March 2016

Simulation of 48-Hour Queue Dynamics for A Semi-Private Hospital Ward Considering Blocked Beds

Wei Chen
University of Massachusetts Amherst

SIMULATION OF 48-HOUR QUEUE DYNAMICS FOR A SEMI-PRIVATE HOSPITAL WARD CONSIDERING BLOCKED BEDS

A Thesis Presented

By

WEI CHEN (CAMI)

Submitted to the Graduate School of the

University of Massachusetts Amherst in partial fulfillment

of the requirements for the degree of

MASTER OF SCIENCE IN INDUSTRIAL ENGINEERING AND OPERATIONS RESEARCH

February 2016

Department of Mechanical and Industrial Engineering

SIMULATION OF 48-HOUR QUEUE DYNAMICS FOR A SEMI-PRIVATE HOSPITAL WARD CONSIDERING BLOCKED BEDS

A Thesis Presented

By

WEI CHEN (CAMI)

Approved as to style and content by:

Hari Balasubramanian, Chair

Chaitra Gopalappa, Member

Joan Roche, Member

Sundar Krishnamurty, Department Head
Department of Mechanical and Industrial Engineering

ACKNOWLEDGMENTS

Foremost, I would like to express my sincere gratitude to my advisor Professor Hari Balasubramanian for the continuous support of my Master's study and research, for his patience, motivation, and immense knowledge. His guidance helped me in all the times of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Master's study. Besides my advisor, I would like to thank the rest of my thesis committee: Professor Joan Roche, and Professor Chaitra Gopalappa for their encouragement and insightful comments.

My sincere thanks also goes to Doug Neal, for offering me the 6-month internship opportunity in his Process Improvement group, empowering me to shadow processes and to build collaboration with interdisciplinary teams at Baystate Medical Center. And special thanks to Patricia Samra, RN, for patiently and persistently supporting my thesis development, offering valuable resources and giving me reality checkups.

Besides, I would like to thank everyone in the Baystate-UMass computer modeling team including Patricia Samra, Doug Neal, Joan Roche, Alicia Meacham, Haiping Li, and Joe Schmidt, Mike Ehresman for sparing time to review my presentations and stimulating interesting discussions. I am also grateful of the emotional support and inspirations I received during challenging times from my former colleagues in Process Improvement under the Division of Health Care Quality, including Dave Hansen, Jessica Fraga, Mary-Ryan Kusiak and Tom Vega. Last, I would like to thank my families and my partner for supporting me spiritually throughout this journey.

ABSTRACT

SIMULATION OF 48-HOUR QUEUE DYNAMICS FOR A SEMI-PRIVATE HOSPITAL WARD CONSIDERING BLOCKED BEDS

February 2016

M.S.I.E.O.R., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by Professor Hari Balasubramanian

This thesis study evaluates access to care at an internal medicine unit with solely semi-private rooms at Baystate Medical Center (BMC). Patients are divided into two classes: Type I patient consumes one bed; Type II patient occupies two beds or an entire semi-private room as a private space for clinical reasons, resulting in one empty but unavailable (blocked) bed per Type II patient. Because little data is available on blocked beds and Type II patients, unit-level hospital bed planning studies that consider blocked beds have been lacking. This thesis study bridges that gap by building a single-stream and a two-stream discrete micro-simulation model in Excel VBA to describe unit-level bed queue dynamics at hourly granularity in the next 48-hour time horizon, using historical arrival rates and census-dependent discharge rates, supplemented with qualitative results on complexity of patient-level discharge prediction. Results showed that while we increase additional semiprivate beds, there was notable difference between the traditional single-stream model and the two-stream model concerning improvement in bed queue size. Possible directions for future research include patient-level discharge prediction considering both clinical and nonclinical milestones, and strategic redesign of hospital unit(s) considering overflows and internal transfers.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Background

Many hospitals are under tremendous pressure to manage discrepancy between capacity and demand, and to balance quality of care and patient throughput. While the overall U.S. hospital capacity has been static or declining, the total hospital admissions have increased or remained the same. The American Hospital Association (AHA) Annual Survey revealed that total U.S. hospital beds decreased from 920.8 thousand in 2012 to 914.5 thousand in 2013. Meanwhile, total admissions in all U.S. registered hospitals increased 2%, from 35.4 million in 2012 to 36.2 million in 2013. American College of Emergency Physicians (ACEP) disclosed that Emergency Department (ED) admissions take up roughly 81.8 % of unscheduled hospital admissions, a sharp increase from 64.5 % in the last decade. As the number of unscheduled admissions has increased sharply, hospitals today are more than ever challenged to ensure timely access to inpatient care.

U.S. hospital overcrowding is increasingly prevalent due to a combination of factors, including declining total U.S. hospital beds, clinical labor shortage and increased usage of hospital emergency departments (ED) as a medical safety net. Hospitals become congested and stress when they encounter more patient arrivals than what they were designed to handle. Hospital congestion may result in long bed queues and deteriorating quality of care. In general, hospital clinicians have some insights about the system dynamics concerning bed crisis, but many lack a computer tool to alert potential bed crisis in advance or to test bed planning alternatives before implementing any changes.

