Predicting Department of Medicine Inpatients' Discharges at US Hospitals

By
Ling Qi
B.S., Polymer Science and Engineering, DongHua University, Shanghai, China, 2010
M.S., Chemical and Biomolecular Engineering, Ithaca, NY, USA, 2012

Submitted to the Department of Mechanical Engineering and the MIT Sloan School of Management in partial fulfillment to the requirements for the degrees of

Master of Business Administration

and

Master of Science in Mechanical Engineering

in conjunction with the Leaders for Global Operations (LGO) Program at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2018

Signature redacted Author
Department of Mechanical Engineering and the MIT Sloan School of Management
May 18, 2018

Signature redacted Certified by
Jacob Cohen, Thesis Supervisor
Senior Associate Dean for Undergraduate and Master's Programs
Senior Lecturer in Accounting and Law, MIT Sloan School of Management

Signature redacted Certified by
Kamal Youcef-Toumi, Thesis Supervisor
Professor of Mechanical Engineering, Department of Mechanical Engineering

Signature redacted Accepted by
Maura Herson
Director, MBA Program, MIT Sloan School of Management

Signature redacted Accepted by
Rohan Abeyaratne
Chairman, Committee on Graduate Students, Department of Mechanical Engineering

Predicting Department of Medicine Inpatients' Discharges at US Hospitals

by
Ling Qi
Submitted to the MIT Sloan School of Management
and the Mechanical Engineering Department
on May 23, 2018, in partial fulfillment of the requirements for the degrees of
Master of Business Administration
and
Master of Science in Mechanical Engineering
in conjunction with the Leaders for Global Operations Program

Abstract

In the last few years, US hospitals have faced severe challenges with bed capacity management that leads to capacity congestion. Delivering patients to the right bed at the right time is very important to patient care quality. However, the current process employs a self-reporting system to receive bed availability from each unit. This method does not provide consistent estimates nor does it provide a standardized, proactive bed capacity management perspective. In addition, the Department of Medicine (DOM) has a very complex patient population, both clinically and non-clinically. Various team structure and uneven distributed bed resources introduce additional challenge on patient discharges.
The project aims to develop a predictive analytics tool that consistently and reliably identifies potential patient discharges in the next 24 hours. The prediction tool allows hospitals to incorporate a more proactive bed capacity management process. Every day, a ranked list with each patient's likelihood to be discharged will be the output. This list guides a more focused conversation within the care team to make patient discharge decisions. In addition, the prediction tool provides a comprehensive summary of barriers to discharge.
In this work, we extended the model developed by Zanger [9] for predicting surgical patients' discharges to medicine inpatients' discharge prediction. By partitioning the training and validation set by the date on 12/31/2017, the current performance for the full model on January 2018 medicine inpatients has a prediction power of ~ 0.74 (Area Under Curve of a Receiver Operating Characteristic curve – AUC ROC there onwards). We further evaluated the model performance for specific patient populations. With patients' Length-Of-Stay (LOS) up to 3 days, the model's performance in terms of AUC ROC can reach ~ 0.8; 0.78 for model with patients' LOS up to 5 days, 0.77 for model with patients' LOS up to 7 days, and 0.72 for model with patients' performance up to 12 days. In addition, the model can capture 57.8% discharges in the next 48 hours, and 33.1% discharges in the next 24 hours.

Thesis Supervisor: Jacob Cohen
Senior Associate Dean for Undergraduate and Master's Programs
Senior Lecturer in Accounting and Law, MIT Sloan School of Management
MIT Sloan School of Management

Thesis Supervisor: Kamal Youcef-Toumi
Professor of Mechanical Engineering, Department of Mechanical Engineering

Table of Contents

List of Figures

List of Tables

This page is intentionally left blank.

1 Introduction

The goal of the project is to develop a predictive model that identifies which Medicine patients are likely to be discharged from the hospital in the next 24 hours, hence enabling a more proactive process of managing the hospital's bed capacity for the Admitting Department on bed planning.

1.1 The Department of Medicine

This project is focused on the General Internal Medicine Division (hereafter referred to as 'General Medicine'). The DOM, and General Medicine in particular, care for an inpatient population that is clinically heterogeneous, with many patients arriving to the hospital with complicated, multi-system ailments, and psychosocial complexities. The diagnostic uncertainty associated with newly admitted patients and the diversity of clinical needs result in an environment in which patient length-of-stay (LOS) at the hospital can range from a day or two to several months (long LOS patients). This environment can be sharply contrasted with surgical services where many admissions are elective and patients' diagnoses are largely known before admission. Thus, paths of treatment are clear and LOS is more predictable for patients in surgical services.

The beds can be classified as 'regionalized' or 'non-regionalized' beds. Regionalized beds are pools of beds located on a floor that is staffed by nurses and physicians of the DOM, and are used solely to care for medicine inpatients. Non-regionalized beds are pools of beds on floors that belong to other (mostly surgical) services and are staffed by nurses from these services. The placement of General Medicine patients in non-regionalized beds occurs because the demand for General Medicine beds frequently exceeds the allocated regionalized capacity, while other services often have excess capacity on their assigned floors [3]. General Medicine patients who are assigned to non-regionalized beds are cared by the regionalized nurses and physicians from DOM that do not work on the floor but only manage the care of the DOM patients assigned to beds on the corresponding floor. Patients that are expected to require a high physician workload due to their acuity or complexity are classified as 'Level 1' and must be placed in a regionalized setting so that their physicians are consistently in close proximity. All other less acute patients are referred to as 'Level 2' patients, and may be placed in both regionalized or non-regionalized settings [3].
General Medicine patients are cared for by two different types of physician teams: (i)Teaching teams; and (ii) Hospitalist teams. Teaching teams are composed of resident physicians in training and their supervising Attending Physicians. Hospitalist teams are made up of only post-residency physicians. Teaching teams staff only regionalized beds on DOM floors, whereas, hospitalist teams can cover regionalized beds or non-regionalized beds on multiple floors of other services. Variation in experience level leads to variation in decision-making in treatment, diagnoses, and patients' discharge.

1.2 Project Overview

1.2.1 Problem statement

Currently, the bed assignment process is managed by the Admitting Department at US hospitals. The current system is based on a self-reporting system that is highly decentralized, and lacks clear prioritization rules and complete information. In addition, the misalignment in the intraday timing of admissions and discharges also poses additional challenges to the bed assignment process - discharges generally do not occur until late in the day, therefore care providers prioritize teaching activities and inpatient care for newly admitted patients and patients likely to stay in the hospital throughout the morning over discharging patients.

1.2.2 Project goals

Given the complex nature of DOM's patient population and hence the increased difficulty in prediction, the project is scoped within patients from regionalized floors.
Expanded based on a predictive algorithm developed for surgical floors to predict daily patients' discharges [9], this project focuses on predicting patients' discharges on regionalized floors at the Department of Medicine. The goal of this particular project is to predict patient discharges in the next 24 hours, in order to allow a more proactive bed management system. Every day, the algorithm predicts patient discharges in the next 24 hours, and generates a ranked list with each patient's likelihood to be discharged. This list also enhances communication across different team members by sharing discharge barriers in a centralized, transparent way. This list also guides a more focused conversation within healthcare team to make patient discharge decisions.

1.2.3 Project approach

To make decisions on who are the patients that are ready to be discharged in the next 24 hours, we employ the same data sources used by care teams when evaluating the patient's progress, as recorded in the hospital's Electronic Medical Records (EMR). Those sources of data include, among others: demographic information regarding the patient (age, gender, etc.), hospitalization information (admission date, etc), and clinical and administrative information (vital signs, lab results, functional assessments, etc.).

We collaborate with care teams to apply clinical insights to extract clinical interpretation from the raw data. We categorize any clinical or administrative event associated with a patient based on two concepts 1) 'Clinical Milestones' - clinical events that correlate with patient progression toward discharge, such as stable vital signs, can move independently, etc.; 2) 'Barriers to Discharge' - any clinical or administrative event that may postpone a patient's discharge, such as pending lab tests, or pending bed offers from a facility post-acute care settings.

We use a Multi-Layer Perception Classier (neural network) for the prediction model. The algorithm is trained on 35,212 medicine patients, and validated on about 2,764 medicine patients.

This project is built upon prior work developed by LGO' 18 student Jonathan Zanger on surgical patients' discharges prediction algorithm [9]. Although both projects share the same goal - predicting patients' discharges in the next 24 hours, surgical patients and General Medicine patients are distinctly different and hence significant modification is required for the surgical model to use for predicting discharges for General Medicine inpatients.

In the surgical model, Zanger studied all surgical inpatients and all surgical procedures that existed in surgical services. To predict surgical patients' discharges, Zanger collaborated with surgeons to derive clinical milestones and barriers discharges and apply these clinical insights to develop an algorithm to predict who are the patients that are ready to be discharged from surgical services in the next 24 hours. To adapt the prediction algorithm used for surgical patients' discharges to medicine patients' discharges, three major changes have been made. 1) Patient population: in the medicine model, the patient population is confined to regionalized medicine units. 2) Clinical milestones: due to the fact that medicine patients exhibit a wide range of diagnoses and multiple ailments, it is impossible to group medicine patients the same way as grouping surgical patients by procedure group. In order to be generalized enough to cover a significant portion of medicine patients, milestones of top 40% diagnoses in DOM were extracted and further integrated into the existing algorithm. We employed an iterative process given the complex nature of medicine patients. Extensive case reviews with healthcare teams were conducted to capture the missing clinical milestones or barriers to discharges during the initial development stage to improve the algorithm's prediction power; 3) In-depth processing of case management notes: key phrases from case management notes are extracted as input features to the model to help improve the model prediction power. On medicine units, case managers are mainly responsible for developing discharge plans for all the patients. Case management notes are major sources of information about non-clinical barriers to discharge. We used Natural Language Processing techniques along with extensive case reviews to incorporate as many indicative phrases on patients' discharge readiness signaling information into the model.

In addition, since medicine patients are more complex than surgical patients (heterogeneity in diagnoses, unknown pre-arranged discharge location, etc.), more in-depth analyses and various scenarios have been investigated to further gain a better understanding of DOM patients and discover opportunities to improve the discharge process. For example, analyses have been conducted to understand the challenges of discharging caused by discharge location as well as length-of stay (LOS). Similarly, to understand how LOS affects model prediction power, different experiments have been conducted to confine medicine patients' LOS to five days, seven days and twelve days to gain more in-depth understanding on the challenges in discharging patients as their LOS increases.

One of the key metrics used to gauge model prediction performance is Area-Under-Curve (AUC) of ROC (Receiver Operating Characteristics) curve. We report AUC ROC on both random partition model (80% test population vs. 20% validation set), and on partition-by-date model – January 2018 model performance (partitioned by 12/31/2017; trained on patient population admitted from 05/01/2016 ~ 12/31/2017, and tested on patient population admitted from 12/31/2017 ~ 01/31/2018). Moreover, a daily report on patient discharge accuracy is also presented to get a granular understanding of daily patients' discharge prediction on general medicine floor.

1.3 Results

The baseline model currently has a prediction power (AUC ROC) of 72.1% when tested on January 2018 patients. Due to the complexity in discharging General Medicine patients, the model performs best when patients are discharged to discharge locations termed as "Home Care Services", worst at "Facilities". In addition, the model's prediction power deteriorates as patients' LOS (Length-Of-Stay) increases. After adding enriched information from case management to the model, the model's performance increases to about 74%.

We also studied model performance when restricting to a certain range of LOS. The model takes the patients' information up to their X^{th} day of hospitalization. Further analysis demonstrated that the model can reach a prediction power of 80% when applied to patients' LOS within 3 days, 78% within 5 days of LOS, 77% within 7 days of LOS (72% when LOS is over 12 days). These results highly overlap with observations from a hospital's administration, that the more complex (clinically and psycho-socially) the patients are, the longer they stay at the hospital, and in turn the harder it is to predict their discharge.

1.4 Thesis Outline

The thesis is organized as follows. Chapter 2 provides an overview of relevant studies in the existing literature. It includes hospital capacity management and application of machine learning in hospital operations. Chapter 3 presents a current state analysis of DOM patient discharge process based on shadowing and in-person interviews with healthcare teams. Operational challenges associated with the current process will be highlighted and be revealed for future improvement opportunities. Chapter 4 presents model methodology and evaluates model performance and sensitivity analysis. Chapter 5 provides in-depth discussion on model performance. Chapter 6 summarizes the conclusions and recommendations drawn from this body of work.

2 Literature Review

Emergency Department (ED) overcrowding is widespread in US cities and has reportedly reached crisis proportions. Trzeciak and Rivers [10] have studied the complex causes and potential solutions for the overcrowding crisis. Not only did they emphasize that the ED is a vital component of America's health care "safety net" and overcrowding in ED treatment areas threatens public health by compromising patient safety and jeopardizing the reliability of the entire US emergency care system, but they also highlighted that the main cause for ED overcrowding is due to inadequate inpatient capacity for a patient population with an increasing severity of illness, and it will require multidisciplinary system-wide support to resolve ED overcrowding. According to Hoot and Arnosky [11], three major factors are distilled as major causes to ED overcrowding: Input factors, throughput factors and output factors. Input factors reflected sources and aspects of patient inflow, such as non-urgent visits, frequent-flyer patients, and influenza season. Throughput factors reflected bottlenecks within the ED, primarily due to inadequate staffing. Output factors reflected bottlenecks in other parts of the health care system that might affect the ED, typical causes lie in inpatient boarding and hospital bed shortage. In addition, surveys from major stakeholders and health care providers also identified other factors such as increasing patient volume and acuity, shortage of treatment areas, nursing staff and ancillary services also account for ED overcrowding. In this study, Hoot and Arnosky proposed potential solutions to ED overcrowding, which include: 1) Increase resources: additional personnel, observation units and hospital bed access; 2) Better demand management: non-urgent referrals, ambulance diversion and destination control; 3) Operations research: crowding measures and Queuing theory.

Bazzoli et al. [12] noted that although additional capacity might be needed in some markets, better management of existing resources could be a more effective solution. Hospital beds are the most important resources, and indeed ED overcrowding was largely caused by an inability to assign a patient to the right bed as Proudlove, Gordon and Boaden [13] discussed an interesting topic in a recent article: can good bed management solve the overcrowding in accident where the Emergency Department Bed Management (BM) forms an important part of operational capacity planning and control, a wider activity concerned with the efficient use of resources. Outside the health context, the production/operations function of an organization is concerned with activities such as scheduling and work flow to enable throughput to meet demand, and minimize work in progress and maximize resource utilization. The authors listed the impacts that bed management can have on an acute care settings, among which, one key impact is on BM managing of supply through discharge management. The general thrust of many initiatives is to foster a 'discharge focus' in medicine. Case management teams are involved in targeting delayed discharges and bed blocking, and finding expediting discharges during periods of pressure at the admissions end. Many have also set up discharge lounges (sometimes combined with an admission lounge to form a "transit" lounge) to attempt to lessen the effects of discharges happening later in the day than admissions. The success of a discharge lounge seems to depend heavily on its location and facilities, and the enthusiasm of discharge staff 'pulling' patients through the process to the lounge. Other initiatives include nurse led discharge and a system, whereby each patient is assigned a discharge status.

Leveraging Operational Research expertise, various approaches have been attempted over the past decades in search for a better system of bed management. For example, Harper and Shahani [14] developed a simulation model to better manage bed capacity in hospitals. Their model considers various types of patient flows at the individual patient level, and resulting bed needs over time. The consequence of changes in capacity planning policies and management of existing capacities can be readily examined. Their work also highlighted the need for evaluating hospital bed capacities in light of both bed occupancies and refused admission rates. They pointed out that the relationship between occupancy and refusals is complex and often overlooked by hospital managers. The challenge on managing bed resources also give rise to inventions. Roscow, Adam and Roth [15] created an enterprise-wide hospital bed management system, an integrated health care delivery network with enabling software and network technology to maximize bed resources, manage varying census levels and avoid patient diversions through real-time monitoring, automation and communication. It is an easy-to-use business intelligence application that is designed to allow administrators, clinicians and managers to easily access, analyze and display real-time patient and bed availability information from ancillary information systems, databases and spreadsheets. It enables users to see trends and relationships in hospital bed management data directly from their desktop personal computers.
Intelligent bed management system is indeed very necessary, however, the reality is that the bed assignment process is a reactive, instead of a proactive process. This means that bed managers are notified after a bed is available. Due to various constraints that a new inpatient may bear, ideal patient-bed match does not always occur. For example, a patient who requires an individual private room might not have a bed at the particular moment he/she needs; even it occurs, the matching could be a 'local optimal'. New directions of research have evolved to investigate approaches that can reduce the hospital overcrowding without adding additional capacity. As mentioned previously, throughput and output factors contribute to ED overcrowding. Therefore, an area of research has been dedicated to better discharge management. De Grood [16] et al. reviewed past literature on discharge prediction and its effective use in hospital capacity management. Discharge Prediction (DP) in their review paper refers to a family of operational techniques, which involve assigning a predicted date of discharge to patients upon their admission to hospital. These predictions are made by the medical team based on a patient's clinical status at the time of admission and are typically updated throughout the hospital stay. Patient care services and operations can then be aligned around this date, with the goal of minimizing delays and inefficiencies during the patient's stay, reducing their Length-of-say (LOS) and helping to alleviate overcrowding through improved patient flow. DP offers greater control over the efficiency of the discharge process. It can theoretically improve both throughput and output by aligning clinical and operational services during a patient's hospital stay and during discharge planning. The intent is that

the resulting efficiencies will reduce LOS, thereby increasing the bed capacity available to meet admission demands and improving overcrowding. Although hospital administrators find DP (Discharge Planning) very attractive and sees its potential in alleviating hospital overcrowding without the increased operating costs incurred by adding staff and bed, the specific contribution of DP itself remains unclear. Moreover, although it is in use in many hospitals, the most effective way to use DP is unknown [16].

