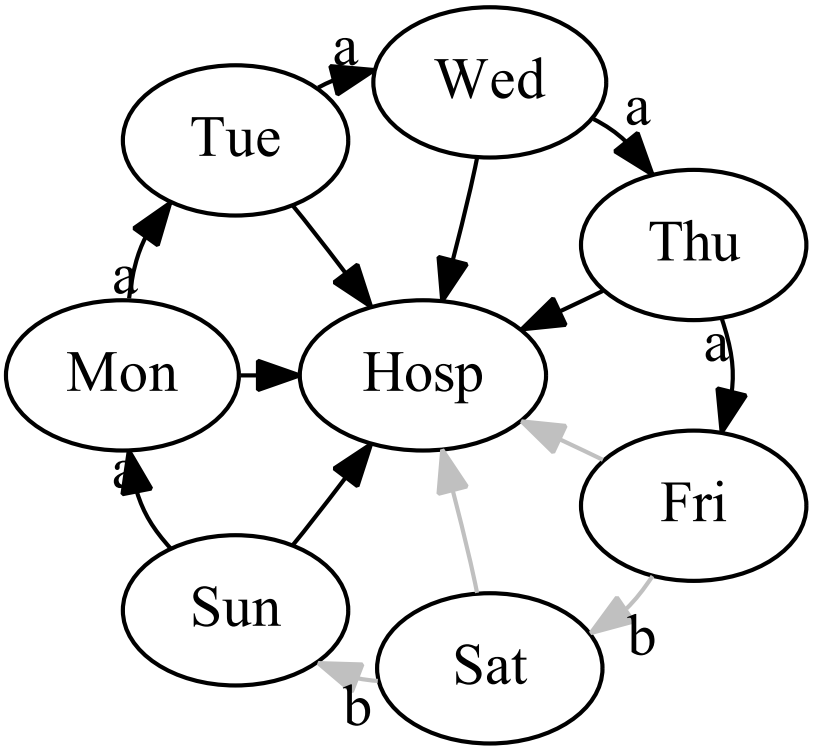

Figure 4.5: An absorbing Markov model of readmission probability.

$$\mathbf{M} = \begin{array}{c} \\ Sun \\ Mon \\ Tue \\ Wed \\ Thu \\ Fri \\ Sat \\ Hosp \end{array} \begin{array}{c} \begin{array}{cccccccc} Sun & Mon & Tue & Wed & Thu & Fri & Sat & Hosp \end{array} \\ \begin{pmatrix} 0 & a & 0 & 0 & 0 & 0 & 0 & 1-a \\ 0 & 0 & a & 0 & 0 & 0 & 0 & 1-a \\ 0 & 0 & 0 & a & 0 & 0 & 0 & 1-a \\ 0 & 0 & 0 & 0 & a & 0 & 0 & 1-a \\ 0 & 0 & 0 & 0 & 0 & a & 0 & 1-a \\ 0 & 0 & 0 & 0 & 0 & 0 & b & 1-b \\ b & 0 & 0 & 0 & 0 & 0 & 0 & 1-b \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

By representing a Markov model as a matrix M of transition probabilities, M^n represents the probability of starting in position i and ending on state j by step n. The sum of each row of M^n is exactly 1, for any value of $n \geq 0$.

To find the average number of transitions (days) before reaching the absorption state (being readmitted), we first find the probability of not being readmitted after any given number of days n. If Q is the transition matrix M with the row and column for the absorbing state removed, then the sum of the probabilities from position i to any other state j in the matrix IQ^n represents the probability of not being absorbed by transition n.

$$Q = \begin{pmatrix} 0 & a & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & a & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & b \\ b & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

We can then calculate a matrix N where the expected number of visits to state j starting in state i before absorption is represented by the ith row and jth column.

$$\begin{aligned} N &= \lim_{n\to\infty}(I + Q + Q^2 + \cdots + Q^n) \\ &= \lim_{n\to\infty}(I - Q)^{-1}(I - Q)(I + Q + Q^2 + \cdots + Q^n) \\ &= \lim_{n\to\infty}(I - Q)^{-1}((I - Q) + (Q - Q^2) + (Q^2 - Q^3) + \cdots + (Q^{n+1})) \\ &= \lim_{n\to\infty}(I - Q)^{-1}(I - Q^{n+1}) \\ &= (I - Q)^{-1}(I - 0) \\ &= (I - Q)^{-1} \end{aligned}$$

Where I is the identity matrix with 7 rows and columns, and N is the inversion of the matrix $(I - Q)$:

$$I - Q = \begin{pmatrix} 1 & -a & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -a & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -a & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -a & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -a & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -b \\ -b & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

We found the inverted matrix to be:

$$(I - Q)^{-1} = \frac{1}{(1 - a^5b^2)} \begin{pmatrix} 1 & a & a^2 & a^3 & a^4 & a^5 & a^5b \\ a^4b^2 & 1 & a & a^2 & a^3 & a^4 & a^4b \\ a^3b^2 & a^4b^2 & 1 & a & a^2 & a^3 & a^3b \\ a^2b^2 & a^3b^2 & a^4b^2 & 1 & a & a^2 & a^2b \\ ab^2 & a^2b^2 & a^3b^2 & a^4b^2 & 1 & a & ab \\ b^2 & ab^2 & a^2b^2 & a^3b^2 & a^4b^2 & 1 & b \\ b & ab & a^2b & a^3b & a^4b & a^5b & 1 \end{pmatrix}$$

Since N represents the expected number of visits to state j starting in state i before absorption, the expected number of visits to all states given a starting state j is simply the sum of the rows (also given by $(I-Q)^{-1}c$) where c is a column vector of ones:

$$M = \frac{1}{(1 - a^5b^2)} \begin{pmatrix} a^5b + a^5 + a^4 + a^3 + a^2 + a + 1 \\ a^4b^2 + a^4b + a^4 + a^3 + a^2 + a + 1 \\ a^4b^2 + a^3b^2 + a^3b + a^3 + a^2 + a + 1 \\ a^4b^2 + a^3b^2 + a^2b^2 + a^2b + a^2 + a + 1 \\ a^4b^2 + a^3b^2 + a^2b^2 + ab^2 + ab + a + 1 \\ a^4b^2 + a^3b^2 + a^2b^2 + ab^2 + b^2 + b + 1 \\ a^5b + a^4b + a^3b + a^2b + ab + b + 1 \end{pmatrix}$$

The absolute difference between the time-to-readmission between Friday and Wednesday is:

$$
\begin{aligned}
RD &= \frac{(a^4b^2 + a^3b^2 + a^2b^2 + ab^2 + b^2 + b + 1) - (a^4b^2 + a^3b^2 + a^2b^2 + a^2b + a^2 + a + 1)}{(1 - a^5b^2)} \\
&= \frac{ab^2 + b^2 + b - a^2b - a^2 - a}{1 - a^5b^2} \\
&= \frac{(a+1)(b+1)(b-a)}{a^5b^2 - 1}
\end{aligned}
$$

Since $\frac{(a+1)(b+1)(b-a)}{a^5b^2-1}$ will always exceed 0 if $a > b$ (and a and b are between 0 and 1), then a lower probability of admission on the weekend implies a longer time-to-readmission (a lower hazard) for Friday discharges as compared to Wednesday discharges, the opposite of what we observed in our descriptive analysis.

4.5.3 Survival analysis

Table 4.1 shows the fitted coefficients of both the time-varying and the time-fixed models. In the time-fixed model, Friday discharges have significantly higher hazard of readmission than Wednesday discharges [HR:1.03 (1.02-1.05)], but Monday, Thursday, and holiday discharges also had significantly higher hazards than Wednesdays. Also, those admitted on the weekend, or on the day after a holiday, had a significantly higher hazard than Wednesday admissions.

In the model with time-varying variables, the time-varying variables had much stronger effects than the time-fixed variables. Both weekend variables and holiday variables had lower hazards of readmission, while Mondays and the day after holidays had a higher rate of admission. Friday discharges continued to have a higher hazard of readmission than Wednesday discharges, but Saturdays and the day after holidays also had significant effects on readmission. None of the admission variables had any effect on readmission in the time-varying model.

Variable	Time-fixed HR (95% CI)	Time-varying HR (95% CI)
Admission		
Sunday	1.01 (1.00—1.03)	1.00 (0.97—1.04)
Monday	1.01 (0.99—1.02)	0.99 (0.96—1.02)
Tuesday	0.99 (0.98—1.01)	0.98 (0.95—1.01)
Wednesday	(Reference)	(Reference)
Thursday	0.99 (0.98—1.00)	1.00 (0.97—1.03)
Friday	1.00 (0.99—1.02)	1.00 (0.97—1.03)
Saturday	1.01 (1.00—1.03)	1.01 (0.97—1.04)
Holiday	1.01 (0.99—1.03)	0.97 (0.93—1.02)
Post-holiday	1.03 (1.01—1.06)	1.00 (0.94—1.06)
Discharge		
Sunday	0.97 (0.96—0.99)	0.98 (0.94—1.01)
Monday	1.01 (1.00—1.02)	1.02 (0.99—1.05)
Tuesday	1.00 (0.99—1.01)	1.01 (0.98—1.04)
Wednesday	(Reference)	(Reference)
Thursday	1.01 (1.00—1.02)	1.02 (1.00—1.05)
Friday	1.03 (1.02—1.05)	1.04 (1.01—1.07)
Saturday	1.01 (0.99—1.02)	1.05 (1.01—1.08)
Holiday	1.03 (1.01—1.05)	1.04 (0.99—1.10)
Post-holiday	0.99 (0.96—1.02)	1.07 (1.00—1.14)
Time-varying (post-discharge)		
Sunday	—	0.76 (0.74—0.79)
Monday	—	1.05 (1.02—1.08)
Tuesday	—	0.99 (0.96—1.02)
Wednesday	—	(Reference)
Thursday	—	1.00 (0.97—1.03)
Friday	—	0.98 (0.95—1.01)
Saturday	—	0.76 (0.73—0.78)
Holiday	—	0.91 (0.87—0.95)
Post-holiday	—	1.06 (1.00—1.12)

Table 4.1: The effect of time-related variables on readmission. The hazard ratios and associated 95% confidence intervals are presented for the day-of-week variables for two Cox proportional hazards models of readmission. The time-fixed model includes indicator variables for the day of discharge and the day of admission, using Wednesday as the reference day of week. The time-varying model adds a set of time-varying indicator variables to represent the current day of week after discharge. Each model controlled for the variables described in table 4.2.

Variable	Time-fixed HR (95% CI)	Time-varying HR (95% CI)
Demographic and healthcare use		
Age (years)	1.01 (1.01—1.01)	1.01 (1.01—1.02)
Sex (male)	1.12 (1.12—1.13)	1.13 (1.11—1.15)
Previous readmissions (each)	1.07 (1.07—1.07)	1.07 (1.07—1.07)
Outpatient prescriptions (each)	1.04 (1.04—1.05)	1.05 (1.04—1.05)
Major Diagnostic Category		
Alcohol/drug-related disorders	1.38 (1.28—1.48)	1.40 (1.17—1.67)
Blood and immunological disorders	1.47 (1.43—1.51)	1.53 (1.43—1.65)
Burns	0.74 (0.56—0.98)	0.76 (0.32—1.84)
Circulatory system	1.27 (1.25—1.29)	1.29 (1.25—1.34)
Digestive system	1.20 (1.18—1.22)	1.22 (1.17—1.28)
Ear, nose, mouth and throat	1.07 (1.03—1.10)	1.04 (0.96—1.12)
Endocrine system	1.30 (1.27—1.34)	1.32 (1.23—1.41)
Eye	0.89 (0.85—0.94)	0.97 (0.86—1.08)
Factors influencing health status	1.22 (1.18—1.26)	1.24 (1.15—1.34)
Hepatobiliary system and pancreas	1.21 (1.18—1.24)	1.20 (1.14—1.27)
Human immunodeficiency virus	1.76 (0.79—3.92)	1.08 (0.15—7.66)
Infectious and parasitic diseases	1.14 (1.11—1.18)	1.22 (1.13—1.31)
Injuries and poison	1.13 (1.09—1.18)	1.09 (0.99—1.20)
Kidney and urinary tract	1.24 (1.22—1.27)	1.26 (1.20—1.33)
Mental diseases and disorders	1.05 (1.03—1.07)	1.07 (1.01—1.13)
Musculoskeletal system	(Reference)	(Reference)
Myeloproliferative neoplasms	2.60 (2.51—2.69)	2.57 (2.37—2.79)
Nervous system	1.04 (1.02—1.06)	1.07 (1.02—1.12)
Reproductive system	1.16 (1.12—1.21)	1.19 (1.09—1.31)
Respiratory system	1.57 (1.55—1.60)	1.60 (1.54—1.67)
Skin and subcutaneous tissue	1.19 (1.16—1.22)	1.22 (1.15—1.29)

Table 4.2: The effect of non-time-related variables on readmission. The hazard ratios and associated 95% confidence intervals are presented for the demographic and disease-related variables for two Cox proportional hazards models of readmission. These models included the variables described in Table 4.1

Table 4.2 shows the other variables that were fit with the time-varying and time-fixed models. Nearly all of the included variables were significant. The coefficients were roughly the same for all non-time related variables in both the time-fixed and time-varying models.

4.6 Discussion

In this study, we used different modelling techniques applied to empirical data to estimate the effects of the day of week of discharge (and admission) on readmission. In a survival model with only time-fixed covariates, we found that Friday discharges tend to be readmitted sooner than Wednesday discharges. This association is observed despite there being fewer admissions and discharges and weekends, which according to our Markov model, should result in Friday discharges being readmitted *later* than Wednesday discharges. Using a survival model with time-varying covariates, we controlled for the varying probability of admission on different days of the week, and found that controlling for the effect of the probability of admission enhanced the strength of the effect of Friday discharge on readmission.