Additionally, there is great variability among hospitals in terms of bed capacity, ward design and patient mix. This makes it difficult to find a common solution that is suitable for a wide range of healthcare systems.

1.2 Motivation

Motivated by hospital and emergency room overcrowding, this thesis is an empirical study that looks at bed planning challenges at Baystate Medical Center (BMC), which is a 716-bed tertiary-care teaching hospital in Springfield, Massachusetts. BMC has the second-busiest emergency department in Massachusetts. Overall, BMC saw a total of 33,944 annual admissions. Daily mean number of patients that were admitted through the ED was 81 in 2014. According to the U.S. National Hospital Ambulatory Medical Care Survey (NHAMCS) 2010, the percentage of ED visits resulting in hospital admission was 13.3% on average. However, BMC saw approximately 27.8% ED patient visits resulting in hospital admission in fiscal year 2014. BMC is facing the challenge of providing patients who are admitted through the ED with timely access to inpatient beds.

For some BMC general medicine units, over 90% of the patients admitted came through the ED. Current BMC bed board system does not provide bed queue prediction for the next two days. The hospital desires more robust unit performance evaluations and short-term prediction of bed queues to alert surge capacity planning for adult medicine units so that clinicians can be more proactive about future stress scenarios.

1.3 Problem Description

Let MU denote an acute-level medicine unit. MUs are downstream acute care units that are usually patients' last stop before they leave the hospital. Looking at the big picture, MU congestion may lead to increased number of patients waiting for MU beds while consuming other beds in the upstream care units such as the ED or Intensive Care Units (ICU). See the figure below for patient flows concerning acute care units.

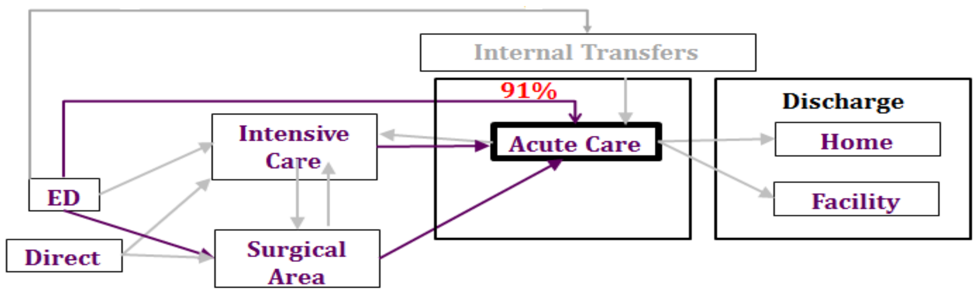


Figure 1 Hospital Patient Flows Concerning Acute Care Units

MU encounters patients of many different ages, with a wide variety of diagnoses and sometimes complex social needs. A majority of the patients have admitting diagnoses that do not require surgical intervention. The wide variety of admitting diagnoses may include but not limited to pneumonia, COPD, mini-stroke, tuberculosis, gastrointestinal complaint, or altered mental status. Every day, the case manager of a medicine unit is challenged to provide care coordination for adults with complex care needs and social needs. This thesis study focused on modeling one particular adult medicine unit, because this unit had not changed significantly in terms of patient mix, patient volume, bed capacity or ward design in the past five years.

The chosen MU in this study has 34 shared beds or 17 semi-private treatment rooms. Special circumstance arises when a patient uses a semi-private room as a private space due to clinical reasons such as end-of-life comfort measures, behavioral health issues, or having diagnosis in infectious disease. As a result, the other bed in this semi-private room will be temporarily blocked and is unavailable to new admissions. This phenomenon contributes to the variability in MU performance. A graphical representation of a MU is displayed below.

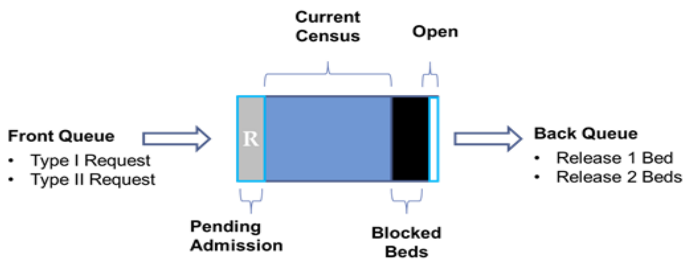


Figure 2 Graphical Representation of a MU

In the figure above, white, grey and black colors represent beds that are available, occupied or blocked respectively. The timestamp for arrival is the time that a patient completes the admission process. The timestamp for discharge is the time that a patient leaves the unit. Here are some key operational definitions: Type I patients do not use semi-private rooms as privates; Type II patients do resulting in blocked beds. The selected MU for this study only has shared rooms. Type II patients will in reality consume 2 beds, using shared rooms as privates resulting in a blocked bed per Type II patient. Patient census represents the actual number of patients in a unit. Blocked Beds are the number of empty but unavailable beds associated with Type II patients. Open Bed is the number of beds available for new admissions.