Predicting the number of people who will be discharged in the next day can be approached in several ways. One approach would be to calculate each patient's expected length-of-stay (LOS) and then use the variation around that estimate to calculate each day's discharge probability. Several studies have attempted to model hospital length of stay using a broad assortment of methodologies, but a mechanism to accurately predict this outcome has been elusive with Verburg et al [17]. concluding in their study's abstract that "...it is difficult to predict length of stay..."). A second approach would be to use survival analysis methods to generate each patient's hazard of discharge over time, which could be directly converted to an expected daily risk of discharge. However, this approach is complicated by the concurrent need to include time-dependent covariates and consider the competing risk of death in hospital, which can complicate survival modeling. A third approach would be the implementation of a longitudinal analysis using marginal models to predict the daily probability of discharge but this method quickly overwhelms computer resources when large datasets are present. Very recently, Walraven and Forster [17] published their research on TEND Model (Tomorrow's Expected Number of Discharges) predicts the number of patients who were discharged from the hospital the next day. They identified all patients greater than 1 year of age admitted to multisite academic hospital between 2013 and 2015. In derivation of patients, they applied survival-tree methods to patient-day covariates (patient age, sex, comorbidities, location, admission urgency, service, campus and weekday), and identified risk strata having unique discharge patterns. Discharge probability strata for the previous 6 months was summed to calculate each day's expected number of discharges. The results they revealed in their study included 192,859 admissions. The daily number of discharges varied extensively. They identified 142 discharge risk strata. In the validation patients, the expected number of daily discharges strongly predicted the observed number of discharges. The relative difference between observed and expected number of discharges was small. They are conducting further studies to determine if this information improves hospital bed management.

This page is intentionally left blank

3 Current State Analysis

The main goal of this chapter is to provide an in-depth understanding of the current patient discharge process for patients cared by the DOM, particularly on the General Medicine floors. Section 3.1 discusses the unique staffing structure of the care teams (3.1.1) on Medicine floors, the allocation of beds on a general medicine floor (3.1.2), as well as the complex nature of medicine patient population (3.1.3). Next, Section 3.2 discusses the current patient discharge process of patients on medicine floors and related workflows. Section 3.3 presents the main categories of reasons that can delay discharges, both intra and inter-day.

Chapter 3 is based on the integration and analysis of information obtained from shadowing and observations of key stakeholders (physicians, nurses, case managers and the Admitting Department), in-person interviews, focus group meetings, case reviews with care teams, as well as data analysis to harness insights and to corroborate the anecdotal comments gathered from interviews and group meetings.

3.1 General Medicine Practice

General Medicine is a sub-specialty within DOM, that is responsible for nine hospital floors, and accounts for over 40% of the hospital's total inpatient population. General Medicine cares for complex patients with both clinical and psycho-social conditions. Furthermore, its unique staffing structure and the allocation of beds adds further complexity to its operations. In terms of staffing, General Medicine floors are broadly divided into two types of practices: (i) teaching floors; and (ii) hospitalist floors (see section 3.1.1 for more details). General Medicine patients are typically placed in either regionalized beds and non-regionalized beds. Regionalized beds refer to a pool of beds that are located on a single floor and reserved exclusively for General Medicine patients. Moreover, these beds are covered by a local team of General Medicine nurses, and a teaching or a hospitalist physician team. In contrast, non-regionalized beds are located on floors that belong to other services, and are shared with other practices, mostly surgical services, and they are not exclusively dedicated to the use of General Medicine patients. Medicine patients are assigned to these beds only if they are not used by the owner service. These floors are staffed by nurses from the service to whom they belong. However, medicine patients on these floors are cared for by physician teams from the General Medicine that are responsible to cover the non-regionalized beds.

3.1.1 Staffing structure

Two types of care teams exist within General Medicine. Teaching teams are composed of medical residents that are supervised by attending physicians. Each team provides care to the patients on a specific regionalized floor, at the same time conducting teaching activities as part of the residency program. Each team works on a specific physical floor and their members rotate in and out as part of the corresponding residency program. In contrast, hospital teams are staffed with hospitalists who are experienced physicians who rotate their service schedule to do patient daily rounding. Hospital teams can cover a pool of regionalized beds on a specific floor, or cover non-regionalized beds that are on several non-medicine floors throughout the hospital. The care team on each General Medicine floor includes physician teams, nursing teams, case management teams as well as social workers. Figure 1 shows the assignment of each team on each unit, both teaching teams and hospitalists teams. Figure 2 shows the physician team structure, both teaching teams and hospitalists teams.

3.1.1.1 Teaching Teams

Teaching teams are composed of resident physicians who are enrolled in medicine residency training. The first-year resident physicians are referred to as 'interns', and the second-year resident physicians, referred to as 'JAR', supervise the junior resident physicians ('interns'). Each teaching team is also staffed with two attending physicians who impart medicine practice to the medical residents, as well as to supervise the entire team. The attending physicians have the ultimate responsibility for the clinical care and decision-making. Members of the teaching teams rotate in-and-out every two to four weeks.

Teaching teams can be further categorized into Bigelow teams and Flex teams. Bigelow teams refer to teaching teams who care for regionalized units, Flex teams care for pools of beds located on two regionalized floors that are physically close to each other. Bigelow teams are made up of three interns, two JARs (Junior Assistant Residency), and two attending physicians. They care for at most 16 patients on a single floor. Moreover, they take up to four admissions each day, and no overnight admissions are allowed. Compared to Bigelow teams, Flex teams are smaller in size but also more complex in coordination - they care for patients located in two different but adjacent floors, each of which is shared with a Bigelow team. Flex teams are four-person teams, with two interns, one JAR, and one attending physician. The capacity for Flex team is the same as for Bigelow teams. Because patients are not located on one floor, each time when a new admission occurs, the JAR has to coordinate with two floors on bed availability. As a consequence, it adds significant operational challenges for JARs staffing with Flex teams.

3.1.1.2 Hospitalist Teams

Hospitalist teams are different from teaching teams in that they consist of board-certified internal medicine physicians. In 2015, several changes were implemented to re-structure hospitalist teams at the hospital. Each hospitalist team covers nine to eleven patients at a time. In contrast to the teaching teams, where residents provide most of the direct patient care and complete record-keeping tasks with senior physician's supervision, hospitalist physicians are responsible to conduct all of these tasks. Some hospitalist teams are assigned to a regionalized floor, others cover non-regionalized beds. In 2015, General Medicine patients were placed on 13 different non-regionalized bed pools on different floors, most of which are on surgical floors. The use of non-regionalized beds is driven by volume of General Medicine patients that almost always exceeds the effective capacity of the regionalized beds, and is done in scenario in which the local service cannot fill its beds.

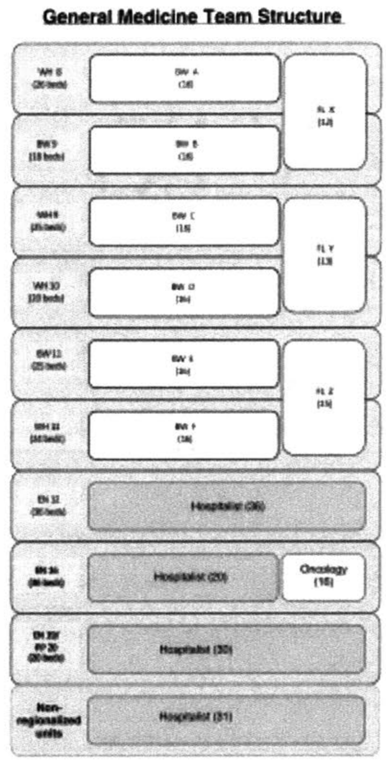


Figure 1 Various team structures and their assigned units are shown here. Maximum number of patients can be cared for by each team is shown in the parenthesis [4] Wh, En, Pps are the building name. Wh 9 means 9^{th} floor of Wh building.

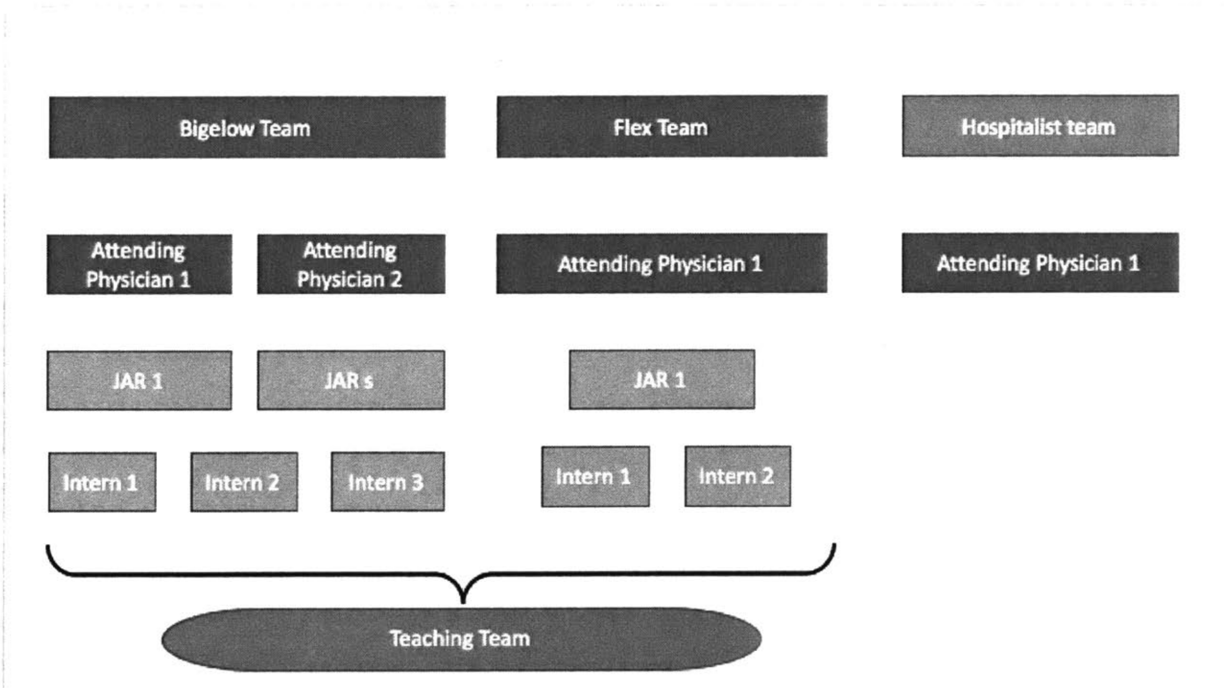


Figure 2 Physician team structure: Teaching teams and hospitalist teams. Note, the nurses are the same nurses staffed to the particular floor

3.1.1.3 Nursing Teams

Nurses form an important part of the patient care team. In the context of General Medicine, nursing teams are always floor-based. This means, for example, patients on Wh 8, will be cared for by the Wh 8 nursing team, regardless of whether they are covered by a teaching or a hospitalist team. This is also true for patients placed on non-regionalized units that are assigned to services other than General Medicine; they will be cared for by the nursing team that is local to that floor.

There are various roles within the nursing team on each floor. These roles include the staff nurses who directly care for the patients, an attending nurse who assists with discharges and removing roadblocks to patient care, and a nurse supervisor who oversees the unit operations. Another role that is of particular interest in the context of this project is the resource nurse, who is responsible for reviewing and approving patient placements on the unit for new admits.

The nurse supervisor is responsible for creating nurse staffing schedules to make sure that all patients on the unit get the nursing care they need. However, the uncertainty in the number of patients' admissions/discharges and patients' complexity leads to frequent situation where the nurse staff is not sufficient to cover all the physical beds in the unit. For example, this can occur when an unexpected number of nurses are absent or when the acuity of the patients on the unit is unusually high (or a combination of the two). When this occurs, the nurse supervisor will communicate the situation to the Admitting Department who will close an appropriate number of beds until more staffing is available or the units' acuity decreases [4].

3.1.1.4 Case Management Teams

Discharge planning is performed by a dedicated team of registered nurses serving as case managers. In their functional DCP roles, case managers also assume other more general roles, such as assessors, planners, facilitators, and advocates.
Case management at hospitals aids with utilization review (UR) and discharge planning (DCP). Both of these tasks are critically important in helping to control hospital expenses, while more importantly, ensuring that the quality of patient care is not compromised. At the most basic level UR is focused on ensuring efficient, effective, and appropriate capacity and care resource management. DCP's focus is on ensuring safe and timely transitions of patients discharged from hospitals to post-acute care settings that meet patient needs. These settings comprise of various facility types and home-care plans. The DCP and UR roles are mutually reinforcing. Since the focus of this project is on the discharge process, we next discuss the DCP processes and practice of the case management teams.

To meet patients' continuous needs after leaving the hospital, discharge planners, in consultation with the care team (physicians and nurses primarily), patient's family, payer representatives, and potentially additional actors, develop a discharge plan that ensures a safe transition from the hospital to a facility that provides sub-acute level of care. While there may be common themes among discharge plans for patients, each plan is tailored specifically to a patient. The variation in patient needs and non-routine aspects of discharge planning introduce challenges to the development and implementation of the discharge planning. These challenges might be clinical and maybe uncorrelated with a patient's clinical complexity.

3.1.2 Bed Resources

Patient beds are scarce resources in any hospital. Due to the high demand of General Medicine inpatients' admissions, beds are assigned to patients based on their illness acuity (level 1: more acute patients, and level 2: less acute patients; specific assignment is based on physician's judgement, the author does not possess the knowledge on how to distinguish level 1 vs. level 2 patients.). The location of the beds across all the units and the various care levels needed for patients require different care team structure to accommodate such differences. The uneven distribution of beds (regionalized vs. non-regionalized beds) inevitably added more challenges to the care team in terms of coordination, communication and logistics.

3.1.3 Patient Population

General Medicine cares for an inpatient population that is heterogeneous by nature. This heterogeneity contributes to the complexity of the clinical care for these patients required across the continuum of care. One implication is that medicine inpatients at the hospitals exhibit a widely variable length-of-stay (LOS). Figure 3 shows the distribution of LOS of patients admitted in September 2017. As one can see the LOS varies from two days to over 50 days. The factors that contribute to patients' complexity can be broadly categorized into clinical factors and psycho-social factors. The latter are not clinical but nevertheless still critical to the care teams' work responsibility. The following section will discuss both the clinical and non-clinical characteristics of General Medicine patients.

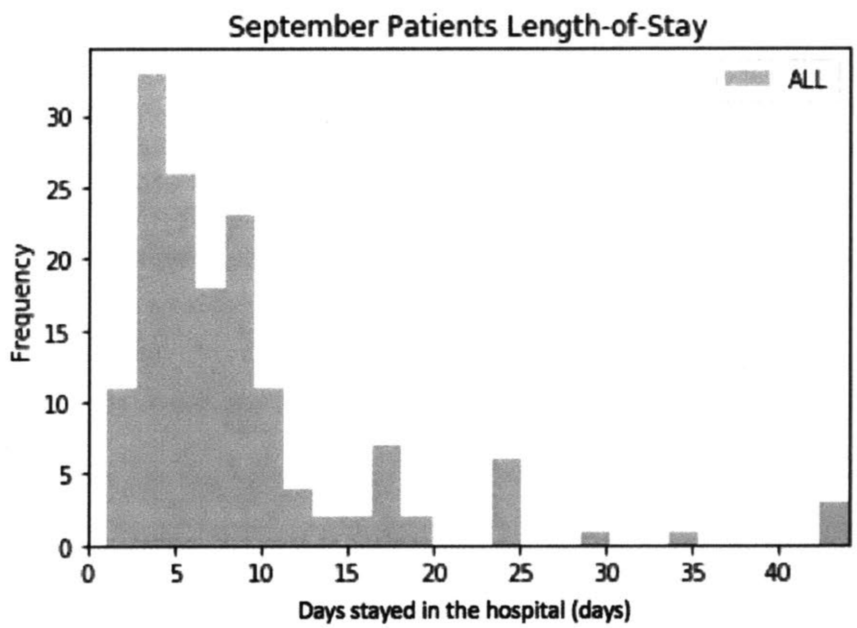


Figure 3 LOS of General Medicine patients on regionalized units, September 2017. A wide range of LOS of patients' hospitalization is observed.

3.1.3.1 Patients' clinical complexity

At General Medicine, the clinical complexity can be seen from the total amount and the heterogeneity of admission diagnoses, as well as the average number of diagnoses of each inpatient. In particular, patients can be diagnosed with multiple diagnoses per hospitalization episode. For example, based on the data from 05/01/2016 to 08/06/2017 on regionalized units, there are 21,415 unique admission diagnoses in total. In this time period, 9,973 inpatients were admitted to regionalized units, which accounts for a total of 13,145 hospitalization episodes. Within this time span, the average number of admission to General Medicine per patient was 1.3, and the average admission diagnoses per hospitalization was 1.6. To be more specific, Figure 4 shows that about 60% of inpatients had only one diagnosis, about 25% of the inpatients had two diagnoses, and another 10%

of inpatients had three diagnoses. In addition, the distribution is very right skewed whereas the largest number of diagnoses per hospitalization is 28.

