Factors associated with emergency department length of stay in Montreal – results of a quasi-experimental study and electronic medical record data analysis

Wajiha Masood

Department of Family Medicine

McGill University, Montreal

August 2020

Supervised by: Dr. Tibor Schuster

Co supervisor: Dr. Peter Nugus

A thesis submitted to McGill University, in partial fulfilment of the requirements of the degree of Master of Science in Family Medicine.

ABSTRACT

Introduction to the Problem:

Canada has the longest ED (emergency department) waiting time in the developed world and the province of Quebec has the longest mean ED waiting time in Canada. This problem is worsening for the developed world, partly because of a rapidly growing elderly population with chronic illnesses. ED crowding – often resulting from inadequate primary care services, especially for an increase of ageing in the population, and the "boarding" of patients admitted to the hospital for whom no in-patient beds are available – is associated with decreased quality of care, an increase in medical errors and has a positive association with increased mortality for patients present in the emergency department. The ED can serve an important role in providing "integrated care" between primary care and hospitalized care, and between generalist and specialist medical and surgical services. Efficiency – of "flow" – is an important aspect of integrated, or coordinated, care. ED crowding causes less efficient care that may lead to complications that require longer hospital stays and are more expensive to care for. The problem with which this thesis deals is that such a complex, inter-connected system as healthcare rarely presents systemic interventions that provide the opportunity to ascertain whether or how ED flow can be improved. This study takes advantage of a rare large-scale natural experiment-which would be characterized as an "extreme case" involving a major hospital re-location in Montreal,Quebec, Canada. The Royal Victoria Hospital was relocated from its original location to a new location in April 2015 and merged with four other hospitals to form a "super hospital" at the new site. Research analyzing electronic medical record (EMR) data using advanced statistical methods including machine learning for association with crowding is on-going. This study aims to utilize the EMR data to evaluate if there is a difference in patients' ED LOS (length of stay), before versus after the move of the ED from the original site to the new site. A secondary aim of this study is to

explore the most important predictor variables that relate to the length of stay at the ED including age and other socioeconomic variables, assessing access to primary care physician (co primary study end point by using modern statistical methods and machine learning.

Study Design and methods:

The study applied a Quasi Experimental design to utilize the data collected routinely in the ED, before and after the relocation. Data were collected from 01/06/2014 till 25/04/2015 for the original site (pre-relocation) and 26/04/2015 till 31/03/2016 for the current site (post relocation). The data were statistically analyzed through R studio version 3.6.2. After using advanced statistical methods including multiple regression modeling and machine learning modeling, information about the waiting time at the two sites and the effect of different factors/covariates on the waiting times were collected. Descriptive and inferential analysis were used to compare patient- and ED-process data before and after the re-location.

Results:

In the univariate analysis (no adjustment for covariates), there was an estimated mean difference of 0.55 hours (95% CI -0.79; -0.30 hours) in the LOS between the original site and the current site. In the multivariable analysis adjusting for the potential confounding covariates site, age, gender, primary care md description, referring MD description, arrival mode description, triage priority; the mean ratio of ED LOS (new site / original site) was estimated to be 0.98 [95%CI:0.97 to 1.00], i.e. suggesting relative differences in ED LOS between sites of between zero and 3%. The machine learning method that used random forest explained 21% of the variation in ED LOS due to covariates, with age more than 61 being the most important predictor factor for longer ED LOS.

Implications and contribution:

This thesis showed that it is possible and desirable to understand and evaluate organizational re-structuring / interventions, to explore their effect on efficiency of ED and predicting ED LOS based on covariates. In an inter-connected, dynamic system, we rarely have the opportunity to identify a purposeful systemic intervention to improve the ED efficiency. Health organizational interventions including site relocation results in some changes in the ED functioning and ED population dynamic, which allows for an evaluation of the effect of these changes on ED length of stay and ED efficiency.

This thesis has shown that a systemic, purposeful intervention like site relocation and restructuring allows for focused study and discernment of variables contributing to improved ED efficiency, especially when coupled with modern machine learning techniques. Such techniques surpass conventional statistical modeling because they have the potential to provide greater insight for predicting ED LOS and for improving the efficiency of health services. These results may help in identifying factors that reduce ED LOS and may help in building an algorithm to potentially build a tool to predict ED waiting times based on relevant factors.

ABSTRAIT

<u>Introduction au problème</u> **:**

Canada a le temps d'attente au service d'urgence le plus long des pays développés et la province de Québec a le temps d'attente avec le plus long moyen au service d'urgence. Ce problème s'aggrave dans le monde développé, en partie à cause d'une population de personnes âgées en croissance rapide souffrant de maladies chroniques. Le surpeuplement des urgences - qui résulte souvent de services de soins primaires inadéquats, en particulier pour une augmentation du vieillissement de la population, et le « internement » des patients admis à l'hôpital pour lesquels aucun lit d'hospitalisation n'est disponible - est associé à une diminution de la qualité des soins, une augmentation des erreurs médicales et a une association positive avec une mortalité accrue pour les patients présents à l'urgence. Le service d'urgence peut jouer un rôle important dans la fourniture de « soins intégrés » entre les soins primaires et les soins hospitaliers, et entre les services médicaux et chirurgicaux généralistes et spécialisés. L'efficacité - du « flux » - est un aspect important des soins intégrés ou coordonnés. Le surpeuplement des urgences entraîne des soins moins efficaces qui peuvent entraîner des complications nécessitant des séjours hospitaliers plus longs et plus coûteux à soigner. Le problème que traite cette thèse est qu'un système aussi complexe et interconnecté que les soins de santé présentent rarement des interventions systémiques qui offrent l'opportunité de déterminer si ou comment le flux ED peut être amélioré. Cette étude tire parti d'une rare expérience naturelle à grande échelle - qui pourrait être qualifiée de « cas extrême », impliquant un déplacement majeur de l'hôpital à Montréal, Québec, Canada. L'hôpital Royal Victoria a été déplacé de son emplacement d'origine à un nouvel emplacement en avril 2015 et a fusionné avec quatre autres hôpitaux pour former un « super hôpital » sur le nouveau site. Des recherches sur l'analyse des données des dossiers médicaux électroniques (DME) à l'aide de méthodes statistiques avancées, y compris l'apprentissage automatique pour l'association avec le surpeuplement, sont en cours. Cette étude vise à utiliser les données du DME pour évaluer s'il

existe une différence dans la durée de séjour des patients au SU (durée du séjour), avant et après le déplacement du SU du site d'origine vers le nouveau site. Un objectif secondaire de cette étude est d'explorer les variables prédictives les plus importantes liées à la durée du séjour à l'urgence, y compris l'âge et d'autres variables socio-économiques, en évaluant l'accès au médecin de soins primaires (co-critère principal de l'étude en utilisant des méthodes statistiques modernes et une machine apprentissage.

***Conception de l'étude et méthodes* :**

L'étude a appliqué une conception quasi expérimentale pour utiliser les données collectées régulièrement dans l'urgence, avant et après le déménagement. Les données ont été collectées du 01/06/2014 au 25/04/2015 pour le site d'origine (pré-relocalisation) et du 26/04/2015 au 31/03/2016 pour le site actuel (post relocation). Les données ont été analysées statistiquement via la version 3.6.2 de R studio. Après avoir utilisé des méthodes statistiques avancées, notamment la modélisation de régression multiple et la modélisation d'apprentissage automatique, des informations sur le temps d'attente sur les deux sites et l'effet de différents facteurs / covariables sur les temps d'attente ont été collectées. Des analyses descriptives et inférentielles ont été utilisées pour comparer les données des patients et des urgences avant et après le déplacement.

Résultats :

Dans l'analyse univariée (pas d'ajustement pour les covariables), il y avait une différence moyenne estimée de 0,55 heure (IC à 95% -0,79 ; -0,30 heure) dans la DS entre le site d'origine et le site actuel. Dans l'analyse multivariée, ajustement pour les covariables confondantes potentielles site, âge, sexe, description MD des soins primaires, description du médecin traitant, description du mode d'arrivée, priorité de triage; le rapport moyen ED LOS (nouveau site / site d'origine) a été estimé à 0,98 [IC 95%: 0,97 à 1,00], c'est-à-dire suggérant des différences relatives de ED LOS entre les sites comprises entre zéro et 3%. La méthode d'apprentissage automatique qui utilisait la forêt aléatoire expliquait 21% de la variation de la ED LOS due aux covariables, l'âge de plus de 61 ans étant le facteur prédictif le plus important pour une ED LOS plus longue.

Implications et contribution :

Cette thèse a montré qu'il est possible et souhaitable de comprendre et d'évaluer les restructurations / interventions organisationnelles, d'explorer leur effet sur l'efficacité de la dysfonction érectile et de prédire la durée de vie de la dysfonction érectile sur la base de covariables. Dans un système dynamique interconnecté, nous avons rarement l'occasion d'identifier une intervention systémique ciblée pour améliorer l'efficacité des urgences. Les interventions organisationnelles de santé, y compris le déménagement des sites, entraînent certains changements dans le fonctionnement du SU et la dynamique de la population du SU, ce qui permet une évaluation de l'effet de ces changements sur la durée du séjour au SU et l'efficacité du SU.

Cette thèse a montré qu'une intervention systémique et ciblée telle que la délocalisation et la restructuration de sites permet une étude ciblée et le discernement des variables contribuant à l'amélioration de l'efficacité de l'ED, en particulier lorsqu'elle est associée à des techniques modernes d'apprentissage automatique. Ces techniques dépassent la modélisation statistique conventionnelle parce qu'elles ont le potentiel de fournir un meilleur aperçu pour prédire ED LOS et pour améliorer l'efficacité des services de santé. Ces résultats peuvent aider à identifier les facteurs qui réduisent ED LOS et peuvent aider à la construction d'un algorithme pour potentiellement construire un outil pour prédire les temps d'attente ED en fonction de facteurs pertinents.

Table of contents

ABSTRACT 2
ABSTRAIT 5
Table of contents... 9
List of Abbreviations...10
List of Figures 11
List of Tables 12
Acknowledgement...13
Preface:..15
Chapter 1 16
INTRODUCTION: 16
Chapter 2 27
LITERATURE REVIEW:...27
Chapter 3 39
Study Design and Research Methodology...39
4. RESULTS:...49
Chapter 5 66
Discussion: 66
References: 74
Appendix A…………………………………………………………………………….89
Appendix B…………………………………………………………………………….90

List of Abbreviations

ED: Emergency department

EMR: Electronic Medical Record

IRB: Institutional Review board

LOS: Length of stay

RVH: Royal Victoria Hospital

LWBS: Left without being seen

MUHC: McGill University Health Center

EDIS: Emergency department information system

List of Figures

1. **Figure 1……………………………………………………………………….page no. 18**
2. **Figure 2………………………………………………………………………page no. 34**
3. **Figure 3………………………………………………………………………page no.50**
4. **Figure 4A……………………………………………………………………page no.51**
5. **Figure 4B……………………………………………………………………page no.52**
6. **Figure 5………………………………………………………………………page no.56**
7. **Figure 6………………………………………………………………………page no.58**
8. **Figure 7………………………………………………………………………page no.64**
9. **Figure 8………………………………………………………………………page no.65**

List of Tables

1. Table A……………………………………………………………………………………Page no.41
2. Table B……………………………………………………………………………………Page no.43
3. Table C……………………………………………………………………………………Page no.45
4. Table D……………………………………………………………………………………Page no.46
5. Table 1……………………………………………………………………………………Page no.55
6. Table 2……………………………………………………………………………………Page no.57
7. Table 3……………………………………………………………………………………Page no.60
8. Table 4……………………………………………………………………………………Page no.61
9. Table 5……………………………………………………………………………………Page no.61
10. Table 6……………………………………………………………………………………Page no.62
11. Table 7……………………………………………………………………………………Page no.64

Acknowledgement

I have been very fortunate in having a lot of people around me who helped me through my journey of completing my thesis. My heartfelt gratitude goes first to my supervisor Dr. Tibor Schuster, whose help, support and advise was most crucial in the successful completion of this thesis. His mentorship, professionalism and flexibility in accommodating me played the key role in completion of this thesis. I consider myself very fortunate for being under his supervision where I not only learned about statistics, machine learning methods, coding and research methods in general but also developed a love for health care research. Dr. Schuster's in-depth knowledge and passion about applying modern machine learning modeling on big data was truly inspirational and played a cardinal role in completion of this thesis. With his support, my poster got accepted for NAPRAG annual conference-2018 and I was offered CIHR travel award. I truly feel that without his unwavering support, positivity and encouragement, it would have been very hard for me to complete my research and degree.

Secondly, I would like to acknowledge the support and help of Dr. Peter Nugus, who made this project possible with his immense knowledge and ethnographic research back ground in the emergency departments and other hospital and community care settings. His insight and published work on workplace and organizational processes, care coordination, and work culture in complex organizations added to the context and depth of this thesis immensely. His advice about research writing especially helped me in getting through my qualitative course work and in writing this thesis.

I am also grateful to Dr. Bertrand Lebouché for being part of the thesis committee and for his valuable advice and guidance. I would also like to acknowledge the help and support of Ms.

Genevieve Gore, who helped me with the literature search. I would also like to acknowledge the help of Mr. Jamie DeMore and Ms. Sherrie Child for providing administrative support.

I would like to acknowledge the funding and awards that I received during this period including STP- Schuster - CRC/Schuster/X $15,000 (05-OCT-2017), Grad Excellence Award-00220 $2,505.90 (29-AUG-2017), Grad Excellence Award-90029 $2,506.00 (17-JAN-2018), Schuster Grad Award $2,930.00 (05-OCT-2017), Grad Excellence Award-00220 $2,506.00 (31-AUG-2018) and GREAT Award-00220-MD $302 (24-APR-2019), given to me very generously by my supervisor Dr. Tibor Schuster and McGill department of Family Medicine.

I am also very grateful to my friends Asma Aqib, Jaynal Abedin, Lubna Javed and Farah Ahmed, for their endless help and support with my thesis work.

Lastly, my heartfelt gratitude is for my husband, my children, my mother and my mother in law for their unwavering support and love through this journey.

Preface:

This thesis is written in the light of the guidelines set by McGill University for thesis submission at Master's level. It is written in the traditional format.

Dr. Wajiha Masood (student) was primarily responsible for writing the thesis. Dr. Tibor Schuster (Supervisor) and Dr. Peter Nugus (Co-supervisor) supervised and guided Dr. Wajiha Masood in the research conduct, statistical analysis and were responsible for the overall supervision of the study. Dr. Peter Nugus (the lead investigator of a larger research project encompassing the sub study described within this Master's thesis) provided the data for the research analysis after Institutional Review Board (IRB) approval. An extensive literature search was conducted prior to data analysis. The thesis employs three main methodological approaches in a complementary fashion: 1) A literature review on Emergency Department (ED) crowding and increased length of stay of patients in EDs. 2) The use of a quasi-experimental study design to make inference about factors potentially affecting ED visit length of stay. 3) Analysis of electronic medical records (EMR) data reflecting ED processes and patient profiles and identification of factors associated with prolonged ED stays using modern machine learning prediction models.

Chapter 1

INTRODUCTION:

Emergency department (ED) crowding and prolonged patient waiting times are a global problem (1,2,3,4,5,6,7,8). This problem is worsening for the developed world, partly because of a rapidly growing elderly population with chronic illnesses (4,5,6,7,8).

Care Setting

This thesis addresses research questions in the context of care provided by EDs. Contextually, the ED belongs to a sector of care known as the "acute care sector". To clearly distinguish the different care settings applicable or adjacent to the presented research, specific working definitions of the various settings are provided. Acute care is defined as "services including all promotive, preventive, curative, rehabilitative or palliative actions, whether oriented towards individuals or populations, whose primary purpose is to improve health and whose effectiveness largely depends on time-sensitive and, frequently, rapid intervention" (9). The components of acute care are as follows:

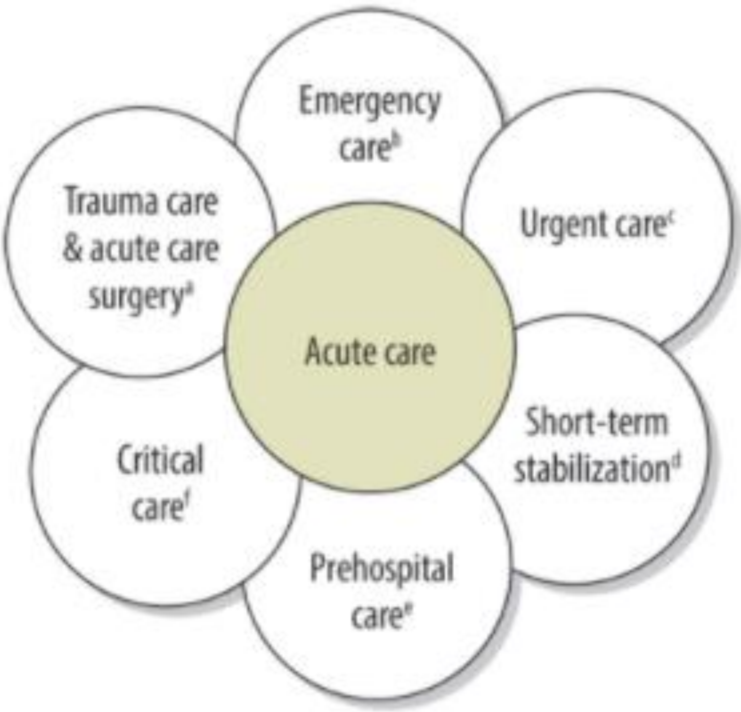
a) Trauma care and acute care surgery: Care of patients with acute surgical needs, for example, acute appendicitis or strangulated hernias (9).
b) Emergency care: Care of patients with acute life/limb-threatening medical and potentially surgical needs, such as acute myocardial infarctions or acute cerebrovascular accidents (9).
c) Urgent care: Ambulatory/outpatient care in a unit delivering medical treatment outside a hospital premises, mostly on an unscheduled, walk-in basis. Examples include evaluation of an injured ankle or evaluation of fever in a child (9).

d) Short term stabilization: Provision of care to individuals with acute needs before starting definitive treatment. Examples include administering intravenous fluids to a seriously injured patient before transfer for surgery (9).

e) Pre-hospital care: Treatment given in the community before the patient's arrival at a formal health-care facility providing definitive care. Examples include care by paramedics (9).

f) Critical care: The care of patients with life-threatening illnesses needing definitive treatment under vigilant observation. Such patients are usually care for in intensive care units (ICUs), and, if ED patients, would be expected to be transferred to an ICU. Examples are patients with respiratory distress requiring intubation and ventilation (9).