For the MU in this study, average patient length of stay (LOS) was approximately 5 days. Between May 2010 and April 2011, patient admission source for this MU by ED, Direct Admit and Other was roughly 91%, 8% and 1% respectively. During period of time, a total of 1,963 patients were discharged from this unit. Besides, a total of 141 patients were admitted to this unit but transferred elsewhere. The chosen MU has 34 shared beds or 17 semiprivate treatment rooms. On average it had 5.37 discharges per day. Average midnight census in year 2013 was roughly 30.3 patients. Patient occupancy was around 89%. Average number of beds that were empty but unavailable ranged from 0 to 8 with a mean at 2.8 in 2013 and 2014. In short, on average this MU had 0.9 open beds, while the other 33.1 beds were unavailable for new patients historically.

1.4 Objective

The specific challenge for my thesis is to develop a Monte Carlo simulation of unit-level patient flows to evaluate bed queue dynamics in the near future based on current state of the system. Through qualitative observations, clinician interviews and designing a data-driven micro-simulation, this thesis study aims to produce insights that hopefully guide care coordinators to proactively manage bed resources under uncertainty.

The scope of the study is to convert a bed request simulation, a census-dependent discharge simulation based on one year's historical data and clinicians' expert knowledge into a 48-hour bed queue prediction for a MU at BMC. The focus of my work is on prototyping the decision aid tool for nurses by providing the logic flow to perform this conversion, driven by a discrete micro-simulation model built in Microsoft Excel VBA.

Because there is little data available on blocked beds associated with Type II patients, studies of this phenomenon have been lacking. This thesis study bridges the gap by closely looking at MU considering blocked beds (MUBB), and by comparing the short-term queue dynamics in traditional MU model with the MUBB model.

Specific aims are as follows: (1) to identify the necessary specification for building a prototype bed crisis alert tool; (2) to demonstrate whether a traditional single-stream MU model is sufficient, or a two-stream MUBB model behaves differently and is worthwhile; (3) to test a few simple bed planning schemes under stressed scenarios. Simulation outputs include time-dependent performance measures such as bed queue size and bed utilization.

1.5 Significance

To judge the value in predictive stress testing a hospital unit and in considering blocked beds for medicine units with shared beds, it is necessary to first quantify the value of a hospital bed. According to the U.S. Department of Health and Human Services, national average per diem cost for a hospital bed was $1,817 in year 2009. Conservatively, assuming that fixed cost accounts for 40% of the total per diem cost of a hospital bed, that is roughly $726.8 fixed cost per hospital bed per day. Without considering seasonality, suppose that on each day the average number of blocked beds is approximately 3, and then the average yearly cost associated with blocked beds would be $795,846. Yet this amount does not account for the indirect cost of patients boarding in upstream care units such as an ED, an ICU, or a step-down or an intermediate transitional care unit before accessing a bed on a general medicine floor.

To conceptually discuss the indirect cost accrued due to a congested downstream care unit, it is important to acknowledge the interdependency between various hospital units again. Suppose that an upstream care unit such as an ED or an ICU sees increased patient arrivals, many downstream medicine units should expect some increase in patient demand as well. Note that clinically resources are required to admit or discharge a patient in a downstream care unit. Under stressed scenarios, the critical clinical resources might have to be allocated thinly favoring the sickest patients first, indirectly contributing to congestion in the patient discharge process at a downstream care unit. Suppose that for 2 or 3 days a downstream care unit stops discharging patients, the upstream unit such as an ED or an ICU would likely see increased number of patients boarding. Here, boarding patients are defined as those remaining in an upstream care units waiting for open beds or the actual transfer after having been clinically approved to transfer to another unit. Hourly bed queue size is the key unit performance measure examined in this thesis study, and it may capture the number of patients boarding in upstream units.

Unit-level capacity or process challenges may impact hospital-wide patient flows. All in all, it is valuable to analyze hospital downstream unit performance; and it is financially worthwhile to study blocked beds.

CHAPTER 2

LITERATURE REVIEW

2.1 Simulation and Modeling of Hospital Patient Flows

Factors such as time-sensitive decision making, uncontrollable emergency admissions and unique incentive schemes characterize patient flows in hospital queuing network, which differs from traditional customer routing in other service systems. This thesis primarily deals with patient flows through the adult general medicine units as opposed to special wards such as the maternity wards; hence I focus on literature concerning patient flows through general inpatient units. A few past researches modeled hospital-wide patient flows as queuing networks (Armony et al. 2013, Ozen et al. 2014). Armony et al. (2013) utilized Exploratory Data Analysis (EDA) to study the special features of the hospital queuing networks. This research discussed the role of information availability in ED-Inpatient routing decisions and suggested a few areas of future research including how to manage ED admission overflows while fairly distributing case workload among the general inpatient wards, and ways to measure ward workload and overflows as well as its relationship with patient turnover rates.