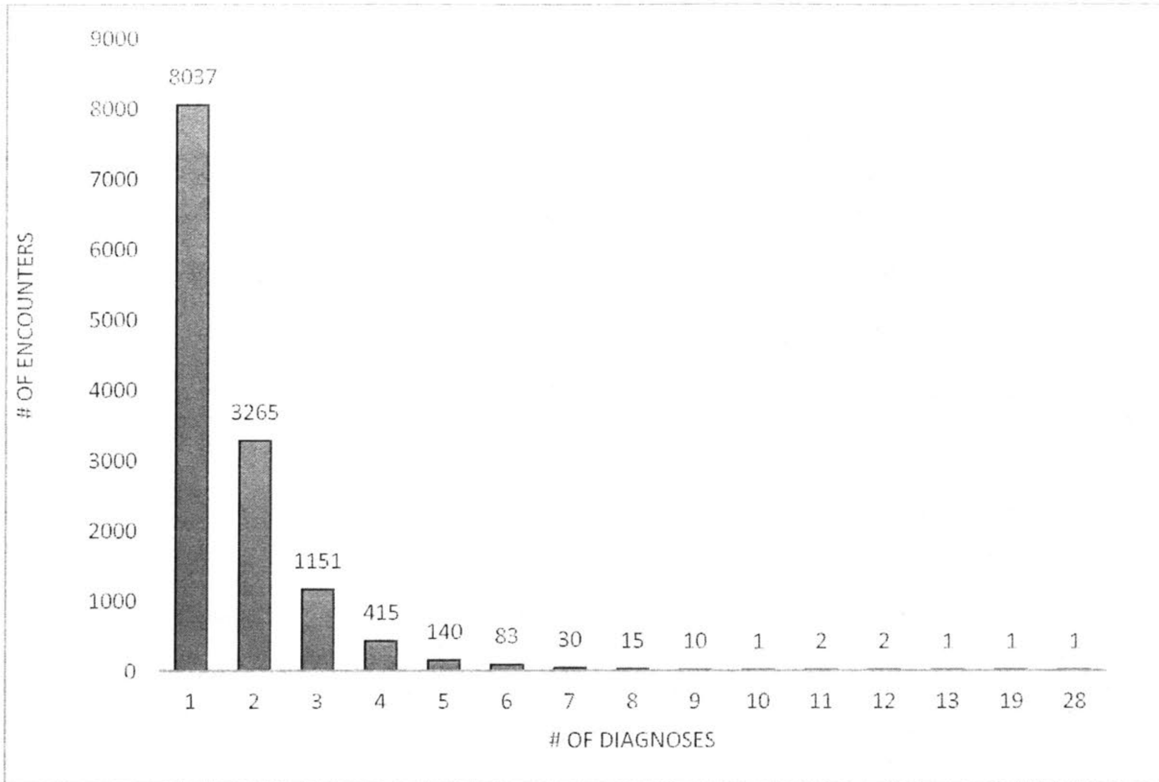


Figure 4 Patients' diagnoses distribution on regionalized units. Data from May 1st 2016 – September 1st 2017. Each diagnosis is based on each patient's hospitalization.

Table 1 demonstrates the heterogeneity in terms of the most common diagnoses (top 40%) presented by General Medicine patients on regionalized floors. The most common diagnosis, Pneumonia, only accounts for 5% of the total inpatients. Other top common diagnoses account for less than 5% of total diagnoses. The very large number of patients' diagnoses implies that there are likely to exist many care paths with different characteristics and timeline, explaining the wide range of LOS. Another typical characteristic is the clinical instability of the patients. Often times, patients are seemingly ready to be discharged next day but unexpectedly experience exacerbation overnight and must be treated again, therefore delay the planned discharges.

Diagnosis name	Count by diagnosis	Percentage of Total Diagnoses
Pneumonia	687	5.2%
Fever	438	3.3%
Shortness of breath	419	3.2%
Abdominal pain	391	3.0%
Altered mental status	374	2.8%
Alcohol withdrawal	371	2.8%
Cellulitis	323	2.5%
Chest pain	290	2.2%
Hypoxia	253	1.9%
AKI	242	1.8%
COPD exacerbation	225	1.7%
Hyponatremia	222	1.7%
CHF	220	1.7%
GI bleed	209	1.6%
Syncope	195	1.5%
Anemia	189	1.4%
Pancreatitis	171	1.3%
UTI	156	1.2%

Table 1 Top 40% of diagnoses of General Medicine patients on regionalized units. Each diagnosis corresponds to a unique diagnosis ID based on ICD 10 coding. Data source based on data from May 1st 2016 – September 1st 2017

3.1.3.2 Non-clinical complexity

The clinical complexity and instability of medicine patients often compound with other non-clinical factors that caused patients to stay in General Medicine floors for prolonged periods. In order to be discharged from the hospital, the patients not only need to be clinically ready, but the patients should also have other non-clinical factors fully addressed. These factors can range from logistics issues, such as transportation and family availability, to post-acute care destination readiness, such as availability in rehabilitation facilities, nursing homes or home services. Any of the above, has to be resolved to obtain a successful and safe discharge. As a result, a patient could stay at a General Medicine unit for a prolonged period of time despite being medically cleared to be discharged. This is a highly undesirable situation in that patients staying at the hospital for nonclinical reasons are susceptible to hospital-acquired infections and also take away beds from other patients, who have real clinical needs.

Securing a post-acute placement for a patient typically takes substantial planning, coordination, and is contingent upon multiple parties' agreement, many of which are outside of the hospital. These tasks require dedicated efforts from the case management team to facilitate this process, but they are understaffed. It is contingent upon the patient's insurance plan, family/patient preference and acceptance from the post-acute care facilities. A more detailed description of the case management discharge planning process is introduced in Section 3.2. Table 2 summarizes a list of clinical and non-clinical factors that have to be aligned for a patient to be able to successfully discharged from the hospital.

Clinical Factors	Psycho-social factors - Patient	Psycho-social factors - Family	Psycho-social factors - Environment
Symptoms	Psychiatric diagnoses	Social supports	Physical barriers
Demographics	Mental health concerns	Caregiver perception of need	Residence - home, facility, homeless
Diagnoses	Substance abuse	Caregiver coping ability	Need for transition - living/care arrangement
Functional	Cognitive ability	Caregiver absence	Financial resources and barriers
Limitations	Coping ability	Caregiver willingness	Mode of communication
Injuries	Adherence assessment	Communication deficits	Benefit coverage
Impairments	Literacy	Caregiver stress	Health professionals' Involvement
Sensory deficits	Knowledge/Comprehension	Caregiver prior learning	Ability of health professionals to communicate with patient/family
Treatments	Motivation/ Readiness to charge	Caregiver past achievements	Health professionals' knowledge of services and resources
	Change in role demands	Changes in role demand	Changes in role demand
	Health beliefs	Caregiver health benefits	Caregiver health benefits
	Culture and migration background	Caregiver literacy	Access to resources
	Primary language		
	Stress		
	Resilience		
	Adjustment to conditions		
	Prior learning		
	Educational Level		
	Past Achievements		
	Communication deficits		
	Risk behaviors		
	Past utilization of services		
	Willingness to accept help		
	Occupation		
	Spirituality		
	Access to care		
	Advance directives		

Table 2 Major factors that need to be aligned before a successful patient discharge can occur. The factors are broadly categorized as Clinical and Non-Clinical factors. Physician and nursing teams are mainly responsible for a patient's clinical factors. Non-Clinical factors fall in the realm of case management, and social workers' responsibility. Adapted from [20]

3.2 Current Discharge Process

A typical discharge process for any a newly admitted, unassessed patient can be broken down into three phases: (i) Admit window; (ii) Pre-discharge window; (iii) Active discharge window. The following section is organized to follow the breakdown of a typical discharge window. Section 3.2.1 and Section 3.2.2 discuss the process and activities associated with the Admit window, and Discharge window, respectively.

3.2.1 Admit Window

Case management responsibilities

The first step related to the discharge process involves assessment of the patient. Case managers are required to perform a high-risk screen for each patient within 24 business hours, with the only exception of patients admitted and discharged over the weekend. The high-risk screen can be placed in the context of the typical phases of a discharge planning cycle consisting of assessment, intervention, planning, implementation, and evaluation. The information from high risk screening is entered into EPIC Hyperspace (a software system that care teams use to input patient information), and can be accessed via EDW (Enterprise Data Warehouse). If a patient is identified as high-risk, an initial assessment is then conducted and is required to be completed within 48 business hours of admission of all high-risk patients. Since discharge planning takes a significant amount of time and effort, identifying which patients will require case management intervention helps to ease discharge planning. The assessment may involve examination of medical records, meeting with the patient or family, and communication with other actors involved in the patient's care prior to hospitalization.

From the discharging perspective, when a patient is identified as high-risk this triggers another type of screening, the initial assessment (IA). The IA is required to be completed within 48 business hours of admission for all high-risk patients. The assessment may require examining medical records, meeting with the patient or family, communication with other case managers, communication with other actors involved in the patient's care prior to hospitalization, or any combination of the above. The information from the IA allows the case manager to be a more effective advocate and facilitator for the patient in the intervention phase of case management. To explain, some case managers use the initial assessment to identify consultation and services that the patient may require, such as nutrition, physical therapy, social work, occupational therapy, or speech language pathology, for example.

The initial assessment also allows the case manager to be a more active participant and advocate in patient discussions with other members of the care team, such as in multidisciplinary rounds. Finally, the initial assessment can facilitate rapport-building with the patient and/or family that may pay great dividends in later phases of discharge planning. Following the 'Admit Window' there are three general categories of patients:

1) Patients not meeting high-risk criteria;
2) Patients meeting high-risk criteria but, upon initial assessment, do not require case management intervention;
3) Patients meeting high-risk criteria and who, upon initial assessment, may require case management intervention.

After the newly admitted patient has been assessed, the patient undergoes the treatment stage primarily carried out by physicians and nursing teams. However, case managers still assess each patient on a daily basis and present updated patients' needs during multidisciplinary rounds. They do so for the following purposes:

1) To determine whether a patient continues to require hospital level of care;
2) To guide the care team in terms of what case management need to do to move a patient along the next level of care whatever that level may be;
3) To anticipate discharge needs and any issues related to discharge from the moment a patient is admitted.

<u>Physicians and nursing teams' responsibilities</u>

During the patients' hospitalization, physician and nurses are mainly responsible for caring for the patients. Physicians round (meaning, they go to patients' rooms to visit the patients on a daily basis, hospital speak) on patients' conditions every day, and update each patient's condition in EPIC Hyperspace. They also place new orders (lab tests, imaging orders, consult orders, etc.) according to patients' recovery progress. In terms of discharging, physicians send out a list of expected discharges via email as primary methods of communicating across floors. This list includes information on expected date of discharge, discharge location, and current barriers to discharge.

Nurses, particularly staff nurses, on the other hand, communicate with physicians on the holistic wellness of the patients, such as, is the patient still in pain, or is the patient still confused. This information helps the physicians to gauge overall readiness to safely discharge the patient.

Case management, nurses, and physicians communicate on a daily basis through multi-disciplinary rounds. The rounds usually occur around noon, after physicians visited the patients. The purpose of the rounds is to facilitate communication across different function and ensure each member of the care team has the same level of the patients' conditions.

3.2.2 Discharge Window

Case management responsibilities

Case mangers begin active discharge planning when, based on information from the care team (physicians, attending nurses, physical therapists, etc.), the patient's probable discharge needs become clear. This clarity allows the CM to identify options available for the patient upon discharge. The CM then meets with the patient to review options and receive the patient's preferences. It is important to note that patient preferences can, in some instances, be determined or influenced by family members, health care proxy, or legal guardian preferences, particularly when the patient lacks capacity to make independent decisions, or the patient has effectively ceded decision-making authority to others. Case managers conduct the following activities during the discharge window:

1) Place referrals to appropriate agencies or facilities
2) Assure any patient/family teaching is completed
3) Determine if there are any specific needs related to medications a patient may require
4) Assure patients' transportation method home and confirm after care is in place

However, often the timeline of the discharge plan is a moving target. It is influenced by changes with patients' conditions, and hence a new plan needs to be developed from the very beginning to accommodate the new needs accordingly. Figure 5 demonstrates the major categories of patient needs that have to be aligned and fulfilled before an actual discharge can occur. Each step, in fact indicates work required from the case managers to communicate and coordinate with other entities to synthesize information required at each decision point. This process is not only time-consuming as it involves multi-party coordination, but it also requires a significant amount of rework once one factor is not met; Sometimes it may require case managers to restart the entire planning process. Moreover, timing is critical in this process as patients' clinical conditions may change, which requires other factors to be tailored to the new condition; in some cases, even a patient's preference or insurance can change over the course of a patient's hospitalization. As one can see, the process is highly dynamic and subject to significant likelihood of rework. Worth noting here is that, case managers usually have on average 10+ cases, each of which could be at a different step of the discharge process.

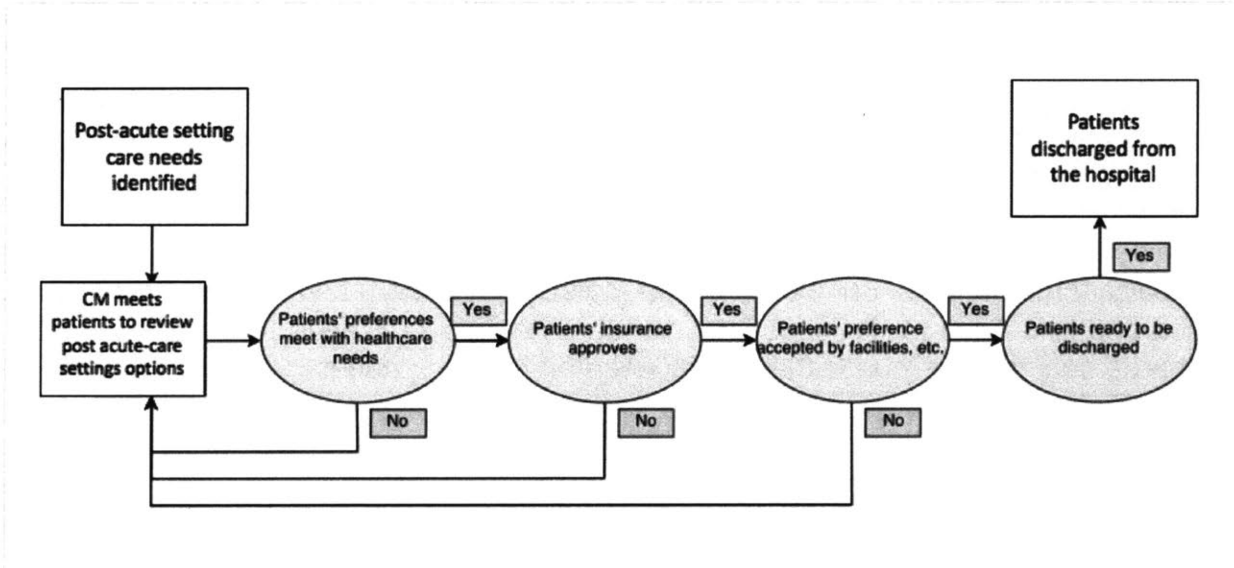


Figure 5 Key milestones case managers are responsible for to implement patients' discharge plan.

3.3 Key challenges during discharge process

Discharging some of the medicine inpatients can be very challenging in that there are too many factors that have to be aligned to have a cohesive plan (Table 2), and any single factor can change rapidly the patient's hospitalization. These sudden changes usually cause delay in executing the discharge plan; sometimes the changes maybe so disruptive that case managers need to restart the plan development process. The common causes that disrupt an existing plan can usually be categorized into the following categories:

1) Clinical factors

   Patients' clinical conditions can exacerbate overnight right before the planned discharge date. If this happens, case managers usually have to wait until patients' conditions stabilize and may need to develop a new plan that is suitable to the patients' new care needs. The evolution of patients' clinical situations may have a ripple effect that triggers changes on logistical planning or psycho-social factors.

2) Psycho-social factors

   Psycho-social factors are very common with medicine inpatients. Table 2 provides a complete list of psycho-social factors that case managers usually encounter when they develop discharge plan. These factors are also highly variant, and may not be correlated to patients' clinical complexity.

3) Logistical factors

   Logistical factors usually delay patients' discharges. They are highly unpredictable and largely dependent on third parties. Typical examples include rejection from referrals to facilities, or no available bed offerings from facilities; no proper transportation methods arranged on the day of discharge; or no adequate setup for home care services, etc.

In addition to the complexity of patients, process-dependent issues also bring challenges to the discharge process. For example, lacking transparency in communicating expected discharges and barriers, lacking clear standards and prioritization for discharges, and discharge occurs after rounds that are focused on teaching duties.

In summary, the planning and execution of discharging a patient is a complex process, especially for General Medicine patients. We would like to design a tool for the care teams and the Admitting Department to overcome some of the above-mentioned challenges such as: built-in variation in decision making process caused by different team structures and uneven bed resources distribution, multi-party communication challenges on patients' readiness to be discharge, reactive bed assignment process, etc. In a grand scheme, the tool is designed to provide a centralized view of barriers-to-discharge for any patient and a set of clear prioritization guidance on patients' readiness to discharge. This information can eventually be beneficial to multiple units within Department of Medicine, for example, a systematic way to prioritize workflow to ensure no delays occur in the process of discharging patients. Chapter 4 and 5 present the methodology as well as the preliminary results.