Our time-fixed model of the effect of discharge day-of-week was comparable to the model published by van Walraven and Bell [14] in that Friday discharges appeared to have a shorter time-to-readmission than Wednesday discharges (although we did not censor 30 days after discharge). In the time-fixed model, perhaps due to the greater precision of our model, we also found that those discharged on Mondays, Thursdays, or holidays, also had a slight but significantly shorter time-to-readmission than Wednesday discharges.

Our descriptive analysis indicated that patients are discharged preferentially on weekdays as compared to weekends, and most frequently just before the weekend. Usually, being discharged just before the weekend means being discharged on Friday, but when other holidays occur near the weekend, the day-

of-the week can be different. For example, the day before Good Friday, the discharge rate significantly increased, while the Good Friday discharges are very low, suggesting that patients are being preferentially discharged before holidays, to avoid discharge during holidays.

Our descriptive analysis also found that admissions were decreased on the weekends. There are effectively three categories of hypotheses for the low probability of admission on weekends:

1) Patient healthcare-seeking behaviour may change on the weekends. Some patients who have paid sick leave may be motivated to "miss work" by delaying healthcare until the next weekday. However, in our cohort, we only included those at least 65 years of age, a mostly retired population. Patients may also have less access to public transit and other services that would help them access care on the weekend. Finally, patients have less access to primary care services on the weekend [89]. On one hand, this should increase healthcare-seeking on weekends, because of a lack of options. On the other hand, patients may make the decision to seek healthcare at a hospital after a primary care visit, because of physician recommendation.

2) Physicians don't admit as often on the weekends. Although for some medical emergencies, admission to a hospital is certain, in most cases, emergency physicians themselves affect the probability of admission; physician experience, age, and risk preferences can affect the probability that a patient will be admitted [81–84]. Also, less staff work on holidays and weekends, and teaching hospitals are typically staffed by medical residents [67]. When the hospital has less staff in various services (such as radiology), it is difficult to do consultations, which may decrease the probability of admission.

3) Health status worsens on weekends. Finally, it may be possible that certain activities improve health on the weekends, leading to less emergencies. However, this seems unlikely because on weekends people over 65 tend to increase activity, which may lead to more acute changes in health, rather than less. In fact, mortality due to homicide, suicide, and motor vehicle accidents have been

shown to increase on the weekends [90, 91].

After controlling for the effect of admissions, the effect of Friday discharges *increased*. This suggests that Friday discharges do have a shorter time-to-readmission than Wednesday discharges, and that the effect is reduced if the effect of the probability of admission on weekends is not controlled. Epstein [44] found a correlation between readmission and admission, but concluded that this correlation can account for differences in readmission rates between hospitals. In our study, we found that the differences in admission patterns can attenuate estimates of differences in care unless they are properly controlled. Because Friday is the most common discharge day, hospital staff may rush the discharges, leading to less discharge instructions, and perhaps poorer outcomes after discharge [92]. Many patients over 65 years of age depend on long-term care and social support services, which typically do not accept patients on the weekend. Also, patients who have their care delayed till the following Monday may have worse outcomes [14].

The decreased availability of weekend services may be reducing admissions and discharges on the weekend, but it also may have a more direct effect on patient outcomes. Other work has found that increased weekend services may ameliorate patient flow, perhaps improving care and saving money in the long-term. Wong [68], using a dynamic simulation model of patient flow, found that increased weekend discharges would lead to decreased congestion at the emergency department. Bell [66] suggested that if hospitals operated at 50% of weekday capacity on the weekends (they currently operate at about 12%), the total volume of weekly procedures would increase by 14%. Varnava [93] found that discharge decisions for those with AMI were affected inappropriately by the day of the week, leading to clinically unnecessary increased length of stay. Another study of length of stay [94] found that 24% of medically unnecessary patient-days involved an inability to access medical services on the weekend. Our work provides indirect evidence that the lack of weekend services may be disrupting patient flow, negatively affecting readmissions.

Our Markov model showed that ignoring admission probabilities can lead to an underestimate of the effect of discharge day on readmission. But the probability of admission is just one aspect of the dynamics of patient flow; changes in access to primary care, long-term care, and social services can all change the time-to-readmission. In future work, we plan to incorporate access to and care at these services to understand the dynamics of patient flow.

The admission-discharge-readmission path of patient care is naturally cyclic, but regression techniques are not best suited for cyclic modeling. Regression techniques can crudely account for first- or second-order cyclic effects (has recently been readmitted, has recently been readmitted before that readmission), but there is no way to elegantly model a truly cyclic dynamic. In system dynamics modeling, the unit of analysis is a continuous quantity, and is modeled like fluids moving through a system of pipes and tanks (or like the probabilities through a Markov model). Some have argued that a system dynamics perspective in public health, which can incorporate the cyclic nature of patient flow, may help us to draw inference on patient flow dynamics, and to identify effective interventions [68, 95, 96]. In future work, we plan on using system dynamics modeling to further explore the dynamics of readmissions.

In our study of the day of week of discharge, we found that a seemingly simple effect of discharge day on readmissions hides the rather complex dynamics of patient flow. In future work, we plan to explore the nature of patient flow between hospital, primary care, and long-term care more closely, to develop effective interventions that can reduce the risk of readmission associated with Friday discharge.

Chapter 5

Hospital readmissions and targeted maximum likelihood estimation

5.1 Preamble

For the US hospital readmission laws to be effective, they must penalize only those hospitals where high readmission rates can be attributed to poor quality of care, and not to differences in patient health. Despite the importance of risk adjustment, and the availability of administrative databases, the risk models only adjust for a few well-known confounders like age and sex.

In the first manuscript, I developed predictive models of readmission risk to target transitional care interventions, and in the second manuscript, I focused on drawing inference about the day-of-week of discharge on readmission. In the following manuscript, I use TMLE to use predictive models to take advantage of the rich confounder data, and draw inference on the effect of hospital quality of care on readmission.

5.2 Abstract

Background: Hoping to improving quality of hospital care, the US and other jurisdictions financially penalize hospitals with poor (confounder-adjusted) 30-day readmission rates. Although hospital administrative data are information-rich, confounder adjustment tends to be crude. Non-parametric machine learning techniques can take advantage of these rich data to predict readmission, but cannot isolate the independent effect of hospitals on readmission risk.

Research Design: To estimate the effect of care at different hospitals on 30-day readmission risk, we used TMLE, which allowed us to use a non-parametric machine learning technique (random forest) to take advantage of the rich confounder data. We used an 11-year cohort of 65-year-old patients from 20 hospitals in Montreal, Canada, and developed three models to estimate the marginal readmission risk at each of the hospitals after hospitalization for heart failure, AMI, and pneumonia. We controlled for hundreds of confounders including outpatient drug prescriptions, medical procedures, and diagnoses. We compared the TMLE-estimated risk to a logistic regression model similar to one currently used to penalize hospitals.

Results: Within each hospital, crude readmission risk varied widely across the twenty hospitals for AMI 2 525 / 15 746 (16%), heart failure 5 520 / 24 847 (22%), and pneumonia 3 183 / 20 421 (16%). In the logistic regression model, the odds ratio ranged from 0.95-1.02 for AMI, 0.92-1.04 for heart failure, and 0.96-1.04 for pneumonia. When we applied TMLE, the odds ratio ranged from 0.57-2.30 for AMI, 0.50-1.85 for heart failure, and 0.47-1.55 for pneumonia.

Conclusion: Our results suggest that currently used techniques to financially penalize hospitals for readmission risk will underestimate the differences between hospitals. Using TMLE, we took advantage of rich confounder data, and revealed wide differences in quality of care between hospitals.

5.3 Introduction

In the US and other jurisdictions, administrators have sought to improve quality of care by financially penalizing hospitals with poor readmission rates [25]. To avoid penalizing high quality of care at hospitals that admit sicker (more likely to be readmitted) patients, readmission rates are adjusted for patient-level confounding [2]. Although hospital readmission rates are typically only adjusted for a few well-known confounders such as age, sex, previous readmissions, and summarized comorbidity scores [31], healthcare administrative data are often information-rich, including drug prescriptions, diagnoses, and medical procedures.

Some epidemiologists have argued for the application of machine learning techniques to handle newly available, information-rich data sources [97]. Non-parametric machine learning techniques can accurately discriminate patient readmission risk using hundreds of variables in a computationally efficient way [77]. Non-parametric models also allow us to avoid specifying a functional form, making it easier to detect complex relationships like multi-way interactions. However, most machine learning techniques were developed for prediction rather than inference; we cannot use them alone to isolate (target) effect measures of specific variables, such as care at a particular hospital, on readmission risk.

TMLE is a doubly-robust causal inference technique that allows the use of machine learning technique to estimate target parameters of interest [15]. In TMLE, two (possibly non-parametric) models are developed: one to estimate the probability of exposure, and another model to estimate the probability of the outcome. These two probabilities are combined in a parametric model with only the parameter of interest. In this way, the discriminative power of non-parametric models can be used to extract estimates of parameters of interest.

Although some studies have used the rich confounder data in combination with machine learning techniques to predict hospital readmissions [98, 99], no study

to our knowledge has used these data to draw causal inference on the effect of quality of care on readmissions. In this study, we sought to estimate the independent effect of hospital care on 30-day readmission for twenty Montreal hospitals, within three different admission diagnoses (pneumonia, heart failure, and acute myocardial infarction). We used a non-parametric machine learning technique, (random forest [100]), with TMLE to take advantage of the rich confounder data and minimize bias in our estimate of readmission risk.

5.4 Methods

5.4.1 Study design

We used a cohort extracted from a Canadian provincial (Quebec) administrative database of hospitalizations, obtained from the RAMQ. We enrolled patients into this cohort on the month that two conditions were satisfied: 1) they had at least one diagnosis of a respiratory illness (the exact list of respiratory ICD-9 codes is given in Table 6.1 in the Appendix) between January 1st, 1996 and March 31, 2006 (the study period), while living in the 2006 census metropolitan area of Montreal, and 2) were at least 65 years of age. We used this cohort because it contains the majority of 65-year-old patients who were hospitalized in the region during the study period.

From among this cohort, we selected hospital discharges for those who had accrued at least one continuous year in the cohort preceding the day of admission. We restricted our data to only the discharges from the twenty hospitals with the most discharges of patients 65 years of age or older within the study period; the twenty hospitals accounted for 75% of all such discharges. We only selected hospital discharges which resulted from hospital stays of at least one day. Therefore, the earliest possible hospital discharge was January 2, 1997.

From among the identified hospital discharges, we selected only those with one

of three high-volume admission diagnoses with high rates of hospital readmissions: pneumonia, AMI, and heart failure. We identified each of the admission diagnoses using ICD-9 codes; for pneumonia we used codes ranging from 480-487, for heart failure we used all 428 codes, and for AMI we used all 410 codes. The following methods were applied individually to all three disease subsets.

5.4.2 Hospital readmissions

The unit of analysis was the hospital discharge; a person could be discharged multiple times. A hospital readmission was defined as an emergency hospital admission to any Quebec hospital in the 30 days following a discharge. A person who died or had a non-emergency readmission in the 30 days following discharge was considered not readmitted.

5.4.3 Confounders and risk factors

For each hospital discharge, we collected plausible confounders that measured states at the time of, or prior to, admission. We used the demographic characteristics including age at time of admission (in years), sex, birth year-month. We also used the number of previous readmissions (within the preceding year) and the admission diagnosis (as measured by the specific ICD-9 code). We also included the day of week of discharge, which has been previously shown to have an association with readmissions [14], and the month of discharge, because we hypothesized that readmission risk would vary by season in Montreal.

Additionally, for each discharge, we collected the Quebec hospital diagnoses, Quebec hospital procedures, and drugs dispensed outside of the hospital but inside Quebec, in the year preceding the admission. The hospital procedures were recorded in the CCP system. Hospital diagnostic codes were coded using the ICD-9 system. Finally, drugs which were prescribed and dispensed outside the hospital, and were being taken on the day of admission were also recorded

for each patient in the *code commune* system, which categorizes drugs based on the chemical compound. To ease computation, before fitting any model, we removed any diagnosis, procedure or drug that occurred less than 30 times among all discharges. We chose 30 because it appeared to be a natural breakpoint; if the number of variables included is a function f of the threshold, then the first derivative of f dropped at 30 for all three disease categories.

We believed that residential location would strongly affect the probability of admission to the hospital nearest that census tract. We included it in our models because we also expected it to crudely approximate a (expected) confounder: socio-economic status. We used the residential postal code at the time of admission to assign each patient in the cohort to a census tract, as defined by the 2006 Canadian census. (Census tracts contain between 2,500 and 8,000 people, and, at the time of their creation, are demarcated so as to maximize homogeneity of socioeconomic characteristics.) [101]

5.4.4 Statistical analyses

For each discharge i, we sought to estimate the effect of each of the twenty hospitals $A \in \{a_1, \ldots, a_{20}\}$ on 30-day readmission (Y), accounting for the vector of confounders (W). To estimate this risk, we used targeted maximum likelihood estimation, which consisted of several steps. We fit a model of the exposure $g = Pr(A|W)$ (using random forest described below). Next, we estimated of a model of readmission risk based on the confounders W and the variables for each of the hospitals $Q = Pr(Y = 1|A, W)$. We then calculated $h_a(A, W)$ (sometimes referred to as the clever covariate) described in equation 5.1.

$$h_a(A, W) = \frac{I(A = a)}{g(a|W)} \tag{5.1}$$

(where I is the indicator function which evaluates to 1 when its argument is true, and 0 otherwise), and estimated all ϵ_a in the fluctuation function de-

scribed in equation 5.2.

$$Y_i = expit(logit(Q(Y_i|A_i, W_i)) + \sum_{j=1}^{20} \epsilon_{a_j} \times h_{a_j}(A_i, W_i)) \tag{5.2}$$

We estimated all twenty ϵ_a by regressing the 30-day readmission outcome Y (with a logit link function) onto $h_a(A, W_i)$ (with no intercept) offset by the inverse logit of the initial estimate of readmission risk $Q = (Y|A, W)$. Finally, for each discharge, we computed the estimated risk of 30-day readmission for all twenty counterfactual conditions (the risk of readmission for every discharge as if they had attended different hospital) using Equation 5.3.

$$Q^*_{ai} = expit(logit(Q(Y|a, W_i)) + \frac{\epsilon}{g(a|W)}) \tag{5.3}$$

For each hospital, we then calculated the mean readmission risk (Q^*_a) and associated odds ratio.