Figure 1: Components of Acute Care (10)



Although these aspects of acute care are aligned with ED care, the focus of this thesis is specifically on care provided within the ED.

Significance and Background of Emergency Department Crowding

Despite the central role of EDs as point of contact for general care, patients' waiting times for treatment at EDs has been a significant issue worldwide (1). Prolonged waiting times in the ED are reported to be associated with decreased patient satisfaction and trust (11,12), and decreased quality of care (13). Longer stays at EDs were found to be associated with significant objective clinical outcomes like mortality (14), as well as with major clinical processes of care, such as time to treatment for time-sensitive conditions such as pneumonia.

Longer ED visits have also been shown to be associated with increased costs for care (15,16). Increased ED waiting time has an economic cost for the healthcare system and the patients, because caring for non-emergency cases in ED is more expensive. It places significant economic burden on the public health care system, given that these visits cost 5 to 10 times more than what the cost would have been if those patients had access to a primary care/family doctor (17,18,19). For patients, the hours spent in the ED waiting to see an ED doctor translate into lost productivity and lost wages (20,21,22,23). In Australia, in a large single site ED study, the risk ratio of mortality at 10 days in patients presenting at times of ED crowding was found to be 1.34 (95% CI = 1.04 to 1.72) (24). In another three hospitals, the hazard ratios for mortality at 2, 7, and 30 days, after adjusting for age, diagnosis, referral source, urgency and mode of transport to hospital, were 1.3, 1.3, and 1.2 for patients admitted during crowded periods compared with uncrowded conditions (measured as hospital bed occupancy) (25).

Canadian context:

In general, as ED services are typically urgent or emergent in nature, anticipated waiting times for case consideration are relatively short compared to other non-urgent medical conditions. Therefore, performance measures become more important in the context of ED, as it is expected to affect patient outcomes.

There are different targets for performance measures in ED in different countries. In the US a target of 4-6 hours waiting time has been used (11) compared to the UK where a mandated maximum of 4 hours is length of stay in ED is instituted with the intention of improving the quality of ED care and efficiency. It has been reported in Canada that the waiting time in the ED before being treated is longer than in other developed countries, (1) and was 19% longer than the US (11). The Canadian Institute for Health Information's statistics in 2013–2014 revealed that for more than 10 million reported ED visits, the sample comprising 60% of all ED visits in Canada, nine out of ten ED visits were completed in 7.5 hours or less after adjusting for age. Admitted patients spent nearly five times longer in the ED than non-admitted patients, about nine out of ten waited 21.4 hours or less in the ED for an inpatient bed (26). Canada also has the highest rate for patients' utilization of the ED for various health care needs (4,5,6,7,8). This problem is further complicated by the fact that Canadian patients suffer additionally because of capacity restraint and a relatively limited number of medical and surgical specialists, who are relied upon not only for in-patient admission, but to consult regarding and confirmation, decisions to discharge patients directly from the ED (7) (27).

Provincial context:

Among Canadian provinces, Quebec has the longest mean ED waiting time (20). Prolonged ED waiting times were also reported to lead to frustration among patients, and ED health care providers in a variety of occupational roles (28,29,30). The severity of the problem is reflected by the fact that, a number of cases of patient deaths, while waiting for hours to be seen by a physician in the ED, have been reported in Quebec (28,29,30).

The province of Quebec also ranks the lowest among Canadian provinces with regard to the coverage of the population by a primary care physician (31). It has been shown that continuity of care and having access to a primary care physician, is associated with decreased ED use, especially by the elderly population (21,32). This effect is more visible in the urban areas, in comparison with rural areas (21,32). The decreased use of emergency services has also shown to be associated with reduced health care costs (21,32). Therefore, not only is ED waiting times a significant problem worldwide; the problem is particularly acute in Quebec, making it an ideal setting to explore innovations or interventions to redress this serious threat to the quality and safety of patient care.

Integrated care:

Research on geriatric care, compared across three markedly different healthcare systems, including Canada, showed that, despite considerable expenditure on health services per capita in developed countries, fragmented and uncoordinated care continues to characterize the care of those with multiple needs, especially older people (12). Care coordination has been defined as "the deliberate integration of patient care activities between two or more participants involved in a patient's care to facilitate the appropriate delivery of health care services" (33).

The complexity of chronic illness, and the frequent involvement of multiple specialties and facilities make care coordination, or integration, a foundation of high-quality health care (34).

The ED plays an important role in integrated care. “Integrated care”, the effective interplay between professionals and services providing different functions, in the ED leads to decreased hospitalizations, and reduces in hospital mortality (35). Efficiency is an important component of integrated care. Efficient care is defined as ‘‘avoiding waste, in particular waste of equipment, supplies, ideas, and energy’’ (36). The ED plays a challenging and vital role in integrating different specialized services in the hospital. In addition, its appeal as a provider of unplanned access to medical services ensures it provides a link between the community (primary care) and the hospital (secondary and tertiary care) (37). Better integrated care, as the effective interplay between professionals and services providing different functions, has been shown to decrease hospitalizations, and to reduce in hospital mortality (35).

Factors associated with ED crowding and prolonged ED LOS:

A major reason for crowding is the utilization of ED services to provide care to admitted patients for increased time duration. (38,39). This practice is known as “boarding” (39). Boarding results from inefficient care integration between the departments and capacity restraints in the in-patient department (27). The increasing demand on ED caused by the increase of aging population and boarding of patients results in crowding in the ED, which affects the quality of care (27,40,41).

The increasing demand on ED services is caused partly by the increase of the aging population in the developed countries, and also partly by the under-supply of in-patient beds, under-supply of specialists, and under-supply of primary care services precluding the need for some ED visits, and which would allow more efficient discharge of in-patients, freeing up in-patient beds. Therefore, the ED also received non-urgent ED visits, defined as visits for conditions for which a delay of several hours would not increase the likelihood of an adverse outcome (42).

ED crowding has been shown to affect the quality of care delivered by EDs (40,41). A higher rate of non-urgent ED visits may be an indicator of inadequate primary care access or despite being of emergent nature, preventable, that is, could have been prevented by earlier and timely access to primary care services (42). Elderly people and other patients with inadequate primary care coverage often use the ED as a safety net for non-urgent care, as a point of service, further contributing to the exacerbation of the situation of crowding and prolonged patient waiting times (4,5,6,7,8,32,43).

Therefore, health system efficiency, in the form of effective utilization, and efficient "flow" through the ED to the hospital or for discharge, is an important component of integrated care (27). "ED "flow" is important because the ED has an important connecting function in the broader health system. This puts a premium on LOS in the ED as an indicator of both efficiency and quality of care, precisely because a central role of the ED is to clinically categorize and disperse patients. In other words, the central indicator of not only the efficiency, but also the quality of ED care, is how well patients are moving through it, rather than staying in a holding pattern" (27).

Efficient care is defined as ''avoiding waste, in particular waste of equipment, supplies, ideas, and energy'' (36). ED crowding causes less efficient care that may lead to complications that require longer hospital stays and are more expensive. A significant contributor to ED crowding and resultant prolonged length of stay is so called "boarding" of patients – housing patients who have been formally admitted to an in-patient department, but for whom an in-patient bed is not yet physically available. Boarding results from inefficient care integration between the departments and capacity restraints in the in-patient department.

EMR data and machine learning:

There has been increasing research and policy focus on finding ways to avoid unnecessary ED visits (44). The increasing availability of electronic medical records with sufficient data from millions of patients, along with advancement in modern machine learning provides a unique platform for development of new risk prediction models that outperform standard statistical models. In the context of ED, it might be due to the fact that ED admissions represent complex relationships between predictors and outcome variables, that limits the utility of standard statistical methods (44). Machine learning models have outperformed standard statistical models, especially in settings where clinical data have abundant and more complex relationships between variables (44,45,46).

Motivation for the present research study

Integrated care in the hospital has not yet been adequately researched (47). Given the complexity of integrated care and its relevance to ED crowding and increased length of stay, it is valuable to take an "extreme case" example (48) of care complexity, and discern lessons about integrated care from a major hospital (and ED) restructuring. Accordingly, this inter-disciplinary study examines an immense health service relocation, restructuring and re-organization in Montreal, Quebec, of a measure rarely witnessed. Costing $2.3 billion, having been planned for 17 years, and involving 7,000 staff and up to 2 million patients, four McGill University hospitals and institutes had, from April 2015, been merged into a single facility. Despite the need to integrate complex services for greater efficiency and effectiveness, little is known in management and organizational literature, much less in health services literature, about what constitutes a more or less effective and efficient restructuring to improve the effectiveness and efficiency of health services. Although most research has concentrated on perceptions and activities prior to relocation and/or re-structuring, there is little research to provide guidance on new opportunities for collaboration and work's efficiency provided by the new environment. The research study conducted as my master's thesis project is important because it will provide new information about what organizational factors make ED care complex, and the organizational facilitators and impediments to provide care for complex patients in ED, especially given its important integrative role in facilitating the efficient flow of patients through the ED to the hospital or for discharge home or to another facility.

Considering the complex and ubiquitous problem of ED crowding and prolonged lengths of stay and the need for better integration of care, this study uses the extreme case example of structural changes in a major urban hospital ED pre-and post-relocation to a new site in relative close proximity to the original site. It utilizes the routinely collected EMR data to analyze the association of site-specific and patient level factors, as well as the relocation itself, with ED length of stay.

Conceptual framework:

The most widely accepted conceptual framework used to explain (over)crowding in ED context is the input–throughput–output model (49). Input factors are related to the demand for ED services, throughput factors are related to the ED processes of evaluation and treatment, and output factors are related to ED disposition. Crowding is mostly caused by disruption at points along this process (49). It is pertinent to understand that this model is a good fit for EDs that have adequate capacity for their catchment population. If a catchment population has an ED that does not meet its capacity requirements than crowding is bound to happen even if all the processes in the ED are handled efficiently.

Different strategies have been used to reduce patients' waiting time in the ED with varied results. There is a paucity of information on studies designed to examine the restructuring of hospitals to improve the quality of efficiency of care including reduction in ED LOS, possibly due to the massive and complex nature of such a relocation. The complexities involved make it generally difficult to disentangle effects of relocation (e.g. changes in coverage area & population) and effects of restructuring of health services before and after relocation. Thus, it is not surprising that there seemed to be no empirical research designed to investigate the effect of relocation on the ED efficacy when hospitals relocate and bring different services together.

In theory, studying the effects of relocation (and associated changes in health services and organization) would require a well-designed randomized experiment, i.e. randomizing a large number of comparable health service units to a pseudo study intervention "relocation", and then comparing health service and patient outcomes between units that remained on the same site and those who were relocated. This design would be less prone to time-dependent confounding than naïve pre-post study designs. Unfortunately, such a randomized controlled trial is, in practice, infeasible. However, a natural pre-post design may still offer valuable evidence on the "effect" of relocation if i) the pre and post comparison period can be considered exchangeable with regard to external contemporary factors possibly affecting the primary outcome indicators, ii) if the change of environment (i.e. coverage area and patient population) largely overlap and iii) measurement of relevant outcome indicators and potentially influencing factors remains the same (i.e. no systematic change in EMR data collection and processing).

For this reason, a quasi-experimental (pre-post) design was deemed as appropriate design for studying the effects of ED health service changes and associated changes in ED patient lengths of stays, in the context of the relocation of the Royal Victoria Hospital (RVH) from the original site at Pine avenue to the new site at Glen as part of McGill University Health Center "Super Hospital" in April 2015.

Relevance and the Importance:

This study will help to discern indicators of service efficiency of hospital ED services that are, directly or indirectly associated with structural or organizational changes induced by the relocation of a large Canadian metropolitan ED. The findings of this project will help policy makers, health managers to better understand the problem of prolonged ED stay lengths for alleviation of this problem. It will do this by accomplishing two objectives. The first objective is to determine if there is a difference in LOS before versus following the re-location of this hospital. The second objective is to determine the most important predictor variables that relate to the LOS.

Chapter 2

This chapter of the thesis contains a comprehensive literature review that explores known reasons for ED crowding and prolonged ED LOS, its implications for patients, service providers and the health system as well as suggested solutions.

LITERATURE REVIEW:

ED crowding is a global problem affecting key stakeholders of the health system including patients, health care providers and hospitals (22,23,50). ED crowding is defined as "a situation where the demand for emergency services exceeds the ability to provide care in a reasonable amount of time" (51)

2.1.2 ED crowding severity and the Canadian context:

In the developed world, people living in Canada wait the longest in the ED, before being seen by an ED doctor (4). This observation has been partially attributed to the fact that ED utilization in Canada is significantly higher than in other developed countries, which amplifies crowding (4). The relatively high utilization rate of ED services in Canada is at least partly related to the fact that ED patients in Canada have the highest proportion of ED usage for services that could have been provided by a general physician, family doctor or primary care team (4). This finding signifies that there is considerable room for improvement, with regards to the accessibility of primary care for the Canadian population.

2.1.2.1 *Provincial context*:

Quebec fares the worst in Canada, when it comes to ED crowding and waiting times (20). To provide an overview, a report published by Quebec's health and welfare commissioner in 2016 provided the following statistics; 35 percent of the people in Quebec EDs waited for five or more hours during their last visit (52). This is in contrast to nearly zero percent of the

Dutch, and just 5 percent of the German populations encountering similarly long wait times in the ED (52). In terms of other provinces, neighboring Ontario reported that only 15 percent of the people experienced a similarly long waiting time in the ED (52).

The following statistics convey the extent of the problem of ED crowding and waiting times in Quebec for the year 2015-2016: 1.5 million ED visits in the province (out of 3.2 million annual visits which roughly equals to 45%) maximum wait time prescribed by the Minister of Health and Social Services, amounting to 13 million extra hours spent waiting for care (20). When the average hourly wage in Quebec is taken into consideration, this amounts to over $300M, excluding the additional costs to the health system (20).

The estimated percentage of most urgent cases that are of a relatively urgent or emergent nature is about 40%, while the less urgent cases that could often have been dealt with elsewhere amount to about 60% of the total ED visits (20). The average duration of ED stays for all patients including outpatient and stretcher patients is about 9 hours.

Another factor, as mentioned, is the so-called “boarding” of patients, i.e. the formal stay of a patient at the ED department while awaiting physical transfer to another medical service unit. In Quebec, in university hospitals, the average duration of ED LOS of patients requiring hospitalization varies from 14 to 37 hours, depending on the hospital (20). It is noteworthy that specialist referrals required for some of the ED patients for diagnosis and further management including admission caused the longest delays in emergencies in university hospitals, with 64% of hospitals reporting an average response time of greater than 4 hours for such patients, for an emergency consultation request (20). To understand the severity of the problem it must be taken into account that this is better than the average province-wide waiting time for a specialist consultation in ED, which is greater than 8 hours (20).

2.1.3 ED crowding impact on health care outcomes

ED crowding has numerous adverse consequences that affect patients and health care workers, along with an economic impact resulting from overutilization of ED medical services (22,23,50), where overutilization is defined as unnecessary and avoidable use of ED (22,23,50).

Impact on patients:

The seriousness of the problem of crowding in ED becomes starkly apparent when taking into consideration its association with decreased quality of care and an increase in medical errors. (31). Furthermore, ED crowding has shown to be positively associated with increased mortality for patients present in the ED (31). There is evidence that ED crowding serves as a marker for inadequate care and poor outcomes for ED patients who might require hospitalization (31). Patients also suffer emotionally as they do not find their experience of ED care fulfilling and some leave without being seen (31).

Impact on Staff:

Health care workers are also affected as they experience reduced work satisfaction, leading to decreased productivity (54). One study in the US, examining retrospective chart review data from 2005-2008 at a level 1 trauma center ED, also found an increased threat to staff safety, showing a significant association between ED crowding and violence towards staff (55).

Economic impact:

When EDs are used as a point of primary or general care rather than for emergencies, the associated costs for care increase (15,16). The related economic burden is significant as such ED visits are estimated to cost 5 to 10 times more than what the costs would have been if those patients had accessed a primary care/family doctor instead (34,35,36).