4 Methodology

Chapter 4 presents an overview of the project methodology. Section 4.1 presents the raw data that are used to generate the inputs for the model. These include: (i) Clinical and administrative data (section 4.1.1); (ii) Demographic information (section 4.1.2); (iii) Environmental information (section 4.1.3). Section 4.2 discusses the significant factors that need to align along the patients' pathway-to-discharge. These factors can be broken down into two categories: (i) Clinical milestones (section 4.2.2); and (ii) barriers to discharge. The raw data (section 4.1) are then processed into features, which are the actual inputs for the model. In section 4.3, we briefly discuss how raw data are converted into features. Section 4.4 includes architecture that constructs the model algorithm. Section 4.5 reports the key performance metrics that are used to measure the prediction model's performance. Figure 6 shows the model architecture, from input layer, clinical insights layer, algorithm layer to output layer.

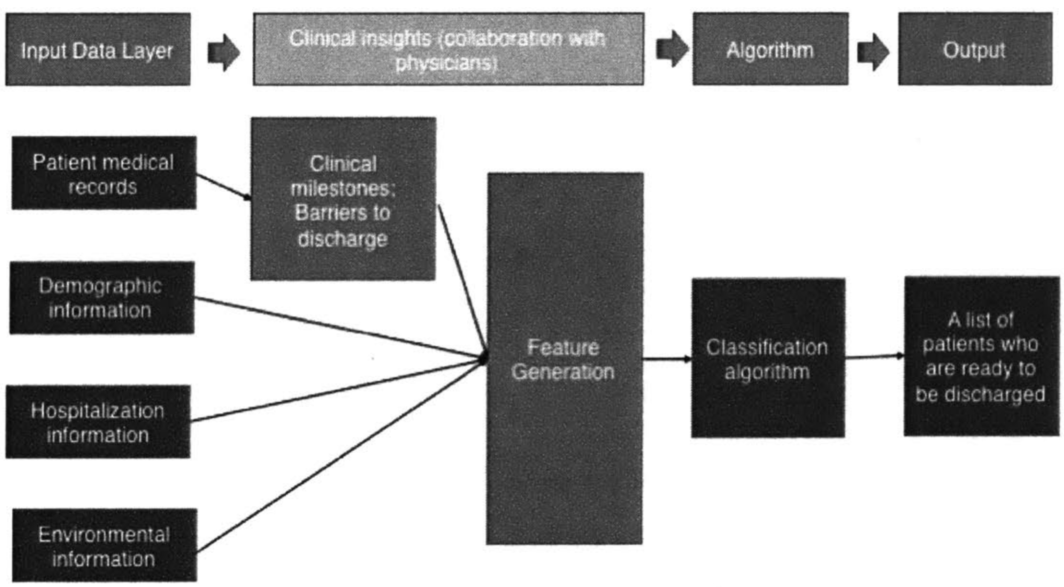


Figure 6 Building blocks of prediction model architecture [9]

4.1 Raw data overview

Partners HealthCare implemented an electronic health record system at the hospital, in partnership with Epic, the industry-leading provider of health information technology. Epic Hyperspace is the software system that records all patients' healthcare data and is the interface that care team members use to input the patient information. On a daily basis, care teams visit the inpatients on the unit, review their historical data, and update the input with new observations and place new clinical orders to diagnose or treat the patients. When it comes to patient discharges, all the patients' data serve as signals of patients' clinical readiness to be discharged. For example, physicians (JARs in Teaching teams primarily) review each section of the data and generate a list of anticipated inpatient discharges with a succinct summary of patients' barriers to discharge, disposition location and estimated days to discharge.

In this research, when evaluating medicine inpatients' readiness for discharge, we access the same data sources used by care teams. All the data is backed up every day to Partner's Enterprise Data Warehouse (EDW). The EDW system aggregates data from different source systems at the hospital, including EPIC, and allows the users to view and filter the data collected across the system. The EDW system is essentially a database where one can access the stored data through SQL queries. In this project, we continue to use the software interface developed by Zanger [9] to communicate with EDW servers. We set the starting date of the data on 05/01/2016 as it was when EDW is effective and available with high quality data (the data before 05/01/2016 is of poor quality, therefore not ideal for developing the model). EDW is synched with EPIC Hyperspace daily, therefore it allows us to have the most up-to-date data on a real-time basis.

The major sources of data include: demographic information of the patient (age, gender, marital status, income level, etc.), hospitalization information, as well as clinical and administrative information (vital signs, lab results, medications, functional assessments, care providers' notes, etc.). The information altogether signals a patient's readiness for discharge, as it demonstrates the patient recovery progress, therefore it is important to understand what raw data are being used.

In section 4.1.1.1, we present the Electronic Medical Records (EMR) data that is used to track the patients' recovery progress. EMR data is generated during a patient's hospitalization, hence they are dynamic and evolves over time. In section 4.1.1.2, we present the data that records the patients' demographic information. This information represents a patient's attributes, hence static. In section 4.1.1.3, we present environmental information that is determined based on the timing of admission.

4.1.1 Clinical and Administrative Data

Clinical and administrative data are very important input to assess a patient's readiness for discharge. This set of data evolve as the patient's clinical condition improves. The different types of clinical and administrative data included in this model are: 1) Medication administration; 2) Lines, drains & airways (LDAs); 3) Procedure orders; 4) Lab results; 5) Flowsheet records; and 6) care providers' notes (case management primarily).

<u>Medication Administration</u>
Medications refer to drugs prescribed to patients during their hospitalization. Common families of medications include: antibiotics, anti-inflammatory drugs, narcotics, IV steroids, and IV fluids, etc.

chlordiazePOXIDE (LIBRIUM) 25 MG capsule [392537252]			Order Details
Dose: **25 mg**	Route: **Oral**	Frequency: **2 times daily PRN for anxiety**	
Dispense Quantity: **5 capsule**	Refills: **0**	Fills Remaining: --	

Figure 7 Screenshot of medicine prescription order from EPIC Hyperspace

A medication administration record is created on EPIC's database any time a medicine is given to a patient. Figure 7 is an example of a medication order. Each medication record includes the medicine type, dosage, route, dispense quantity and usage frequency, as well as a timestamp data when it is prescribed to the patient. This information helps us to identify whether a patient is on a medication treatment that prevents him or her from being discharged.

<u>LDAs (Lines, Drains and Airways)</u>

LDAs are medical devices that are used to deliver or remove fluids from the body. The typical applications of LDAs to patients' treatment includes: delivery of intravenous fluids, drainage of urine from the urinary bladder, and delivery of oxygen into the lung, etc. LDAs are important information that signals patients' readiness for discharge, especially, the type of LDAs, the placement timestamp data and removal timestamp data are of primary interest for our research (*Figure 8*). Patients who have LDAs on them are usually not ready to be discharged until they have been removed.

[REMOVED] PICC Single Lumen Right Basilic			
Placement date:	**10/03/16**	Removal date:	**10/12/16**
Placement time:	**1209**	Removal time:	**1420**

Figure 8 Screenshot from EPIC Hyperspace: PICC line removal order

<u>Procedure Orders</u>

Procedure orders are used by care providers to place treatment instructions for patients, often by requesting other care providers' services in the hospital. Since there is often a wait time associated with procedure order placement, therefore it could delay a patient's discharge. We currently capture the timestamp data of each procedure orders placement. Typical types of procedure orders are currently being used by the model are:

Consult orders

Consult orders are placed by care providers for patients when a clinical issue has reached the bounds of knowledge, experience, or comfort zone of the team or physician [18]. They need additional advice on diagnosis or management from domain specialists. The model currently includes 60 consult order types and the common types are: physical therapy consults, occupational therapy consults, nutrition services; in medicine services, psychiatric consults are also very common.

Diet orders
Patients' diets vary according to their clinical status. A diet order includes the timestamp data of when the diet order is placed as well as the type of diet given to the patient.

Imaging orders
Imaging orders are commonly seen among medicine patients. They are used by care providers for diagnosis, staging and monitoring of the patients. A patient cannot be discharged until the imaging order has been completed. In this model, we capture the placement timestamp and order completion timestamp for all the imaging orders used in the hospital: X-Ray, MRI, CT, ECHO, US, breast imaging.

Laboratory orders
Lab tests are requested through placement of procedure orders. Similar to imaging orders, care providers rely on lab tests to diagnose, monitor and screen a patient. A patient often cannot be discharged until a lab test has been completed. We currently capture the orders of multiple lab tests, including: RBC, Plasma, Cryoprecipitate, Platelets, and pathology.

<u>Laboratory Results</u>
Laboratory results are accessible as part of the hospital's EMR system. The records typical include the lab type, the status of the lab test, the result of the test, and the timestamp data of the lab tests (Figure 9).

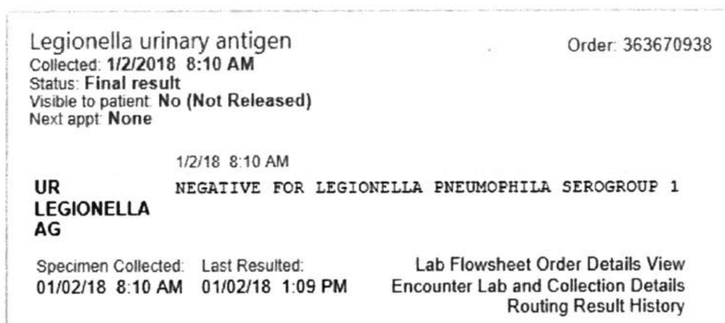


Figure 9 Screenshot from EPIC Hyperspace: Lab test result - legionella urinary antigen

Currently, the model includes the following lab types: 1) Sodium; 2) Potassium; 3) Creatinine; 4) Glucose; 5) White Blood Cell Count; 6) Hemoglobin; 7) The Prothrombin time (INR); 8) Troponin; 9) Lactate; 10) The N-terminal prohormone of brain natriuretic peptide (NT-proBNP or BNPT); 11) Toxicology; 12) Legionella; and 13) Drug-screening test.

<u>Flowsheet Records</u>
EPIC's 'Doc Flowsheets' are used to document various types of data in a patients' chart, including: admission data, trending information (e.g. vital signs), and ongoing assessments and care provided. Data maintained as flowsheet records is accessible to the care providers using 'Flowsheet Templates', an interface which allows to view a

collection of flowsheet records together. Flowsheet templates are generated to specific purpose, such as: assessments, vital signs or I/O.
We currently capture two types of flowsheet records:

Vital Signs
Vital signs are clinical measurements that indicate the state of a patient's essential body functions. Our model captures multiple vital signs, including: body temperature, pulse rate, respiration rate, and blood pressure. Some of the vital signs are documented manually into EPIC by the care providers, while others are imported automatically by an electronic measurement devices.

Care Providers' Assessments
Care providers use EPIC's flowsheet interface to document their evaluation regarding the patient's clinical status. When using flowsheet records for documenting assessments, care providers describe the patient's status by filling in a table containing several items. Each item evaluation is completed using multiple choice of pre-defined values. The use of structured data fields with a limited set of options makes it easier for the model to extract meaningful information systematically compared to free-text evaluations. Currently, we capture assessments recorded by: (1) Bedside Nurses; (2) Case Managers; (3) Physical Therapists; and (4) Speech-Language Pathologists. An overview of the roles of the care provider listed above, and of the assessments captured by the model, is provided following the introduction to the 'Care Providers Notes' records type.

Care Providers' Notes
Care providers use digital notes to document the care delivered and the clinical events relevant to diagnosis and treatment for a patient during the course of hospitalization. Notes are commonly used to communicate findings, opinions and plans between members of the care team, allow care providers to compare past status to current one, and allow retrospective review of case details.
Our model uses care providers' notes for two different applications. First, we capture free-text notes of bed managers and physical therapists as additional inputs to the model, that complete missing information from the structured assessments. Second, we use the creation time of notes to identify when a consultation from specialist has occurred, and connect it to the consultation order (that is captured in 'procedure orders' records).

4.1.2 Demographic Information

According to interviews conducted with care team members, patient demographic characteristics, such as age, income level, or marital status have different impacts on patients' readiness for discharge. We therefore have included these data sources into the model for prediction purpose. Demographic information is different from clinical data in that it is static information, and usually does not evolve during a patient's hospitalization. The major categories of demographic information are:

Age
Patients' age is a significant factor when it comes to patients' recovery progress.

Gender
Gender, according to clinical observation, plays a key role in patients' recovery.

Marital Status
Marital status is a very important factor in the context of discharging medicine patients. A patient that has strong family support has less challenges in post-acute care settings than a patient who lives alone. We included the indicators of whether a patient is single, married, divorced or widowed.

Approximated Income
Approximated income level is estimated on a zip-code level based on the US median household income from the 2006 - 2010 census. The financial status of a patient has significant impact on patients' discharge readiness. A homeless patient who has no insurance is far more challenging to discharge than a patient with insurance that has great coverage. International patients or homeless patients have no zip-code information recorded.

4.1.3 Environmental Information
The model also takes account of the weekend effect into consideration, specifically the days of the week, and the vacation schedule. This information is important because the hospital is staffed differently during weekdays versus weekends. The staffing change also impacts discharge plan (case managers absent, facilities closed for inquiries, lab technicians cannot return lab results, etc.) This behavior influences the care teams' decision-making process before and after the vacation days as well.

4.2 Patients' Pathway-to-discharge – significant factors alignment
4.2.1 Patient Population Scoping

The patient population of interest determines the key inputs for the model. This includes patient specific information, clinical and non-clinical milestones, and barriers to discharge. However, the heterogeneity of medicine patients' diagnoses makes grouping patients in a straightforward fashion more challenging than in surgical services, e.g., by procedure groups. Therefore, we adopted an iterative development process. The initial population of interest is from regionalized units of General Medicine patients. Our rationale is to reduce the complexity caused by the staffing structure and bed resources allocation (some beds are located at two adjacent floors vs. some beds are located on one floor) at DOM. The current patient population included in the current model are patients from the following regionalized units: Wh 8, Wh 9, Wh 10, Bw 9, Bw 11, En 12, En 16, En 19, Pps 20.

4.2.2 Clinical Milestones of Recovery

We term the events that correlate with patient's progression toward recovery and discharge as 'Milestones'. The concept of 'Milestones' is initiated from post-operative recovery milestone research. Each milestone on the patient's recovery progression can be seen as a large step forward towards the patient's discharge. Regardless of the diagnoses, some clinical milestones are shared. These include: 1) beginning of oral diet; 2) discontinue intravenous narcotics; 3) stable vital signs; 4) mobilize the patient out of bed; 5) discontinue epidural medications; 6) consult physical therapy; 7) remove LDAs; 8) complete lab tests and procedure orders.

The status of a clinical milestone could be either: 1) not-executed yet; 2) pending; or 3) executed. These statuses are determined using the entry data (timestamp data) from the providers. Take an MRI order as an example, once a procedure order is placed, the clinical milestone status in our model changes from 'not-executed yet' to 'pending'. Once the result record is updated, the clinical milestone status changes from 'pending' to 'executed'. Note, first MRI or ECHO are very popular tests to order in internal medicine services; second, if an order is a rare event, we will still include it, because the default state of the status is '0', therefore, it will have neutral effect on prediction.

4.2.3 Barriers to Discharge

Barriers to discharge refer to any clinical or administrative event that leads to a delay in a patient's discharge (note, the delay can happen to any patient, even when LOS is as short as two days, it is uncertain on how long a typical barrier to discharge will prolong a patient's stay). Barriers to discharge are determined based on pre-defined formulas to examine the values of barrier elements. These formulas are developed based on clinical practice. Barriers to discharge exist in both clinical and psycho-social factors. For example, a positive lab test result is considered as a barrier to discharge on the day of prediction, and when the test result turns negative is considered as resolution in barrier to discharge. Another typical example of a psycho-social barrier is when a patient is rejected by a rehab facility or do not have adequate insurance coverage required by the facility. The barriers to charge are extended and added to the medicine model, nevertheless, the process of identifying barriers to discharge is an iterative process. Physicians reviewed each false positive case to exhaustively search for the missing barriers to discharge, and they were added to the model.

Clinical Factors

In this research, we have identified over a hundred potential barriers to discharge, and they can be categorized as follows.

- Medication administration
- Abnormal lab results
- Abnormal vital signs
- Awaiting consult appointments

- Newly placed procedure orders/lab tests
- Presence of LDAs
- Psycho-social factors
- Gaps in functional and mobility abilities
- Abnormal bedside events

In our research, we extended the surgical model [9] to include the new barriers to charge based on the top 40% diagnoses, Table 3 summarizes the new barriers to discharge for the medicine prediction model. The list was generated after the physician reviewed the barriers to discharge for each of the top 40% diagnoses. For example, if a patient is on nebulized medication, then the patient is not ready to be discharged until the patient is no long on nebulized medication. The barrier to discharge evolves with the patient's condition or depending on other facilities at the hospital (lab test, MRI/Echo procedure facility, etc.)