Ozen et al. (2014) conducted an empirical study that quantifies the impact of various discharge windows to mitigate inpatient bed congestion pertaining to Baystate Medical Center. A hospital-wide simulation of multi-server queuing network was developed considering arrivals of multiple patient classes. Early in the day discharge policy resulted in a lower improve in timely access to inpatient beds than prioritizing discharges in units with longer admission queues. Discharge represents a care transition process that is complex, and understanding the clinical and social context of the discharge process is much needed to design, simulate and further test the discharge prioritizing scheme. And in order to implement a prioritized discharge policy, a unit-level dashboard on queue sizes is needed to provide dynamic prioritization criteria for discharge planning. Additionally, the unscheduled or random arrivals of patients into the ED and inpatient department have mostly been modeled as homogeneous and nonhomogeneous Poisson arrivals (Armony et al 2013, Bhattacharjee & Ray 2014, Ozen et al. 2014).

Helm et al. (2009) built an accessible flexible hospital system-level patient flow framework in C++ with the hope to allow others transforming it into a stochastic simulation model to help managers stabilize occupancy. The study was motivated by the benefits of stabilizing system-wide hospital occupancy, which was said to reduce delays in care and lead to simplified resource planning endeavors. The researcher assumed that 47% inpatient admissions were comprised of emergency department admissions. Historical data from an actual hospital was used to generate two hospitals: one uses front loaded scheduling practice without daily control thresholds, while the other uses level loaded scheduling that allows divergence or cancellations. The results suggested that the latter hospital had superior system performance. In reality, emergency patient arrivals are uncontrollable in large teaching hospitals and may make up a high percentage of all hospital admissions. Each hospital unit or service line may have unique characteristics that may allow or prohibit this type of admission control scheme.

Kim et al. (2014) conducted an empirical study to identify various non-standardized admission strategies that were used to manage patients inflowing to a hospital's ICU. The researcher gathered patient-level dataset of over 190,000 hospitalizations across 15 hospitals, interviewed physicians to gather expert knowledge about performance of several admission schemes, and employed econometric analyses to evaluate patient outcomes of various admission strategies. The goals included reducing readmissions and hospital length-of-stay, and a simulated hospital ICU with 21 beds was used to evaluate 4 admission strategies, some of which had objective component and discretionary components, or threshold levels of admission, or bed capacity and the expected rerouting costs. This study provides a guideline for future research on managing unit-level admission overflows.

2.2 Discharge Simulation

The theoretical study done by Chan et al. was motivated by the phenomenon that patients who are clinically ready for discharge may still occupy a bed for a varying period of time, which may be called 'avoidable days' in hospital practices. The researchers focused on modeling the patient discharge process as a theoretical queuing system with time-varying arrival rates, where the servers represent patients who completed their clinical journeys and are awaiting for 'discharge readiness' inspection by clinicians. Using both theoretical and numerical analysis, the results of this study suggests that optimizing discharge readiness inspection time significantly improve system performance under stressed arrival scenarios; spreading out patient 'discharge readiness' inspections throughout the day with uniform distribution seems to yield good system performance in stability, and patient throughput. This study establishes an ideal state of the hospital system in theory, suggesting discharge planning schemes that may be difficult to implement in reality, but it can serve as an upper threshold of hospital system performance.

Khurma et al (2013) created a simplified simulation model of the discharge process at a regional premier tertiary acute care hospital, aimed to explore hospital discharge planning changes that may reduce discharge delays. The most important problem identified by the researcher was “Awaiting post-discharge facilities”, which contributed to 41% of discharge delay days. Patient length of stay (LOS) was broken down into two elements: acute care days and Alternate Level of Care (ALC) days. First, Khurma identified the patients requiring placement in long term care (LTC) to be the most persistent category of patients contributing to ALC days, and the top ranked medical units sending most patients to LTC. The simulation model mimicked the discharge process. It has five time milestones: Admission – Referral to Social Work (RefSw) – Involvement of Social Work (InvSw) – Completion of Placement Application (Appl) – Discharge (D/C). Sometime in between the Appl-D/C phase, the patient converts from acute care status to the ALC status. The data on the duration of each phase was collected in days and fitted to a probability distribution based on Anderson Darling test in Minitab. Inconsistency was found in the time taken to complete these process intervals Ad-RefSW and RefSW-InvSW. The simulation was run with a sample size of 152 patients and 608 patients. In the current state model, mean times for Ad-RefSW and Ref-InvSW were 3.52 days and 2.18 days. In the future state model, the maximum time allowed for the aforementioned two processes were set to 3 and 2 days. The results showed a statistically significant 4.5-day reduction in the median LOS (from 35 to 30 days) among patients who wait ALC days. In a word, long term care facilities and complex continuing care facilities cause more persistent ALC cases and longer ALC days. It is critical to quantify this delay accurately to help with the discharge predictions and capacity planning (Khurma 2007).