2.1.4 ***Importance of integrated care in decreasing ED LOS:***

The EDs play a challenging and vital role in integrating different specialized services in the hospital, as mentioned in the introduction. Traditionally, integrated care has been viewed as a linear process in ED, but a more comprehensive approach to understanding the working of ED is through complex adaptive systems approach (47) as ED work is fundamentally dynamic, comprising numerous strong links of multiple dynamic agents interacting in real time with "emergent, dispersed and highly decentralized behaviors". (47).

This makes care coordination and integration a key part of an efficient ED. It is pertinent to note that many older people, and other patients, frequently use emergency services for non-urgent care, as a point of service, further exacerbating the situation of crowding and prolonged patient waiting times in ED (4,5,6,7,8). When this is viewed in the context of the complexity of chronic illness and the frequent involvement of multiple specialties and facilities, it makes care coordination, or care integration, a highly important foundation of high-quality health care (34).

Another important component of integrated care is efficiency. An efficient care system strives to reduce waste, including overuse of services, waste of time spent in waiting (for lab results, consults, transfers, bed availability) (34).

Fragmented and less than optimal integration of services results in inadequate care, which may make the disease less manageable and result in avoidable complications requiring urgent/ED care – further promoting ED crowding. Furthermore, ED visits are more expensive to care for, compared to what the cost would have been, if the integrated provision of care was adequate in preventing the complication in the first place (34,36).

A major reason for crowding is the utilization of ED services to provide care to patients boarding in the ED. (38,39). Boarding again puts emphasis on the importance of integrated care as it frequently results from inefficient care coordination between the departments and capacity restraints in the respective in-patient department.

Despite the cardinal value of integrated care in a high-quality health care system, integrated care in the hospital ED has not yet been adequately researched (47). This becomes highly relevant in the case of Quebec where the elderly population with multi morbidity is growing faster than anywhere else in Canada (56,57). This population needs more frequent access to primary care and because of the gaps in primary care in Quebec, this growing population's main point of care becomes the ED (4,32).

It has been shown that continuity of care and having access to primary care physician is associated with decreased ED use and lower costs (4), especially by elderly people, and the difference is more pronounced in the urban areas (32). Furthermore, it has been shown that reducing ED load is possible by extending and broadening the scope of services offered by the primary care (58). The province of Quebec has the lowest rank in terms of coverage of population by primary care doctors (31). Nearly 25 percent of the population in Quebec, and

32 percent of the population, in the city of Montreal alone, are without a family doctor (31). Due to this, a significant portion of the population starts treating the hospital ED as the point of care, rather than using it as a unit for acute emergent care. This is reinforced by studies from patients' perspective that poor accessibility to and continuity of primary care, leads them to use ED for their non-acute care needs (60,61).

The prevalent utilization of ED for non-acute care partly explains why more than half of the consultations in the ED are for non-urgent (62). This has been put forward as one of the main reasons for crowding (17,18,19). The impact on ED operations of both the primary and acute care sectors underscores the important role of the ED in providing integrated care

2.2 Reasons behind ED crowding

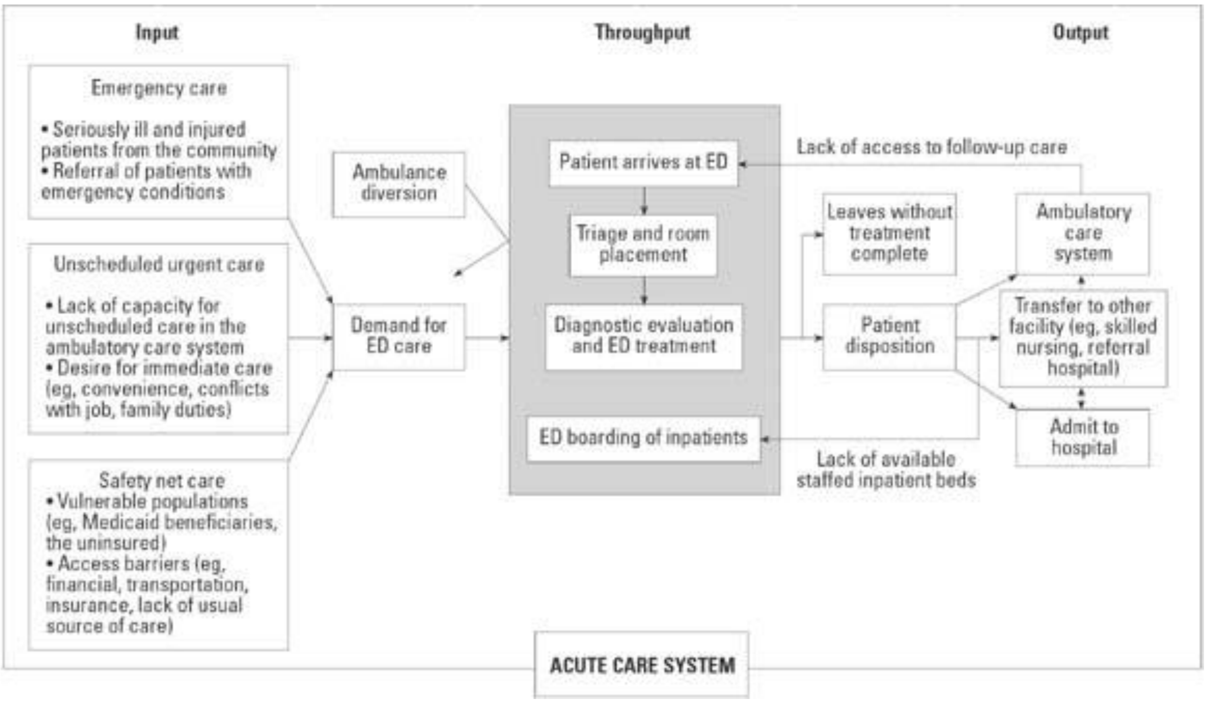
Crowding in the ED reflects problems elsewhere in the healthcare model. A health care model or model of care is defined as the manner in which health services are delivered". As such, it is a way to "outline [*sic*] best practice care and services for a person, population group or patient cohort as they progress through the stages of a condition, injury or event" (63). To understand the reasons for crowding in the hospital ED, the input-throughput-output model is routinely used, which is explained in the text below.

2.2.1 ED crowding input throughput output model

Overutilization of the ED reflects problems elsewhere in the system with regards to accessibility to primary health services for urgent needs (64,65). To understand the complex problem of ED crowding, this study uses the input –throughput-output model (49). Input factors are related to the demand for ED services, throughput factors are related to the ED processes of evaluation and treatment, and output factors are related to ED disposition. ED crowding is caused by disruption at points along this process. (66,67,68,69,70).

Figure 2 below explains in detail the following: 1) Input factors including Emergency care, unscheduled/urgent care, and safety net care, 2) Throughput factors involving patient arrival, triage, diagnostic workup/treatment, and boarding of patients, 3) Output factors involving patient disposition to admission in hospital, ambulatory care system and transfer to other facility,

Figure 2 The input-throughput-output conceptual model of ED crowding (67)
(source: Annals of Emergency Medicine 2003 42173-180DOI: (10.1067/mem.2003.302))



Input factors entail the reasons for patient inflow to the ED, for example, non-urgent visits due to lack of access to primary care provider, or due to fragmentation of services for elderly patients suffering multi morbidity (66). Crowding occurs as a result of continuous increase in patient volumes and subsequent longer ED stays (64,65). Contributing factors include inadequate access to primary healthcare and fragmented care (71,72,73)

Throughput factors:

Throughput factors are related to capacity/operations including operational bottlenecks and inefficiencies within the ED, for example, suboptimal staffing, or space constraints (66). It is also related to inability of healthcare system broadly to train and retain ED staff to deal with increasing patient loads (71,72,73).

Output factors:

Output factors are indicators of bottlenecks in other departments of the hospital – or the health care system broadly – that impact the ED workflow, for example, bed availability within inpatient departments and transfers to those departments (66,68,74).

There has been an increased interest by the public and the government on the issue of bed boarding or boarding of patients in ED which results from delays in transfer of admitted patients out of the ED (66). This interest is driven by the fact that it is now recognized that bed boarding causes disruptions in ED flow and efficiency, and is associated with adverse outcomes for the patients, besides being one of the chief reasons behind ED crowding and prolonged LOS (66, 68, 74,75,76)

2.3 Possible solutions:

It is important to understand that a "one size fits all" solution is not likely to be successful as the reasons for crowding are varied and contextually unique to the environment in which the crowding occurs; therefore, the solution need to be innovatively designed case by case (77). Every ED has its own unique input out throughput characteristics that need to be taken into account when considering a dynamic solution to ED crowding and prolonged LOS.

A recent systematic review analyzing ED crowding compiled a list of proposed adaptations to ease patient load in ED and manage crowding, as discussed below (77).

Input solutions: GP (general practitioner) led walk in care centers, co-located GP in the ED, extended GP opening hours, social interventions including awareness campaigns and redirection to alternative facilities for less urgent presentations (77).

Throughput solutions: Split ESI (emergency severity index) 3 on presentation, earlier physical assessment, including physician supported triage, separate fast track/flexible care area for patients with lower urgency conditions, shorter turnaround time for laboratory tests/diagnostics, ED nurse flow coordinator, nurse initiated protocols, earlier inpatient consultation by the inpatient doctors, increased ED bed numbers, increased ED staff (77).

Output solutions: Active bed management, leadership program/support, implementation of nationally mandated, timed patient disposition targets, ED staff direct admit rights, admitting team prioritize ED admissions, alternative admission policies, increased inpatient beds and staff (77).

The findings of this recent systematic review suggest that there is an incompatibility between the causes of crowding identified through research and solutions that were implemented to address them. There is a need to focus on which interventions worked in which contexts, keeping in mind the distinct causes of crowding pertinent to that environment (77).
It was also identified in the systematic review that elderly patients with multi morbidity represent an increasingly significant contributor of ED crowding. Thus, comprehensive,

continuous, efficient and coordinated primary care is an increasingly essential part of a high-quality health care system (77).

Evidence suggests that the problem and solutions of ED crowding lie mostly outside the ED (inadequate access to primary care, shortage of in patients' beds resulting in boarding) and therefore the solutions, need to tackle the whole of system for finding acceptable, feasible and sustainable solutions to ED congestion (77). Patients with chronic illnesses requiring coordinated ongoing care have been found to receive better and more comprehensive care in countries with strong primary care networks. However, fragmentation in transitional care for patients after being discharged from the hospital remains a problem in all countries (6).

2.4 EMR data and machine learning:

A number of studies have demonstrated that machine learning techniques are superior to conventional statistical methods and clinical assessment tools, and highly promising for predicting the risk of hospital readmission (78,79), recurrence of ED visits (80), or risk of mortality form sepsis in ED patients (81). It has been observed that machine learning models outperform conventional statistical models by learning from the data, with minimal need for variable transformation or model structure as they are able to automatically grasp and learn from a priori unknown interactions present in the data (44).

Another advantage of utilizing electronic health record data through machine learning modeling is that it can be updated as new local real-world data is acquired, ensuring the results remain relevant, with better applicability of findings (82). Perhaps, the ED environment has the most pressing need in the health care system for utilization of such methods to reduce the uncertainty (83).

A key point that can be learned from local "big data" through machine learning, is how to differentiate between a generalizable method of using machine learning to generate locally relevant results using local data versus a generalizable predictor model of an outcome that was modelled using a cohort data that might not be a true representative for other set ups (81). The machine learning method enables local health care systems to generate their own predictor model that sensitive and relevant to their local population, based on the data collected in their system, given each health care system has a unique set of population with its unique phenotypic expression of disease (81). This may serve as a powerful tool in improving and enhancing patient care (81).

2.5 Motivation for the present study:

The motivation for this study is to understand the complexity of the ED crowding problem and its association with primary care coverage, by using an extreme case example (48). The study utilizes a rare phenomenon of hospital relocation and restructuring, analyzing EMR data collected before and after the move to assess ED LOS, using modern machine learning modeling to explore the association of covariates and ED LOS and their interactive dynamic with each other.

Specific objectives

Primary objective:

1) To evaluate if there is a difference in patients' ED length of stay before versus after the move of the RVH ED from the original site to the new site?

Secondary objective:

2) To explore the most important predictor variables that relate to the length of stay at the ED including site (relocation of hospital ED), age and other socioeconomic variables, and level of access to primary care physician (co primary study end point)

Research Questions:

1. Is there a difference in patients' expected ED length of stay before versus after relocation of the ED of the Montreal Royal Victoria Hospital from its original site to a new site in 5 kilometers proximity?

2.What are the statistically most important factors associated with the length of stay at the ED including the site (relocation of hospital ED), demographic variables sex and gender, ED arrival and triage classification as well as variables measuring access to a primary health care provider?

Chapter 3

Study Design and Research Methodology

In the following chapter, the research methodology including the quasi-experimental study design applied in my master's project are outlined and justified.

3.1 Study Design

Major hospital restructuring and relocation represents a natural pre-post experiment which allows, under specific assumptions, the assessment of the effects of factors altered by the respective structural and organizational change on patients' length of stay in the ED, using routinely collected data.

The specific assumptions necessary for justifying the utility of a pre-post design are:

i) The pre and post comparison period can be considered 'exchangeable', i.e. no major external policy changes (globally and locally) have occurred alongside the transition and relocation of the ED, implying that pre and post measurement period are comparable except for possible (expected) ED-internal differences.
ii) The change of environment (i.e. coverage area and patient population) largely overlap for both original and relocation site
iii) Measurement of relevant outcome indicators and potentially influencing factors remains the same (i.e. no change in EMR data collection and processing).

The overarching research question addresses a question of impact, in particular, what is the impact of relocating hospital ED services on general health services performance and in particular, the expected length of ED stays. In order to make inference about such an impact, ideally, a randomized controlled trial would be conducted where confounding biases would be minimized. However, due to the natural constraints of the research setting, an RCT was not feasible, and hence the study was conducted using a quasi-experimental pre-post design. This design was deemed appropriate and cost effective for answering the research questions. The design utilizes the natural arrangement, in which the EMR data were routinely collected, reflecting the structural and organizational changes associated with the ED site relocation. The study uses longitudinal EMR data, routinely collected for ED patient populations under both location settings.

3.2 ED Location Settings:

The Royal Victoria Hospital (RVH) is a major university hospital in Montreal, Quebec, Canada. It is affiliated with McGill University. It was completed in 1893, at the original location (old site) at 687 Pine Avenue and was relocated to the so called "Glen site" and merged into a "super hospital" (a conglomerate of 4 hospitals and their institutions) on 26th of April 2015. The RVH ED has been one of the most frequently accessed EDs in the city of Montreal, pre-and post-relocation with more than 30,000 ED patients per year.

Table A below shows the number of patient visits over the duration of study. The number of patient visits was comparable.

Table A: patient visit details at the two ED sites

Site	No. of patient visits	Timeline	No. of days	Average no. of patient visits /day
Original site	32424	01/06/2014-25/04/2015	328	99
Current site/New site	36506	26/04/2015-31/03/2016	340	107

3.3 Study Data Source:

The primary data source of this study were emergency medical records (EMRs) of the Royal RVH ED. The EMRs contained routinely collected data for all patients who register in the ED between 1st of June 2014 till 31st of March 2016.

Prior to data acquisition, lead researcher, Dr. Nugus, consulted with the Quality, Patient safety, and Performance Department, and the Institution Review Board (IRB) to seek ethical clearance in order to obtain permission from Medical Records Professional Services of the McGill University Health Center (MUHC). After IRB clearance was obtained, the de-identified data was provided to Dr. Nugus in the form of a Microsoft Excel file. The data were held securely on a password-protected computer in the department of Family Medicine, to which Dr. Nugus and Dr. Schuster and I had access.

The data were derived from "Med-urge", an ED information system (EDIS), the primary software that recorded and stored the EMR data, and is used in this study. Med-urge is a patient tracking system, covering patient demographics, clinical information, admissions, transfers, and discharges. It includes precise time tracking information entered by clinical staff in real time. Med-urge reduces data entry duplication and improves data quality.

3.4. Data extraction and pre-processing:

The study included a selection of variables from ED EMR data of all patients who presented at the ED of the Royal Victoria Hospital from 1st of June 2014 till 31st of March 2016. Double entries under the same unique medical visit number were excluded from the study database. The final study database included eight sub datasets (triage, stretchers, demographics, pec (French “prise en charge”: person in charge), consultations, tests, diagnosis and admission). In addition, records for each of the sub datasets were stored in for different file sets (A to D), covering different study periods:

A: 01-06-2012 to 31-05-2013 (1 year) B: 01-06-2013 to 31-03-2014 (10 months)

C: 01-06-2014 to 31-05-2015 (1 year)

D: 01-06-2015 to 31-03-2016 (10 months).

Table B provides an overview of the eight sub datasets and their respective variables. The variables in red were included in the data analysis. These variables were selected for analysis as they were baseline ED admission characteristics which were of main interest for the study as the aim was to predict ED stay time upon arrival.