Medication (complete)	IV antibiotics, steroids, diuretics
	Benzodiazepines IV or oral: Ativan, Librium
	Intramuscular and oral phenobarbital
	Other meds: aspirin 325, sublingual nitroglycerin
	Nebulized medications
	Colonoscopy prep medications: NuLYTELY/GoLYTELY
	IV fluids
Procedure Orders (complete)	Pending pathology
	Blood transfusions
	Pending procedures: particularly EGD and colonoscopy
	Echo, stress testing, cath
Lab results (complete)	Lactate
	Pending microbiology (blood, urine culture)
	Toxicology
	NT - BNP
	Urine - Leginella

Table 3 Newly generated milestones and barriers to discharge based on top 40% General Medicine diagnoses

According to clinical practice, each barrier to discharge is processed into dynamic features based on 1) definition of anomalies 2) the progression of the daily measurement during the course of patient's hospitalization. (See section 4.3)

Psycho-social and Logistical Factors - Text processing of case management notes
Case managers record patients' psychosocial information in both structured and unstructured data. While structured data can be easily retrieved through EDW and be processed into dynamic features, unstructured information exists in free text form (case management notes). In order to incorporate this information, we employed the natural language processing (NLP) technique to extract key phrases that are indicative of patients' discharge readiness into the existing model.

Python – RAKE (Rapid Automatic Keywords Extraction) [19] is used to parse case management notes. RAKE is a simple Python library that selects key phrases based on given criteria (number of words in a phrase, number of times that this phrase has appeared in the text, and minimum length of each word). Each phrase parsed by RAKE is sorted by its designated score computed by RAKE. Over 700,000 free texts from General Medicine case management have been downloaded from EDW and processed with RAKE using various parameters.

Highly scored words may not necessarily indicate patients' discharge readiness. In order to ensure that each key phrase also makes clinical sense, we have interviewed case managers to incorporate their expertise into the selection of key phrases (shown in Table 4) that will be finally adopted into the prediction algorithm.

substance	dementia	agitation
methadone	suboxone	homeless
psychiatri -	mental	behavior
Out-of-state	cognitive deficit	pending/awaiting

Table 4 Key phrases from case management notes that are added to existing model

The phrases are then incorporated into the current model on a five-point scale ('-1', '-0.5', '0', '0.5', and '1') where '-1' indicates fully not ready to be discharged, '1' means fully ready to be discharged, '0' means neutral effect on discharge. The difference between '0.5' and '1' is based on case management experience of the difficulty of discharging. We consulted case managers on what phrase deserves the most attention when they decide on patients' readiness to discharge. There is no strict correlation on the number of times a word appears to the score we give in the model. Again this is where we rely on clinical experience to develop the model.

4.3 Features Generation

A feature vector is created for each medicine patient for everyday during their hospitalization, from the first day of admission until the day before discharge. In our current design, the daily feature vector represents the state of the patient based on all the information recorded on EPIC at 23:59 that day. Figure 10 shows a simplified version of a feature vector for a patient [9].

Days Since Surgery	Esophagectomy Surgery	Age	IV Narcotics	PO Narcotics	Foley Catheter	PT Order Placed	PT Evaluation	Vital Signs Stable	Discharged Next day?
0	1	36	-1	0	-1	0	0	-1	-1
1	1	36	-1	0	-1	0	0	-1	-1
2	1	36	0	-1	-1	1	-1	1	-1
3	1	36	0	-1	1	1	-1	1	-1
4	1	36	0	-1	1	1	1	1	1

Figure 10 A simplified feature vector of a 36 y.o patient discharged 5 days after surgery [9].

The vector included three types of features: (i) demographic and hospitalization patient information; (ii) environmental information (date and time); and (iii) 'Milestones to Recovery' and 'Barriers to Discharge' events originated from the EMR data. Among these three types of features, additional processing (section 4.3.1, section 4.3.2 and section 4.3.3 present the methodology on converting) is required to convert 'Milestones' and

'Barriers to discharge' events into numerical values for the use of a feature vector. We present the three components of this process as follows: section 4.3.1 Numerical value generation for current event status; section 4.3.2 Event timing capturing; section 4.3.3 Feature 'expiration timeout'.

4.3.1 Numerical value generation for current event status

A feature with positive value signals progression toward discharge, while a feature with negative value indicate that discharge is not ready yet, and progress needs to be made. We assign values from '-1' to '+1' to each milestone and barrier-to-discharge event based on its status.

The milestone events can be described as one of the following status and is assigned a value to represent its corresponding state: 1) 'Not Executed Yet' - an event that is in 'Not Executed Yet' is assigned with a value '0'; 2) 'Pending' - a 'pending' milestone event is assigned with a value of '0.5'; 3) 'Executed' - this state of the milestone is assigned with a value of '1'. Each event status is associated with a specific timestamp, therefore, based on the timestamp data and the status of the event, we can assign values to these events and track their progress over the patients' hospitalization.

The 'Barriers to discharge' are categorized with one of the following states: 1) 'Unknown / Never Existed' - such an event is assigned with a value of '0'; 2) 'Existing' - such an event is assigned with a value of '-1'; 3) 'Resolved' - once a barrier to discharge is resolved, its status is updated to '+1'.

In some cases, the assignment of a numerical value to an event is done in a customized way when we want to capture the significance that the existence of a certain event has on a patient's discharge. More specifically, we assign multiple levels of values for the 'Exists' state, and keeping 'Unknown / Never Existed' state value a 0 and the 'Resolved' state value as '1'.

4.3.2 Event timing capturing

It is important to understand when the status of an event has changed, in addition to being able to capture the current status of a 'milestone' or 'barrier to charge' event. When generating the feature vector, we use the time passed since the 'milestone' or 'barrier to discharge' status changed as an input to the model. In case of 'Barriers to discharge', two separate values are used to capture both the time that has passed since the barrier has occurred, and the time that passed since it was resolved.

4.3.3 Feature 'expiration timeout'

When a significant amount of time passes since a barrier or milestone event status has been captured, the impact of this event on patient's likelihood to next-day discharge is negligible. In this scenario, we want to 'reset' the event's status as if the patient had never experienced this event (‘Never Existed’ in case of barriers to discharge events; ‘Not Executed yet’ in case of milestone events, i.e., the value associated with this particular event is set back to zero). For each feature, we assigned a 'timeout' period after which the feature is restored to its default value. The value of the 'timeout' property was determined based on its clinical nature, and how frequently its data-sources are updated on EPIC. Using events expiration is particularly relevant for milestone events. For example, recall that PT evaluation is considered a milestone on the pathway to recovery, and that we use the timing of a consult order placement as an indicator for progress towards discharge. However, if the patient is still admitted a long period of time after the consult order was placed (more than a week), it is likely that the patient would require a revised PT evaluation prior to discharge. Using 'expiration time', changes the status of the milestone to 'Not Executed Yet', allowing the model to treat a new PT evaluation orders as an indicator for progress towards recovery. For non-consult orders (consult orders refer to PT, diet orders, etc., while non-consult orders refer to LDA placement/removal, lab test result etc.), we treat these differently, and will not enforce an expiration time on the event because if LDAs are not removed, the status will always be ‘Not Executed Yet’ until it’s removed, then it is changed to ‘Executed’)

4.4 Architecture for Deep Learning Neural Networks

In this work, we employ the Python – scikit learn library [20] for developing the deep learning algorithm. A Multi-Layer Perceptron (MLP) classifier is used for classification purposes. Note, given prior development [9], MLP gives the best performance compared to other learning techniques such as logistic regression, random forest, etc. MLP’s superior performance is attributed to the non-linearity it introduces to the input features. In addition, the classifier itself is a back-propagation based learning algorithm that corrects weights of each edge as training occurs. The architecture of the current neural network has one input layer, one hidden layer with 23 - 30 neurons per layer, and one output layer. The stochastic gradient descent algorithm is used to minimize loss function. Ridge regularization is used to prevent overfitting. Non-linearity is introduced through the ReLU (Rectified Linear Unit) activation function. Every time when new input features are introduced, the hyper-parameters of neural networks are re-run to optimize the prediction accuracy. The hyper-parameters to be adjusted are the following: the number of neurons in the hidden layer, learning rate, regularization parameter, and error tolerance for the loss function. Grid search has been used to find optimal values for the hyper-parameters for the neural nets that delivers the best Area Under the ROC Curve performance. (Note, Grid search simply means that we try a combination of hyper-parameters and identify a set of those that gives the highest AUC ROC)

4.5 Key Performance Indicators (KPIs)

This section introduces all the metrics that are used to evaluate our model. Detailed results are presented in Chapter 5. Two major categories of metrics are used: 1) section

4.5.1 discusses Receiver Operating Characteristic (ROC) curve; 2) section 4.5.2 discusses about Daily precision and sensitivity rates: list length accuracy discharges and individual discharge probability by threshold are used to gather comprehensive information about the reliability of the model. Figures shown in the following section are referenced based on the results produced by Zanger's [9] surgical model, and are used as examples for this section.

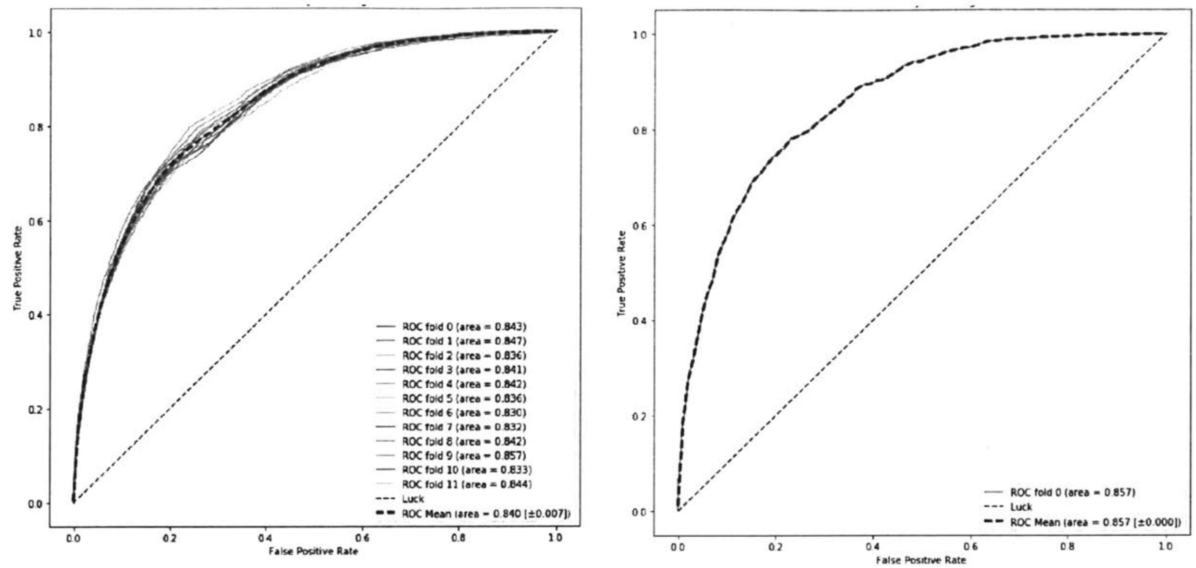
4.5.1 ROC Curve

In statistics, a receiver operating characteristic curve, i.e., a ROC curve, is a graphical plot that illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied.

The ROC curve is created by plotting the true positive rate against the false positive rate at various threshold settings. Ideally, we want to maximize the true positive rate while minimizing the false positive rate - the model should minimize the number of 'missed discharges' and at the same time minimize the number of 'false positives'.

The Area Under the Curve of the ROC curve (AUC ROC) is equal to the probability that the classifier will rank a randomly chosen patient who would be discharged within 24 hours instance higher than a randomly chosen patient who would stay at the hospital in the next 24 hours. An AUC ROC of 1 represents a perfect classification; an area of 0.5 represents an uninformative classifier.

In our research, two types of ROC curves interest us: ROC based on random sampling – which gives us the confidence of the model's general performance, shown in Figure 11 (a). The other type is the ROC of a particular month, shown in Figure 11 (b), which gives us a sense of how good the model is at predicting patient discharges at a particular month.



a) b)

Figure 11 ROC curves. a) ROC curves for randomly split training/test set; b) Training/test split based on a specific date. Results are from surgical model [9]

4.5.2 Daily Precision and Sensitivity Rates

Based on the input feature vectors and prediction model, each patient is computed with a score between 0 and 1, where 0 indicates a patient is definitely not expected to be discharged in the next 24 hours, whereas 1 means a patient is ready to be discharged. This score is computed using an embedded function from Python Sklearn package. Note, these scores DO NOT indicate the discharge likelihood, they are simply a score that has been normalized by Python Sklearn when it classifies whether a patient is ready to be discharged or not, but they are not the same as discharge probability.

We can compare the estimated discharge date with the actual discharge date based on these criteria. The outcome of the result falls into one of the following four scenarios:

1) **True Positives (TP):** Patients who are predicted to be discharged within the next 24 hours indeed *are* discharged. Simply, the algorithm predicts correctly. We also call these *Captured Discharges*.
2) **False Positives (FP):** Patients continue their hospitalization despite being predicted to be discharged within the next 24 hours. In short, according to our threshold, we mis-predicted them.
3) **True Negatives (TN):** Patients who are not predicted to be discharged within the next 24 hours and continue to stay in the hospital.
4) **False Negatives (FN):** Patients who are actually discharged in the next 24 hours, but according to our criteria they still should stay in hospital. We term these cases as *Missed Discharges*.

In the model where training and validation sets are partitioned by a specific date, we also generate a monthly statistic. In order to decide whether a patient is expected to be discharged in the next 24 hours, we have employed two criteria to classify: 1) Top N list based approach 2) Threshold based approach. Patients whose scores satisfy the criteria are classified as 'ready to be discharged in the next 24 hours', otherwise they continue to stay in the hospital.

Top N list based approach: patients whose scores are among the top N highest are classified to be discharged in the next 24 hours. In this work, we have used Top 10 list and Top 30 list to evaluate the model accuracy - patients score the top 10 or 30 highest scores are classified to be discharged in the next 24 hours. In this research, we look at the top 10 and top 30 patients with the highest probability to be discharged that are predicted by the model, shown in Figure 12. As a continuation of surgical prediction model work [9], we continue to use the same metrics to measure the prediction accuracy. However, this criterion can be adjusted given the specific DOM operating situation. We categorized our prediction results into three categories: 'True prediction', indicated by 'green'; 'Single day error', indicated by color 'blue', this means that we missed the patients' discharge date by 24 hours; 'More than one day error', indicated by color 'red', this means that we missed the patients' discharge date by more than 24 hours.

Threshold based approach: we set an absolute threshold to classify whether the patient is expected to be discharged in the next 24 hours. Patients whose scores are higher than the threshold are expected to be discharged, while those whose scores are lower are expected to stay. In this work, we have used two thresholds, 0.5 and 0.7, to evaluate the model accuracy - a patient whose score is higher than 0.5 (or 0.7) is expected to be discharged, and if lower, he/she will continue to stay in the hospital. The threshold level implies the level of confidence of the number of patients who will be discharged in the next 24 hours (Figure 13).

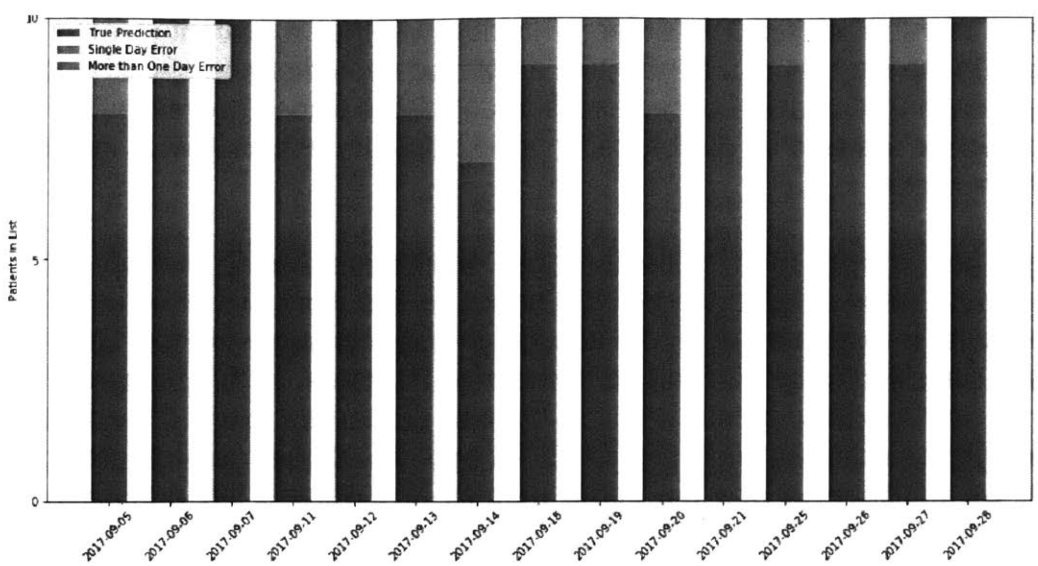


Figure 12 Daily list of patient discharges for a particular month - top 10 patients list, results based on surgical model [9]

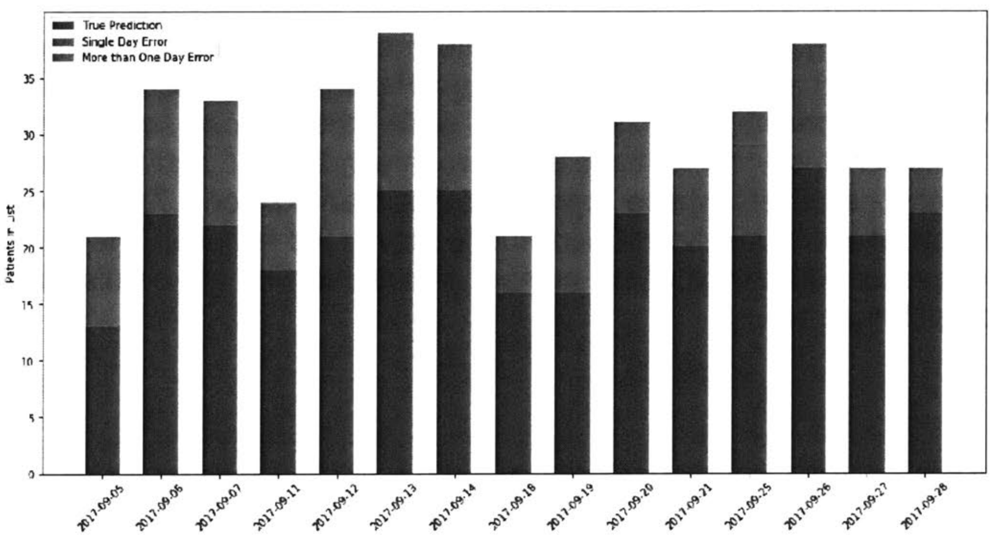


Figure 13 Daily list of patient discharges for a particular month - 70% confidence level that the patients on the list are likely to be discharged, results from surgical model [9]

Choosing a proper threshold is a balancing act. A high threshold - in our case, refers to Top 10 list or a threshold of 0.5 confidence level - typically results in a high proportion of True Positives, but more 'Missed Discharges'. A low threshold, on the other hand, typically leads to more False Positives but less amount of 'Missed Discharges'.
Given this list and a chosen threshold, we measure:

1. **Precision**: the fraction of true discharges among the list of patients predicted to be discharged. The fraction of false positives (predicted to leave but ended up staying) out of the list of patients predicted to be discharged is (1- Precision)

2. **Sensitivity**: The true positives rate. The portion of discharges predicted correctly among all discharges.
3. **False Positive Rate**: The portion of patients that are predicted to be discharged but ended up staying. We break false positives into two categories: 1) patient discharged within less than one day delay; 2) patient discharged after more than one day.