To estimate both models $g(A_i|W_i)$ and $Q(Y_i|A, W_i)$, we used a random forest, a non-parametric model based on decision trees [100]. Decision trees use the independent variables (W_i) to repeatedly split data into partitions that are as homogeneous as possible with respect to the outcome of interest (specifically measured with the Gini impurity index [102]). Random forest improves decision trees by using bootstrap aggregation (bagging); multiple decision trees are grown on bootstrap replicates (sampled with replacement) to avoid overfitting. Additionally, within each tree, only a sample of the covariates is used (in our case we used a square root of the number of variables included in the mode, rounded down).

For both models $g(A_i|W_i)$ and $Q(Y_i|A, W_i)$, we arbitrarily chose to grow 1200 trees, and then measured the accuracy as a function of the number of trees to ensure that growing further trees would be unlikely to improve accuracy. Because the model was used solely to estimate the ***probability*** of admission to spe-

cific hospitals (and not to predict exactly which hospital was attended), when calculating the Gini impurity index to build the trees we configured the model to favor calibration over discrimination: we weighted each of the twenty predicted hospitals by the inverse of the proportion of discharges at that hospital. When measuring the accuracy for each discharge, we only used trees for which the discharge was "out-of-bag", that is, we only used trees for which the bootstrap sample did not include the discharge.

To describe importance of the covariates in both models $g(A_i|W_i)$ and $Q(Y_i|A, W_i)$, for each variable, we measured the decrease in the Gini impurity index for each partition in which the variable was used, in every tree. A low Gini (i.e. higher decrease in Gini) means that a particular predictor variable plays a greater role in partitioning the data into the defined classes. We plotted the densities of variables with four different classes (census tract, procedure, diagnosis and drug) at different levels of Gini decreases.

Random forest classifies each item by majority vote: each tree in the forest assigns each discharge to a specific class. Although the vote proportion is between zero and one, it is not calibrated well as a probability. To calibrate the vote proportion, we used Platt scaling [103] (logistic regression of the outcome (Y_i) on to the vote proportion).

When the probability of exposure $g(A|W)$ is very low, that discharge would receive a large weight in estimating Q^*. For any $g(A|W)$ below some fixed value δ, we set $g(A|W)$ to δ, a common technique in TMLE [15]. We recomputed our analyses at 31 different values of δ, ranging from 10^{-2} to 10^{-5}, decreasing the exponent at intervals of 0.1.

Finally, we compared our results of our analysis with a logistic regression for 30-day readmission. In this model, we included only the age, sex, number of previous admissions, and the Charlson comorbidity score (Elixhauser version) [104], along with indicator variables representing the hospitals themselves.

5.4.5 Software

The data were cleaned and prepared for statistical analysis using the Postgres relational database (version 9.3.6). We implemented our models using the R statistical package (version 3.1.1), [85] using the "bigrf" package (version 0.1.11) to grow the random forests [105]. We plotted our figures using the "ggplot2" package (version 1.0.1) [88]. All the code to develop used to process our data, fit our models, and typeset this article is available for download at Github.

5.5 Results

Over the course of January 2, 1996 to March 31, 2006, 482 064 people were entered into our cohort. Among these, 16 521 were ever admitted for pneumonia, 13 884 were ever admitted for AMI, and 15 822 were ever admitted for heart failure. People ever admitted for pneumonia had a mean (median) 1.2 (1) pneumonia admissions, heart failure patients had a mean (median) 1.6 (1) heart failure admissions, and AMI patients had a mean (median) 1.1 (1) AMI admissions. In total, we analyzed 20 421 pneumonia discharges, 15 746 AMI discharges, and 24 847 heart failure discharges.

The accuracy of the random forest model (for both models g and Q) did not appear to improve significantly beyond 125 trees (see Figure 5.1). In Figure 5.3 we plot the importance of variables (as measured by the Gini impurity index) in the random forest models for four variable classes, for all disease subsets for both the g and Q model. Although census tracts were found to be important in prediction of hospital choice, the other three variable classes were had a high density of important variables as well. The prescription drugs in particular had a high proportion of important variables, and generally the lowest proportion of unimportant variables. For the Q model, the variable density appeared bimodal within variable importance for all four variable classes. Additionally, the pre-admission drug prescriptions appeared to be strongly important in predict-

ing readmission for all pneumonia, heart failure, and AMI admissions.

The predicted probability of admission to any particular hospital ($g = Pr(A = a|W)$) was less than 5% in 88% of cases (across all disease subsets and hospitals). We set δ (the lower bound of g when used to fit the ϵ values for Q^*), to two different values, 10^{-2} and $10^{-2.5}$. Across all disease subsets and hospitals, 39% of discharge/hospital combinations and a g less than 10^{-2}, and 4% had a g less than $10^{-2.5}$. Figure 5.2 describes the histogram of g when it is below 0.05 for each disease/hospital combination separately.

The unadjusted proportion of patients readmitted in 30 days varied across hospitals for each disease subset (Tables 1-3). The linear correlation between the proportion of deaths during hospital stay and the proportion readmitted was (0.19, -0.55, -0.28) among AMI, heart failure, and pneumonia admissions respectively. Using a model that adjusts for a few well-known confounders, for AMI, heart failure, and pneumonia respectively, one, three, and five hospitals had significantly different odds than the reference hospital. Notably, the significant odds ratios are all relatively small, with point estimates ranging from 0.92 - 1.04. In contrast, in the TMLE models, at both values of δ, for all admission diagnoses, nearly all of the hospitals had significantly different odds than the reference hospital.

In some hospitals and disease subsets, the parameter δ, (the lower bound on the probability of exposure $g(A|W)$) had a considerable effect on the marginal risk and the associated odds ratios. For example, for AMI (shown in Table 5.1), the marginal risk for hospital 17 increases by six percent when δ decreases from 10^{-2} to $10^{-2.5}$. In Figure 5.4, we display the marginal risk for each of the twenty hospitals and disease subsets as a function of the parameter δ. For many hospitals, the effect was quite strong; for pneumonia admissions, hospital 16 went from having the second-lowest marginal risk when $\delta = 0.1$ to having the highest marginal risk when $\delta = 0.025$.

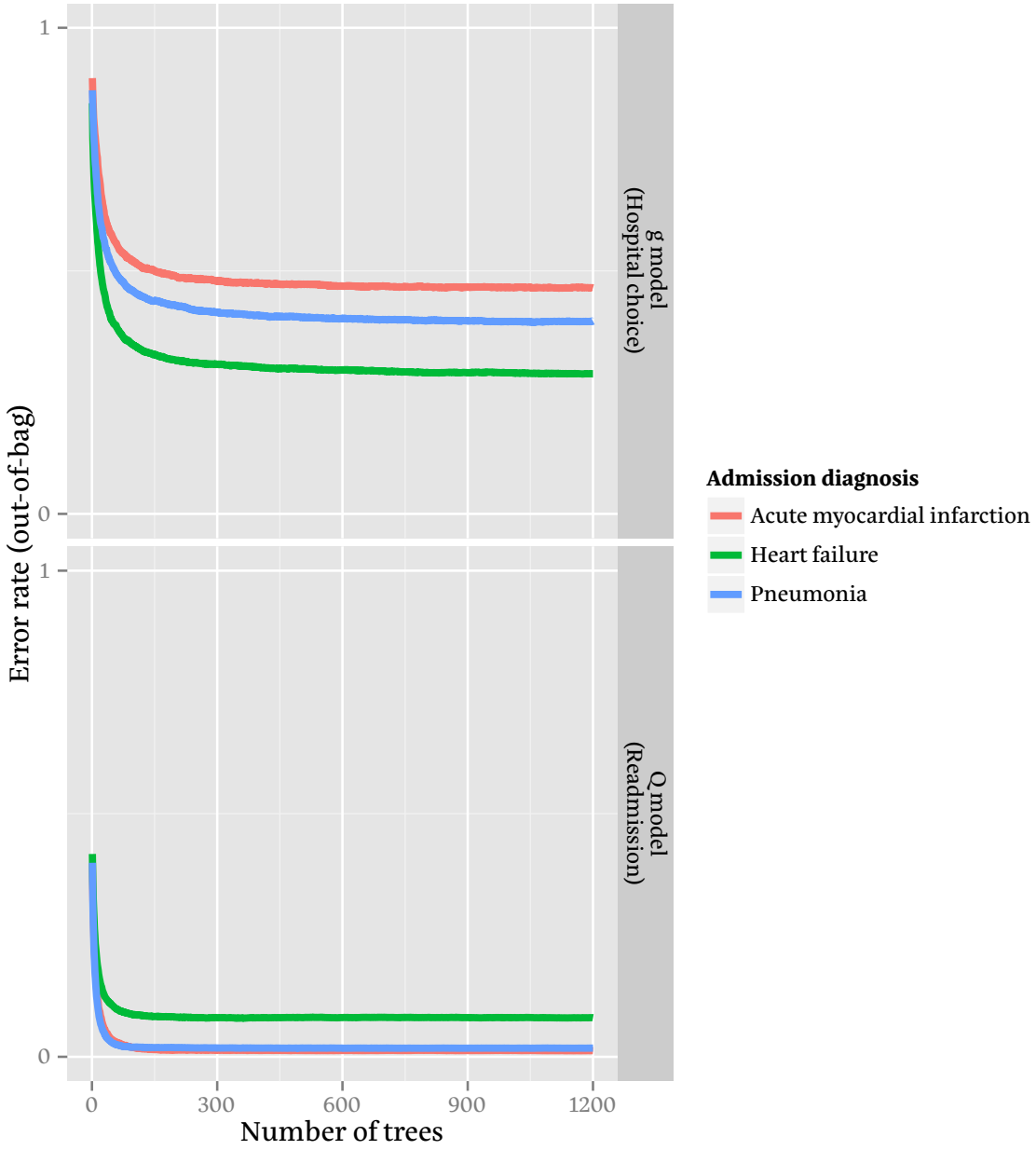

Figure 5.1: Error rate for both random forest models of hospital choice (g) and readmission (Q) as a function of the number of trees grown. For each admission, only out-of-bag trees were used to predict the given outcome.

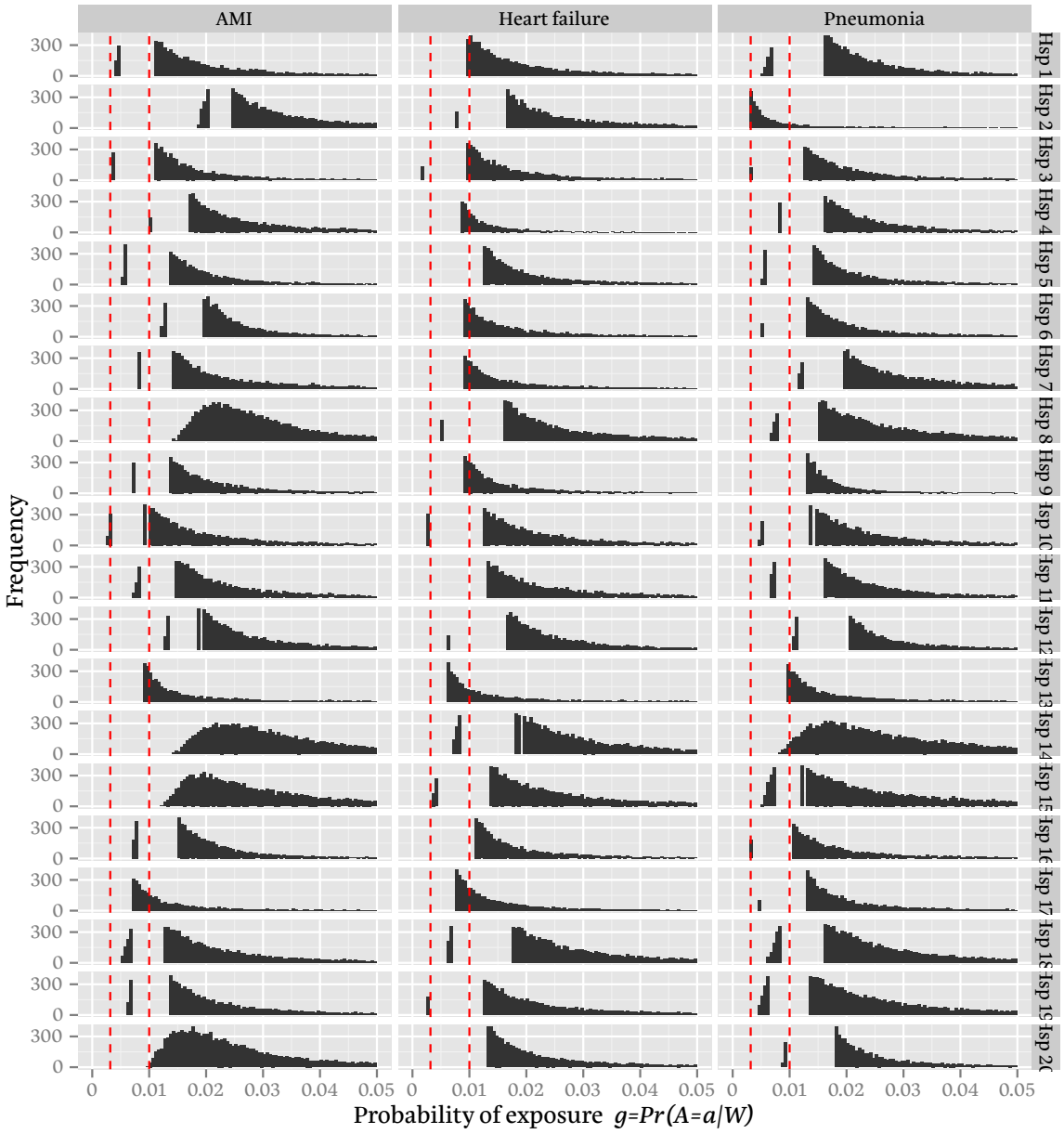

Figure 5.2: Histogram of the probability of exposure (g). The range of probability on the horizontal axis is restricted to (0,0.05). The bin width is 0.005. The dotted red lines indicate the two values of δ used in Tables 1, 2, and 3.