Table B: overview of 8 sub datasets and their variables

Triage	Stretchers	Demographics	PEC (prise en charge/person in charge)	Consultations	Tests	Diagnosis	Admission
Dossier	Dossier	Dossier	Dossier	Dossier	E l c	Dossier	Dossier
Visit number	Visit number	Visit number	Visit number	Visit number	E l c desc	Visit number	Visit number
Episode creation date	Episode creation date	Episode creation date	Episode creation date	Episode creation date	Start date	Episode creation date	Episode creation date
Episode start date	Episode start date	Episode start date	Episode start date	Episode start date		Episode start date	Episode start date
Episode end date	Episode end date	Episode end date	Episode end date	Episode end date		Episode end date	Episode end date
Triage diagnosis	Stretcher code	Length of stay	Datetime of PEC	Consultation request date		Diagnosis	Hospitalization request date
Triage priority	Stretcher code desc	Age	MD of PEC	Requester code		Diagnosis desc	MD requesting
Triage priority desc	Stretcher start date	Gender	MD of PEC desc	Requester desc		Diagnosis Type	MD requesting desc
First triage desc	Stretcher end date	Gender desc	MD specialization of PEC	Consultation code		Diagnosis Type desc	Requested service
Last triage desc	Time on stretcher	Postal code	MD specialization of PEC desc	Consultation code desc			Requested service desc
Triage start date		Primary care md desc	First PEC	Consulting MD code			Request cancellation date
Triage end date		Referring md desc	Last PEC	Consulting MD desc			Reason of cancellation
		Arrival mode		Start date			Reason of cancellation desc
		Arrival mode desc		Completion date			Allocated bed
		Provenance		Completion status			Allocated bed date
		Provenance desc		Completion status desc			
		Code of origin institution					
		Code of origin institution desc					
		Visit reason					
		Visit reason desc					
		Orientation					
		Orientation desc					
		Institution of transfert					
		Institution of transfert desc					

Handling of data redundancies:

There was an overlap in data recordings provided in file C. The date of relocation of the RVH to the new site was 26th of April 2015. To separate the data of the two sites, the data were arranged by date in file C and then the new site/current site timeline data (data entries after 25th of April 2015) were cut from file C (creating a new file CA) and pasted to the file D (creating a new file DA). The revised files CA and DA were re-checked and had the following new timelines, reflecting the change in location of the hospital:

CA: 01-06-2014 to 25-04-2015 (10 months and 25 days=328 days, N=32424)

DA: 26-04-2015 to 31-03-2016 (11 months and 5 days=340 days, N=36506

The revised CA and DA files were selected for the analyses, for the following reasons:

1) To keep the timelines of data collection as similar and as close together as possible: the Quebec government has been making a series of changes to reduce ED waiting over the years (39). Furthermore, an older data set for one site might be more likely to show longer durations of ED stays, introducing time-dependent bias, as the government policies were gradually implemented over time.

2) To keep the sample size as similar as possible: Some of the files had data entry issues (non-uniformity of data entries). For example, in the triage data set, some records showed two different entries for the same unique patient visit/dossier number, based on attendance. As the data set comprised a large number of entries, manually resolving such data issues would have been very time consuming. Moreover, the observed inconsistencies in some of the sub data files were not directly related to the objectives of the study. Therefore, the data

containing variables of interest relating to the objectives of the study from the datasets were extracted and used in the analysis.

Variable selection:

From the 25 variables available in the sub dataset "Demographics", the following 6 variables as shown in table C were considered relevant for the statistical analysis:

Table C: Variables selected from demographic data set

	VARIABLE	UNIT	COMMENT/DESCRIPTION
1	Age	Years	
2	Gender	Male/Female/Others	
3	Length of Stay	Hours	Starting from the time of patient registration till the patient physically left the ED
4	Primary Care MD Description	Yes/No	Primary care doctor coverage, as reported by the patient
5	Referring MD description	Yes/No	Yes, if patient referred by an MD (referral letter?)
6	Arrival Mode description	Ambulance, Walking, Police, Armored Vehicle, Heli, Others	Based on how the patient presented to ED. The patients arriving via ambulance were triaged first and then registered, other patients were first registered and then triaged.
7.	Visit reason description	Chest pain, Shortness of breath, Abdominal pain, Generalized weakness, Minor complaints (+/- Lower limb complaints)	The patient visits were broadly classified into categories, based on their chief presenting complaint

From the sub dataset "Triage", the variable Triage Priority was selected for analysis and merged with the respective demographic data using the dossier and visit number as key variable. Table D below provides the details of variable

Table D: Description of Variable Triage priority

Triage Priority	Level	Time description
1	Category 0	dead on arrival
2	Category 1	Triage priority 1 minute
3	Category 2	Triage priority 15 minutes
4	Category 3	Triage priority 30 minutes
5	Category 4	Triage priority 60 minutes
6	Category 5	Triage priority 120 minutes

3.5. Statistical Analysis

Statistical Software

Statistical analyses were performed using the statistical software package R (version 3.6.2) (84).

Descriptive Analyses

Standard descriptive statistics were provided by study group (site) and the overall study population to appropriately describe variable distributions. Frequency distributions (counts and %) were provided for categorical data. For numerical data, the mean, median, standard deviation, inter quartile range as well as range (minimum, maximum) were reported.

Supplementary analysis:

Even though this was not part of the study objectives, in view of the complex and worsening ED crowding and increased ED LOS particularly in Quebec, some contingency tables were explored. This included comparison of patients with or without access to primary care doctor and presentation of minor complaints, comparison of patients with or without an MD referral and triage category at presentation, comparison of patients' presentation description and triage category.

Analysis of the primary study objective

The primary outcome variable of the study was the ED length of stay (LOS). LOS was calculated using the total waiting time in the ED, and defined as the period between the recorded time of patient registration and the recorded time the patient left the ED (physically transferred to the inpatient department, transferred to another institution, discharged home, or left without being seen (LWBS). The primary exposure variable of the study was site i.e. “original site” versus “new site” after relocation). As the primary outcome variable is a measurement of duration, the variable distribution was strictly positive and right-skewed. For inferential purposes, the naïve (univariate) difference in mean ED stay times (between sites) were calculated with 95% confidence interval. In addition, to account for the skewness of the data, a non-parametric difference in location parameters (median of the differences between repeated random samples from both sites) and a respective 95% confidence interval was calculated following the method proposed by Hollander and Wolfe, 1973 (85). In order to account for potential confounding covariates in the inferential analysis, a log-linear model for ED stay time was fitted including the variables site, age, gender, primary care MD, referring MD, arrival mode and triage priority. The model also included a three-way interaction between age, gender and access to a primary care provider to account for the different implied frailties. The effect parameter of interest was the difference in logarithmic ED stay times between the two sites. Through back-transformation of this parameter estimate using the exponential function, the model-based expected relative difference in ED stay times was extracted with 95% confidence interval. As the sample size for this study was relatively large, statistical testing was deemed non-informative and the study results were primarily based on confidence intervals and their consistency with clinically meaningful site differences in ED stay times.

Analysis of the secondary study objective:

To explore the statistically most important predictor variables for the primary outcome variable ED LOS, the machine learning methods random forests and regression trees were applied (86). In contrast to conventional statistical (regression) modeling, machine learning methods rely less on manual model specifications. Instead, a large space of possible explanatory black-box models is explored using the data in hand, and the resulting model or prediction algorithm is the one that has demonstrated to perform best on (typically repeatedly drawn) independent sub samples of the data. Put in simple words: machine learning means 'learning about the best possible explanation (prediction) of the outcome, using the data in hand'.

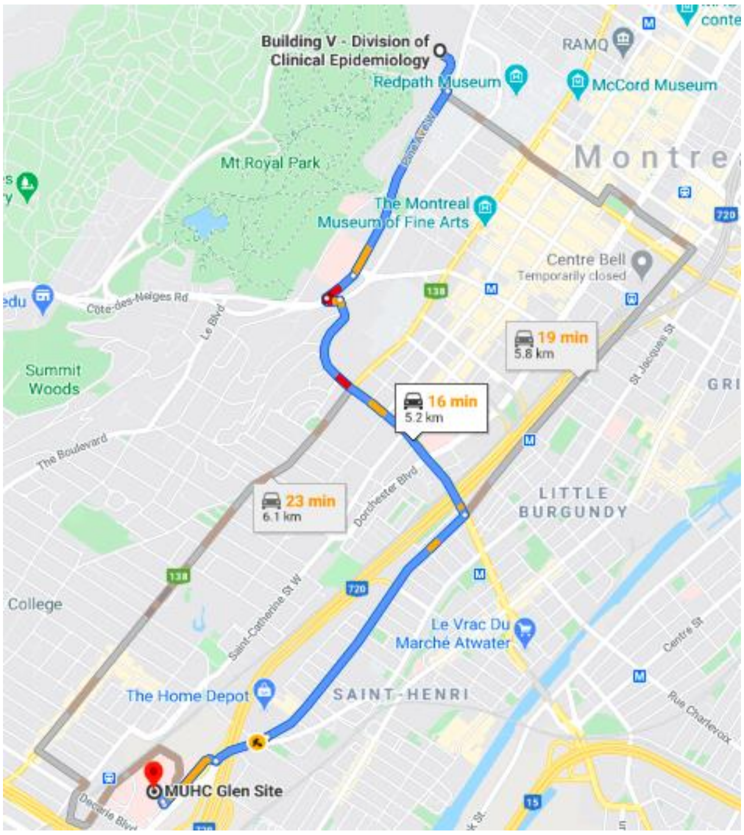
Random forests are a popular subclass of machine learning and imply repeated sampling of the available data and repeatedly fitting of so-called regression trees to each data sample. A regression tree is a linear regression model with high-level interaction terms, also implying optimal cut-off search for continuous or discrete numerical predictor variables to build respective interaction levels. Within a random forest, hundreds of regression trees are fitted, and the utility of each variable used in the forest is assessed through its aggregated (average) explanatory value. More specifically, to measure the statistical importance of a single predictor variables, the percentage increase in the mean square error (MSE) of the random forest model associated with permuting (i.e. basically "dropping") the respective variable is computed. Variables that lead to a larger increase of MSE are more important in explaining the outcome of interest than variables that lead only to small or no increases of the MSE once they are permuted. Conventional regression models have only limited capability to assess variable importance as the interpretation of regression coefficients strongly depends on independence assumptions (no multicollinearity) that are often violated. Furthermore, standard regression models are sensitive to misspecification and do typically not achieve the level of prediction accuracy of machine learning methods (86).

4. RESULTS:

4.1 Relocation and context of earlier research

As stated, the re-location (also hereafter referred to as "the move") of the Royal Victoria Hospital from the original site at Pine Avenue to the new site at the Glen location in Montreal, Quebec, Canada, was planned for 17 years and cost $2.3 billion. It involved four McGill University hospitals and institutes, 7,000 staff and up to 2 million patients, that were merged into a single "super hospital" in April 2015. The RVH moved from a downtown proximity to a nearby location 5.2 kilometers away (please see the attached map in Figure 3).

Figure 3: Map of Montreal showing the distance between the original site at Pine avenue and the new location at "Glen", Montreal, Quebec, Canada (88).



4.2 Population characteristics catchment area original vs. new site

Since the hospital moved to a nearby location, a major shift in catchment population is not expected. However, it is interesting to broadly note the differences in the median income of population in 2015, surrounding the original site and the new site. The attached heat map of Montreal in figure 4 a and 4 b display the median income for 2015, and does not show any major disparity in the household income, indicating that the population within the catchment area between the two sites did not differ significantly, when median household income is taken into account.

Figure 4A: Median household income in Montreal 2015 (89)
(source: http://neighbourhoodchange.ca/documents/2018/02/income-map-montreal-2015.pdf)

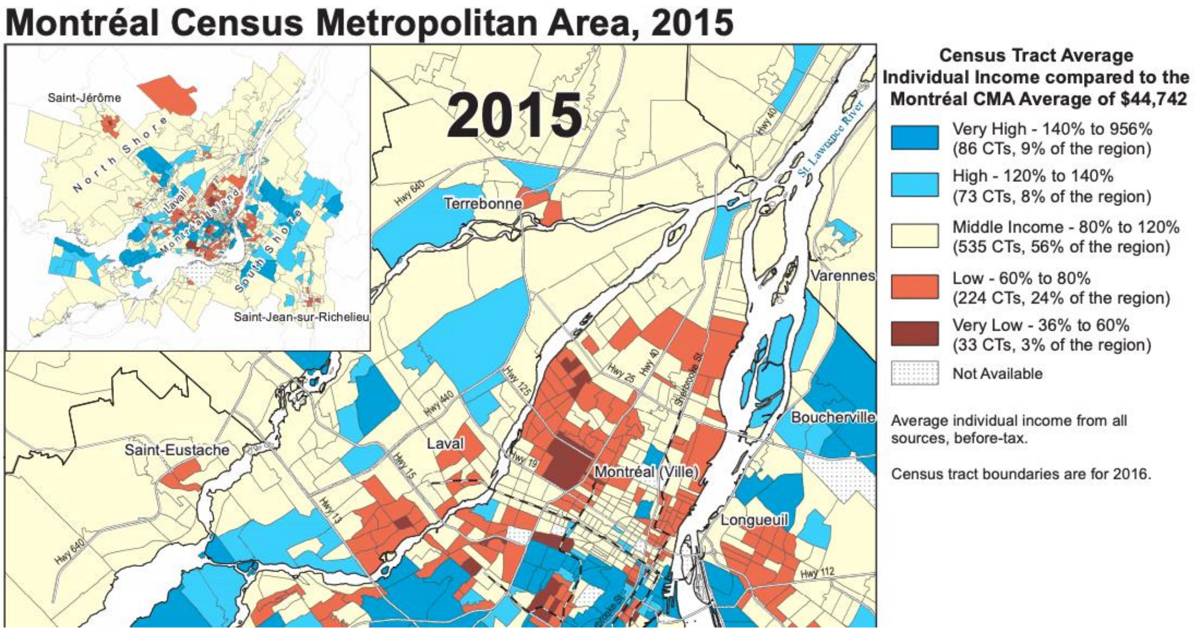
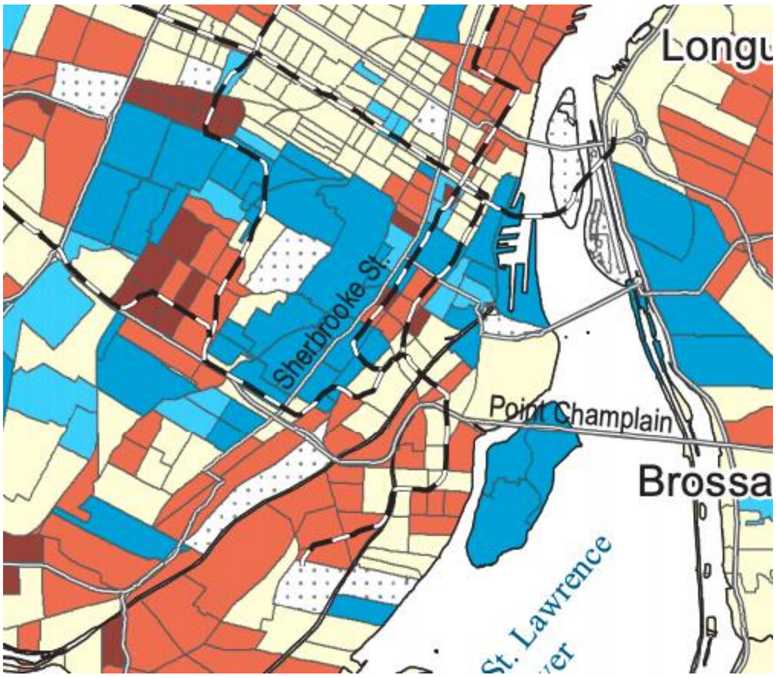


Figure 4B: Median household income 2015 (zoomed in view) (89)
(source: http://neighbourhoodchange.ca/documents/2018/02/income-map-montreal-2015.pdf)



4.3. ***Structural changes associated with the relocation of RVH ED***

Qualitative results from earlier research done at Master's thesis level, on the topic "the coordination of flow across practice boundaries: The collaborative work of Emergency and Internal Medicine" (90) has identified the following changes in the structure of RVH ED and Internal Medicine department after the move;

1) Departmental changes: The Geriatrics department stopped admitting patients at the new site.

2) Capacity changes: The Internal medicine beds were reduced by 109 beds at the new site.

3) Shift changes: The internal medicine (IM) doctors have changed the way they organize their shifts in the ED. At the old site, every day a new IM team would look after ED patients; however, at the new site the scheduled IM team was designated to look after the ED patients for the entire week.

4) Changes in Access: The new site gave metro/train access to the patients, unlike the old site, and a higher percentage of patients started presenting to the ED as walk in patients. The old site had a higher number of patients accessing through the ambulance service.

4.4 Descriptive Comparison of the ED patient populations before and after relocation

4.4.1 Baseline characteristics and arrival data

With the help of the available EMR data, I was able to retrieve information of ED cases of the RVH from the original site (1/06/2014-25/04/2015) and at the new site (26/04/2015-31/03/2016) which naturally formed the two primary comparison groups within this study. The description of the study population is shown in Table 1. The number of observations for the original site is 32424 and for the new site is 36506.

As explained in the methodology section, the number of observations is higher for the current site as has 340 days of observation (107 patient visits per day) compared to 328 days of observations from the original site (99 patient visits per day). Hence, there was about a 10% increase in patient visits at the new site, in comparison to the original site.

The gender distribution was similar at both sites, comprising 54% of females and 46% males. The age distribution was also comparable with the median age of 48 years and mean age of 49 for the original site and median age of 47 years and mean age of 49 for the new/current site.