4.5.3 Extension of Model Performance Evaluation

The key metrics described in 4.5.2 are used to evaluate the overall model performance. In addition, the complex nature of medicine patients requires us to evaluate the model from multiple perspectives that are beyond the general, full model performance. In this research, we have extended our model performance evaluation to the following scenarios:

1) **Prediction model for patients' LOS up to X^{th} days (X less than 12 days)**
LOS of medicine patients has a wide span - it can vary from as short as two days to as long as a couple of months. Currently, the model includes patients with all LOS. In this work, we narrowed down the scope of the model so that it includes patients' LOS up to X^{th} day (X less than 12 days). The benefit of doing so is to allow us separate the complex patients' scenarios from the model, therefore allow us to improve the model performance in a more focused way. In Chapter 5, we report the model performance (primarily using AUC ROC as a performance metric) with patients' LOS up to 3 days, 5 days, 7 days and 12 days. We trained the model with patients' data up to their X^{th} day of LOS, and test the model on patients within their X^{th} day of LOS. The model will not be used to predict patients' discharges after their X^{th} day of LOS in the hospital.

2) **Prediction model for patients with specific discharge location**
As described in Chapter 3, discharging medicine patients is a daunting task, primarily due to the fact that medicine services are not pre-planned like surgical procedures. Therefore, the discharge location is unknown at the time a patient is admitted, and it takes great efforts from the case management team to develop a discharge plan and to finally implement it. The complexity of discharging a patient is not just correlated with the clinical complexity of a patient, rather it is a combination of both clinical factors and psycho-social factors. In this research, we break down the typical discharge locations in three major categories: home, home with services and facilities (See Chapter 5 for more details). We group patients according to these three locations, and run models separately on the following two scenarios to understand a more nuanced model performance: 1) patients discharged home with home care services or home; 2) patients discharged to facilities.

In this chapter, we present the model architecture, from raw data source, significant factors along the patients' path-to-discharge, feature generation, neural network architecture, to the key performance metrics. In the next chapter, we report the model prediction performance based on the key performance metrics discussed in section 4.5.

5 Results and Discussion

In this chapter, we examine the model performance in predicting medicine inpatients' discharges in the next 24 hours in terms of AUC ROC (a key metric for model prediction power), as well as daily precision and sensitivity rates. Section 5.1 presents the model's performance when the training and test population is partitioned by 12/31/2017. In this section, we also report the difference in model performance when using different sets of case management notes (section 5.1.1.1). In section 5.2, we evaluated two variations of the model: (i) the model is restricted to patients whose LOS is less than X days (X can be 3, 5, 7, 12 days); (ii) the model is restricted to patients with specific discharge location. The motivation behind a restricted version of model performance evaluation is to understand whether the model can be more applicable in less complex situation.

5.1 January 2018 Model Performance

The model was trained (meaning, the model learns the underlying behavior of the dataset) using a population of 35,212 medicine inpatients (hospital encountering) hospitalized from 05/01/2016 to 12/31/2017, and was tested on medicine inpatients hospitalized during January 2018 (01/01/2018 ~ 01/31/2018), with a population size of 2,674. The use of a point in time to split the data enables us to mimic the model's behavior as if it were operational: the predictor is trained on the data of medicine inpatients hospitalized in regionalized units until a certain point in time, and used on currently admitted patients. In addition, it allows us to analyze the accuracy of the model for all medicine inpatients hospitalized at the hospital for a given day, and to measure how many discharges the model has correctly captured that day.

5.1.1 ROC Curve

The ROC curve for January 2018 is shown in Figure 14. AUC ROC shows that the prediction power of the current model is 0.739.

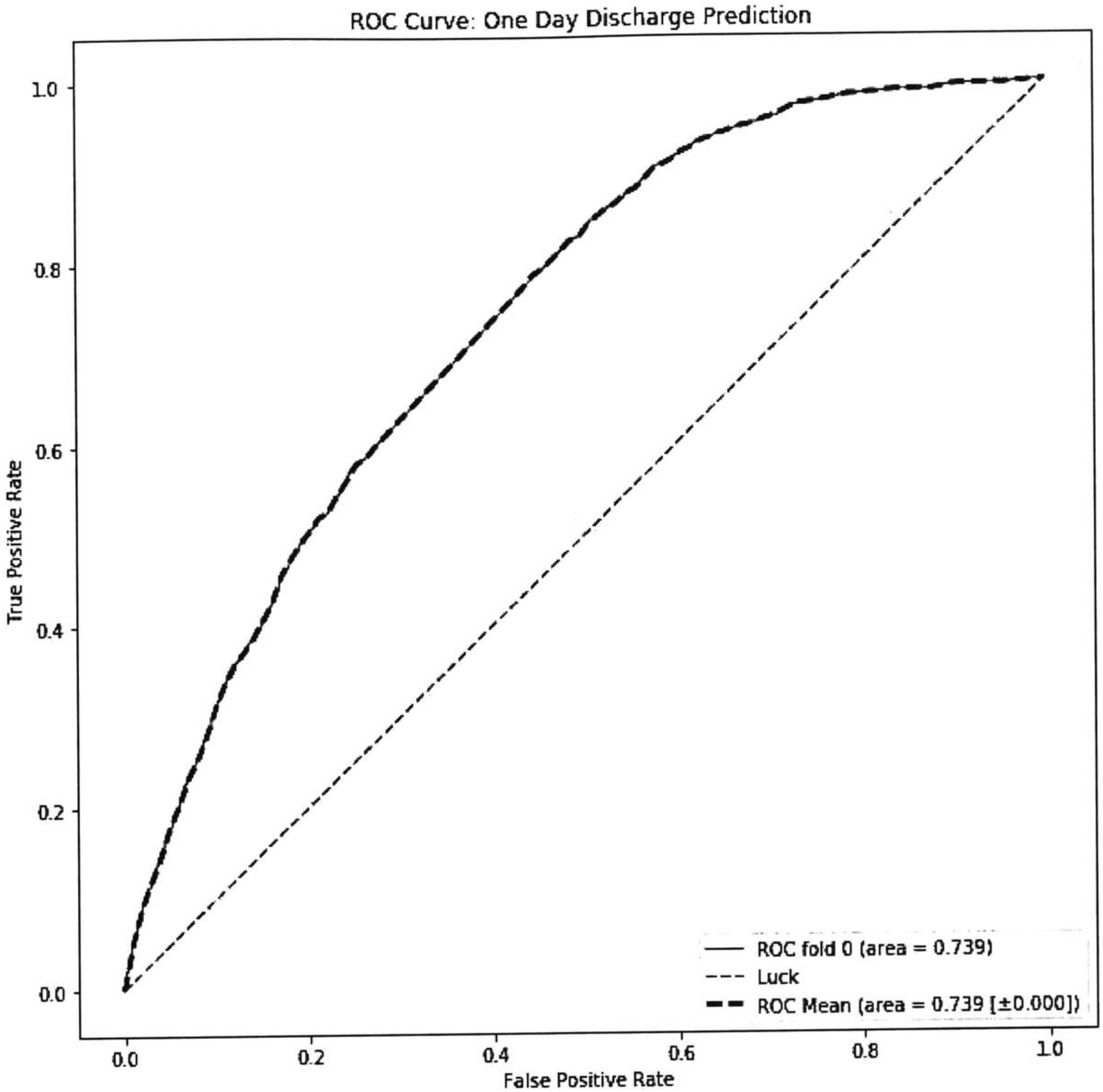


Figure 14 ROC curve for January 2018 medicine inpatients. Timeframe: 05-01-2016 ~ 01/31/2018. Training and validation population: medicine inpatients on regionalized floors.

Note, the monthly prediction of AUC ROC in Zanger's work is 0.857 [9]. The significant difference in AUC ROC for these two populations highlights the key distinction when discharging surgical and medicine patients. Surgical services are mostly pre-planned, elective procedures, whereas the majority of the patients admitted to medicine floors arrive from ED. These patients' discharge plans need to be worked out in parallel to the diagnosis process during their hospitalization at the hospital. Therefore, discharging medicine patients involves more uncertainty and adaptation to dynamic change.

5.1.1.1 Comparison to model without CM notes and with surgical CM notes

To capture the dynamic nature associated with discharging medicine patients, we rely heavily on the information recorded in case management (CM) notes. In 5.1.1.1, we compare the performance of the current model to that without CM notes, and to the one using key phrases from surgical case management notes (note, since our model is extended based on the surgical model, at first, we did not improve surgical CM notes, but later we realized that medicine CM notes contain more signaling information on patients' discharge readiness, therefore we extended the phrases based on text processing of medicine case management notes). *Table 5* shows the model performance (AUC ROC) in these three scenarios.

Model Description	*AUC ROC*
Model A - With medicine CM notes	0.739
Model B - With surgical CM notes (see note above)	0.731
Model C - Without any CM notes	0.721

Table 5 Model performance comparison under three scenarios: 1) model with CM notes tailored to medicine inpatients 2) model with surgical CM notes; 3) model without any CM notes. AUC ROC is based on January 2018 medicine inpatients validation set, and it is used to compare model performance in these three scenarios.

Comparing Model A with Model C, one can see an improvement in model performance of ~ 2%. This demonstrates the amount of information that CM notes contain that can lead to signaling of discharge readiness. Despite a smaller improvement, Model A, with more pertinent information for medicine patients performs better than the model with surgical CM notes.

5.1.2 Daily Precision and Sensitivity Rates

AUC ROC is an effective indicator for a model's performance, however, it is very general. A more specific understanding on a model's performance on a daily basis can facilitate daily operation management, therefore our research has also provided two other approaches to evaluate the model performance - 1) Top N ranked list; 2) Threshold based list.

5.1.2.1 Model Accuracy by Daily Top N Ranked List

The top N ranked list is generated based on the score associated with each patient's daily discharge likelihood. Patients whose scores are amongst the top N highest are predicted to be discharged the next day. A top 10 ranked list, for example shown in Figure 15, is a list that is compiled based on the top 10 scores of all medicine inpatients during weekdays in January 2018. Similarly, Figure 16 shows the top 30 ranked list of medicine inpatients in January 2018. These results are generated from January 2018 model (test set:

regionalized medicine inpatients hospitalized during 05/01/2016 ~ 12/31/2017; validation set: regionalized medicine inpatients hospitalized during 01/01/2018 ~ 01/31/2018) Our model's accuracy is derived by comparing the actual discharge date of the patient with the predicted discharge date. A true prediction is shown in green. It means that the predicted discharge date is exactly the same as the actual discharge date, whereas blue means that the patient is discharged in the next 48 hours and red means the patient is not discharged with 48 hours of the predicted date. Based on Figure 15 and Figure 16, Table 6 summarizes the percentage of discharges captured in the next 24 hours, the next 24 hours to the next 48 hours, and in the next 48 hours for January 2018's prediction for medicine inpatients, given the top 10 and top 30 criteria.

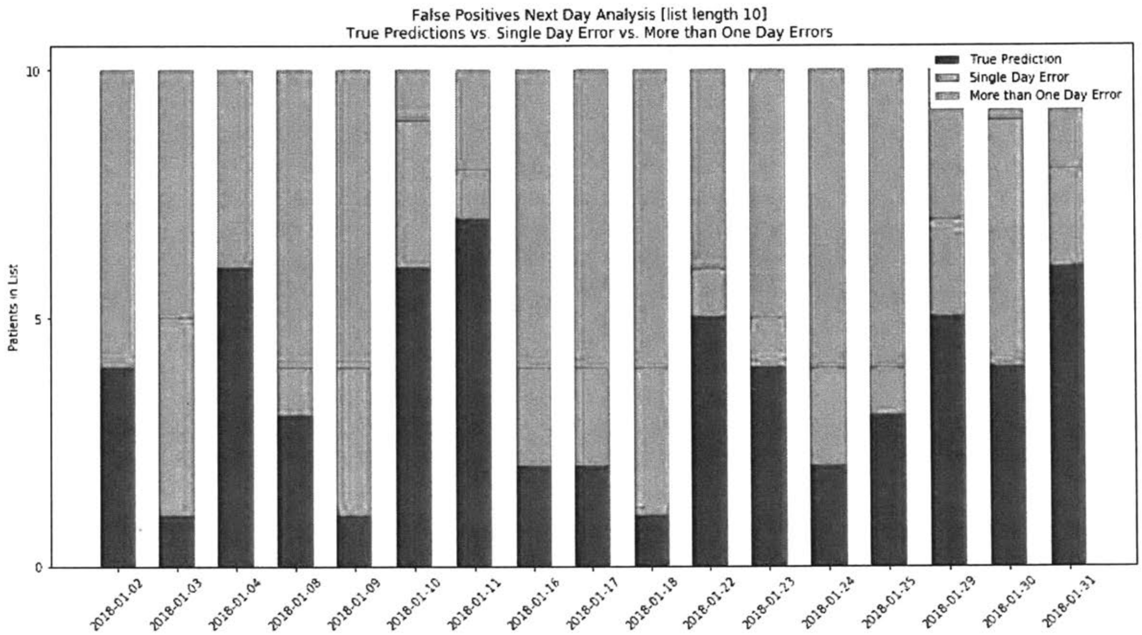


Figure 15 Model accuracy based on Top 10 list. January 2018, weekdays

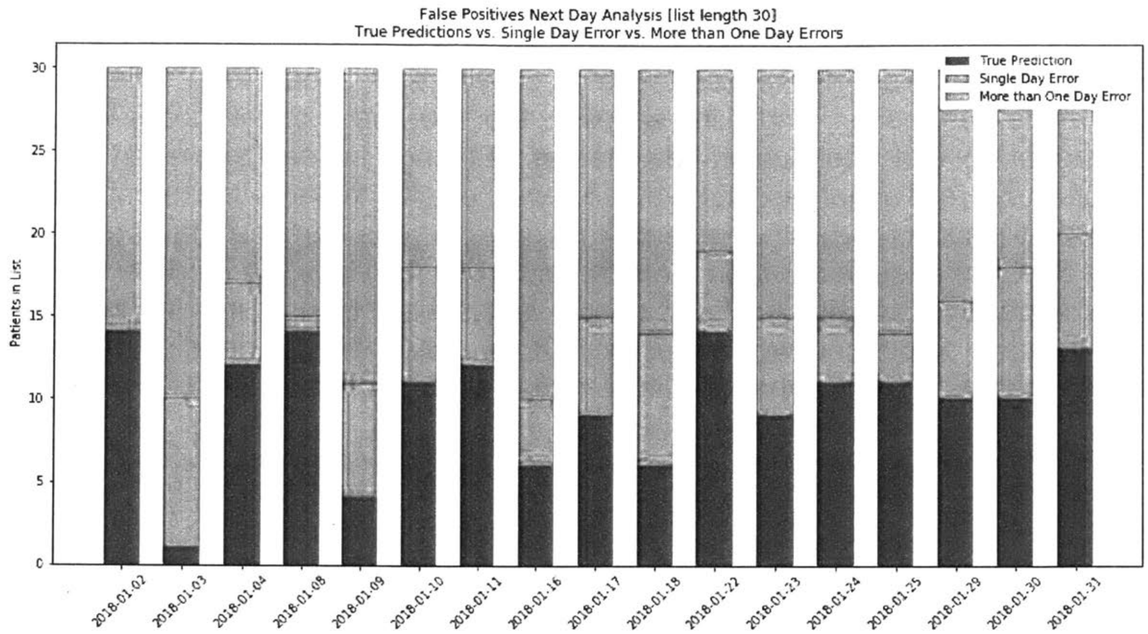


Figure 16 Model accuracy based on Top 30 list. January 2018, weekdays

Top N Criteria	*Discharges in the next 24 hours (%)*	*Discharges in the next 24 to 48 hours (%)*	*Discharges in the next 48 hours (%)*
Top 10	33.1%	24.6%	57.8%
Top 30	30.6%	18%	48.6%

Table 6 Model's capability to capture discharges in the next 24 hours, 24~48 hours, and in the next 48 hours, using both Top 10 list and Top 30 list criteria

From Figure 15 and Figure 16, one can see that there exists a level of variability in our model to capture correct prediction of discharging medicine patients. We believe this is consistent with the fact that medicine patients are more heterogeneous in their clinical and psycho-social conditions, which could cause inconsistency in discharge predictability.