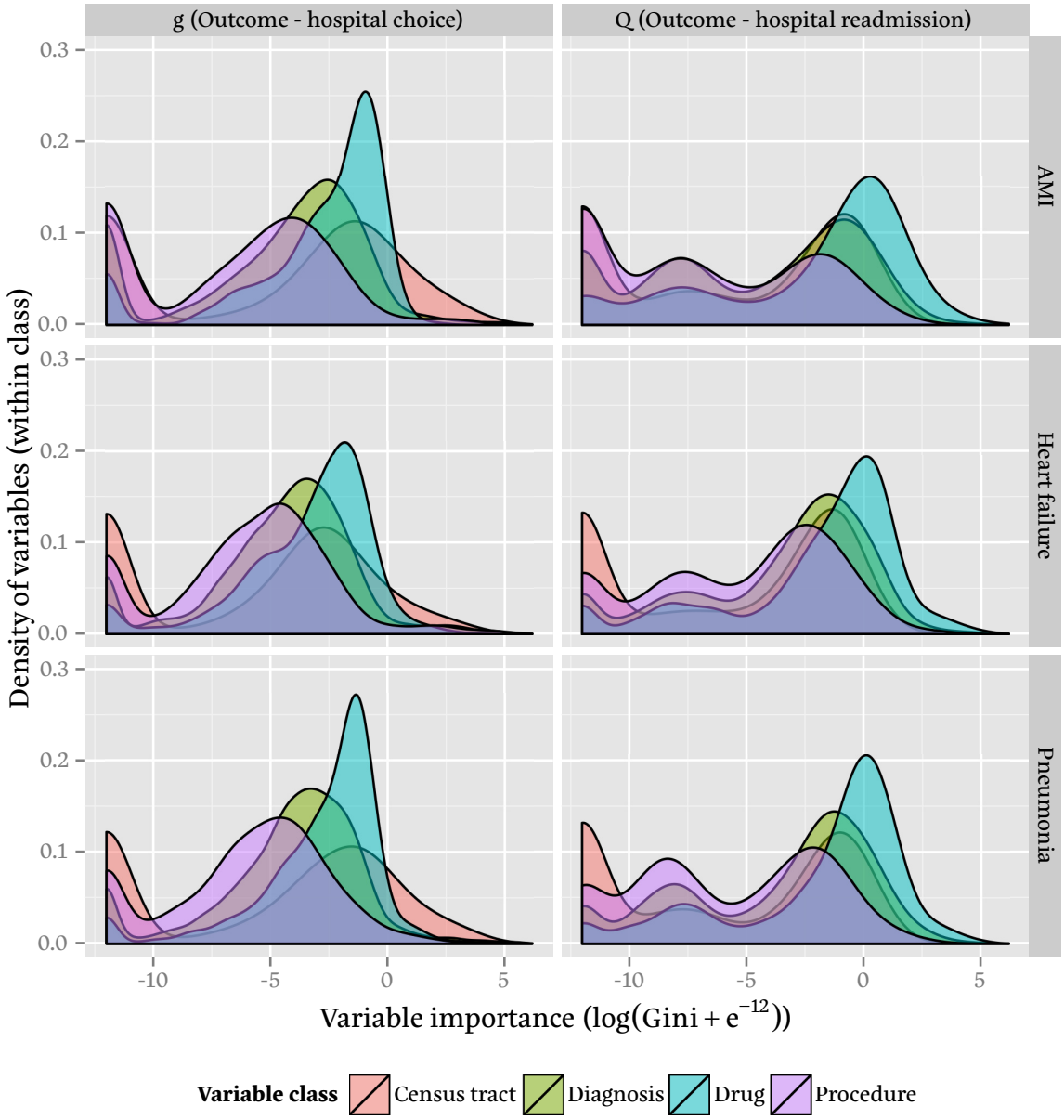

Figure 5.3: Variable importance by model and variable class. For each random forest classifier, the variable importance was measured by the decrease in the Gini impurity index when that variable splits a node. The horizontal axis within each panel is displayed on a log_e scale. Some variables had exactly zero importance; to avoid evaluating the logarithm of zero, we added a small constant (e^{-12}) to the measure of variable importance. The vertical axis in each panel represents the variable density at the corresponding level of variable importance. To transform the individual variable importances into a continuous density, we smoothed using a Gaussian kernel density estimator, using Silverman's 'rule-of-thumb' [106] to select the bandwidth. The density is measured separately for each class; the area under each variable class curve is exactly one.

Hsp.					Logistic regression		TMLE ($\delta = 10^{-2}$)		TMLE ($\delta = 10^{-2.5}$)	
	Admitted	Died (%)	Discharged	Readmitted (%)	Odds ratio (95% CI)	Marginal risk	Odds ratio (95% CI)	Marginal risk	Odds ratio (95% CI)	Marginal risk
1	763	112 (15)	651	105 (16)	0.98 (0.95-1.01)	0.16	0.86 (0.83-0.89)	0.16	0.77 (0.74-0.81)	0.14
2	1557	148 (10)	1409	191 (14)	0.97 (0.95-1.00)	0.15	0.85 (0.82-0.87)	0.15	0.85 (0.82-0.87)	0.15
3	606	83 (14)	523	84 (16)	0.98 (0.95-1.02)	0.16	1.01 (0.97-1.05)	0.18	1.09 (1.04-1.14)	0.19
4	1022	125 (12)	897	136 (15)	0.97 (0.94-1.00)	0.15	0.72 (0.69-0.74)	0.13	0.72 (0.69-0.74)	0.13
5	729	150 (21)	579	98 (17)	0.98 (0.95-1.02)	0.16	0.75 (0.72-0.77)	0.14	0.73 (0.71-0.76)	0.14
6	826	119 (14)	707	106 (15)	0.98 (0.94-1.01)	0.15	0.57 (0.54-0.60)	0.11	0.57 (0.54-0.60)	0.11
7	1491	241 (16)	1250	216 (17)	0.99 (0.96-1.01)	0.16	1.04 (1.01-1.06)	0.18	1.03 (1.01-1.06)	0.18
8	1270	198 (16)	1072	138 (13)	0.95 (0.92-0.98)	0.13	0.69 (0.67-0.71)	0.13	0.69 (0.67-0.71)	0.13
9	780	152 (19)	628	130 (21)	1.01 (0.97-1.05)	0.19	0.54 (0.51-0.56)	0.10	0.52 (0.50-0.54)	0.10
10	778	124 (16)	654	123 (19)	1.01 (0.97-1.05)	0.19	1.19 (1.15-1.23)	0.20	1.27 (1.22-1.31)	0.21
11	705	125 (18)	580	97 (17)	0.99 (0.96-1.03)	0.17	0.89 (0.85-0.92)	0.16	0.90 (0.86-0.94)	0.16
12	1284	266 (21)	1018	166 (16)	0.99 (0.96-1.02)	0.16	0.90 (0.88-0.93)	0.16	0.90 (0.88-0.93)	0.16
13	739	86 (12)	653	110 (17)	0.99 (0.95-1.02)	0.16	1.19 (1.16-1.23)	0.20	1.22 (1.18-1.27)	0.21
14	1307	184 (14)	1123	210 (19)	(Reference)	0.18	(Reference)	0.18	(Reference)	0.18
15	1152	168 (15)	984	129 (13)	0.97 (0.95-1.01)	0.15	0.70 (0.68-0.73)	0.13	0.70 (0.68-0.73)	0.13
16	408	70 (17)	338	43 (13)	0.97 (0.93-1.01)	0.15	0.84 (0.80-0.88)	0.15	0.84 (0.80-0.89)	0.15
17	807	123 (15)	684	134 (20)	1.02 (0.99-1.06)	0.20	1.76 (1.72-1.81)	0.27	2.30 (2.23-2.37)	0.33
18	894	144 (16)	750	116 (15)	0.98 (0.95-1.01)	0.16	0.91 (0.87-0.94)	0.16	0.91 (0.87-0.95)	0.16
19	499	94 (19)	405	50 (12)	0.95 (0.91-0.99)	0.13	0.57 (0.53-0.61)	0.11	0.57 (0.53-0.61)	0.11
20	1025	184 (18)	841	143 (17)	0.99 (0.96-1.02)	0.17	1.05 (1.02-1.09)	0.18	1.05 (1.02-1.09)	0.18

Table 5.1: Risk of 30-day readmission after admission for acute myocardial infarction (AMI) in twenty Montreal hospitals. The proportion of those who were readmitted within 30 days is caluculated using the number discharged alive as the denominator. The confidence intervals for the odds ratios for the parameters in the logistic regression model were calculated using the profile likelihood method.[107] The marginal risk for the odds ratios was calculated by using the regression model to calculate the mean predicted probability of readmission for every admission, except individually fixing the hospital attended to one hospital. The parameter δ represents the lower bound on the probability of exposure to that hospital (g); we display odds ratios and marginal risks for two versions of the TMLE model with varying levels of δ .

Hsp.					Logistic regression		TMLE ($\delta = 10^{-2}$)		TMLE ($\delta = 10^{-2.5}$)	
	Admitted	Died (%)	Discharged	Readmitted (%)	Odds ratio (95% CI)	Marginal risk	Odds ratio (95% CI)	Marginal risk	Odds ratio (95% CI)	Marginal risk
1	1229	141 (11)	1088	248 (23)	1.00 (0.97-1.03)	0.22	0.61 (0.59-0.63)	0.11	0.50 (0.48-0.53)	0.09
2	2071	166 (8)	1905	441 (23)	1.02 (0.99-1.05)	0.24	1.13 (1.11-1.16)	0.19	1.13 (1.11-1.16)	0.19
3	1243	134 (11)	1109	285 (26)	1.03 (1.00-1.07)	0.25	0.71 (0.69-0.72)	0.13	0.52 (0.50-0.54)	0.10
4	1076	122 (11)	954	214 (22)	1.01 (0.97-1.04)	0.23	1.06 (1.04-1.09)	0.18	0.92 (0.89-0.96)	0.16
5	1550	181 (12)	1369	288 (21)	0.99 (0.96-1.02)	0.21	0.71 (0.69-0.72)	0.13	0.58 (0.56-0.60)	0.11
6	827	107 (13)	720	128 (18)	0.97 (0.94-1.00)	0.19	0.73 (0.70-0.75)	0.13	1.08 (1.03-1.14)	0.18
7	2917	386 (13)	2531	666 (26)	1.04 (1.02-1.07)	0.26	1.63 (1.61-1.66)	0.25	1.67 (1.64-1.71)	0.26
8	1456	197 (14)	1259	232 (18)	0.97 (0.94-1.00)	0.19	0.72 (0.70-0.74)	0.13	0.68 (0.66-0.70)	0.12
9	881	111 (13)	770	157 (20)	0.98 (0.95-1.02)	0.20	1.27 (1.25-1.29)	0.21	1.18 (1.16-1.20)	0.20
10	1410	149 (11)	1261	311 (25)	1.01 (0.99-1.05)	0.23	0.66 (0.65-0.68)	0.12	0.57 (0.55-0.60)	0.11
11	1297	153 (12)	1144	258 (23)	1.01 (0.98-1.04)	0.23	0.90 (0.88-0.92)	0.16	0.86 (0.83-0.88)	0.15
12	1323	162 (12)	1161	192 (17)	0.92 (0.89-0.95)	0.13	0.79 (0.76-0.81)	0.14	0.76 (0.74-0.78)	0.14
13	1231	102 (8)	1129	262 (23)	1.00 (0.97-1.03)	0.22	0.94 (0.93-0.96)	0.16	0.91 (0.87-0.95)	0.16
14	2110	234 (11)	1876	424 (23)	(Reference)	0.22	(Reference)	0.17	(Reference)	0.17
15	1389	190 (14)	1199	203 (17)	0.97 (0.94-1.00)	0.19	0.74 (0.72-0.77)	0.13	0.81 (0.79-0.84)	0.14
16	681	94 (14)	587	111 (19)	0.98 (0.94-1.01)	0.20	0.75 (0.73-0.78)	0.14	0.84 (0.80-0.87)	0.15
17	1438	139 (10)	1299	328 (25)	1.04 (1.01-1.07)	0.26	1.50 (1.48-1.53)	0.24	1.85 (1.80-1.90)	0.28
18	1984	212 (11)	1772	438 (25)	1.03 (1.00-1.06)	0.25	0.76 (0.74-0.77)	0.14	0.74 (0.72-0.76)	0.13
19	932	99 (11)	833	163 (20)	0.98 (0.95-1.01)	0.20	0.88 (0.86-0.90)	0.16	0.81 (0.79-0.84)	0.14
20	1048	167 (16)	881	171 (19)	0.99 (0.96-1.02)	0.21	1.25 (1.22-1.27)	0.21	1.20 (1.17-1.23)	0.20

Table 5.2: Risk of 30-day readmission after admission for heart failure in twenty Montreal hospitals. The columns in this table are described in Table 5.1.