There was a slight difference evident in the primary care coverage of patients visiting the two sites. Only 17 % of the patient visits had recorded access to primary care coverage at the original site, which increased to 24% of patient visits with recorded primary care coverage at the new site. The patient visits comprising an MD (medical doctor) referral (that is being referred by a doctor to the ED) showed similar proportions. About 53% of the visits at the original site had been referred by a doctor. The percentage of referrals at the new site was similar with about 55% of the visits being referred.

The arrival mode of the patients at both the sites revealed walking (self-presentation) as the most common mode, comprising 66.2% of the patients at the original site and 75% of the patients at the new site. This was followed by arrival through an ambulance, through which 32% of the patients were brought in at the original site and 23% at the new site. Together, these two modes comprised approximately 98% of the visits to the ED at both the sites. Other modes of arrival with a negligible percentage included arrival accompanied by police, armored car (brought from prison/jail) and rescue/retrieval helicopter, respectively.

Triage priority of the patient visits was also comparable at the two sites. The highest number of patient visits belonged to triage category 3, approximately 38% at both the sites. This was followed by triage category 4 (34% at the original site and 35% at the new site) at both sites, followed by triage category 5, comprising approximately 12% at both the sites, and then category 2, comprising 15% of the total visits at the original site and 14% at the new site, followed by category 1 at both the sites with approximately similar percentages.

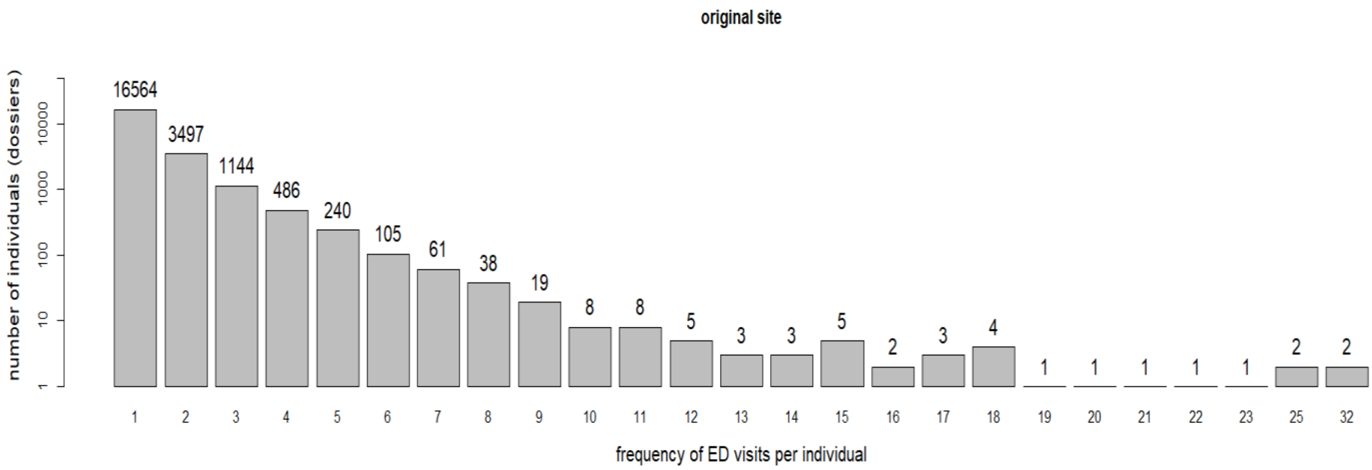
TABLE 1 – Demographics and arrival characteristics of patients admitted to the ED at the original site (01/06/2014-25/04/2015) and new site (26/04/2015-31/03/2016)

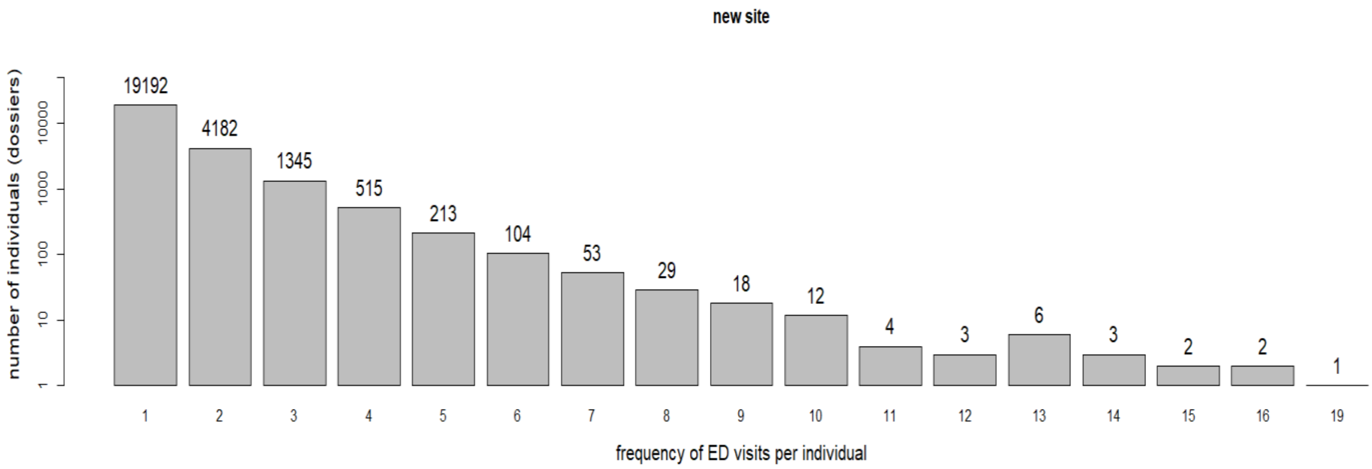
	New site **(N=36506)**	**Original site** **(N=32424)**	**Overall** **(N=68930)**
Age (years)			
Mean (SD)	48.6 (20.0)	49.4 (21.3)	49.0 (20.6)
Median [IQR]	47.0 [32.0, 64.0]	48.0 [31.0, 66.0]	47.0 [31.0, 65.0]
Min, Max	0, 116	0, 115	0, 116
Gender			
Female	19771 (54.2%)	17708 (54.6%)	37479 (54.4%)
Male	16644 (45.6%)	14637 (45.1%)	31281 (45.4%)
Others	91 (0.2%)	79 (0.2%)	170 (0.2%)
Has a primary care provider			
Missing	367 (1.0%)	182 (0.6%)	549 (0.8%)
No	27379 (75.0%)	26779 (82.6%)	54158 (78.6%)
Yes	8760 (24.0%)	5463 (16.8%)	14223 (20.6%)
ED visit with referral			
Missing	365 (1.0%)	182 (0.6%)	547 (0.8%)
No	16083 (44.1%)	14976 (46.2%)	31059 (45.1%)
Yes	20058 (54.9%)	17266 (53.3%)	37324 (54.1%)
Arrival mode to ED			
Ambulance	8389 (23.0%)	10512 (32.4%)	18901 (27.4%)
Walking	27486 (75.3%)	21471 (66.2%)	48957 (71.0%)
Police	13 (0.0%)	16 (0.0%)	29 (0.0%)
Helicopter	0 (0%)	1 (0.0%)	1 (0.0%)
Other	617 (1.7%)	421 (1.3%)	1038 (1.5%)
Armored car	1 (0.0%)	3 (0.0%)	4 (0.0%)
Triage Priority			
0	9 (0.0%)	11 (0.0%)	20 (0.0%)
1	395 (1.1%)	352 (1.1%)	747 (1.1%)
2	4991 (13.7%)	4785 (14.8%)	9776 (14.2%)
3	13769 (37.7%)	12362 (38.1%)	26131 (37.9%)
4	12798 (35.1%)	11046 (34.1%)	23844 (34.6%)
5	4544 (12.4%)	3868 (11.9%)	8412 (12.2%)

4.4.2. Frequency of recurrent ED admissions

The frequency of ED admissions per patients' dossier (unique medical record ID) was determined by linking the admission numbers under the respective record number (frequency of dossiers at the original site n=22204 vs frequency of dossiers at the new site n=25684). The site-specific frequency distributions are displayed in figure 5. The graphs show that 75% of the patients only presented once during the study duration at both the sites. Approximately 16% of the patients presented twice at both the sites. About 9% of the patients had more than two visits to ED recorded during the study duration at both the sites.

Figure 5 - describing number of ED visits per individual (EMR dossier)





4.4.3 Difference in ED length of stay (LOS)

The results regarding the primary study endpoint – the difference in ED LOS before and after relocation of the ED department (old vs new site) – are displayed in table 2. In the univariate analysis (no adjustment for covariates), there was an estimated mean difference of 0.55 hours (95% CI -0.79; -0.30 hours) in the LOS between the two sites. In other words, the original site has 0.55 hours longer mean LOS compared to the new site. The univariate median difference between the site-specific LOS was estimated 0.13 hours (95% CI 0.05; 0.2) in favor of the original site. In the multivariable analysis adjusting for the potential confounding covariates site, age, gender, primary care md description, referring MD description, arrival mode description, triage priority; the mean ratio of ED LOS (new site / original site) was estimated to be 0.98 [95%CI:0.97 to 1.00], i.e. suggesting relative differences in ED LOS between sites of between zero and 3%.

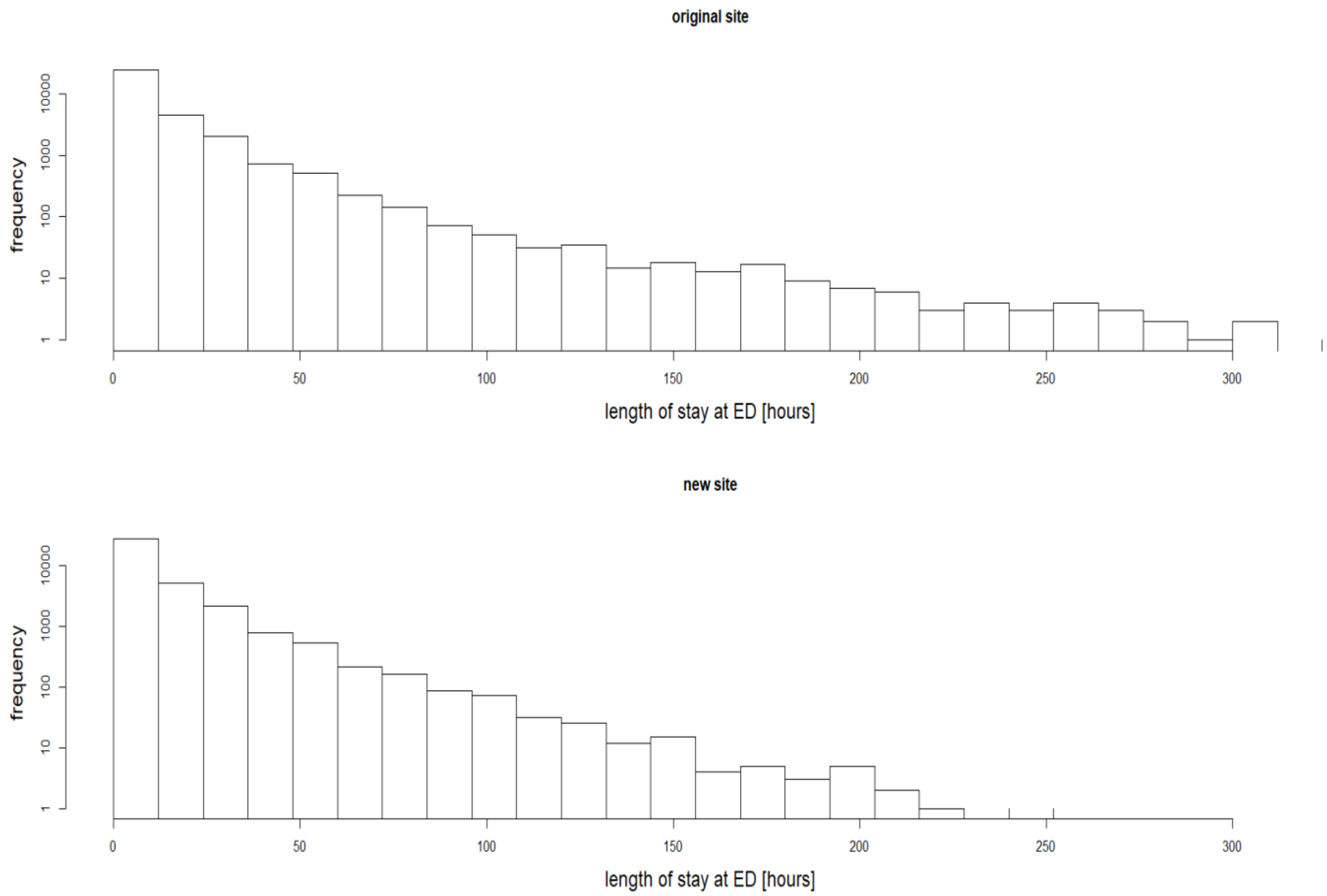
Table 2: ED length of stay (LOS) in hours by site and estimated univariate / adjusted differences in ED LOS parameters with 95% confidence intervals

	new site (N=36506)	original site (N=32424)	Overall (N=68930)
Length of stay [hours]			
Mean (SD)	11.5 (15.5)	12.0 (17.6)	11.7 (16.5)
Median [IQR]	7.02 [4.10, 12.0]	6.65 [3.87, 12.6]	6.85 [3.98, 12.2]
Min, Max	0, 858	0, 461	0, 858
Univariate difference in means	**-0.55**	**[-0.79; -0.30]**	
Univariate difference in location*:	**+0.13**	**[+0.05; +0.20]**	
Adjusted** mean ratio of ED LOS (new / original)	**0.98**	**[0.97; 1.00]**	

*a nonparametric estimator (and 95% confidence interval) for the median difference in ED LOS sampled from both sites (87)
** log-linear regression analysis adjusting for the variables: site, age, gender, having a primary care provider (M.D.): yes/no, having a referring doctor (M.D.): yes/no, arrival mode to ED and triage priority.

In figure 6, the histograms of ED lengths of stay hours are displayed for both sites. It is important to note that the y-axis (frequency) is log-scaled; hence the tails of the distributions showing low counts have actually much less weight in the analysis as visually suggested by the figure.

Figure 6 - Histogram of length of stay in hours by site (frequency axis log-scaled)



4.4.4 ED visit reason

Reasons for ED visits encompassed a long list, attached in the appendix of this thesis (see appendix B). The most common visit reason both the sites was abdominal pain, with 9% of patient visits falling under it at original site, and 10% at the new site. This was followed by shortness of breath, comprising 5% and 6 % of the patient visits, respectively. Minor complaints comprised the third most common visit reason to ED, with 4% and 5%, respectively. The fourth most common ED visit reason was chest pain with cardiac features comprising 4% of patient visits at each site.

4.4.4.1 ED visit reason and Triage category

A contingency table was generated to display frequency of triage categories by ED visit reason. Table 3 provides details of the 4 most common ED visit reasons versus the triage priority. According to this table, 66% of the patients presenting with abdominal pain belonged to triage category 3 or lower, a similar percentage was observed for presentations that indicated shortness of breath (63% <= triage category 3). About 98% of the presentations for chest pain with cardiac features belonged to triage category 3 or lower. Regarding presentations having minor complaints, only 17% of the presentations belonged to triage category 3.

Table 3: Contingency table between triage category and most common ED visit reasons

	Triage category 0	Triage category 1	Triage category 2	Triage category 3	Triage category 4	Triage category 5
Abdominal pain	0	17 (0.3%)	1041 (17%)	3101 (49%)	1691 (27%)	439 (7%)
SOB	0	168 4%	520 (14%)	1738 (45%)	1406 (37%)	14 (0.4%)
Minor complaints	0	1 0.03%	43 (1.4%)	504 (16%)	953 (31%)	1554 (51%)
Chest pain with cardiac features	0	59 0.02%	2498 (88%)	275 (10%)	3 (0.1%)	1 (0.03%)

4.4.4.2 ED Minor complaints diagnosis and primary care coverage

A second contingency table was generated between the patients presenting with minor complaints versus primary care coverage (see table 4). It showed that out of the total patients presenting to ED without primary care access (54148), 4.4% presented for minor complaints, this percentage was 4.7% in the group that had primary care coverage (14220).

Table 4: Contingency table showing primary care coverage versus presentations with minor complaints

Primary care coverage	ED visit reason minor complaints
Missing	36
No	2357 / 54148 (4.4%)
Yes	662 / 14220 (4.7%)

4.4.4.3 Triage category and Primary care coverage:

A third contingency table was generated (see table 5 below) to see presentations to ED based on Triage categories versus presence or absence of primary care coverage. It showed an approximately similar percentage of patients presenting within each triage category from the two comparison groups (patients presenting with primary care coverage to ED versus patients presenting to ED without primary care coverage).

Table 5: Cross table showing primary care coverage vs triage categories

	Triage category 1	Triage category 2	Triage category 3	Triage category 4	Triage category 5	
Primary care MD: No	608/54148 (1.1%)	7651/54148 (14%)	20320/54148 (37.5%)	18905/54148 (34.9%)	6664/54148 (12.3%)	54148
Primary care MD: Yes	130/14220 (0.9%)	2109/14220 (14.8%)	5680/14220 (39.9%)	4708/14220 (33%)	1593/14220 (11.2 %)	14220
	738	9760	26000	23613	8257	

4.4.4.4 Triage category and Referring MD

Finally, a fourth contingency table was generated to explore the association between triage priority and referring MD: yes/no. Table 6 below shows the details that 54% of the patients belonging to triage category 3 or lower were referred to ED by a physician. About 45% of the patients referred by the physicians to ED belonged to Triage category 4 and 5. Results regarding the patients who did not have an MD referral, showed 48% of the presentations belonging to Triage category 4 and above and Approximately 52% of the presentations belonging to Triage Category 3 or lower.