Table 7 shows the average percentage of captured medicine patients discharges with respect to total number of discharges in January 2018, using both Top 10 and Top 30 criteria.

Top N Criteria	*Average captured discharges of total medicine patients' discharges (%)*
Top 10	14.3%
Top 30	33.7%

Table 7 Average captured discharges of total medicine patients' discharges (%), using both top 10 list and top 30 list

Figure 17 demonstrates the average percentage of captured discharges of medicine patients as the list length increases. Given the complexity of medicine patients, the percentage of captured discharges tends to converge at about 70%, whereas in the surgical model, as the list length increases to top 100, the surgical model was able to capture about 95% of total surgical discharges. Although increasing the list length can help capture more discharges, the trade-off we need to make is the number of false positives that incur as the list lengthens. In Figure 18, one can see that the number of false positives increase faster than the number of true positives. As a consequence, the ratio of true positives to false positives decreases as the list lengthens: in Figure 19, the true positives to false positives ratio drops to ~ 35% at list length of 100 patients comparing to that of ~ 40% when the list length is at 20 patients.

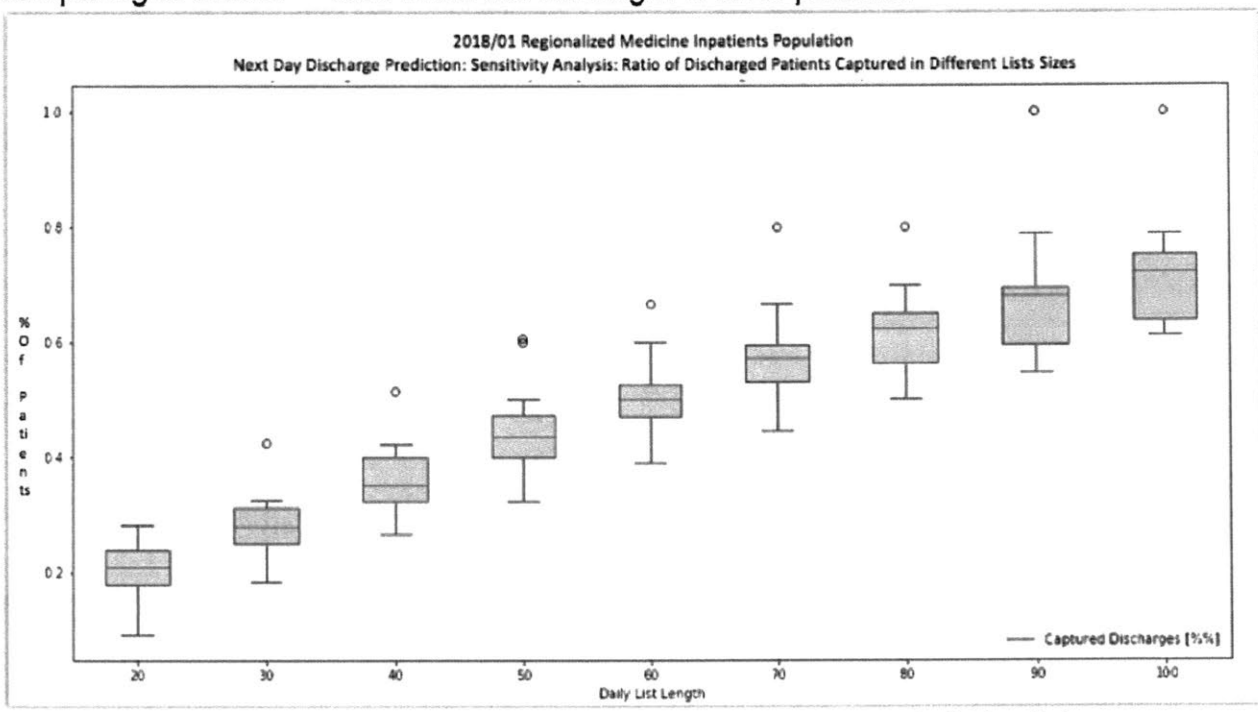


Figure 17 Next day discharge prediction, model's captured discharges as the list length increases

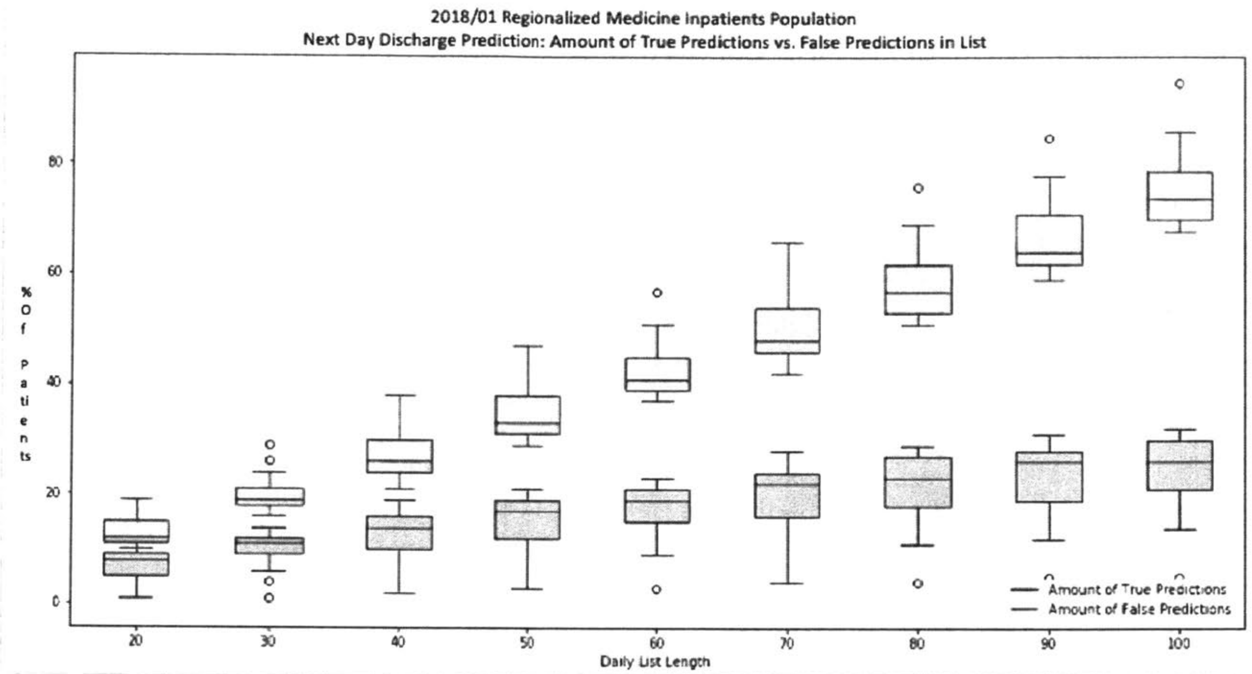


Figure 18 Amount of true predictions and false predictions as the list length increases

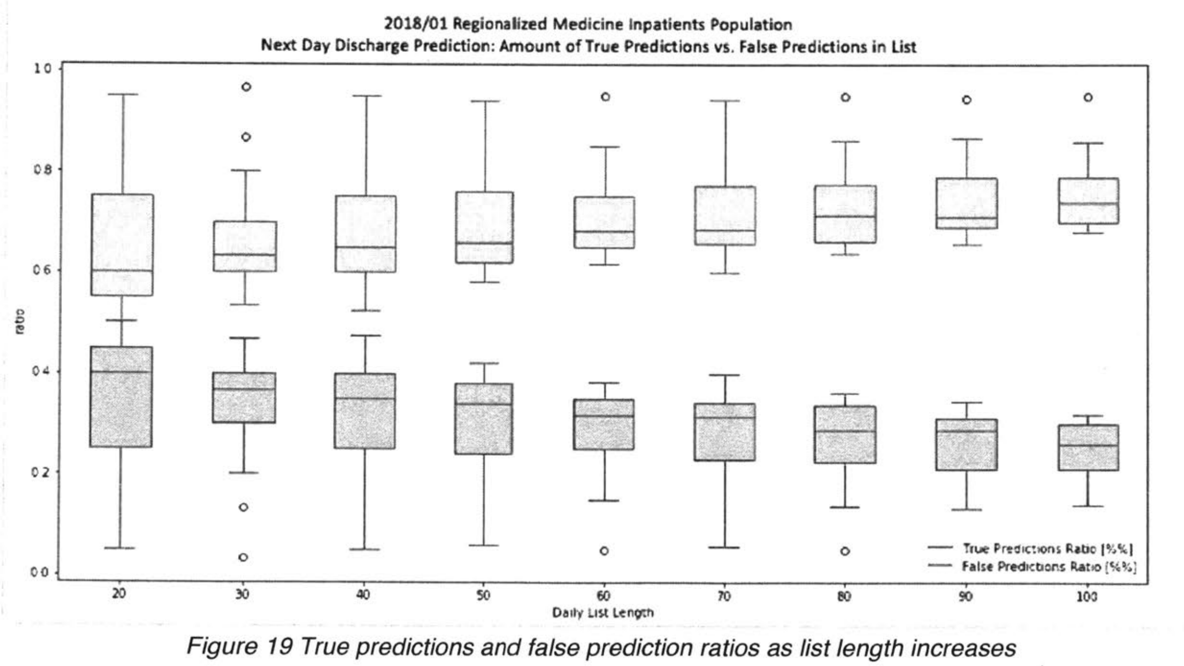


Figure 19 True predictions and false prediction ratios as list length increases

5.1.2.2 Model Accuracy by Threshold

Another way to evaluate model performance is to measure the model accuracy (the score calculated by the algorithm) against a chosen absolute threshold. To be consistent with the criteria used in surgical model performance evaluation, we continue to choose 0.5

and 0.7 as our two criteria to decide whether a patient is going to be discharged in the next 24 hours. For example, if a patient's score is higher than the 0.5 threshold, we classify the patient's discharge as occurring in the next 24 hours, otherwise, the patient continues to stay. Figure 20 and Figure 21 shows the model accuracy using threshold values 0.5 and 0.7. However, these two criteria are not the best fit to measure the model performance for medicine patients because there is a higher level of uncertainty in correctly capturing the medicine patients' prediction. A lower level of confidence level might be needed.

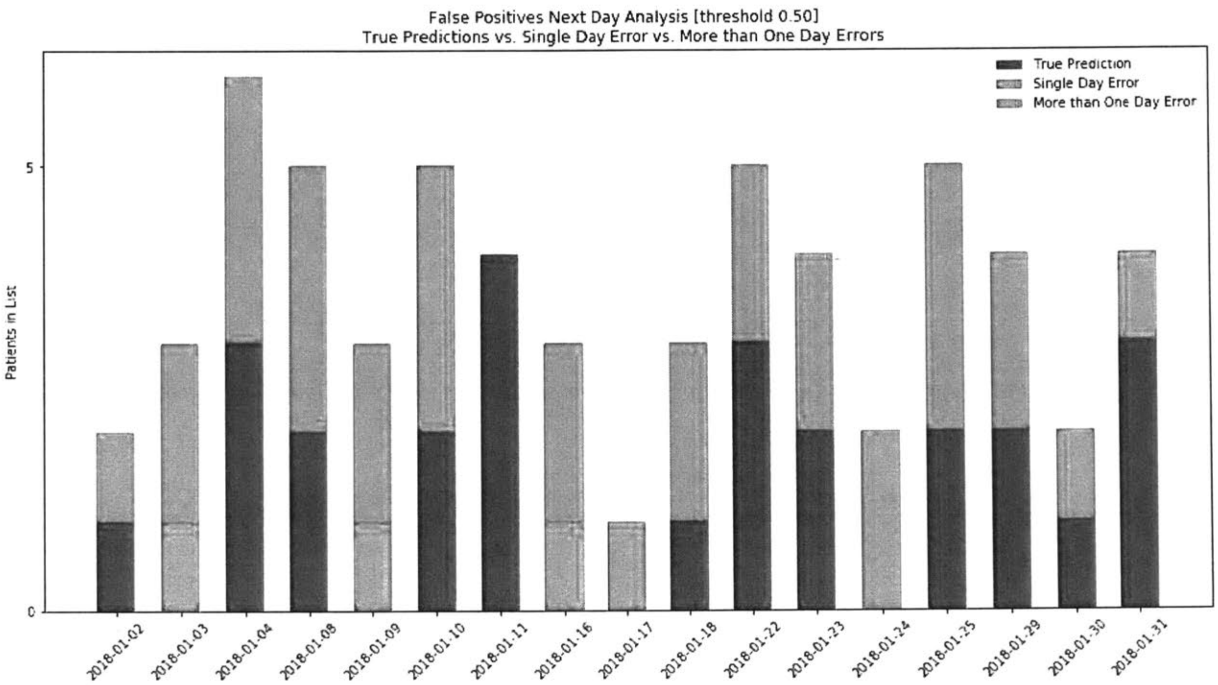


Figure 20 Model accuracy using a score threshold of 0.5: January 2018, weekdays

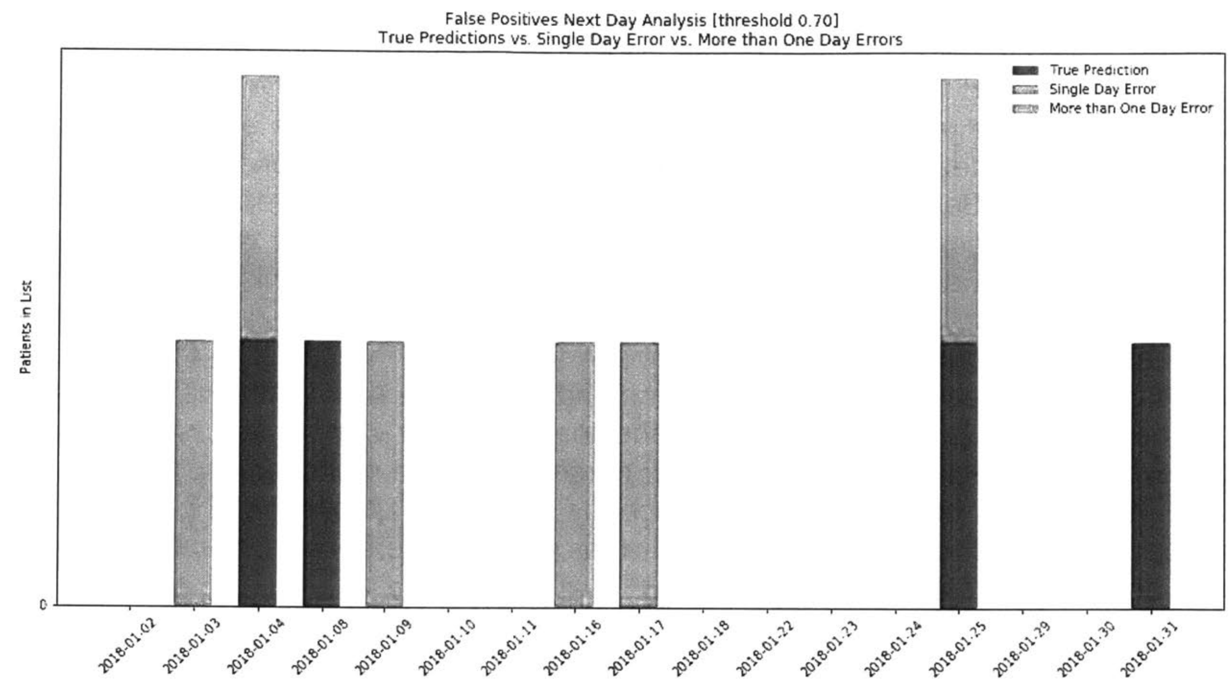


Figure 21 Model accuracy using a score threshold of 0.7: January 2018, weekdays

5.2 Model performance extended evaluation

5.2.1 Model performance with respect to discharge location

Different from surgical services where most surgical procedures are pre-planned and elective, the majority of the DOM patients arrive from the ED. In the context of patient discharge planning, medicine patients' disposition location bears more uncertainty given there is no prior information of the patients' insurance, family preferences or post-acute care settings acceptance. Therefore, discharge planning has to begin after the patient is admitted into the hospital and is led by case management team. As mentioned in Chapter 3, significant amount of time efforts is required for a case manager to develop a cohesive discharge plan. Coordination and communication with multiple parties can be a formidable task to manage and is subject to more uncertainty in discharging given the number of parties involved. One hypothesis is that the more coordination needed to arrange patients' discharge poses more potential delay to the process. We plotted model performance with respect to various disposition locations. Figuring out which location to discharge a medicine inpatient requires the most efforts and it is the primary concern especially in the context of discharging - medicine inpatients usually do not have a planned discharge location when they are admitted; in contrast, surgical inpatients usually have pre-planned discharge locations before their surgeries. Only when a disposition location is secured can a patient be ready for discharge. Again, we use the data from July 2017 (07/01/2017~07/31/2017). Typical discharge locations of General Medicine patients are 'Home or Self Care', 'Home Health Services', 'Skilled Nursing Facilities', 'Rehab Facilities', 'Long Term Care', 'Psychiatric Hospital', 'Short Term Hospital', 'Left Against Medical Advice', and 'Hospice'. Other than 'Home or Self Care', 'Home Health Services', in this research, we refer the rest of the above-mentioned locations as 'Facilities'. The accuracy of the model prediction is plotted against these three categories - 'Home or Self Care', 'Home Care Services', 'Facilities'. Shown in Figure 22, the current model performs the best in the category of 'Home or Self Care'. For the month of July 2017, patients that are discharged to 'Facilities' are the most challenging to predict. This result, while not counterintuitive, suggests a certain level of correlation between the amount of coordination it takes to discharge a patient and the difficulty of predicting the discharge. The less parties involved, the less clinical and administrative complexity the patients are, the better the model's performance is. In addition, this result is consistent with interviews with key stakeholders such as physicians, nurses and case managers. In many cases, discharge withhold or postponement depends on the availability of the facilities. In addition, it also depends on other factors such as patient preferences, patients' healthcare needs and insurance approvals.