Hsp.					Logistic regression		TMLE ($\delta = 10^{-2}$)		TMLE ($\delta = 10^{-2.5}$)	
	Admitted	Died (%)	Discharged	Readmitted (%)	Odds ratio (95% CI)	Marginal risk	Odds ratio (95% CI)	Marginal risk	Odds ratio (95% CI)	Marginal risk
1	1184	176 (15)	1008	159 (16)	1.00 (0.98-1.03)	0.15	1.23 (1.18-1.27)	0.15	1.21 (1.17-1.26)	0.15
2	199	11 (6)	188	31 (16)	1.02 (0.97-1.08)	0.17	1.09 (1.07-1.12)	0.14	1.25 (1.17-1.34)	0.16
3	1085	132 (12)	953	160 (17)	1.01 (0.98-1.04)	0.16	0.83 (0.80-0.87)	0.11	0.82 (0.78-0.87)	0.11
4	863	91 (11)	772	113 (15)	1.00 (0.97-1.03)	0.15	0.85 (0.81-0.88)	0.11	0.84 (0.81-0.88)	0.11
5	923	147 (16)	776	143 (18)	1.04 (1.01-1.07)	0.19	0.96 (0.93-1.00)	0.12	0.95 (0.91-0.99)	0.12
6	788	136 (17)	652	89 (14)	1.00 (0.96-1.03)	0.14	0.89 (0.85-0.94)	0.12	0.91 (0.86-0.96)	0.12
7	2194	228 (10)	1966	328 (17)	1.03 (1.00-1.05)	0.17	1.33 (1.29-1.37)	0.16	1.33 (1.29-1.37)	0.16
8	1485	243 (16)	1242	173 (14)	0.99 (0.97-1.02)	0.14	0.97 (0.93-1.01)	0.12	0.97 (0.94-1.01)	0.13
9	990	166 (17)	824	158 (19)	1.04 (1.01-1.08)	0.19	1.30 (1.25-1.35)	0.16	1.28 (1.23-1.33)	0.16
10	1214	139 (11)	1075	181 (17)	1.01 (0.99-1.04)	0.16	1.45 (1.40-1.51)	0.18	1.46 (1.40-1.51)	0.18
11	892	147 (16)	745	119 (16)	1.02 (0.98-1.05)	0.16	1.39 (1.34-1.44)	0.17	1.40 (1.35-1.46)	0.17
12	1102	185 (17)	917	91 (10)	0.96 (0.93-0.98)	0.10	0.47 (0.44-0.50)	0.06	0.47 (0.44-0.50)	0.06
13	1914	204 (11)	1710	325 (19)	1.03 (1.00-1.05)	0.18	0.80 (0.77-0.83)	0.10	0.84 (0.79-0.89)	0.11
14	1980	278 (14)	1702	263 (15)	(Reference)	0.15	(Reference)	0.13	(Reference)	0.13
15	1365	179 (13)	1186	163 (14)	1.00 (0.97-1.03)	0.15	0.86 (0.83-0.90)	0.11	0.85 (0.81-0.89)	0.11
16	541	77 (14)	464	46 (10)	0.96 (0.93-1.00)	0.11	1.45 (1.39-1.52)	0.18	1.55 (1.46-1.65)	0.19
17	1338	163 (12)	1175	193 (16)	1.02 (0.99-1.05)	0.17	0.86 (0.83-0.89)	0.11	0.79 (0.76-0.82)	0.10
18	1356	168 (12)	1188	200 (17)	1.02 (0.99-1.04)	0.17	1.40 (1.36-1.44)	0.17	1.40 (1.35-1.44)	0.17
19	1020	123 (12)	897	122 (14)	0.99 (0.96-1.02)	0.14	0.94 (0.90-0.98)	0.12	0.98 (0.93-1.03)	0.13
20	1152	171 (15)	981	126 (13)	0.98 (0.95-1.01)	0.13	1.11 (1.07-1.16)	0.14	1.11 (1.07-1.16)	0.14

Table 5.3: Risk of 30-day readmission after admission for pneumonia in twenty Montreal hospitals. The columns in this table are described in Table 5.1.

5.6 Discussion

Using targeted maximum likelihood estimation (TMLE) to adjust precisely for measured confounders, we found that the differences in marginal risk of 30-day hospital readmission in twenty Montreal hospitals were larger than indicated by a logistic regression model that only adjusted for a few confounders. Additionally, our study revealed some practical positivity violations for some hospitals, suggesting that the relative readmission risk may not always be estimable from observed data.

Our study has several strengths. By using a doubly-robust estimation technique, and by accurately adjusting for thousands of plausible confounders, we minimized the bias in our estimates of the effect of hospital care on readmissions. Our work suggests that the difference in bias reduction was not trivial; in assessing the effect of hospitals on readmission, the two models lead us to different conclusions on the differences in quality of care at hospitals. Also, since we did not have to restrict our cohort to a single healthcare insurance network, we had a large cohort of patients from all socioeconomic classes. Because we had complete access to all hospital visits in the province, we could accurately measure which patients were readmitted.

Other hospital readmission studies have applied statistical and machine learning algorithms to readmission data to develop predictive models [73, 98], including one using the data used in this study [99]. Most studies, including our own, found relatively poor accuracy. No study to our knowledge has used machine learning algorithms to draw causal inference on target parameters. Predictive models of hospital readmissions may not be very accurate, but they can can improve our ability to draw inference on target parameters.

Some authors [31, 108] believe that by using readmission rates as a quality metric, we assume that readmissions are preventable. Hoping to develop a quality metric that compares preventable readmissions, some researchers have attempted to identify which individual readmissions were preventable. Some

studies have clinicians classify individual readmissions as preventable [109–111], despite evidence that clinicians cannot reliably measure preventability [46]. Other studies use pairs of admission/readmission diagnosis codes that identify "potentially preventable" readmissions [48]. However, the proportion of those actually preventable among the "potentially" preventable differs among hospitals [112], meaning that potentially preventable readmissions are not an adequate proxy for preventable readmissions [108].

But to estimate the effect of an exposure (like hospital care) on an outcome (like readmission), we do not need to identify exactly which individuals would not have had the outcome if they were not exposed [113]. Some readmissions are unpreventable: no matter where they were treated, they would be readmitted. If patients were randomized among different hospitals, the number of unpreventable readmissions would be (asymptotically) the same among all hospitals, and any difference in readmission rates would be the "preventable proportion". Since the patients were not randomized to each hospital, we attempted to recreate that situation by controlling for confounding. Assuming that we have adequately controlled for confounding, we have estimated the independent effect of each hospital on readmission risk, without identifying whether *individual* readmissions were preventable.

In this study, practical positivity violations occur when large subgroups of the hospitalized patients are rarely admitted to specific hospitals. Practical positivity violations can bias our estimates of the parameter of interest, because our risk estimates are heavily dependent on the few admitted patients from certain subgroups, and on the precision of our estimate of the probability of their attendance. For some hospitals, our estimates for marginal risk were sensitive to the parameter δ, which set a lower bound on the probability of exposure g, suggesting practical positivity violations. We believe that the discovery of practical positivity violations is an important finding: observational data may not provide us with enough information to meaningfully compare certain hospitals.

To avoid adjusting for a variable that is a component of our exposure (hospital

care), we did not adjust for hospital length of stay [50]. Similarly we excluded all diagnoses and procedures that occurred during the hospital admission, because these covariates were components of our exposure.

The major competing risk for 30-day hospital readmission is death, but others include moving outside the study area, or admission to a hospital for a non-emergency reason. In our analysis, we did not account for these competing risks. If patients died within 30 days of discharge more often at one hospital than another, we could have biased our estimate of readmission risk. Similarly, if patients died during the hospitalization more often at some hospitals than others, it could have created a selection bias (left censorship) in which hospitals with better care were discharging sicker (but still living) patients, who would be more likely to be readmitted. Also, there is no special significance of 30 days in readmission, except for the fact that it is (recently) widely used as a cutoff. In future work, we plan to account for both left censorship and competing risks in a model that estimates the effect of hospital care on time-to-readmission.

Entry to our cohort was dependent on having one diagnosis of a respiratory illness in an inpatient or outpatient setting. Respiratory illness was defined rather broadly, including extremely common diagnoses such as “cough”. We expect that the majority of 65-year-old patients who would be hospitalized would have at least one respiratory illness diagnosis in an outpatient setting. We cannot, however, exclude the possibility that parameter estimates were affected by selection bias with respect to the full population of 65-year-old patients.

The effect of hospital care on readmissions is confounded by a vast spectrum of health-related states of the admitted patients. In the absence of a clear theoretical basis of the structure of that confounding, we can 1) identify relatively few, well-understood and measurable confounders to include in our model, or 2) forgo any theoretical understanding of the structure of confounding, and attempt to identify the broadest measurable set of even faintly plausible confounders. The first option has some advantages: in a situation where data collection is expensive, it may not be plausible to measure thousands of variables.

Additionally, by reducing the confounders to a well-understood few, the model gains credibility because it can be shown that the confounders are having the expected effect. Non-parametric techniques such as random forest don't allow us to look (easily) at the individual effects of the confounders, and even in a parametric model it would be difficult to analyze thousands of variables. We summarized the densities of the effects a few classes of variables in Figure 5.3, but this still does not allow the variable-by-variable analysis typical in an epidemiologic study. Also, by including many confounders we also risk *inducing* bias, such as the M-bias [114, 115]. However, the recent availability of large scale healthcare administrative data has put us into the situation where the cost of data collection is relatively low. By using machine learning techniques like random forest, we also automatically fit multi-way interactions that we would be unlikely to explore in a model fit "by hand". Finally, because the structure of the confounding is unclear, we cannot assess if M-bias is present, and some research suggests that the scale of M-bias may be small when compared to traditional confounding [116]. We argue that in this situation, where we have a large data set, thousands of measurable confounders, and little understanding of the structure of confounding, the second option is more appropriate.

The unit of analysis in this study was the discharge, but each discharge was "clustered" within a patient. The expected within-cluster homogeneity could have biased our estimates of variance, and our parameter estimates. However, because the number of clusters (unique patients) was relatively high when compared to the sample size (the number of discharges), we do not expect that our parameter or variance estimates to be biased very strongly.

Beside random forest, we could have used many other machine learning techniques on these data, many of which we explored in other work [99]. Also, some ensemble machine learning techniques, (in particular SuperLearner [117] which is commonly used with TMLE), are available, that combine any number of other machine learning techniques. We found that in these data, ensemble learning techniques were too computationally expensive. We selected random

forest because of its relative simplicity, and because our variables were nearly all binary, for which decision trees are particularly suitable.

Calibration of the random forest vote proportions in the Q model strongly affected estimates of our parameter of interest in the Q^* update step. Other articles using non-parametric techniques typically combined them with other models in an ensemble learner (like SuperLearner). The final step in (many) ensemble learners is to combine all the probability estimates in a parametric model, which would effectively calibrate the probabilities. In our study, a single, non-parametric technique was used, so an additional, separate calibration step was necessary to convert the ranking scores into a probability estimate.

Hospital readmissions can be a relatively crude proxy for quality of care, but they can still provide valuable insight. In a seminal research article on quality of care measures, Donabedian writes: "But how precise do estimates of quality have to be? At least the better methods have been adequate for the administrative and social policy purposes that have brought them into being. The search for perfection should not blind one to the fact that present techniques of evaluating quality, crude as they are, have revealed a range of quality from outstanding to deplorable." [118] Our work suggests that, when finely adjusted for confounding, hospital readmissions reveal wide differences in hospital quality of care.

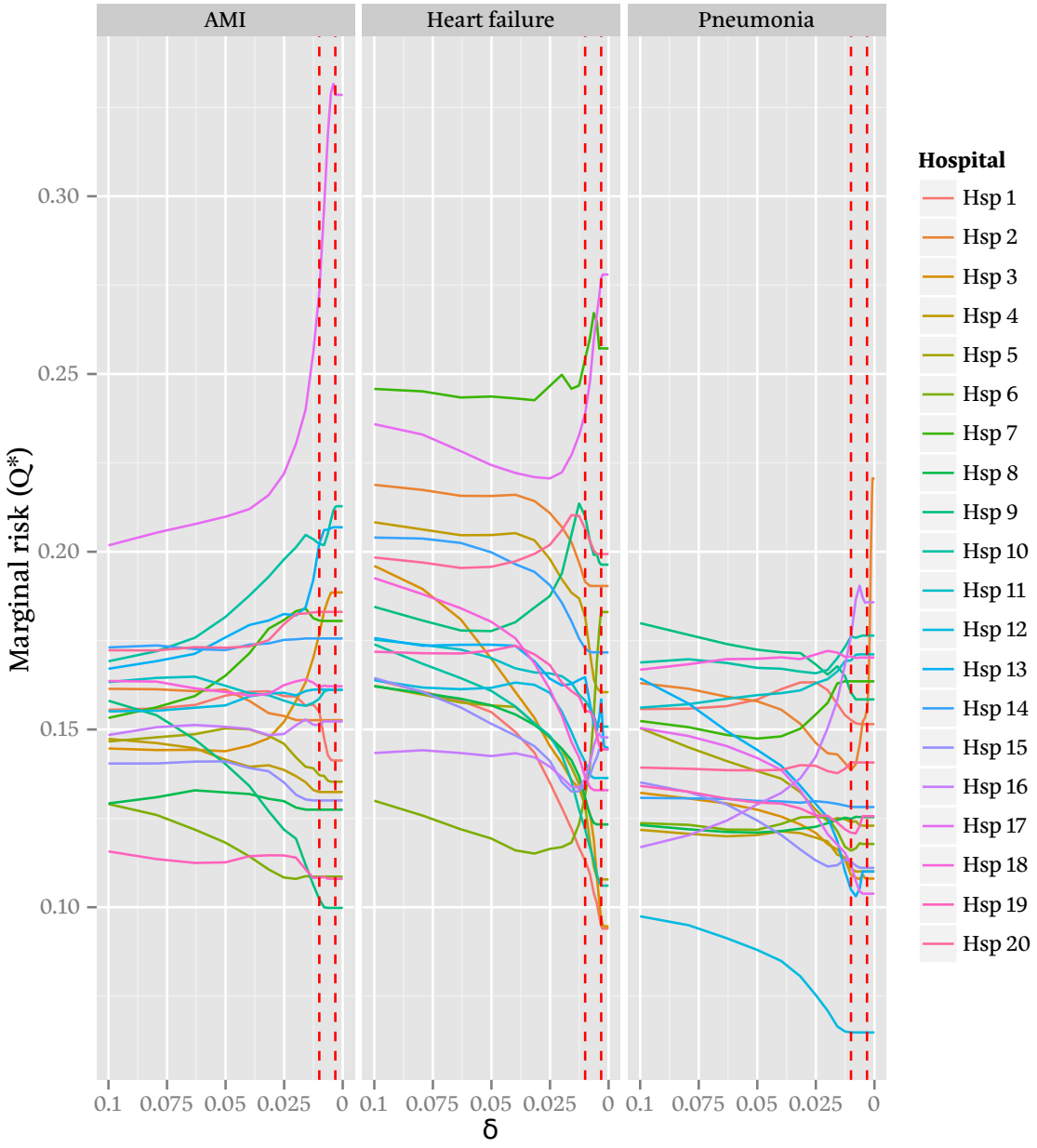

Figure 5.4: Effect of δ (the bound on $g(A|W)$) on the marginal risk (Q^*). The vertical axis represents the marginal risk as calculated by the TMLE model. The marginal risk (Q^*) was evaluated at 31 levels of δ, from 10^{-1} to 10^{-5}, (the exponent decreasing by 0.1). Note that the scale of the horizontal axis decreases from left-to-right. The hatched vertical lines mark the two levels of delta displayed in Tables 5.1,5.2 and 5.3.