2.3 Health Care Forecasting and Modeling Approaches

Xu & Chan (2014) built a predictive model of Emergency Department (ED) arrivals to of patient arrivals to the Emergency Department (ED) with the hope to help manage ED congestion by creating proactive diverging policies using future patient arrival information and interfere before the ED gets highly congested. The results showed that proactive divergence policies yielded improvement in patient waiting times over standard practice. Errors in predictive information were quantified as ‘noise tolerance’ to ensure that proactive policy outperforms the standard policy (by 15%) in patient waiting times given the same number of patient census in the ED.

Researchers in a wide range of disciplines have shaped forecasting practice today. Forecasting from an OR perspective has its unique advantages. Fildes et al (2008) listed the fields that have interested OR researchers including computationally intensive models and applications in operations and marketing. Traditionally, OR methods are applied to model inventory policies and predict the value of shared information across the supply chain. After reviewing many approaches and methodologies in forecasting, the researcher concludes that OR discipline contributes to forecasting practice through developing models that integrate new forecasting methods and the specific organizational context to be applied (Filders 2008).

Overall, the health care operations research (OR) modeling approaches can be broadly classified as analytical, simulation and statistical or empirical. Two past thorough literature review articles have looked at the frequency of use of simulation and modeling in health care and the specific and the level and specific domains of application. The researcher reviewed international research journals utilizing simulation and modeling in healthcare, 82% of which was published between 1990 and 2007. It was found that the use of simulation as primary methodology and qualitative study as secondary method has been on the rise (Brailsford et al. 2009). Also, there is a review of all papers presented over the past 35 years at the "Operational Research Applied to Health Services (ORAHS)" meeting platform, disclosing that the unit-and-hospital-level studies have become dominant, while the regional-and-national-level studies decreased. There was also a notable increase at the patient-and-provider-level because of recent interest in care pathways (Brailsford & Vissers 2011). All in all, various computer simulation modeling techniques have been applied to healthcare problems regarding patient flow modeling and operational performance analysis.

CHAPTER 3

METHODOLOGY

3.1 Discrete Micro-Simulation Model

Discrete micro-simulation was suitable for this study, because the hourly arrival and discharge rate was relatively small; sometimes nothing happened in a particular hour. Firstly, a simple MU dashboard was created in Excel VBA. This simple MU model did not classify patients by the type of room they requested. To unravel the complex clinical and non-clinical factors affecting patient-level discharge timings, I conducted qualitative observations of interdisciplinary care team rounding patients at BMC in two acute internal medicine units in February 2015. Next, the two-stream MUBB model considering blocked beds was developed with census-dependent discharge rates.

Time was modeled as discrete at the hourly granularity. Once simulation parameters were inputted, the user could press the “Run” button. The model would first identify the appropriate historical mean arrival and discharge rate for the next 48 hours. A total of 1000 scenarios of bed demand and patient discharges would be created. If the unit is full at the end of a time period, patients wait in a queue until a bed becomes available. When a patient seizes a bed during an hour block, he or she will be included in the ending bed census for that hour period. If there is no bed available, the current queue will carry over into the future time periods.

The micro-simulation model was initialized with today’s date and time or user input. The user was asked to specify simulation parameters such as unit capacity, start day midnight census, beginning bed census, beginning bed queue and number of blocked beds. The simple MU model is a traditional hospital ward model that does not consider blocked beds. The two-stream MUBB model considers blocked beds explicitly by simulating two-stream arrivals on a first come first serve basis. There was no particular bed assignment prioritization rule set for Type II patients. For the pseudocode of the simple MU model, please see the Appendix.

3.1.1 Patient Type

Again, the MU under this study was an acute medicine unit with 17 shared rooms and no privates. When a patient demanded a private space, the bed assignment nurse had to mark 1 bed in a shared room as occupied and to block the other in order to legally admit the patient. To simplify this phenomenon, patients were classified into 2 types by the number of beds required for their admission. The prediction model had two streams of patient flows:

- Type I patient consumes 1 bed
- Type II patient consumes 2 beds (using a shared room as a private room)

Historical arrival and discharge data was not available by patient type. Hence, it was assumed that in the worst scenario, 1 in 3 patient arrivals would be Type II patients among all patients arriving in an hour. Discharge data by patient type was not available, so the same patient ratio assumption was made.

3.1.2 General Modeling Logic

The simulation takes snapshots of the system at the end of each hour. At 12am each day, today's discharges will be pre-calculated. For each hour, census-dependent discharges are processed before new Poisson arrivals. It was assumed that the time between bed assignment and when a patient completes admission process would be less than one hour. When there is an open bed, a bed request is satisfied and a bed is occupied immediately within the hour period. Bed requests are classified into two categories. Category 1 Request is from current hour. Category 2 Request is from previous hours. One bed assignment rule is implemented to prioritize Category 2 requests over Category 1 requests if feasible. When many bed requests fall into the same category, then a bed lottery will run to randomly draw patients to be admitted, with all bed requests having equal chance of being selected. Key system performance variables such as bed queue and number of open beds are collected at hourly granularity. For the pseudocode of the two-stream MUBB model, please see the Appendix.