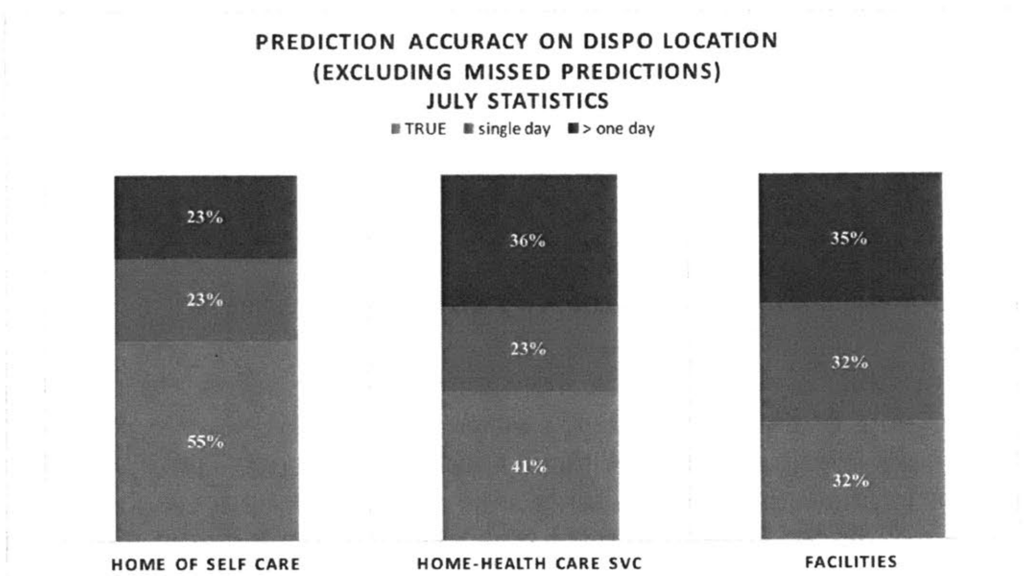


Figure 22 Model performance with respect to patients' discharge location, July 2017 data (07/01/2017 ~ 07/31/2017) – Model C

Based on this insight, our research further evaluates the model performance (with medicine CM notes) with respect to patients' discharge locations. In one group, we grouped the patients who have been discharged to locations as 'Home or self care' and 'Home-health care services'; in the other, we grouped the patients who have been discharged to 'Facilities'. Table 8 shows the training and validation sample size of each group.

	Home or self-care, Home - health services	Only to Facilities
Training set sample size	25,010	9,310
Validation set sample size	1,718	908
Model Performance	0.742	0.727

Table 8 Training set size, Validation set size, and AUC ROC of two models: 1) models with patients discharged to home or home care services; 2) model with patients discharging to facilities – restricted subpopulation based on Model A.

5.2.2 Shorter LOS - Model performance up to certain range

During in-person interviews and Long LOS meetings, a consistent observation of General Medicine population is that the LOS of these patients has a wide range and can last as long as several months. Figure 23 below shows 07/2017 patients' LOS in all regionalized units. The background histogram shows a large range of patients' LOS from as short as two days to as long as over several months. We plotted the correction prediction of next-day patients' discharges against the overall LOS. On the left-hand side, a high degree overlap can be observed between the model's correct prediction and its prediction that has missed one day for shorter LOS. On the right hand plotted is the predictions that model has missed for more than two days. As one can see, the model's failure rate increases as LOS increases.

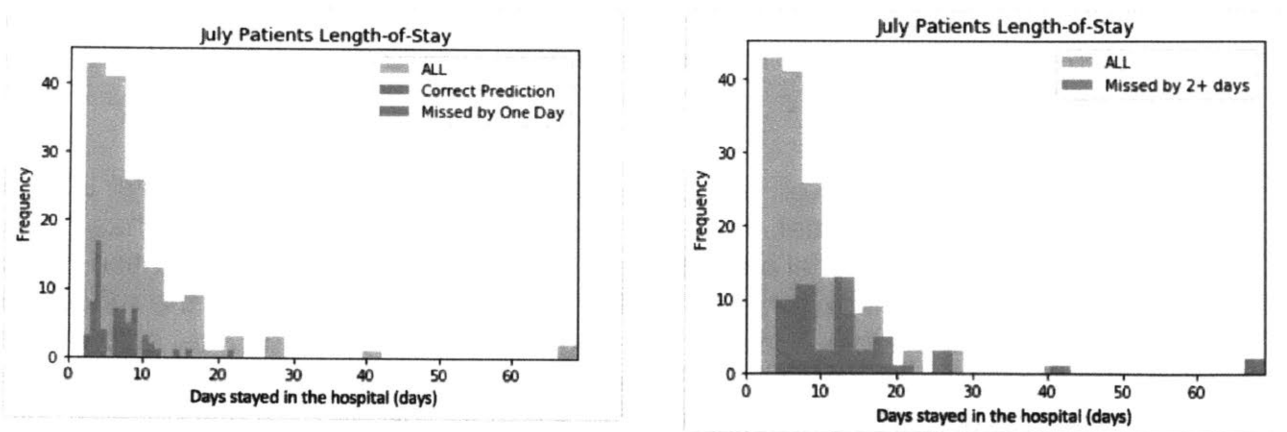


Figure 23 A histogram of LOS of medicine inpatients on regionalized floors. Time frame: 07/01/2017 ~ 07/31/2017

Furthermore, to identify the clear breakdown of the model's with respect to patients' LOS, Figure 24 demonstrates the prediction accuracy on patients' LOS at different ranges: 0-3 days, 4-11 days and 12+ days. The prediction accuracy is broken down into three categories: 100% accuracy, model missed patients' discharge by exactly 24 hours and model missed patients' discharges by 2 days or more. Although it is desirable to have 100% accuracy, some stakeholders pointed out '1 day missed' is an acceptable result for

General Medicine patient population, given the high variability in the discharge process. For the month of July 2017, we can see that the model performs the best when predicting patients with 0 ~ 3 days, 64.7% accuracy; if the prediction result missed by one day is acceptable, then the model is capable of capturing patients' discharges 100% in the LOS from 0 ~ 3 days period (note, this is ad-hoc accuracy check, instead of prediction). However, as LOS increases to twelve days and more, in the month of July 2017, the current model is only capable of capturing 11.4% discharges, and 80% discharges have been missed by more than two days.

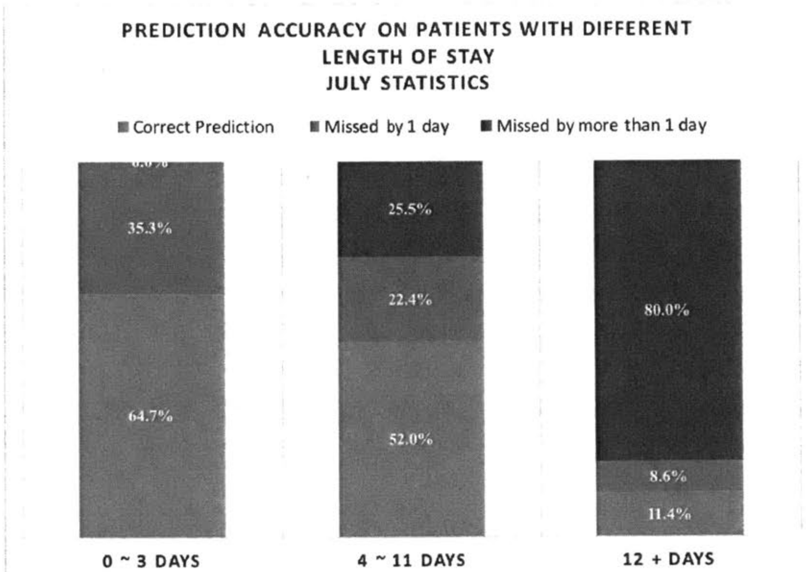


Figure 24 Model performance with respect to LOS: Model performs the best when patients' LOS is between 0~ 3 days. Model performance decreases as patients' LOS increases.

Based on analysis conducted on model performance with respect to patients' LOS, in which the model performs better on patients with short LOS, our research has further extended the analysis into model prediction power at various LOS. The motivation behind this part of analysis is to look for a span of patients' LOS during which the model can generate useful and reliable operational suggestions. By adjusting the algorithm to separate longer LOS from shorter LOS, the model could be useful to reliably predict patients with shorter LOS, therefore providing a tool for the Bed Admitting Department and providing a robust way for healthcare teams to manage bed availability within certain LOS range.

Currently, the results from section 5.1.1.1 are trained and tested on the entire spectrum of patients' LOS (from the minimum LOS to maximum LOS). The following results demonstrated model prediction power at various LOS. The model is trained and tested for patients LOS up to a certain day. For example, if the model predicts patients' discharges up to 7 days, this simply means that we take patients' healthcare data up to their 7th day's stay at the hospital, and train the model using these data; we then test against the patients with their LOS up to 7 days. The implication of this method is that this model can be used for predicting patients' discharges up to patients' X^{th} day's stay at the hospital so that the model can provide high fidelity prediction within a certain range, and stop employing the model thereafter. From the algorithm development perspective, it is also beneficial to improve the model on a narrower scope and expand to its complete complexity.

We evaluated model performance (validation set accuracy, January 2018 model A) with respect to patients' LOS up to 3 days, 5 days, 7 days, 12 days. Table 9 summarizes the findings on model prediction power. The first row 'All LOS - Full Model' represents the baseline, where the model is trained on the full range of patients' LOS. As one can see, as LOS decreases, the model prediction power also increases – from 0.746 for up to 12 days' LOS to 0.799 for up to 3 days' LOS. In addition, one can see that the model without CM notes performs consistently worse. This result again proves that the information from medicine CM notes can signal discharge readiness, and that they may have greater relative added value for longer LOS patients.

Model Description	*Model Performance Without CM notes*	*Model Performance With Medicine CM notes*
All LOS - Full Model	0.721	0.739
LOS up to 3 days	0.786	0.799
LOS up to 5 days	0.770	0.778
LOS up to 7 days	0.750	0.772
LOS up to 12 days	0.723	0.746

Table 9 Model performance with patients LOS up to X^{th} day – restricted to subpopulation of Model A. Medicine CM notes are used in this model. AUC ROC is used to compare the model performance in various scenarios. Time frame: 05/01/2016 ~ 01/31/2018

This page is intentionally left blank

6 Conclusions and Recommendations

6.1 Conclusions

Despite the significant room to further improve the model's performance, the preliminary development of the prediction tool used for predicting medicine patients' discharges has the potential to improve the following areas of the current operations:

1. **Proactive bed management**
   The prediction tool has an output that is a list of patients with ranked order of their discharge likelihood in the next 24 hours. This knowledge of anticipation allows the bed managers at the Admitting Department to react more proactively when assigning beds to incoming patients. The daily bed availability list provides an overview of potential open beds, and it helps for a more optimal way of bed assignment, given more information is available to the bed managers.
2. **Centralized and transparent capacity management with clear priority guidelines**
   The prediction tool uses a uniform decision-making process to predict who are the patients that are going to be discharged in the next 24 hours. This uniformity minimizes the built-in variation in current care teams' decision-making processes, which is brought in by various team structures, bed resources distribution and experience level, etc. The list is ranked using the patient's score, therefore, it also provides a clear prioritization guideline on which patient needs additional efforts from care member to ensure a successful, on-time discharge.
3. **Effective communication among care team members**
   A successful discharge requires excellent coordination among any care teams – physicians communicate the most up-to-date clinical barriers-to-discharge with nursing, and case management team. Currently, this is done using email distribution to communicate barriers-to-discharge with the rest of the floors, and they are not as accurate and as uniform as the prediction tool can offer. Based on such information, case managers can develop a suitable discharge plan for the patient accordingly and adjust timely when a patient's clinical condition changes. Case managers also need to communicate with physicians on the availability/limitation from the post-acute care settings to ensure physicians are properly aware and adjust treatment plan. The amount of communication is significant considering the number of patients each team has to care for. The prediction tool outputs a collection of all possible barriers-to-discharge on a daily basis. By reviewing this feature, each member on the care team can clearly see the updated patients' condition and ensure the communication is on the same page.

6.2 Recommendations

During this study, we have identified a different set of challenges associated with discharging medicine patients. The psycho-social factors could potentially delay a patient discharge, regardless of whether a patient is medically cleared and ready to be discharged to a post-acute care setting. On top of the psycho-social factors, another bottleneck that has been unveiled is the order wait time after it has been placed. The following recommendations have the potential to improve the current process:

1. More anticipation when developing discharge plans: Often times there exists a lag caused by awaiting to hear back from various possible discharge locations, with the updated daily list of the patient, case managers can start to look for the most likely discharge locations ahead of time. By doing so, it has the potential to reduce the wait time of an approval process.
2. More coordination between different units: With clear prioritization rules, the discharge list output by the model can also be used by various other units, such as radiologists performing imaging orders, consult services, lab technicians processing lab tests. By doing so, it can help various other units to prioritize which patients' orders, lab tests or consult services should be processed first so that the patient can be discharged. The wait time for an order/lab test result to come back can be minimized if acting upon the prioritization guidelines.

In general, we recommend the study to be further refined so that the model accuracy can increase. For example, one can further investigate the missing clinical milestones and barriers to discharge to ensure complete features are incorporated; or one can expand to diagnoses beyond the top 40%. In addition, after improving the model, one can also derive a more accurate estimation on how many patients will be discharged the next day. After achieving a satisfactory performance, the hospital can consider a small-scale pilot test to deploy the prediction tool, and study how the tool can benefit care teams as well as other areas of improvement.

Reference

[1] E. Ugarph, “Patient Flow Optimization in the Department of Medicine.”

[2] “About the Department Of Medicine.”

[3] B. Zimmerman, “Mass General forms task force to combat ED overcrowding.”

[4] “Assessing occupancy at the hospital: Meet the Capacity Committee.”

[5] K. Conti, “US hospitals continues to grapple with patient overcrowding.”

[6] J. Zanger, “Patient Discharges Prediction for the Surgical Department at the hospitals.” 2018.

[7] S. Trzeciak and E. P. Rivers, “Emergency department overcrowding in the United States: an emerging threat to patient safety and public health,” *Emerg. Med. J.*, vol. 20, no. 5, pp. 402–405, Sep. 2003.

[8] J.S.Carroll, J.W. Rudolph, “Design of high reliability organizations in health care” Qual Saf Health Care. 2006 Dec; 15(Suppl 1): i4–i9

[9] G. J. Bazzoli, L. R. Brewster, G. Liu, and S. Kuo, “Does U.S. Hospital Capacity Need To Be Expanded?,” *Health Aff. (Millwood)*, vol. 22, no. 6, pp. 40–54, Nov. 2003.

[10] N. C. Proudlove, K. Gordon, and R. Boaden, “Can good bed management solve the overcrowding in accident and emergency departments?,” *Emerg. Med. J.*, vol. 20, no. 2, pp. 149–155, Mar. 2003.

[11] P. R. Harper and A. K. Shahani, “Modelling for the planning and management of bed capacities in hospitals,” *J. Oper. Res. Soc.*, vol. 53, no. 1, pp. 11–18, Jan. 2002.

[12] “Enterprise-wide hospital bed management dashboard system,” 13-May-2008.

[13] A. De Grood, K. Blades, and S. R. Pendharkar, “A Review of Discharge-Prediction Processes in Acute Care Hospitals,” *Healthc. Policy*, vol. 12, no. 2, pp. 105–115, Nov. 2016.

[14] C. W. Van and A. J. Forster, “The TEND (Tomorrow’s Expected Number of Discharges) Model Accurately Predicted the Number of Patients Who Were Discharged from the Hospital the Next Day.,” *J. Hosp. Med.*, pp. E1–E6, Aug. 2017.

[15] S. Rose, D. Engel, N. Cramer, and W. Cowley, “Automatic Keyword Extraction from Individual Documents,” in *Text Mining: Applications and Theory*, pp. 1–20.

[16] F. Pedregosa *et al.*, “Scikit-learn: Machine Learning in Python,” *JMLR*, no. 12, p. 2825–2830, Oct. 2011