Table 6: Contingency table showing triage priority and referral by an MD (yes/no)

	NA (N=547)	No (N=31059)	Yes (N=37324)	Overall (N=68930)
Triage.priority				
0	7 (1.3%)	10 (0.0%)	3 (0.0%)	20 (0.0%)
1	9 (1.6%)	429 (1.4%)	309 (0.8%)	747 (1.1%)
2	16 (2.9%)	4447 (14.3%)	5313 (14.2%)	9776 (14.2%)
3	130 (23.8%)	11141 (35.9%)	14860 (39.8%)	26131 (37.9%)
4	230 (42.0%)	11481 (37.0%)	12133 (32.5%)	23844 (34.6%)
5	155 (28.3%)	3551 (11.4%)	4706 (12.6%)	8412 (12.2%)

4.4.5 Transfer of patients out of the ED

At the end of their ED visit, patients were oriented to the following places: at the original site and the new site, 77 % of patients were either reoriented or returned home. About 16% of the patients were admitted at the original site, followed by 13% at the new site. The proportion of LWBS (left without being seen by physician) patients was 8%, slightly higher for the new site, compared to 5% of patients at the original site.

4*.5 Factors associated with length of ED visit*

As described in the methodology section earlier, to identify the statistically most important predictor variables for the primary outcome variable ED stay time, the machine learning methods random forests and regression trees were applied (78).

4*.5.1 Random Forest Analysis Results*:

Random forest regression modeling was done to explore the association of the covariates site, age, gender, having a primary care provider (M.D.): yes/no, having a referring doctor (M.D.): yes/no, arrival mode to ED and triage priority with ED LOS. Table 7 displays the details of the random forest model specification. Overall, the percentage of the variance explained in the ED LOS due to covariates (a measure of goodness of fit) was 12%. A random forest model with perfect prediction accuracy would show a goodness of fit statistic of close to 100%. Figure 7 shows the percentage increase in MSE in predicting ED LOS after essentially removing each covariate at a time from the random forest. Higher increases of MSE indicate a stronger variable importance.

The results indicate that age has the highest relative importance as a predictor for ED LOS (increase in MSE 21%), followed by "arrival mode" (20.6%), "triage priority" (9.2%) and "referring MD" (6.9%). The variables "primary care MD", gender and site showed only relatively low variable importance (increase of MSE <5%). In terms of the primary objective of the study, the findings of the random forest are largely consistent with the multivariable analysis of the association of site (re-location) and ED LOS. The random forest analysis revealed only a very weak (1.2%) increase of MSE associated with removing the variable site from the model. In fact, site showed the lowest variable importance of all variables included in the model.

Table 7: Random forest regression details

Type of random forest:	Regression (continuous outcome)
Number of trees fitted to the data	500
No. of variables tried at each split	2
Mean of squared residuals	239.2571
% Variance explained	**12.28**

Figure 7: Showing percentage increased in MSE for each covariate in relation to ED LOS.

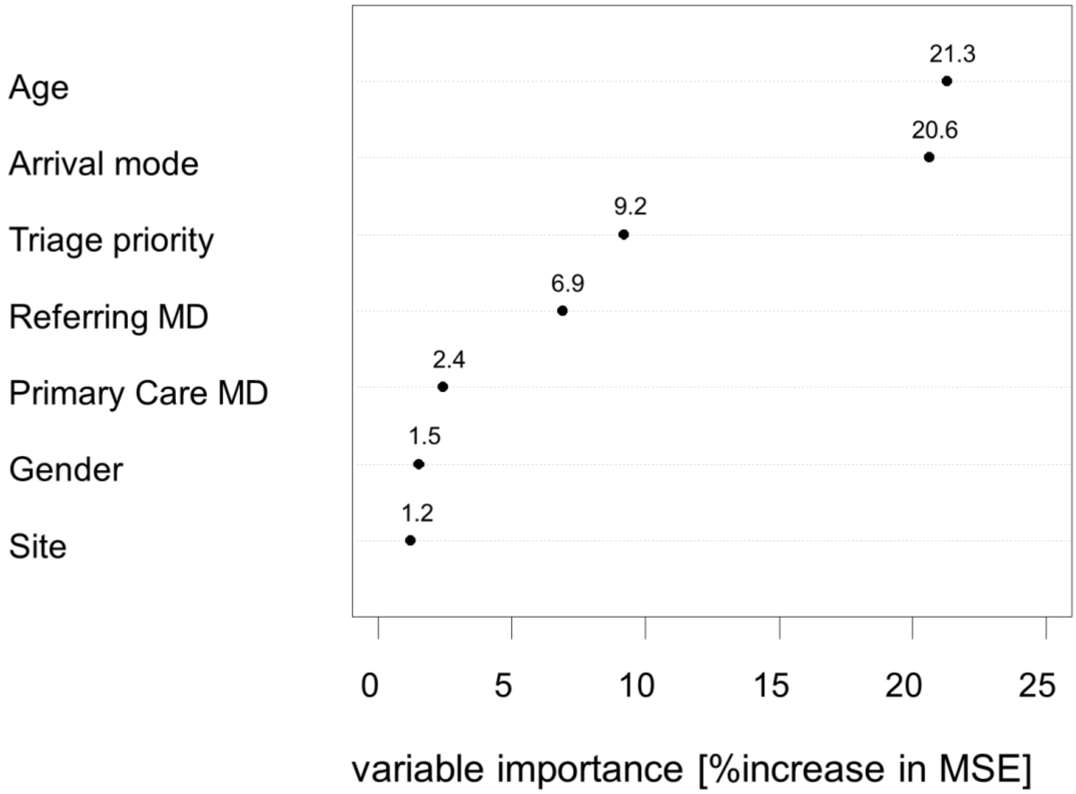
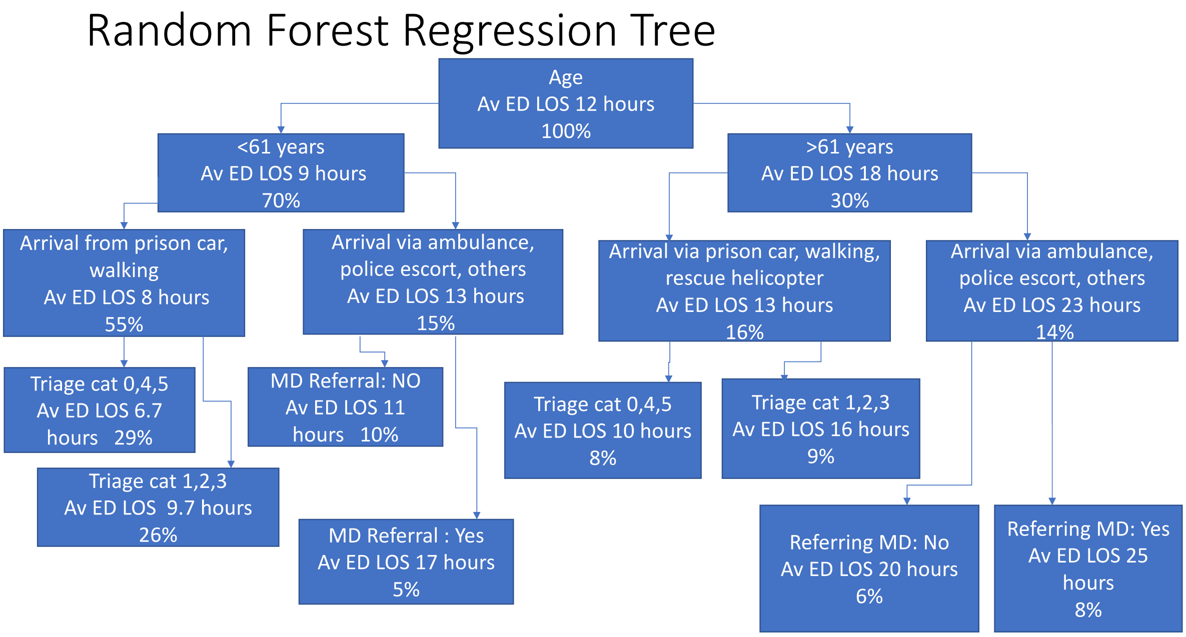


Figure 8: Regression tree with variable average ED length of stay as response variable. Av ED LOS: Average ED LOS, the percentage depicts the representative sample from the data.



The regression tree in figure 8 shows that age is the most important predictive factor for ED LOS, with average ED LOS being 12 hours. Age more than 61 years is the strongest predictor of prolonged ED LOS (30% of the sample) with an average ED LOS of 18 hours. Among those less than 61 years old (70% of the sample) average ED LOS was 9 hours. This group is further stratified by arrival mode, indicating that individuals below 61 years arriving by walk-in or by armored car (55% of the sample) have the shortest predicted length of stay, with average ED LOS of 8 hours, and patients who arrived through ambulance or rescue helicopter (15% of the total sample) had an average ED LOS of 13 hours. Among this sub group, having a referral from a medical doctor was a predictive factor for longer ED length of stay (5% of the total data sample) with average ED LOS being 17 hours.

The longest expected length of stay was for individuals above 61 years old arriving by ambulance, with an MD referral waiting on average 25 hours in the ED, and waiting 20 hours without an MD referral, as well as older individuals arriving through other means and having triage priority of 1, 2 or 3, waiting on average 16 hours in the ED.

The multiple variable regression tree explained 19.3% of the variance of length of stay at the ED.

Chapter 5

Discussion:

The contribution of this thesis was to show that it is possible and desirable to understand and evaluate organizational re-structuring interventions including relocation, to reduce length of stay /increase efficiency of ED with innovative machine learning modeling. The rare context of a major hospital relocation provides a unique insight as it allows for some changes in the ED functioning and ED population dynamic, which makes an evaluation of the effect of these changes on ED length of stay feasible, providing rich data with dense interactions between variables.

The modern machine learning techniques have the potential to provide greater insight than traditional statistical modeling for predicting ED LOS and for improving the efficiency of health services, using local data. These results and methods may help future research in identifying factors that reduce ED LOS and may help in building an algorithm to potentially build a tool to predict ED waiting times based on relevant factors.

From prior to following the re-location – in fulfillment of objective 1 – the study showed in the univariate analysis (no adjustment for covariates), there was an estimated mean difference of 0.55 hours (95% CI -0.79; -0.30 hours) in the LOS between the original site and the current site. In the multivariable analysis adjusting for the potential confounding covariates, the mean ratio of ED LOS (new site / original site) was estimated to be 0.98 [95%CI:0.97 to 1.00], i.e. suggesting relative differences in ED LOS between sites of between zero and 3%.

In response to objective 2- the machine learning method that used random forest explained 21% of the variation in ED LOS due to covariates, with age more than 61 being the most important predictor factor for longer ED LOS (21.3%), followed by “arrival mode” (20.6%), “triage priority” (9.2%) and “referring MD” (6.9).

The data analysis included 32,424 patient visit records for the original site (from 1/06/2014-25/04/2015) and 36,506 (from 26/04/2015-31/03/2016) patient visit records for the new site. The original site on average dealt with 99 patient visits per day, which increased to 107 visits per day at the new site. The baseline characteristics of the study populations were comparable in terms of the age and gender distribution. The prevalence of ED visitors having primary care coverage increased from 17% before relocation to 24% after relocation.

The ED visit frequency data was calculated using dossier number of patients and the number of patient visits attribute to that dossier number. The results showed that about 75% of the patients only attended the ED once during the study. This observation was consistent before and after the re-location. Approximately 16% of the patients visited the ED twice, and the remaining 9% visited the ED more frequently. These numbers indicated that ED patient load could be potentially lowered if recurrent admissions were further reduced.

The arrival mode data showed that walking (self-presentations) was the most common mode of arriving at the ED followed by the ambulance arrival. At the new site, more patients arrived by walking (75%) and a lesser percentage by ambulance (23%), in comparison with the original site (66% walking and 32% ambulance arrivals). The reason for these differences could be the changes in access as the new site has a nearby metro and regional train access while the original site did not have such an access nearby.

The study results showed via the triage priority data that 46 % of the patients presented at the original site with a triage of category 4 or higher, and 47% at the new site presented with a similar triage category. According to the definition of these triage categories, these had less urgent indications and could have potentially been treated in an extended hours primary care practices thus reducing the ED work load and crowding (42,94).

The documented visit reasons were largely similar at the two sites. It was noticeable that the third most common visit reason was minor complaints (4% of the patients presenting at the original site and 5% at the new site). Ideally these patient visits could possibly have been dealt by the primary care service, or extended primary or secondary care, thus reducing the ED congestion in this ED that was promoted as a highly specialized super hospital. A significant finding was that the patients presenting to ED with primary care coverage and without primary care coverage, had a similar percentage of visits for minor complaints (4% approximately). Another finding of interest was that the patient presentations for the 5 triage categories were similar at large when compared between patient group with primary care coverage and patient group without primary care coverage. It was notable, as research indicates that patients without adequate primary care coverage tend to use the ED more as a safety net in comparison with patients having primary care access (43). These patients are more likely to present to ED with lower urgency triage categories and minor complaints (43). This finding may indicate that for this hospital catchment population the primary care coverage had access limitations (time/capacity). Furthermore, it was noted that 45% of the patients referred by physicians to ED belonged to triage category 4 and higher. Ideally some of these patients could have been redirected to urgent care facilities or extended hours GP clinic, thus reducing the ED congestion. This finding may suggest that such alternate care facilities were not be accessible for this catchment population.

The ED length of stay data was positively skewed and had a wide range. The range was even longer at the new site (>0 to 858 hours) compared to the original site (>0 to 461 hours). However, the average length of stay was only marginally shorter at the new site, about 0.55 hours less than at the original site. The results of the covariate-adjusted analysis were consistent with this crude assessment, indicating that the relative difference in mean ED LOS between both sites take does, with 95% confidence not exceed 3%. The observed long ED LOS at individual patient level at both sites are partly due to the boarding of in-patients in the ED – who are formally admitted to an in-patient department, but for whom there is no bed physically available in the corresponding ward or unit. It is expected that part of the reason for the prolonged ED LOS is delay in specialist consultations which is often 4 hours or longer in Quebec (9). However, this or other potential reasons could not be investigated as admission and discharge data from the inpatient department was not part of the study and thus was not available to the research team.

The random forest and regression tree modeling were employed to determine the association of covariates on ED LOS and their possible interactions with each other. The advantage of these methods over conventional statistical models is that they outperform the earlier methods, especially when the data set is large and dense with interactions (44). The machine learning models that we used also did not require manual model specifications in terms of variable transformations and interactions. The random forest results accounted for about 12 percent of variation in the results. The results indicate that age has the highest relative importance as predictor for ED LOS (increase in MSE 21%), followed by “arrival mode”

(20.6%), “triage priority” (9.2%) and “referring MD” (6.9%). The variables “primary care MD”, gender and site showed only relatively low variable importance (increase of MSE <5%). In terms of the primary objective of the study, the findings of the random forest were absolutely consistent with the multivariable analysis of the association of site (re-location) and ED LOS. The random forest analysis revealed only a very weak (1.2%) increase of MSE associated with removing the variable site from the model. In fact, site showed the lowest variable importance of all variables included in the model. Multiple regression tree modeling was used after random forest and the goodness of fit statistic (“r squared”) attributed 19% of the variation in ED LOS due to covariates and their interaction with each other. The observed difference in model accuracy of about 7% percentage points between random forest and regression tree is expected, given that the random forest employs data resampling and assessment of the predictive accuracy in so called “out of bag samples”, whereas the regression tree analysis displays the goodness of fit in the entire dataset. Therefore, the regression tree analysis is more prone to overfitting.

As stated earlier in the introduction and methodology, the impetus for this research was using EMR data and employing machine learning modeling in the background of the unique phenomenon of major hospital relocation and restructure within a 5km radius. The restructuring due to relocation involved departmental changes that dealt with changes in rostering of IM doctors in ED and abolishment of admissions in Geriatrics, as discussed in the results section. It also involved capacity changes including reduction in in-patient beds, but the major organizational processes were not expected to have been overhauled or drastically changed.

It has been demonstrated that despite investment in improving ED services, sustained improvement is not possible without comprehensive extension of primary care services in the province (91). While the Quebec government has invested a lot of money in the restructuring of the hospital services, and has made strides in addressing ED crowding, more is required to streamline the primary care and make it accessible for the whole population (92). Despite continuous provision of funds by the government, the crisis has not been resolved (4).

Anticipated results and interpretations:

Study Limitations

First, waiting time is a proxy outcome for the efficiency of ED care, which is itself linked to patient safety. Ideally, clinical endpoints should be studied, and future investigations should include patient outcomes. Second, the results of the study come from a single ED, which may or may not be generalizable to other urban clinical settings. However, the machine learning methodology maybe applicable and useful in such settings to generate locally relevant results. Third, the EMR data elements available during each ED visit did not include unstructured data (i.e. free text descriptions or comments). Hence, important aspects in the patient history or physical examination may have been missed that may have enhanced the predictive accuracy of the statistical models applied in this research.