3.1.3 Bed Assignment Lottery

The bed placement nurses said that the patients who were admitted to this unit were of similar acuity levels. Hence, it made senses to assume that all patients waiting in line for MU beds should have equal likelihood of getting admitted. Four variables were created for the bed assignment lottery; they were Type I queue, Type II queue, Type I request, Type II request. Note that Category 2 requests represent the queue variables. First, a lottery score between 0 and 1 would be randomly generated for each patient bed request. In order to prioritize Category 2 requests, all lottery scores would be multiplied by 0.1, and Category 2 requests would be rewarded with an addition of 0.5 points. Then, we would sort bed requests favoring larger lottery scores. Read the sorted list from top to bottom to process patient bed requests until open bed is zero. If Type II patient is selected while Open Bed is less than two, then the Type II patient would be sent back to the queue.

3.1.4 Discharge Modeling

Each time the simulation clock ticks 12am, today's daily discharge numbers would be pre-calculated. Simulated discharge was defined as the product of patient census and the eligible discharge percentage. For the purpose of this study, the 12am census on the simulation start day was initialized with mean historical census data by day of week from fiscal year 2013. Historically, few to none of the patients arrived and left a MU unit on the same day. Eligible discharge percentage was derived from mean historical non-sameday (NS) discharge percentage varied with standard deviation. For each simulation day, after pre-calculating today's discharges, empirical hourly discharge distribution would be applied to spread out today's discharges with hourly granularity. As the simulation clock goes forward by hour, the number of open beds is updated with pre-calculated discharges, and then simulated hourly bed requests would be processed to update the hourly bed queue. Since no data on Type II patient discharge was available, it was assumed to census-dependent and follows the NS-discharge percentages. Also, it was assumed that Type II patient uses a shared room as a private room till the end of their hospital stay. In other words, no co-hoarding of patients with the same type of infectious diseases was allowed in the MUBB model.

3.2 Model Specification

3.2.1 Historical Arrivals and Discharges

Mean bed request rate and mean patient discharge rate by hour of day and day of week were estimated from one-year data between April 2010 and April 2011. The historical patient arrivals and discharges varied by hour of day and day of week. Much less patients were discharged on weekends due to factors such as reduced hours of ancillary services, reduced staffing of hospitalists and special consults, limited admission hours at the post-acute care facilities. Fridays saw highest number of discharges, while Tuesdays saw highest number of admissions. For more details, see the figure below.

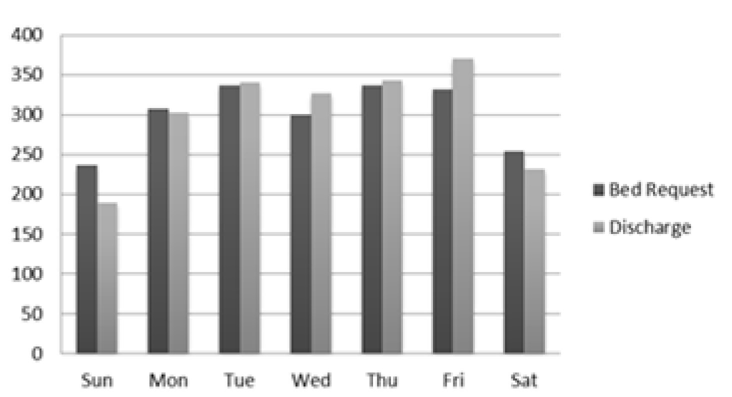


Figure 3 MU Bed Requests vs. Discharge (2010)

There were three types of patient flows through the chosen MU. The first type of flow included patients whom were both admitted to and discharged from this MU. The second type of flow included patients whom were admitted to a different unit but were discharged from this chosen MU. The third type of flow included patients whom were admitted to the selected MU but was transferred elsewhere. The first, second and third type of flows took up 64%, 30% and 7% of all patient cases between April 2010 and April 2011. The third type of flow happened infrequently (7%), and a small proportion of that represented patients whom were downgraded from ICUs. For the purpose of this study, it was assumed that all patients whom were discharged from the chosen MU should have been admitted there, but some were admitted elsewhere initially due to bed request overflows. In short, this study only considered the first and second type of patient flows (93%) through the selected MU. This assumption has some limitations and may overestimate patient arrivals, but the clinicians were comfortable with it for the purpose of this study.

The memory-less property of Poisson process says, the number of arrivals occurring in any bounded interval of time after a point in time is independent of the number of arrivals occurring before that certain point in time. It was assumed that future demand and discharges would be independent of the current ones, satisfying the memory-less property of Poisson probability distribution, which was then used to model patient bed requests. See the two figures below for historically time-varying mean bed requests and discharges by hour of day and day of week, graphed as step functions where the horizontal axis represents the discrete time periods in hour.