Chapter 6

Discussion

In this work, I presented three studies that explored issues in prediction and inference with respect to hospital readmissions. The first manuscript measured the accuracy of model combination techniques to improve the predictive accuracy of hospital readmission risk models. In the second manuscript, I explored how the effect of the day-of-the-week of discharge on readmissions was modified by different probabilities of admission on the weekend. Finally, in the third manuscript, I presented an application of TMLE that allowed me to use machine learning techniques, typically used for prediction, to draw inference on how the quality of care at different hospitals affects readmissions. Here I summarize my most important findings.

Defining readmissions in counterfactual terms was critical for all three manuscripts. I did not need to measure whether individual readmissions were preventable to estimate how the quality of care at different hospitals affected preventable admissions. Because I did not need to measure the individual preventability of readmissions, it made it feasible to use large administrative databases in all three studies.

In the first manuscript (Chapter 3), I compared three different methods of pooling hospital data (local, global, and combined) to develop 30-day readmission

risk models. The “global” model represented the ideal case, where all data from all hospitals were pooled, and a single model was fit to the entire pool. The “local” model represented the situation where no data pooling took place: hospitals would build models only using their own data. The “combined” model pools the models instead of the data; the coefficients from each “local” model are collected and combined into a single model. The “combined” models are advantageous when compared to the “global” models because they don’t require hospitals to directly share patient data, which poses a risk to patient privacy, and they also may be more accurate than the “local” models. I found, as did others[61–65], that the “combined” model (AUC of 0.63 – 0.64) was nearly as accurate as the ideal “global” model (AUC of 0.65). However, I found that the “local” models, which did not pool data at all (AUC of 0.63), were also nearly as accurate as the global model. Over the 11 years of the study period, 7 355 heart failure discharges resulted in a 30-day readmission; if specificity was fixed to 75%, the global model would have detected only 31 more readmissions per year after heart failure than the local model. Our results suggest that a particular hospital's readmission risk models may not benefit from pooling data; hospitals can use their own data to predict readmissions and achieve nearly the same accuracy as models that used pooled data.

One of the main purposes of readmission risk modeling for individuals is to prioritize expensive transitional care interventions. However, existing models simply predict readmission risk, not the expected outcome of any transitional care interventions. It is possible that patients at very high risk of readmission will soon be readmitted even if clinicians apply transitional care interventions. In future work, I plan to explore how a predictive model of readmissions could be combined with data on the effect of ongoing transitional care interventions to better target those interventions. Additionally, if multiple transitional care interventions are available, a model could be developed that would select the optimal intervention based on patient characteristics.

Compared to typical diagnostic models, the predictive accuracy of our readmis-

sion risk models was low (AUC of 0.63 – 0.65). However, prioritizing patients for transitional care interventions does not require near perfect or even high model accuracy; we use these models to separate those at higher risk than the others. Additionally, if we did find a very high accuracy in a readmission risk model, it would suggest one of two things: 1) the readmissions were easy to predict because they were inevitable, meaning that transitional care interventions would be of no effect; 2) the readmissions were not inevitable, but they were easily predictable, suggesting a serious problem in the quality of care. For example, if patients with a specific diagnosis from a certain hospital nearly always were readmitted (implying that we could accurately estimate their risk), but were not admitted at other hospitals (implying that their readmission was not inevitable), it would suggest a serious problem in the quality of care at that hospital.

In the second manuscript (Chapter 4), I examined how admission practices on different days-of-the-week influenced the effect of discharge day on readmission. Other authors [14, 79, 80] had found that Friday discharges were readmitted sooner than Wednesday discharges, and had suggested some targeted interventions such as increased staff on the weekends or allowing admissions into long-term care on the weekend [66–68]. However, the effect of the differing admissions on the weekend had not been investigated. Our descriptive analysis revealed that the probability of admission is greatly reduced on the weekend. Using a Markov model, I found that the lower probability of admission on the weekend would lead to a longer time-to-readmission for Friday discharges. Using a Cox proportional hazards model with only time-fixed covariates, I found that Friday discharges were admitted slightly earlier than Wednesday discharges [HR: 1.03 95% CI: (1.02, 1.05)], replicating the findings of others. However, after adding time-varying covariates to adjust for the weekend effect, U found that Friday discharges were readmitted slightly sooner than in the time-fixed model [HR: 1.04 95% CI: (1.01, 1.07)]. I found that the lower probabilities of admission on the weekend modified the effect of Friday discharges

on time-to-readmission by increasing the time-to-readmission; any model that did not account for the lower probability of admission on the weekend would underestimate the effect of Friday discharge on readmission.

Friday discharges only have a weak (but significant) effect on individual readmission time, but because nearly 20% of hospital patients are discharged on a Friday, an effective intervention has a potentially large population effect. We found that – over 11 years, in 20 hospitals – roughly 130 000 people in our cohort were discharged on a Friday. An intervention that improved the mean time-to-readmission for Friday discharges by 1 day for only 1% of the population would save rougly 1300 in-hospital patient-days. Compared to expensive individual patient interventions, systemic interventions can be a potentially cost-effective means to improve the health care system and patient health.

Systemic interventions on patient flow have the potential to create "virtuous circles" of improvement at many points of care. For example, an intervention which improves long-term care access on the weekends might lead to longer times to readmission, reducing congestion and waiting time in emergency departments, reducing the pressure to discharge patients early, resulting in an even further increase in times-to-readmission. By understanding the system dynamics of the healthcare system, we may be able to identify cost-effective interventions that reduce congestion in different parts of the health care system, reducing waiting times and improving access.

Many jurisdictions financially penalize hospitals with poor readmission rates, but the adjustment for severity-of-illness relies on only a few well-known confounders. In the third manuscript (Chapter 5), I used TMLE in combination with a machine learning technique (random forest) to estimate the differences in the effect of different hospitals on readmissions, taking advantage of the rich confounder data available in administrative databases. The model with only a few well-known confounders estimated a relatively smaller range of odds ratios for the effect of the twenty hospitals (0.95–1.02 for AMI, 0.92–1.04 for heart failure, and 0.96–1.04 for pneumonia) than the TMLE model (0.57–2.30 for AMI,

0.50–1.85 for heart failure, and 0.47–1.55 for pneumonia). Our results suggested that the crudely adjusted hospital readmission rates underestimate the differences of a hospital's effect on quality of care on readmissions.

6.1 Ethics of risk modeling

Although risk models may help target transitional care interventions, and help us model the health care system, their use in health insurance has troubling ethical implications. Recently, large cash prizes were offered in a contest to develop models that accurately predicted the length-of-stay at hospitals [119]. Although hospital administrators could use these models for planning and staffing purposes, insurers are motivated to use the models to price insurance policies. Effectively, those who need insurance for hospital care the most (those at highest risk), will be offered insurance at the highest price. Like other outcomes, readmission risk models could be used to raise the price of insurance for those most in need. Similarly, if the financial penalty is strong enough, a hospital has a perverse incentive to refuse patients at high risk of readmission – a readmission risk model can operationalize these perverse incentives. Some authors suggest that pay-for-performance measures may unfairly penalize those who treat underserved patients[34, 35]. A readmission risk model which identifies certain high-risk ethnicities or social classes may incent hospitals to admit fewer of those groups of people, increasing disparity. In light of these ethical implications, the effect of risk models on vulnerable groups should be closely monitored during implementation.

6.2 Conclusion

In this work, I have made several unique contributions to the study of hospital readmissions. Measuring the preventability of readmissions is costly (requiring

clinical review) and most instruments are unreliable. By defining preventable hospital readmissions in counterfactual terms, it was not necessary to measure the individual preventability of readmissions, making the study of readmissions in administrative databases feasible and valid. In my first manuscript, I have shown that pooling hospital data from several hospitals may not significantly improve the accuracy of predicting hospital readmissions. In the second manuscript, I have shown that inference of the effect of Friday discharge on readmission is modified by the probability of admission on the weekend. In the third manuscript, I have demonstrated that only accounting for a few well-known confounders of the relationship between hospital care and readmissions can result in an underestimate in the differences effect of quality of care on readmissions. Additionally, targeted maximum likelihood estimation provides an effective way to use the rich data on confounders contained within administrative databases, reducing bias in the study of hospital readmissions.

Bibliography

[1] Minges KE, Curtis JP. Living in the Readmission Era. *Circulation: Cardiovascular Interventions*. 2014. 7: 9–10. DOI: 10.1161/CIRCINTERVENTIONS.114.001174.

[2] US Government. *Public Law 111 - 148 - Patient Protection and Affordable Care Act*. 2012.

[3] Government of Ontario. *Excellent Care for All Act, 2010, SO 2010, c 14*. 2012.

[4] Fung CH, Lim YW, Mattke S, Damberg C, Shekelle PG. Systematic review: the evidence that publishing patient care performance data improves quality of care. *Annals of Internal Medicine*. 2008. 148: 111–123.

[5] Centers for Medicare and Medicaid. *QIO Program: Integrating Care for Populations and Communities.* URL: http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityImprovementOrgs/Downloads/QIOIntegrateCare-.pdf.

[6] Axon RN, Williams MV. Hospital readmission as an accountability measure. *JAMA*. 2011. 305: 504–505. DOI: 10.1001/jama.2011.72.

[7] Joynt KE, Jha AK. Who has higher readmission rates for heart failure, and why? Implications for efforts to improve care using financial incentives. *Circulation. Cardiovascular Quality and Outcomes*. 2011. 4: 53–59. DOI: 10.1161/CIRCOUTCOMES.110.950964.

[8] Ross JS, Mulvey GK, Stauffer B, Patlolla V, Bernheim SM, Keenan PS, Krumholz HM. Statistical models and patient predictors of readmission

for heart failure: a systematic review. *Archives of Internal Medicine*. 2008. 168: 1371–1386. DOI: 10.1001/archinte.168.13.1371.

[9] Verhulst L, Reid R. Hold it—my patients are sicker! *British Columbia Medical Journal*. 2001. 43: 328–333.

[10] Jack BW et al. A reengineered hospital discharge program to decrease rehospitalization: a randomized trial. *Annals of Internal Medicine*. 2009. 150: 178–187.

[11] Coleman EA, Parry C, Chalmers S, Min SJ. The care transitions intervention: results of a randomized controlled trial. *Archives of Internal Medicine*. 2006. 166: 1822–1828. DOI: 10.1001/archinte.166.17.1822.

[12] Naylor MD, Brooten D, Campbell R, Jacobsen BS, Mezey MD, Pauly MV, Schwartz JS. Comprehensive discharge planning and home follow-up of hospitalized elders: a randomized clinical trial. *JAMA: The Journal of the American Medical Association*. 1999. 281: 613–620.

[13] Cleland JG, Louis AA, Rigby AS, Janssens U, Balk AH. Noninvasive Home Telemonitoring for Patients With Heart Failure at High Risk of Recurrent Admission and Death. *Journal of the American College of Cardiology*. 2005. 45: 1654–1664. DOI: 10.1016/j.jacc.2005.01.050.

[14] van Walraven C, Bell CM. Risk of death or readmission among people discharged from hospital on Fridays. *CMAJ: Canadian Medical Association Journal = Journal De l'Association Medicale Canadienne*. 2002. 166: 1672–1673.

[15] van der Laan MJ, Rose S. *Targeted Learning: Causal Inference for Observational Data*. Springer Series in Statistics. New York, NY: Springer New York, 2011.

[16] Marciarille AM. Healing Medicare hospital recidivism: causes and cures. *American Journal of Law & Medicine*. 2011. 37: 41–80.

[17] Nightingale F. *Notes on Hospitals*. 3rd ed. Longman Green, 1863. 187 pp.

[18] Starr P. *The Social Transformation of American Medicine*. Colophon books. Basic Books, 1982.

[19] Pearl R. Modern methods in handling hospital statistics. *John Hopkins Hospital Bulletin*. 1921. 32: 184–194.