Potential limitations of the study design also include that the data being used was collected for administrative and organizational purposes; it was not uniquely designed for research. EMR data as a research tool had some limitations particularly because of non-uniform entries of certain variables as discussed in the methodology section earlier. The ED LOS measured through the available data source reflected potentially boarded patients who were physically staying in the ED awaiting bed availability in other in-patient departments of the hospital. It was not possible to separate such patients based on the available EMR data and this might have led to confounding of the actual ED LOS and to associated biases in assessing the predictive value of covariates on ED LOS.

It is also pertinent to mention that machine learning methods, in general as they are considered “block box algorithms”, suffer from issues in terms of interpretation and inferences regarding variable associations that tend to be more complex than logistic regression (93). In the present study, I was able to elucidate the strength and direction of associations of covariates with ED LOS through the interpretation of multivariable regression trees. As random forests build on such trees, the findings present a valid approximation of the implicit covariate structures used in the random forest. As the overall predictive accuracy was rather low ((% of variance of ED LOS: 12% in the random forest analysis and 19% in the regression tree analysis), it must be acknowledged that important predictive factors for ED LOS were not captured within the EMR data used in this study.

Conclusion:

The application of modern machine learning approaches to evaluate the EMR data collected at RVH hospital before (1/06/2014-25/04/2015) and after relocation (26/04/2015-31/03/2016) from the original site at Pine avenue to the current site at Glen, in Montreal, Quebec, Canada, provided insight into factors associated with the ED LOS and their complex interactions. The most prominent variable in predicting ED LOS was age but also being referred from a medical doctor to visit the ED.

The study analysis findings indicate that the relocation of a major ED in Montreal, i.e. the move to a new site in close proximity, did not lead to a clinically relevant change in ED LOS. However, factors associated with ED LOS such as age and MD referrals indicate that a comprehensive and integrated primary health care network, providing extensive coverage especially to vulnerable populations, is of major importance. Such primary care services have the potential to reduce ED crowding and ED LOS, and hence, given the central, inter-connected role of the ED in delivering integrated care, improve the quality and safety of care for ED patients.

References:

1. Institute of Medicine. Future of Emergency Care: Hospital-Based Emergency Care at the Breaking Point. National Academies Press; Washington, DC: 2006.

2. Pines JM, Hilton JA, Weber EJ, et al. International perspectives on emergency department crowding. Acad Emerg Med. 2011;18:1358–1370. [PubMed]

3. Cha WC, Shin SD, Cho JS, et al. The association between crowding and mortality in admitted pediatric patients from mixed adult-pediatric emergency departments in Korea. Pediatr Emerg Care. 2011;27:1136–1141. [PubMed]

4.Roberge, D., Pineault, R., Larouche, D., & Poirier, L.-R. (2010). The Continuing Saga of Emergency Room Crowding: Are We Aiming at the Right Target? *Healthcare Policy*, *5*(3), 27–39.

5. Tsasis P, Bains J. Chronic disease: shifting the focus of healthcare in Canada. Healthcare Quarterly 2009;12(2):e1–11.

6. Schoen C, Osborn R, How SKH, Doty MM, Peugh J. In chronic condition: experiences of patients with complex health care needs, in eight countries. Health Affairs 2008;28:1–16.

7. Schoen C, Osborn R, Doty MM, Squires D, Peugh J, Applebaum S. A survey of primary care physicians in eleven countries. Perspectives on care, costs, and expériences. Health Affairs 2009;28:1171–83.

8. McCusker, J. Closer than you think: linking primary care to emergency department use in Quebec. Dissemination Report. Fonds de la recherche en santé Québec. February 2013

9. https://www.who.int/bulletin/volumes/91/5/12-112664/en/ Last accessed on 20/08/2020

10. https://www.who.int/bulletin/volumes/91/5/12-112664/en/

11. Boudreaux E, O'Hea E. Patient satisfaction in the Emergency Department: a review of the literature and implications for practice. The Journal Of Emergency Medicine [serial online]. January 2004;26(1):13-26

12. Thompson, D. A., Yarnold, P. R., Williams, D. R., & Adams, S. L. (1996). Effects of actual waiting time, perceived waiting time, information delivery, and expressive quality on patient satisfaction in the emergency department. *Annals of emergency medicine*, *28*(6), 657–665. https://doi.org/10.1016/s0196-0644(96)70090-2

13. Schlesinger M, Mechanic D. Challenges for managed competition from chronic illness. Health Affairs Supplement 1993:123–37.

14. Jones S. Why not the best for the chronically ill? Health Insurance Reform Project, Research Agenda Brief. Washington, DC: George Washington University Press; 1996.

15. National Chronic Care Consortium. Integrating care for people with chronic conditions. Bloomington, MN: NCCC CareLink;1996.

16. Institute for Health and Aging, University of California, San Francisco. Chronic care in America: A 21st century challenge. Prepared for The Robert Wood Johnson Foundation; 1996 August.

17. Canadian Institute for Health Information. Understanding emergency department wait times. Ottawa, ON: Canadian Institute for Health Information; 2005. [Accessed 2007 August 22]

18. Association Québécoise des établissements de santé et de services sociaux. Statistiques de l’année financière 2005–2006 tirées du Rapport 2005 sur la méthodologie SIAP. Quebec, QC: Direction de la vigie, du financement et des systèmes, Association Québécoise des établissements de santé et de services sociaux; 2006

19. Association Québécoisedes établissements de santé et de services sociaux. Rapport 2005 sur la méthodologie SIAP. Quebec, QC: Association Québécoise d'établissements de santé et de services sociaux;; 2006

20. Commissaire à la sante et au bien-être Québec. Communique ; Performance du système de santé et de services sociaux québécois Apprendre des meilleurs : étude comparative des urgences du Québec. Available from: http://www.csbe.gouv.qc.ca/fileadmin/www/2016/Communique_Presse/CSBE_Communiqu_Urgences__2016.pdf

21. Gill JM, Mainous AG, III, Nsereko M. The effect of continuity of care on emergency department use. Arch Fam Med. 2000;9: 333-338.

22. Institute of Medicine . Future of Emergency Care: Hospital-Based Emergency Care at the Breaking Point. National Academies Press; Washington, DC: 2006.

23. Pines JM, Hilton JA, Weber EJ, et al. International perspectives on emergency department crowding. Acad Emerg Med. 2011;18:1358–1370. [PubMed] (2a)

24. Richardson DR. Increase in patient mortality at 10 days associated with emergency department crowding. Med J Aust. 2006; 184:213–6.

25. Sprivulis PC, Da Silva JA, Jacobs IG, Frazer AR, Jelinek GA. The association between hospital crowding and mortality among patients admitted via Western Australian emergency departments. Med J Aust. 2006; 184:208–12.

26. Canadian Institute for Health Information (CIHI). Emergency Department Highlights in 2013–2014. Retrieved from CIHI website: https://secure.cihi.ca/free_products/NACRS_ED_QuickStats_Infosheet_2013-14_ENweb.pdf

27. Nugus, P., Holdgate, A., Fry, M., Forero, R., McCarthy, S., & Braithwaite, J. (2011). Work pressure and patient flow management in the emergency department: findings from an ethnographic study. *Academic emergency medicine : official journal of the Society for Academic Emergency Medicine*, *18*(10), 1045–1052. https://doi.org/10.1111/j.1553-2712.2011.01171.x

28. CBC News. Que. hospital probes ER waiting room death 2010 [cited 2017 April 7]. Available from: **http://www.cbc.ca/news/canada/montreal/que-hospital-probes-er-waiting-room-death-1.969523**. Last accessed on 17/07/2020

29. CTV News. Montreal woman dies after waiting 6 hours in ER 2016 [cited 2017 April 10]. Available from: **http://www.ctvnews.ca/montreal-woman-dies-after-waiting-6-hours-in-er-1.575511** Last accessed on 3/08/2020

30. Toronto Sun. Woman dies, forgotten in ER for 13 hours 2013 [cited 2017 April 8]. Available from: **http://www.torontosun.com/2013/09/27/woman-dies-forgotten-in-er-for-13-hours**. Last accessed on 17/07/2020

31. Gladu F-P. Perceived shortage of family doctors in Quebec: Can we do something about it? *Canadian Family Physician*. 2007;53(11):1858-1860.

32. Ionescu-Ittu R, McCusker J,Ciampi A. Continuity of primary care and emergency

33. McDonald, Kathryn M., Vandana Sundaram, Dena M. Bravata, Robyn Lewis, Nancy Lin, Sally A. Kraft, Moira McKinnon, Helen Paguntalan, and Douglas K. Owens. "Closing the quality gap: a critical analysis of quality improvement strategies (Vol. 7: Care Coordination)." (2007).

34. Committee on Quality of Health Care in America, Institute of Medicine (2001) Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academy Pr.

35. Boyle, Adrian, Jonathon Fuld, Vazeer Ahmed, Tom Bennett, and Susan Robinson. Does integrated emergency care reduce mortality and non-elective admissions? A retrospective

analysis. Emergency Medicine Journal (2011): emj-2010

36. Institute of Medicine. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academy Press, 2001.

37. McCusker, Jane, et al. Use of community services by seniors before and after an emergency visit." Canadian Journal on Aging/La Revue canadienne du vieillissement 20.02 (2001): 193-210.

38. United States General Accounting Office. Report to the ranking minority committee on finance, US Senate. Hospital emergency departments. crowded conditions vary among hospitals and communities. Washington, DC: US General Accounting Office; 2003.

39. Carr BG, Hollander JE, Baxt WG, Datner EM, Pines JM. Trends in boarding of admitted patients in us emergency departments 2003-2005. *Journal of emergency medicine*. 2010;39(4):506-511. doi:10.1016/j.jemermed.2008.04.035

40. General Accounting Office. Hospital Emergency Departments: Crowded Conditions Vary Among Hospitals and Communities. Washington DC: General Accounting Office; 2003.

41. Richardson LD, Asplin BR, Lowe RA. Emergency department crowding as a health.

42. DeLia D, Cantor J. Emergency department utilization and capacity: Robert Wood Johnson Foundation. 2009

43. Billings J, Parikh N, Mijanovich T. Emergency department use in New York City: a substitute for primary care? New York: Commonwealth Fund Publications; 2000

44. : Rahimian F, Salimi-Khorshidi G, Payberah AH, Tran J, Ayala Solares R, Raimondi F, et al. (2018) Predicting the risk of emergency admission with machine learning: Development

and validation using linked electronic health records. PLoS Med 15(11): e1002695. https://doi.org/10.1371/journal. pmed.1002695.

45.Choi E, Bahadori MT, Schuetz A, Stewart WF, Sun J. Doctor AI: predicting clinical events via recurrent neural networks. arXiv. 2015 Nov 18 [cited 2017 Jul 28]. http://arxiv.org/abs/1511.05942.

46. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA. 2016; 316(22):2402. https://doi.org/10.1001/jama.2016.17216 PMID: 27898976

47. Nugus, Peter, et al. Integrated care in the emergency department: a complex adaptive systems perspective. Social science & medicine 71.11 (2010): 1997-2004.

48. Yin, R.K. (2014). Case Study: Design & Methods. London: Sage.

49. Asplin BR, Magid DJ, Rhodes KV, Solberg LI, Lurie N, Camargo CA Jr. A conceptual model of emergency department crowding. Ann Emerg Med. 2003; 42:173–80.

50. Cha WC, Shin SD, Cho JS, et al. The association between crowding and mortality in admitted pediatric patients from mixed adult-pediatric emergency departments in Korea. Pediatr Emerg Care. 2011;27:1136–1141. [PubMed (3a)

51. Canadian Association of Emergency Physicians and National Emergency Nurses Affiliation (CAEPNENA) Joint Position Statement on Emergency Department Crowding. Canadian Journal of Emergency Medicine. 2001;3(2):82–84.

52. Authier P, Montreal Gazette. Quebec has worst emergency room wait times, health and welfare commissioner says 2016 June 2 [cited 2017 April 7]. Available from: http://montrealgazette.com/news/quebec/quebec-has-worst-emergency-room-wait-times-health-and-welfare-commissioner-says. Last accessed on 17/07/2020.

53. Boyle P., Pineault R., Roberge D. Assessing Quebec's Multi-Component Program to Reduce Emergency Room Crowding. Canadian Public Policy. 1992;18(2):189–202

54. Stead LG, Jain A, Decker WW. Emergency department over-crowding: a global perspective. *International Journal of Emergency Medicine*. 2009;2(3):133-134. doi:10.1007/s12245-009-0131-x

55. Medley DB, Morris JE, Stone CK, Song J, Delmas T, Thakrar K. An association between occupancy rates in the emergency department and rates of violence toward staff. J Emerg Med. 2012; 43(4):736±44. https://doi.org/10.1016/j.jemermed.2011.06.131 PMID: 22325552

56. Lowe RA, Localio RA, Schwarz DF, et al. Association between primary care practice characteristics and emergency department use in a medicaid managed care organization. Med Care. 2005;43: 792-800

57. Health Council of Canada. Population patterns of chronic health conditions in Canada: a data supplement to why health care renewal matters: learning from Canadians with chronic health conditions. Toronto: Health Council; 2007.

58. Breton M., Camirand F. État de situation dans les salles d'urgences 1991–1995. Quebec: Ministry of Health and Social Services, Government of Quebec; 1995

59. Rowe B.H., Bond K., Ospina M.B., Blitz S., Schull M., Innes G., Afilalo M., Bullard M., Campbell S.G., Curry G., Holroyd B., Toon P., Sinclair D. Emergency Department Crowding in Canada: What Are the Issues and What Can Be Done? Ottawa: Canadian Agency for Drugs and Technologies in Health; 2006. Technology Overview No. 21

60. Cazale L, Dumitru V. Les maladies chroniques au Québec: quelques faits marquants [Chronic diseases in Québec: some notable facts.]. Zoom Santé: Série Enquête sur la santé dans les collectivités canadiennes. Québec: Institut de la statistique du Québec; 2008.

61. Ionescu-Ittu R., McCusker J., Ciampi A., Vadeboncoeur A.-M., Roberge D., Larouche D., Verdon J., Pineault R. Continuity of Primary Care and Emergency Department Utilization among Elderly People. Canadian Medical Association Journal. 2007;177(11):1362–68.[PMC free article] [PubMed

62. Afilalo, J., Marinovich, A., Afilalo, M., Colacone, A., Léger, R., Unger, B. and Giguère, C. (2004), Nonurgent Emergency Department Patient Characteristics and Barriers to Primary Care. Academic Emergency Medicine, 11: 1302–1310. doi: 10.1197/j.aem.2004.08.032

63. https://www.nurses.org.nz/model_of_care(last accessed on 17/8/2020)

64. Levine D. Health Care and Politics: An Insider's View On Managing And Sustaining Health Care In Canada. Montreal, Canada: Vehicule Press; 1 edition (June 1 2015). 240 p.

65. Haggerty JL, Roberge D, Pineault R, Larouche D, Touati N. Features of Primary Healthcare Clinics Associated with Patients' Utilization of Emergency Rooms: Urban–Rural Differences. *Healthcare Policy*. 2007;3(2):72-85.

66. Salehi, L., Phalpher, P., Valani, R., Meaney, C., Amin, Q., Ferrari, K., & Mercuri, M. (2018). Emergency department boarding: A descriptive analysis and measurement of impact on outcomes. *CJEM*, *20*(6), 929-937. doi:10.1017/cem.2018.18

67. Bernstein SL, Aronsky D, Duseja R, et al. The effect of emergency department crowding on clinically oriented outcomes. Acad Emerg Med 2009;16(1):1-10.

68. Hoot NR, Aronsky Systematic review of emergency department crowding: causes, effects, and solutions. Ann Emerg Med 2008;52(2):126-36.

69. Villa-Roel C, Guo X, Holroyd BR, et al. The role of full capacity protocols on mitigating crowding in EDs. Am J Emerg Med 2012;30(3):412-20.

70. Asplin BR, Magid DJ, Rhodes KV, et al. A conceptual model of emergency department crowding. Ann Emerg Med 2003;42(2):173-80.

71. Doan, Q., Genuis, E., & Yu, A. (2014). Trends in use in a Canadian pediatric emergency department. *Canadian Journal of Emergency Medicine, 16*(5), 405-410. doi:10.2310/8000.2013.131280

72. Schull MJ, Lazier K, Vermeulen M, et al. Emergency department contributors to ambulance diversion: a quantitative analysis. Ann Emerg Med 2003;41:467-76, doi:10.1067/ mem.2003.23.

73. Ospina MB, Bond K, Schull M, et al. Key indicators of crowding in Canadian emergency departments: a Delphi study. CJEM 2007;9:339-46.

74. Institute of Medicine of the National Academies. Hospital-based emergency care: at the breaking point. Washington, DC: Institute of Medicine of the National Academies; 2006.

75. Affleck A, Parks P, Drummond A, et al. Emergency department crowding and access block. CJEM 2013;15 (6):359-84.

76. Asaro PV, Lewis LM, Boxerman SB. The impact of input and output factors on emergency department throughput. Acad Emerg Med 2007;14(3):235-42.

77. Morley C, Unwin M, Peterson GM, Stankovich J, Kinsman L. Emergency department crowding: A systematic review of causes, consequences and solutions. *PLoS One*. 2018;13(8):e0203316. Published 2018 Aug 30. doi:10.1371/journal.pone.0203316.