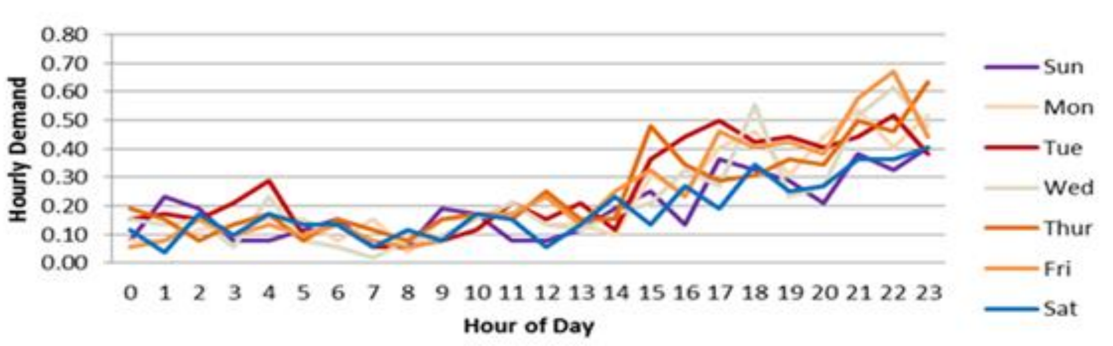


Figure 4 MU Historical Time-Varying Bed Demand

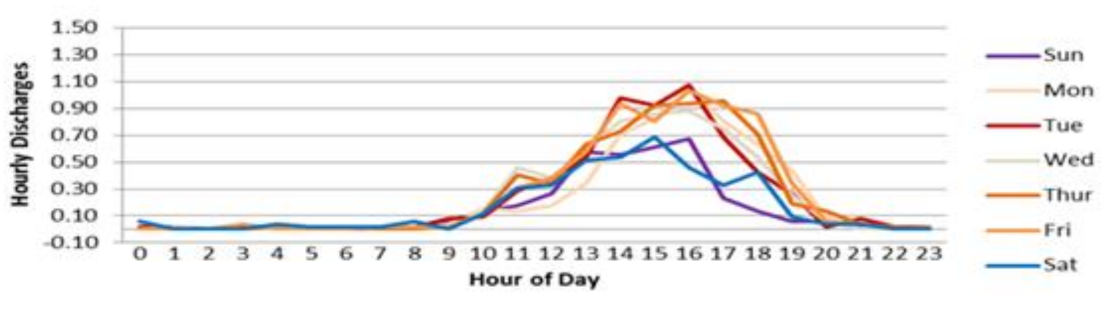


Figure 5 MU Historical Time-Varying Discharges

Historically in the chosen MU, the weekly mean patient discharges was 37.62 patients; the daily average number of discharges was 5.37 patients. Discharges that occurred before 11am, between 11am and 7pm and after 7pm were 3%, 93% and 2% respectively. Roughly speaking, 93% of patient discharges occurred between 11am and 6pm on average weekly. Discharges before 11am was rare, roughly 3%. Discharges after 6pm and 7pm were approximately 6% and 2% respectively. Peak time window for patient discharge took place between 2pm and 6pm, which was 29.61 patients in a given week out of 37.62 patients. That was on average 4.17 patients between 2pm and 6pm daily. See the table below for details.

Table 1 Percentage of Daily Discharges by Hour of Day

	DayOfWeek							
Hour	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Grand Total
0	0%	0%	0%	0%	0%	0%	1%	0%
1	0%	0%	0%	0%	0%	0%	0%	0%
3	0%	0%	0%	0%	0%	1%	0%	0%
4	0%	0%	0%	0%	0%	0%	1%	0%
5	0%	0%	0%	0%	0%	0%	0%	0%
6	0%	0%	0%	0%	0%	0%	0%	0%
7	0%	0%	0%	0%	0%	0%	0%	0%
8	0%	0%	0%	0%	0%	0%	1%	0%
9	2%	1%	1%	0%	0%	0%	0%	1%
10	4%	3%	2%	2%	2%	2%	3%	2%
11	5%	2%	5%	8%	6%	5%	8%	6%
12	8%	3%	6%	7%	6%	6%	8%	6%
13	16%	6%	9%	10%	10%	9%	13%	10%
14	16%	13%	16%	14%	12%	14%	13%	14%
15	17%	15%	15%	15%	15%	12%	17%	15%
16	19%	19%	18%	15%	15%	16%	11%	16%
17	6%	14%	12%	13%	15%	14%	8%	12%
18	4%	11%	7%	9%	11%	13%	10%	10%
19	2%	8%	5%	4%	3%	5%	2%	4%
20	2%	1%	0%	1%	2%	1%	1%	1%
21	1%	1%	1%	0%	1%	1%	1%	1%
22	0%	0%	0%	0%	0%	0%	0%	0%
23	0%	0%	0%	0%	0%	0%	0%	0%
Grand Total	185	290	311	301	324	339	213	1963
By DayOfWeek	9%	15%	16%	15%	17%	17%	11%	100%

Additionally, it was assumed that no patients would arrive and get discharged on the same day. That is all patients would at least stay overnight. Historical percentage of same-day patients was low, about 7 out of a total of 1,963 patient instances in a year. The number of patients who stayed overnight but had a length of stay less than 24 hours were about 130. Together roughly 7% of patients had length of stay less than 24 hours. See the table below for more details concerning MU historical patient length of stay distribution.