[20] United States Bureau of the Census, Hunt W, Hagan A, Pollock H, Furbush E. *Patients in Hospitals for Mental Disease, 1923*. Patients in Hospitals for Mental Disease, 1923 v. 3. U.S. Government Printing Office, 1926.

[21] Abel EK, Fee E, Brown TM. Milton I. Roemer Advocate of Social Medicine, International Health, and National Health Insurance. *American Journal of Public Health*. 2008. 98: 1596–1597. DOI: 10.2105/AJPH.2008.134189.

[22] Myers GW, Roemer MI. Multiple admissions to hospital. *Canadian journal of public health*. 1956. 47: 469–481.

[23] Acheson ED, Barr A. Multiple spells of in-patient treatment in a calendar year. *British Journal of Preventive & Social Medicine*. 1965. 19: 182–191.

[24] Iglehart JK. Medicare begins prospective payment of hospitals. *The New England Journal of Medicine*. 1983. 308: 1428–1432. DOI: 10.1056/NEJM198306093082331.

[25] Committee to Design a Strategy for Quality Review and Assurance in Medicare, Institute of Medicine. *Medicare : a strategy for quality assurance*. Ed. by Lohr KN. Washington, D.C.: National Academy Press, 1990.

[26] Ashton CM, Del Junco DJ, Souchek J, Wray NP, Mansyur CL. The association between the quality of inpatient care and early readmission: a meta-analysis of the evidence. *Medical care*. 1997. 35: 1044–1059.

[27] Centers for Medicare and Medicaid. *Hospital Compare*. URL: http://www.medicare.gov/hospitalcompare.

[28] Baker GR, Avoidable Hospitalization Advisory Panel. *Enhancing the Continuum of Care.* 2011. URL: http://www.health.gov.on.ca/en/common/ministry/publications/reports/baker_2011/baker_2011.pdf.

[29] Goodman DC, Fisher ES, Chang CH. After hospitalization: a Dartmouth atlas report on post-acute care for Medicare beneficiaries. *The Dartmouth Institute, September*. 2011. 28:

[30] Medew J. *Push to fine hospitals*. The Age: Victoria. 2013. URL: http://www.theage.com.au/victoria/push-to-fine-hospitals-20130711-2pt6n.html.

[31] Kansagara D, Englander H, Salanitro A, Kagen D, Theobald C, Freeman M, Kripalani S. Risk prediction models for hospital readmission: a systematic review. *JAMA: The Journal of the American Medical Association*. 2011. 306: 1688–1698. DOI: 10.1001/jama.2011.1515.

[32] Bissey B. Observation, Admission, and RAC—The Next Perfect Storm? *IMA Insights*. 2008. 6: 4.

[33] Hockenberry JM, Mutter R, Barrett M, Parlato J, Ross MA. Factors Associated with Prolonged Observation Services Stays and the Impact of Long Stays on Patient Cost. *Health Services Research*. 2014. 49: 893–909. DOI: 10.1111/1475-6773.12143.

[34] Casalino LP, Elster A, Eisenberg A, Lewis E, Montgomery J, Ramos D. Will pay-for-performance and quality reporting affect health care disparities? *Health Affairs*. 2007. 26: w405–414. DOI: 10.1377/hlthaff.26.3.w405.

[35] Chien AT, Chin MH, Davis AM, Casalino LP. Pay for performance, public reporting, and racial disparities in health care: how are programs being designed? *Medical care research and review: MCRR*. 2007. 64: 283S–304S. DOI: 10.1177/1077558707305426.

[36] Werner RM, Konetzka RT, Kruse GB. Impact of Public Reporting on Unreported Quality of Care. *Health Services Research*. 2009. 44: 379–398. DOI: 10.1111/j.1475-6773.2008.00915.x.

[37] Werner RM. Does Pay-for-Performance Steal From the Poor and Give to the Rich? *Annals of Internal Medicine*. 2010. 153: 340. DOI: 10.7326/0003-4819-153-5-201009070-00010.

[38] Berenson RA, Paulus RA, Kalman NS. Medicare's readmissions-reduction program--a positive alternative. *The New England journal of medicine*. 2012. 366: 1364–1366. DOI: 10.1056/NEJMp1201268.

[39] Medicare Payment Advisory Commission. *Report to the Congress: Promoting Greater Efficiency in Medicine*. 2007. URL: http://www.medpac.gov/documents/jun07_entirereport.pdf.

[40] Kripalani S, LeFevre F, Phillips CO, Williams MV, Basaviah P, Baker DW. Deficits in communication and information transfer between hospital-based and primary care physicians: implications for patient safety and continuity of care. *Jama*. 2007. 297: 831–841.

[41] Kohn K, Corrigan J, Donaldson M. *To Err Is Human: Building a Safer Health System*. Washington, DC: National Academy Press, 1999.

[42] Schneider JK, Hornberger S, Booker J, Davis A, Kralicek R. A Medication Discharge Planning Program: Measuring the Effect on Readmissions. *Clinical Nursing Research*. 1993. 2: 41–53. DOI: 10.1177/105477389300200105.

[43] Fisher ES, Wennberg JE, Stukel TA, Sharp SM. Hospital Readmission Rates for Cohorts of Medicare Beneficiaries in Boston and New Haven. *New England Journal of Medicine*. 1994. 331: 989–995. DOI: 10.1056/NEJM199410133311506.

[44] Epstein AM, Jha AK, Orav EJ. The Relationship between Hospital Admission Rates and Rehospitalizations. *New England Journal of Medicine*. 2011. 365: 2287–2295. DOI: 10.1056/NEJMsa1101942.

[45] Ashton CM, Wray NP. A conceptual framework for the study of early readmission as an indicator of quality of care. *Social Science & Medicine (1982)*. 1996. 43: 1533–1541.

[46] van Walraven C, Jennings A, Taljaard M, Dhalla I, English S, Mulpuru S, Blecker S, Forster AJ. Incidence of potentially avoidable urgent readmissions and their relation to all-cause urgent readmissions. *CMAJ: Canadian Medical Association Journal = Journal De l'Association Medicale Canadienne*. 2011. 183: E1067–1072. DOI: 10.1503/cmaj.110400.

[47] Halfon P, Eggli Y, van Melle G, Chevalier J, Wasserfallen JB, Burnand B. Measuring potentially avoidable hospital readmissions. *Journal of Clin-*

ical Epidemiology. 2002. 55: 573–587. DOI: 10.1016/S0895-4356(01)00521-2.

[48] Halfon P, Eggli Y, Prêtre-Rohrbach I, Meylan D, Marazzi A, Burnand B. Validation of the potentially avoidable hospital readmission rate as a routine indicator of the quality of hospital care. *Medical Care*. 2006. 44: 972–981. DOI: 10.1097/01.mlr.0000228002.43688.c2.

[49] Milne R, Clarke A. Can readmission rates be used as an outcome indicator? *BMJ*. 1990. 301: 1139–1140. DOI: 10.1136/bmj.301.6761.1139.

[50] Ash AS, Fienberg SE, Louis TA, Normand SLT, Stukel TA, Utts J. *Statistical issues in assessing hospital performance: Commission by the Committee of Presidents of Statistical Societies (white paper)*. 2012.

[51] Stone J, Hoffman GJ. *Medicare Hospital Readmissions: Issues, Policy Options and PPACA*. 2010.

[52] Naylor MD, Brooten DA, Campbell RL, Maislin G, McCauley KM, Schwartz JS. Transitional care of older adults hospitalized with heart failure: a randomized, controlled trial. *Journal of the American Geriatrics Society*. 2004. 52: 675–684. DOI: 10.1111/j.1532-5415.2004.52202.x.

[53] Centers for Medicare and Medicaid. *Community-based Care Transitions Program*. URL: http://innovation.cms.gov/initiatives/CCTP/.

[54] South West Local Health Integration Network. *Pilot Project: Improving the Transition from Hospital, Inpatient and Emergency Department, to Home for Seniors Living with Increased Risk*. URL: http://www.southwestlhin.on.ca/~/media/sites/sw/uploadedfiles/Public_Community/Current_Initiatives/Access_to_Care/HRS%20Pilot%20Project%20Summary%20FINAL.pdf.

[55] Brumley RD, Enguidanos S, Cherin DA. Effectiveness of a Home-Based Palliative Care Program for End-of-Life. *Journal of Palliative Medicine*. 2003. 6: 715–724. DOI: 10.1089/109662103322515220.

[56] Gade G, Venohr I, Conner D, McGrady K, Beane J, Richardson RH, Williams MP, Liberson M, Blum M, Della Penna R. Impact of an inpatient pallia-

tive care team: a randomized control trial. *Journal of Palliative Medicine*. 2008. 11: 180–190. DOI: 10.1089/jpm.2007.0055.

[57] Penrod JD et al. Hospital-Based Palliative Care Consultation: Effects on Hospital Cost. *Journal of Palliative Medicine*. 2010. 13: 973–979. DOI: 10.1089/jpm.2010.0038.

[58] Penrod JD, Deb P, Luhrs C, Dellenbaugh C, Zhu CW, Hochman T, Maciejewski ML, Granieri E, Morrison RS. Cost and Utilization Outcomes of Patients Receiving Hospital-Based Palliative Care Consultation. *Journal of Palliative Medicine*. 2006. 9: 855–860. DOI: 10.1089/jpm.2006.9.855.

[59] Tangeman JC, Rudra CB, Kerr CW, Grant PC. A Hospice-Hospital Partnership: Reducing Hospitalization Costs and 30-Day Readmissions among Seriously Ill Adults. *Journal of Palliative Medicine*. 2014. 17: 1005–1010. DOI: 10.1089/jpm.2013.0612.

[60] Woodend AK, Sherrard H, Fraser M, Stuewe L, Cheung T, Struthers C. Telehome monitoring in patients with cardiac disease who are at high risk of readmission. *Heart & Lung: The Journal of Acute and Critical Care*. 2008. 37: 36–45. DOI: 10.1016/j.hrtlng.2007.04.004.

[61] Wolfson M et al. DataSHIELD: resolving a conflict in contemporary bioscience--performing a pooled analysis of individual-level data without sharing the data. *International Journal of Epidemiology*. 2010. 39: 1372–1382. DOI: 10.1093/ije/dyq111.

[62] Wang S, Jiang X, Wu Y, Cui L, Cheng S, Ohno-Machado L. EXpectation Propagation LOgistic REgRession (EXPLORER): Distributed privacy-preserving online model learning. *Journal of Biomedical Informatics*. 2013. 46: 480–496. DOI: 10.1016/j.jbi.2013.03.008.

[63] El Emam K, Samet S, Arbuckle L, Tamblyn R, Earle C, Kantarcioglu M. A secure distributed logistic regression protocol for the detection of rare adverse drug events. *Journal of the American Medical Informatics Association*. 2012. 20: 453–461. DOI: 10.1136/amiajnl-2011-000735.

[64] Rassen JA, Solomon DH, Curtis JR, Herrinton L, Schneeweiss S. Privacy-maintaining propensity score-based pooling of multiple databases ap-

plied to a study of biologics. *Medical care*. 2010. 48: S83–89. DOI: 10.1097/MLR.0b013e3181d59541.

[65] Rassen JA, Avorn J, Schneeweiss S. Multivariate-adjusted pharmacoepidemiologic analyses of confidential information pooled from multiple health care utilization databases. *Pharmacoepidemiology and drug safety*. 2010. 19: 848–857. DOI: 10.1002/pds.1867.

[66] Bell CM. Enhanced weekend service: an affordable means to increased hospital procedure volume. *Canadian Medical Association Journal*. 2005. 172: 503–504. DOI: 10.1503/cmaj.1041063.

[67] Moore JD. Hospital saves by working weekends. *Modern Healthcare*. 1996. 26: 82, 84, 99.

[68] Wong HJ, Wu RC, Caesar M, Abrams H, Morra D. Smoothing inpatient discharges decreases emergency department congestion: a system dynamics simulation model. *Emergency Medicine Journal*. 2010. 27: 593–598. DOI: 10.1136/emj.2009.078543.

[69] Graf C. Functional decline in hospitalized older adults. *The American journal of nursing*. 2006. 106: 58–67, 58–67.

[70] McHugh MD, Ma C. Hospital nursing and 30-day readmissions among Medicare patients with heart failure, acute myocardial infarction, and pneumonia. *Medical care*. 2013. 51: 52–59. DOI: 10.1097/MLR.0b013e3182763284.

[71] Gornick M. Medicare patients: geographic differences in hospital discharge rates and multiple stays. *Social security bulletin*. 1977. 40: 22–41.

[72] Wilhelm EE, Oster E, Shoulson I. Approaches and Costs for Sharing Clinical Research Data. *JAMA*. 2014. DOI: 10.1001/jama.2014.850.

[73] van Walraven C, Dhalla IA, Bell C, Etchells E, Stiell IG, Zarnke K, Austin PC, Forster AJ. Derivation and validation of an index to predict early death or unplanned readmission after discharge from hospital to the community. *CMAJ: Canadian Medical Association journal = journal de l'Association medicale canadienne*. 2010. 182: 551–557. DOI: 10.1503/cmaj.091117.

[74] Tibshirani R. The lasso method for variable selection in the Cox model. *Statistics in medicine*. 1997. 16: 385–395.

[75] Bradley AP. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*. 1997. 30: 1145–1159. DOI: 10.1016/S0031-3203(96)00142-2.

[76] Simon N, Friedman J, Hastie T, Tibshirani R. Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent. *Journal of Statistical Software*. 2011. 39: 1–13.

[77] Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*. 2010. 33: 1–22.

[78] Wiens J, Guttag J, Horvitz E. A study in transfer learning: leveraging data from multiple hospitals to enhance hospital-specific predictions. *Journal of the American Medical Informatics Association: JAMIA*. 2014. DOI: 10.1136/amiajnl-2013-002162.