78. Shameer K, Johnson KW, Yahi A, Miotto R, Li LI, Ricks D, et al. Predictive modeling of hospital readmission rates using electronic medical record-wide machine learning: a case-study using Mount Sinai heart failure cohort. Pac Symp Biocomput. 2017; 22:276–87. https://doi.org/10.1142/9789813207813_0027 PMID: 27896982

79. Jamei M, Nisnevich A, Wetchler E, Sudat S, Liu E. Predicting all-cause risk of 30-day hospital readmission using artificial neural networks. PLoS ONE. 2017; 12(7):e0181173. https://doi.org/10.1371/journal. pone.0181173 PMID: 28708848

80. Pereira M, Singh V, Hon CP, McKelvey TG, Sushmita S, De Cock M. Predicting future frequent users of emergency departments in California state. In: Proceedings of the 7th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics—BCB '16. New York: ACM Press; 2016. pp. 603–10.

81. Taylor RA, Pare JR, Venkatesh AK, et al. Prediction of In-hospital Mortality in Emergency Department Patients With Sepsis: A Local Big Data-Driven, Machine Learning Approach. *Acad Emerg Med*. 2016;23(3):269-278. doi:10.1111/acem.12876

82. Krumholz HM. Big data and new knowledge in medicine: the thinking, training, and tools needed for a learning health system. *Health Aff (Millwood)* 2014; **33**: 1163– 70.

83.Croskerry P. From mindless to mindful practice–cognitive bias and clinical decision making. *N Engl J Med* 2013; **368**: 2445– 8.

84. R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

85. Hollander, M., & Wolfe, D. A. (1973). *Nonparametric statistical methods*. New York: Wiley.

86. Breiman, Leo. "Statistical modeling: The two cultures (with comments and a rejoinder by the author)." *Statistical science* 16.3 (2001): 199-231

87. Hollander, M., & Wolfe, D. A. (1973). *Nonparametric statistical methods*. New York: Wiley. (page 34)

88. https://www.google.com/maps/dir/MUHC+Glen+Site,+Decarie+Boulevard,+Montreal,+QC,+Canada/Building+V+Division+of+Clinical+Epidemiology,+687+Pine+Avenue+West,+Montreal,+QC,+Canada/@45.490338,73.6049138,14z/data=!3m1!4b1!4m14!4m13!1m5!1m1!1s0x4cc9186b1fabd5af:0xcbcdd7bf5d38ada0!2m2!1d73.6007662!2d45.4730245!1m5!1m1!1

s0x4cc91a3991626565:0xdf153e758831776a!2m2!1d-73.5823336!2d45.506541!3e0 (last accessed on 30/08/2020)

89. http://neighbourhoodchange.ca/documents/2018/02/income-map-montreal-2015.pdf (last accessed on 30/08/2020)

90. Banik, R. (2017). *Coodinating flow across practice boundaries : the collaborative work of emergency and internal medicine* (dissertation). McGill University Libraries.

91. Derlet RW, Richards JR. Crowding in the nation's emergency departments: complex causes and disturbing effects. Annals of emergency medicine. 2000;35(1):63-8.

92. Brian Hutchison, Julia Abelson, and Jhon Lavis,Primary Care in Canada, So Much Innovation,So Little Chagnge Health Affairs20, no.3(2001):116-131

93.Vellido A, Martin-Guerroro JD, Lisboa P. " Making machine learning models interpretable." Proceedings of the 20th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN). Bruges, Belgium, 2012.

94. Ambulatory visits to hospital emergency departments. Patterns and reasons for use. 24 Hours in the ED Study Group.*Young GP, Wagner MB, Kellermann AL, Ellis J, Bouley D JAMA. 1996 Aug 14; 276(6):460-5*

Appendix A

IRB
approval-InitApp_29

Appendix B

Table of diagnoses at the two sites:

	new site (N=36506)	original site (N=32424)	Overall (N=68930)
Visit.reason.desc			
	10 (0.0%)	13 (0.0%)	23 (0.0%)
Bizarre behavior	72 (0.2%)	200 (0.6%)	272 (0.4%)
Abdominal mass / distention	74 (0.2%)	80 (0.2%)	154 (0.2%)
Abdominal pain	3483 (9.5%)	2808 (8.7%)	6291 (9.1%)
Abnormal lab values	316 (0.9%)	314 (1.0%)	630 (0.9%)
Abrasion	24 (0.1%)	26 (0.1%)	50 (0.1%)
Allergic reaction	240 (0.7%)	232 (0.7%)	472 (0.7%)
Altered level of consciousness	191 (0.5%)	227 (0.7%)	418 (0.6%)
Amputation	4 (0.0%)	0 (0%)	4 (0.0%)
Anal / Rectal trauma	14 (0.0%)	10 (0.0%)	24 (0.0%)
Anorexia	63 (0.2%)	39 (0.1%)	102 (0.1%)
Anxiety / Situational crisis	292 (0.8%)	431 (1.3%)	723 (1.0%)
Back pain	1241 (3.4%)	967 (3.0%)	2208 (3.2%)
Bilateral leg swelling / Edema	40 (0.1%)	36 (0.1%)	76 (0.1%)
Bite	81 (0.2%)	60 (0.2%)	141 (0.2%)
Blood and body fluid exposure	43 (0.1%)	63 (0.2%)	106 (0.2%)
Blood in stool / Melena	481 (1.3%)	418 (1.3%)	899 (1.3%)
Burn	51 (0.1%)	36 (0.1%)	87 (0.1%)
Cardiac arrest (non traumatic)	63 (0.2%)	28 (0.1%)	91 (0.1%)
Cast check	8 (0.0%)	14 (0.0%)	22 (0.0%)

Chemical exposure	10 (0.0%)	9 (0.0%)	19 (0.0%)
Chemical exposure, eye	17 (0.0%)	14 (0.0%)	31 (0.0%)
Chest pain (cardiac features)	1467 (4.0%)	1369 (4.2%)	2836 (4.1%)
Chest pain (non cardiac features)	912 (2.5%)	704 (2.2%)	1616 (2.3%)
Concern for patient's welfare	18 (0.0%)	36 (0.1%)	54 (0.1%)
Confusion	198 (0.5%)	296 (0.9%)	494 (0.7%)
Constipation	44 (0.1%)	65 (0.2%)	109 (0.2%)
Cool pulseless limb	11 (0.0%)	7 (0.0%)	18 (0.0%)
Cough / Congestion	593 (1.6%)	441 (1.4%)	1034 (1.5%)
Cough and fever	251 (0.7%)	208 (0.6%)	459 (0.7%)
Cyanosis	10 (0.0%)	3 (0.0%)	13 (0.0%)
Dental / Gum problems	108 (0.3%)	104 (0.3%)	212 (0.3%)
Depression / Deliberate self harm	87 (0.2%)	130 (0.4%)	217 (0.3%)
Diarrhea	254 (0.7%)	196 (0.6%)	450 (0.7%)
Diarrhea and fever	47 (0.1%)	44 (0.1%)	91 (0.1%)
Difficulty swallowing / Dysphagia	252 (0.7%)	178 (0.5%)	430 (0.6%)
Diplopia	18 (0.0%)	14 (0.0%)	32 (0.0%)
Direct referral for consultation	378 (1.0%)	331 (1.0%)	709 (1.0%)
Discharge, ear	27 (0.1%)	12 (0.0%)	39 (0.1%)
Dressing change	32 (0.1%)	16 (0.0%)	48 (0.1%)
Ear injury	7 (0.0%)	8 (0.0%)	15 (0.0%)
Earache	220 (0.6%)	148 (0.5%)	368 (0.5%)
Edema, generalized	29 (0.1%)	25 (0.1%)	54 (0.1%)
Electrical injury	12 (0.0%)	8 (0.0%)	20 (0.0%)
Epistaxis	130 (0.4%)	121 (0.4%)	251 (0.4%)

Exposure to communicable disease	6 (0.0%)	9 (0.0%)	15 (0.0%)
Extremity weakness / Symptoms of CVA	211 (0.6%)	311 (1.0%)	522 (0.8%)
Eye pain	201 (0.6%)	180 (0.6%)	381 (0.6%)
Eye trauma	60 (0.2%)	51 (0.2%)	111 (0.2%)
Facial pain (non-traumatic/non-dental)	51 (0.1%)	35 (0.1%)	86 (0.1%)
Facial trauma	117 (0.3%)	100 (0.3%)	217 (0.3%)
Fever	624 (1.7%)	437 (1.3%)	1061 (1.5%)
Flank pain	790 (2.2%)	582 (1.8%)	1372 (2.0%)
Foreign body ear	12 (0.0%)	6 (0.0%)	18 (0.0%)
Foreign body in rectum	1 (0.0%)	0 (0%)	1 (0.0%)
Foreign body, eye	47 (0.1%)	26 (0.1%)	73 (0.1%)
Foreign body, nose	4 (0.0%)	1 (0.0%)	5 (0.0%)
Foreign body, skin	11 (0.0%)	12 (0.0%)	23 (0.0%)
Foreign body, vagina	19 (0.1%)	19 (0.1%)	38 (0.1%)
Frostbite / Cold injury	1 (0.0%)	0 (0%)	1 (0.0%)
Gait disturbance / Ataxia	99 (0.3%)	134 (0.4%)	233 (0.3%)
General weakness	1302 (3.6%)	1281 (4.0%)	2583 (3.7%)
Genital discharge / lesion	29 (0.1%)	35 (0.1%)	64 (0.1%)
Genital trauma	28 (0.1%)	9 (0.0%)	37 (0.1%)
Groin pain / mass	12 (0.0%)	13 (0.0%)	25 (0.0%)
Hallucinations / Delusions	34 (0.1%)	88 (0.3%)	122 (0.2%)
Head injury	475 (1.3%)	510 (1.6%)	985 (1.4%)
Headache	981 (2.7%)	930 (2.9%)	1911 (2.8%)
Hematuria	241 (0.7%)	194 (0.6%)	435 (0.6%)

Hemoptysis	103 (0.3%)	59 (0.2%)	162 (0.2%)
Hiccoughs	5 (0.0%)	1 (0.0%)	6 (0.0%)
Hyperglycemia	95 (0.3%)	100 (0.3%)	195 (0.3%)
Hypertension	220 (0.6%)	178 (0.5%)	398 (0.6%)
Hyperventilation	17 (0.0%)	13 (0.0%)	30 (0.0%)
Hypoglycemia	44 (0.1%)	48 (0.1%)	92 (0.1%)
Hypothermia	2 (0.0%)	1 (0.0%)	3 (0.0%)
Imaging tests	214 (0.6%)	148 (0.5%)	362 (0.5%)
Insomnia	22 (0.1%)	29 (0.1%)	51 (0.1%)
Isolated abdominal trauma - blunt	11 (0.0%)	10 (0.0%)	21 (0.0%)
Isolated chest trauma – blunt	24 (0.1%)	22 (0.1%)	46 (0.1%)
Jaundice	12 (0.0%)	10 (0.0%)	22 (0.0%)
Joint(s) swelling	42 (0.1%)	37 (0.1%)	79 (0.1%)
Labial swelling	37 (0.1%)	25 (0.1%)	62 (0.1%)
Laceration / Puncture	600 (1.6%)	428 (1.3%)	1028 (1.5%)
Localized swelling / redness	773 (2.1%)	555 (1.7%)	1328 (1.9%)
Loss of hearing	27 (0.1%)	16 (0.0%)	43 (0.1%)
Lower extremity injury	760 (2.1%)	498 (1.5%)	1258 (1.8%)
Lower extremity pain	1288 (3.5%)	1019 (3.1%)	2307 (3.3%)
Lumps, bumps, calluses	306 (0.8%)	214 (0.7%)	520 (0.8%)
Major trauma – blunt	5 (0.0%)	4 (0.0%)	9 (0.0%)
Major trauma – penetrating	3 (0.0%)	0 (0%)	3 (0.0%)
Medical device problem	255 (0.7%)	194 (0.6%)	449 (0.7%)
Menstrual problems	89 (0.2%)	115 (0.4%)	204 (0.3%)
Minor complaints NOS	1645 (4.5%)	1410 (4.3%)	3055 (4.4%)

Nasal congestion / Hay fever	23 (0.1%)	15 (0.0%)	38 (0.1%)
Nasal trauma	9 (0.0%)	13 (0.0%)	22 (0.0%)
Neck swelling / pain	62 (0.2%)	52 (0.2%)	114 (0.2%)
Neck trauma	61 (0.2%)	45 (0.1%)	106 (0.2%)
Noxious inhalation	24 (0.1%)	24 (0.1%)	48 (0.1%)
Oliguria	29 (0.1%)	25 (0.1%)	54 (0.1%)
Oral / Esophageal Foreign Body	10 (0.0%)	15 (0.0%)	25 (0.0%)
Other skin conditions	201 (0.6%)	191 (0.6%)	392 (0.6%)
Overdose ingestion	79 (0.2%)	98 (0.3%)	177 (0.3%)
Pallor / Anemia	20 (0.1%)	17 (0.1%)	37 (0.1%)
Palpitations / Irregular heart beat	500 (1.4%)	453 (1.4%)	953 (1.4%)
Penile swelling	35 (0.1%)	12 (0.0%)	47 (0.1%)
Periorbital swelling	28 (0.1%)	27 (0.1%)	55 (0.1%)
Photophobia	104 (0.3%)	44 (0.1%)	148 (0.2%)
Polyuria	71 (0.2%)	44 (0.1%)	115 (0.2%)
Post-operative complications	78 (0.2%)	81 (0.2%)	159 (0.2%)
Pregnancy issues < 20 wks	919 (2.5%)	780 (2.4%)	1699 (2.5%)
Pregnancy issues > 20 wks	37 (0.1%)	60 (0.2%)	97 (0.1%)
Prescription / Medication request	70 (0.2%)	71 (0.2%)	141 (0.2%)
Pruritus	24 (0.1%)	17 (0.1%)	41 (0.1%)
Rash	340 (0.9%)	256 (0.8%)	596 (0.9%)
Re-check eye	27 (0.1%)	40 (0.1%)	67 (0.1%)
Rectal / Perineal pain	105 (0.3%)	47 (0.1%)	152 (0.2%)
Red Eye, discharge	88 (0.2%)	57 (0.2%)	145 (0.2%)
Redness / tenderness, breast	40 (0.1%)	43 (0.1%)	83 (0.1%)

Removal staples / sutures	5 (0.0%)	8 (0.0%)	13 (0.0%)
Respiratory arrest	9 (0.0%)	4 (0.0%)	13 (0.0%)
Respiratory foreign body	12 (0.0%)	13 (0.0%)	25 (0.0%)
Ring removal	2 (0.0%)	1 (0.0%)	3 (0.0%)
Rule out infestation	8 (0.0%)	3 (0.0%)	11 (0.0%)
Scrotal pain and/or swelling	166 (0.5%)	117 (0.4%)	283 (0.4%)
Seizure	229 (0.6%)	498 (1.5%)	727 (1.1%)
Sensory loss / Parasthesias	342 (0.9%)	376 (1.2%)	718 (1.0%)
Sexual assault	3 (0.0%)	2 (0.0%)	5 (0.0%)
Shortness of breath	2130 (5.8%)	1716 (5.3%)	3846 (5.6%)
Social problem	34 (0.1%)	72 (0.2%)	106 (0.2%)
Sore throat	407 (1.1%)	269 (0.8%)	676 (1.0%)
Spontaneous bruising	16 (0.0%)	13 (0.0%)	29 (0.0%)
Sting	2 (0.0%)	6 (0.0%)	8 (0.0%)
Substance misuse / Intoxication	455 (1.2%)	882 (2.7%)	1337 (1.9%)
Substance withdrawal	31 (0.1%)	48 (0.1%)	79 (0.1%)
Suicidal Ideation	129 (0.4%)	361 (1.1%)	490 (0.7%)
Suicide Attempt	47 (0.1%)	78 (0.2%)	125 (0.2%)
Syncope / Pre-syncope	571 (1.6%)	605 (1.9%)	1176 (1.7%)
Tinnitus	22 (0.1%)	15 (0.0%)	37 (0.1%)
Traumatic back / spine injury	95 (0.3%)	82 (0.3%)	177 (0.3%)
Tremors	28 (0.1%)	36 (0.1%)	64 (0.1%)
Unilateral reddened hot limb	49 (0.1%)	37 (0.1%)	86 (0.1%)
Upper extremity injury	703 (1.9%)	551 (1.7%)	1254 (1.8%)
Upper extremity pain	626 (1.7%)	473 (1.5%)	1099 (1.6%)

Urinary retention	243 (0.7%)	155 (0.5%)	398 (0.6%)
URTI complaints	145 (0.4%)	80 (0.2%)	225 (0.3%)
UTI complaints	529 (1.4%)	433 (1.3%)	962 (1.4%)
Vaginal bleed	658 (1.8%)	545 (1.7%)	1203 (1.7%)
Vaginal discharge	44 (0.1%)	43 (0.1%)	87 (0.1%)
Vaginal pain / itch	100 (0.3%)	82 (0.3%)	182 (0.3%)
Vertigo	604 (1.7%)	494 (1.5%)	1098 (1.6%)
Violent / Homicidal behaviour	12 (0.0%)	51 (0.2%)	63 (0.1%)
Visual disturbance	311 (0.9%)	410 (1.3%)	721 (1.0%)
Vomiting and/or nausea	505 (1.4%)	425 (1.3%)	930 (1.3%)
Vomiting blood	84 (0.2%)	68 (0.2%)	152 (0.2%)
Wound check	288 (0.8%)	259 (0.8%)	547 (0.8%)