[79] Mackie AS, Ionescu-Ittu R, Pilote L, Rahme E, Marelli AJ. Hospital readmissions in children with congenital heart disease: a population-based study. *American heart journal*. 2008. 155: 577–584. DOI: 10.1016/j.ahj.2007.11.003.

[80] van Walraven C, Mamdani M, Fang J, Austin PC. Continuity of care and patient outcomes after hospital discharge. *Journal of general internal medicine*. 2004. 19: 624–631. DOI: 10.1111/j.1525-1497.2004.30082.x.

[81] Mutrie D, Bailey SK, Malik S. Individual emergency physician admission rates: predictably unpredictable. *CJEM*. 2009. 11: 149–155.

[82] Ting HH, Lee TH, Soukup JR, Cook EF, Tosteson AN, Brand DA, Rouan GW, Goldman L. Impact of physician experience on triage of emergency room patients with acute chest pain at three teaching hospitals. *The American journal of medicine*. 1991. 91: 401–408.

[83] Coast J, Peters TJ, Inglis A. Factors associated with inappropriate emergency hospital admission in the UK. *International Journal for Quality in Health Care*. 1996. 8: 31–39.

[84] Gaucher N, Bailey B, Gravel J. Impact of Physicians' Characteristics on the Admission Risk Among Children Visiting a Pediatric Emergency Department. *Pediatric Emergency Care*. 2012. 1. DOI: 10.1097/PEC.0b013e318243f8e0.

[85] R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2014.

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

[87] Pinkus A. *YACAS: Yet another computerl algebra system*. Version 1.3.6.

[88] Wickham H. *ggplot2: elegant graphics for data analysis*. Springer New York, 2009.

[89] Glazier RH, Klein-Geltink J, Kopp A, Sibley LM. Capitation and enhanced fee-for-service models for primary care reform: a population-based evaluation. *Canadian Medical Association Journal*. 2009. 180: E72–E81. DOI: 10.1503/cmaj.081316.

[90] Trudeau R. Monthly and daily patterns of death. *Health reports - Statistics Canada*. 1997. 9: 43–52.

[91] Rogot E, Fabsitz R, Feinleib M. Daily variation in USA mortality. *American Journal of Epidemiology*. 1976. 103: 198–211.

[92] Alibhai SM, Han RK, Naglie G. Medication education of acutely hospitalized older patients. *Journal of General Internal Medicine*. 1999. 14: 610–616.

[93] Varnava AM, Sedgwick JEC, Deaner A, Ranjadayalan K, Timmis AD. Restricted weekend service inappropriately delays discharge after acute myocardial infarction. *Heart (British Cardiac Society)*. 2002. 87: 216–219.

[94] Carey MR, Sheth H, Braithwaite RS. A prospective study of reasons for prolonged hospitalizations on a general medicine teaching service. *Journal of General Internal Medicine*. 2005. 20: 108–115. DOI: 10.1111/j.1525-1497.2005.40269.x.

[95] Homer JB, Hirsch GB. System Dynamics Modeling for Public Health: Background and Opportunities. *American Journal of Public Health*. 2006. 96: 452–458. DOI: 10.2105/AJPH.2005.062059.

[96] Midgley G. Systemic Intervention for Public Health. *American Journal of Public Health*. 2006. 96: 466–472. DOI: 10.2105/AJPH.2005.067660.

[97] Glymour MM, Osypuk TL, Rehkopf DH. Invited Commentary: Off-Roading With Social Epidemiology--Exploration, Causation, Translation. *American Journal of Epidemiology*. 2013. 178: 858–863. DOI: 10.1093/aje/kwt145.

[98] He D, Mathews SC, Kalloo AN, Hutfless S. Mining high-dimensional administrative claims data to predict early hospital readmissions. *Journal of the American Medical Informatics Association*. 2014. 21: 272–279. DOI: 10.1136/amiajnl-2013-002151.

[99] Hosseinzadeh A, Izadi M, Verma A, Precup D, Buckeridge D. Assessing the Predictability of Hospital Readmission Using Machine Learning. *Proceedings of the Twenty-Fifth Innovative Applications of Artificial Intelligence Conference*. 2013. 1532–1538.

[100] Breiman L. Random Forests. *Machine Learning*. 2001. 45: 5–32. DOI: 10.1023/A:1010933404324.

[101] Statistics Canada. *2006 Census Dictionary*. 2007.

[102] Gini C. *Variabilità e Mutabilità: Contributo allo studio delle distribuzioni e delle relazioni statistiche.* Bologna: C. Cuppini, 1912. 156 pp.

[103] Platt JC. "Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods". *Advances in large margin classifiers*. MIT Press, 1999: 61–74.

[104] Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity Measures for Use with Administrative Data: *Medical Care*. 1998. 36: 8–27. DOI: 10.1097/00005650-199801000-00004.

[105] Lim A, Breiman L, Cutler A. *bigrf: Big Random Forests: Classification and Regression Forests for Large Data Sets*. R package version 0.1-9. 2014.

[106] Silverman BW. *Density estimation for statistics and data analysis*. Vol. 26. CRC press, 1986.

[107] Cox DR. *The analysis of binary data*. Methuen's monographs on applied probability and statistics. London: Methuen, 1970. 142 pp.

[108] Clarke A. Are readmissions avoidable? *BMJ*. 1990. 301: 1136–1138. DOI: 10.1136/bmj.301.6761.1136.

[109] Witherington EMA, Pirzada OM, Avery AJ. Communication gaps and readmissions to hospital for patients aged 75 years and older: observational study. *Quality and Safety in Health Care*. 2008. 17: 71–75. DOI: 10.1136/qshc.2006.020842.

[110] Stanley A, Graham N, Parrish A. A review of internal medicine re-admissions in a peri-urban South African hospital. *South African Medical Journal = Suid-Afrikaanse Tydskrif Vir Geneeskunde*. 2008. 98: 291–294.

[111] Ruiz B, García M, Aguirre U, Aguirre C. Factors predicting hospital readmissions related to adverse drug reactions. *European Journal of Clinical Pharmacology*. 2008. 64: 715–722. DOI: 10.1007/s00228-008-0473-y.

[112] van Walraven C, Bennett C, Jennings A, Austin PC, Forster AJ. Proportion of hospital readmissions deemed avoidable: a systematic review. *Canadian Medical Association Journal*. 2011. 183: E391–E402. DOI: 10.1503/cmaj.101860.

[113] Hernán MA, Robins JM. *Causal Inference (Draft - May 14, 2014)*. 2014.

[114] Pearl J. Causal diagrams for empirical research. *Biometrika*. 1995. 82: 669–688. DOI: 10.1093/biomet/82.4.669.

[115] Greenland S. Quantifying biases in causal models: classical confounding vs collider-stratification bias. *Epidemiology*. 2003. 14: 300–306.

[116] Liu W, Brookhart MA, Schneeweiss S, Mi X, Setoguchi S. Implications of M Bias in Epidemiologic Studies: A Simulation Study. *American Journal of Epidemiology*. 2012. 176: 938–948. DOI: 10.1093/aje/kws165.

[117] van der Laan MJ, Polley EC, Hubbard AE. Super Learner. *Statistical Applications in Genetics and Molecular Biology*. 2007. 6: 1–21. DOI: 10.2202/1544-6115.1309.

[118] Donabedian A. Evaluating the quality of medical care. *The Milbank memorial fund quarterly*. 1966. 166–206.

[119] Heritage Provider Network. *Heritage health prize: description*. URL: `http://www.heritagehealthprize.com/c/hhp`.

Appendix

Table 6.1: ICD-9 codes for influenza-like illness. The cohort used in these studies used a diagnosis of an ILI as a selection criterion. ICD-9 codes typically use a decimal point, but not in the Quebec administrative databases. The list presented here refelcts the definition of ICD-9 codes as they are used in the Quebec administrative databases.

ICD9	Description
100	Primary tuberculous infection, primary tuberculous complex
101	Tuberculous pleurisy in primary progressive tuberculosis
108	Other primary progressive tuberculosis
109	Primary tuberculous infection, unspecified
110	Tuberculosis of lung, infiltrative
111	Tuberculosis of lung, nodular
112	Tuberculosis of lung with cavitation
113	Tuberculosis of bronchus
114	Tuberculous fibrosis of lung
115	Tuberculous bronchiectasis
116	Tuberculous pneumonia (any form)
117	Tuberculous pneumothorax
118	Other pulmonary tuberculosis

Table 6.1: *(continued)*

ICD9	Description
119	Pulmonary tuberculosis, unspecified
203	Plague primary pneumonic
204	Plague secondary pneumonic
205	Plague pneumonic, unspecified
219	Tularaemia
221	Pulmonary anthrax
249	Glanders
259	Melioidosis
320	Faucial diphtheria
321	Nasopharyngeal diphtheria
322	Anterior nasal diphtheria
323	Laryngeal diphtheria
329	Diphtheria, unspecified
330	Whooping cough, bordetella pertussis (B.pertussis)
331	Whooping cough, bordetella parapertussis (B.parapertussis)
338	Whooping cough, other specified organism
339	Whooping cough, unspecified organism
340	Streptococcal sore throat
529	Chickenpox
	Chickenpox (varicella), uncomplicated
551	Measles, postmeasles pneumonia
739	Ornithosis
741	Specific diseases due to coxsackie virus, epidemic pleurodynia
790	Adenovirus
793	Rhinovirus
798	Other viral infection
799	Viremia, unspecified
	Unspecified viral infection

Table 6.1: *(continued)*

ICD9	Description
830	Rickettsioses, Q-fever
1124	Candidiasis, of lung
1149	Coccidioidomycosis
1150	Infection by histoplasma capsulatum
1151	Infection by histoplasma duboisii
1159	Histoplasmosis, unspecified
1309	Toxoplasmosis
1363	Pneumocystosis
3820	Suppurative and unspecified otitis media, acute suppurative otitis media
3824	Unspecified suppurative otitis media
3829	Unspecified otitis media
4609	Acute nasopharyngitis (common cold), acute nasopharyngitis (common cold)
4618	Acute sinusitis, other
4619	Acute sinusitis, unspecified
4629	Acute pharyngitis, acute pharyngitis
4639	Acute tonsillitis, acute tonsillitis
4640	Acute laryngitis
4641	Acute tracheitis
4642	Acute laryngotracheitis
4643	Acute epiglottitis
4644	Acute laryngitis and tracheitis, croup
4650	Acute laryngopharyngitis
4658	Other multiple sites
4659	Acute upper respiratory infections, unspecified site URTI, unspecified
4660	Acute bronchitis
4661	Acute bronchiolitis
4789	Other and unspecified diseases of upper respiratory tract

Table 6.1: *(continued)*

ICD9	Description
4800	Viral pneumonia, pneumonia due to adenovirus
4801	Viral pneumonia, pneumonia due to respiratory syncytial virus
4802	Viral pneumonia, pneumonia due to parainfluenza virus
4808	Viral pneumonia, pneumonia due to other virus, not elsewhere classified
4809	Viral pneumonia, viral pneumonia, unspecified
4819	Pneumococcal pneumonia
4820	Other bacterial pneumonia, pneumonia due to klebsiella pneumoniae
4821	Other bacterial pneumonia, pneumonia due to pseudomonas
4822	Pneumonia due to haemophilus influenzae (h.influenzae)
4823	Other bacterial pneumonia, pneumonia due to streptococcus
4824	Other bacterial pneumonia, pneumonia due to staphylococcus
4828	Other bacterial pneumonia, pneumonia due to other specified bacteria
4829	Other bacterial pneumonia, bacterial pneumonia, unspecified
4839	Pneumonia due to other specified organism
4841	Cytomegalic inclusion disease
4843	Pneumonia in infectious diseases classified elsewhere, whooping cough
4845	Pneumonia in infectious diseases classified elsewhere, anthrax
4846	Pneumonia in infectious diseases classified elsewhere, aspergillosis
4847	Pneumonia in other systemic mycoses
4848	Pneumonia in other infectious diseases
4859	Bronchopneumonia, organism unspecified
4869	Pneumonia, organism unspecified
	Pneumonia, unspecified
4870	Influenza, with pneumonia
4871	Influenza (flu) NOS
	Influenza, with other respiratory manifestations
4878	Influenza, with other manifestations
4909	Bronchitis, not specified as acute or chronic

Table 6.1: *(continued)*

ICD9	Description
4910	Simple chronic bronchitis
4911	Mucopurulent chronic bronchitis
4918	Other chronic bronchitis
4919	Chronic bronchitis, unspecified
5070	Pneumonitis due to solids and liquids, due to inhalation of food or vomit
5071	Due to inhalation of oils and essences
5078	Pneumonitis due to solids and liquids, other
5110	Pleurisy, without mention of effusion or current tuberculosis
5111	With effusion, with mention of a bacterial cause other than tuberculosis
5118	Pleurisy, other specified forms of effusion, except tuberculosis
5119	Pleurisy, unspecified pleural effusion
5130	Abscess of lung
5131	Abscess of mediastinum
5180	Other diseases of lung, pulmonary collapse
5184	Other diseases of lung, acute oedema of lung, unspecified
5188	Other diseases of lung, other diseases of lung, not elsewhere classified
5192	Other diseases of respiratory system, mediastinitis
7806	Chills
	General symptoms, pyrexia of unknown origin
	General symptoms: fever, not otherwise specified
	Hyperthermia
7841	Symptoms involving head and neck, throat pain
7860	Dyspnoea and respiratory abnormalities
	Shortness of breath
7861	Symptoms involving respiratory system and other chest symptoms, stridor
7862	Symptoms involving respiratory system and other chest symptoms, cough
7865	Symptoms involving respiratory system and other chest symptoms, chest pain
	Pleurodynia

Table 6.1: *(continued)*

ICD9	Description
7953	Nonspecific positive culture findings
V018	Other communicable diseases