Table 2 MU Patient Length of Stay

days	percentage
1	7%
2	18%
3	18%
4	17%
5	11%
over 5	30%

3.2.2 Census-Discharge Percentage via 365-Day MU Reconstruction

In order to obtain historical daily discharge percentage by midnight patient census, I reconstructed the daily MU patient flows with a total of 1,963 MU patient cases that occurred between May 1, 2010 and April 30, 2011. It was done using time stamps Admission Complete Time and Discharge Time. Day 1 was set to May 1, 2010 at 12am. If a patient arrived on May 1, 2010 and was discharged on May 11, 2010, then his or her day index for admission and for discharge would be 1 and 11 respectively. To calculate historical NS-discharge percentage by midnight census, I looped through 1,963 patient instances and counted up number of arrivals, discharges and patient census for each of the 365 days, excluding same-day patient cases. Since the initial unit census on May 1, 2010 was unknown, there was a warm-up period in the 365-day construction of MU as illustrated in the figure below. The warm-up period ends around the 7th or the 8th day.

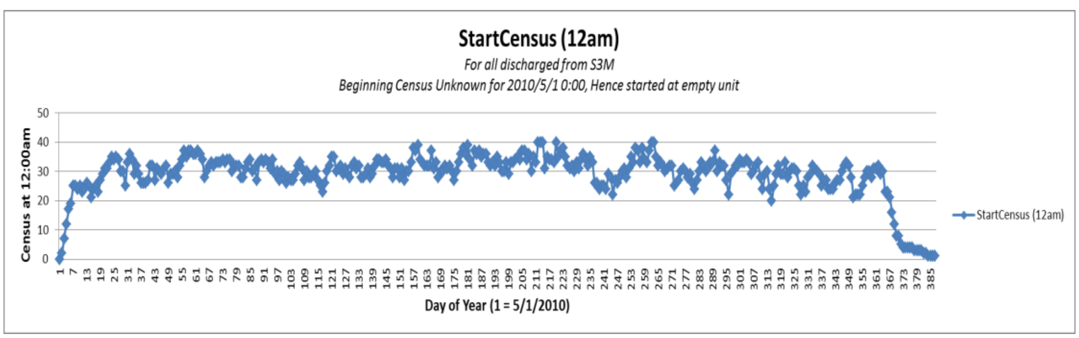


Figure 6 Midnight Census in the 365-Day Reconstruction

3.2.3 Considering Blocked Beds

Again, when a Type II patient uses MU shared rooms as privates, it results in one block bed per Type 2 patient, where the blocked bed is empty but unavailable for new admissions during the entire hospital MU stay of this Type II patient.

Data on number of blocked by 4-hour interval was available for the test MU for the fiscal year of 2013 and 2014. Mean Blocked Beds were 2.1 and 3.3 for 2013 and 2014 respectively. The percentage of semi-private beds blocked ranges from 0 to 8. During the post-Christmas flu season in winter months, roughly 8 out of 34 semi-private beds were blocked on average historically.

3.3 Data Requirements

Regarding data collection, my thesis study used historical patient arrival and discharge rates, timestamps, unit-level patient census and bed capacity from BMC to develop a data-driven simulation model. I used aggregate data reports that were developed by BMC data managers in year 2011 and were approved by Baystate Health IRB for this thesis study. In addition, the hospital administration provided MU-specific data concerning patient census and blocked bed by 4-hour interval from the fiscal year of 2013 and 2014 for validation purpose. No data concerning Type II patient arrival or discharge rates was available. My thesis study used retrospective data from exiting patient databases and clinicians' expert knowledge. Specific information required for this thesis study included but was not limited to:

- Number of inpatient bed requests for every particular hour of a particular day from both ED and other sources

- Number of inpatient discharges for every particular hour of a particular day
- MU patient census and blocked beds by day of week and by 4-hour block
- MU bed capacity and alternative unit designs proposed
- Factors impacting patient timely discharge observed through qualitative studies

3.4 Building Monte Carlo Simulation in Excel VBA

In contrast to an enterprise-level information system that is often costly and would take several years to implement and configure, this thesis study used computer simulation and qualitative observation to provide a general prototype bed prediction tool that could be immediately accessible to clinicians, and would be separate from the Electronic Medical Record system. The goal was to build a tool that would be easy to use for hospital clinicians, data managers and process improvement coordinators. The core Monte Carlo micro-simulation modeling logic was implemented using Microsoft's Visual Basic for Applications (VBA) in the Windows version of Excel. I enabled Analysis ToolPak VBA among Excel Add-Ins or ATPVBAEN.XLA. and used an ATP function in the VBA Editor to simulate Poisson arrivals and directly write outputs onto the spreadsheets. Data inputs and outputs was read from or written directly onto Excel spreadsheets. The model used either historical data or user knowledge about the current state of the system to evaluate the near-future performance of a general medicine unit.

3.5 Computational Complexity

With the use of Excel VBA each group of 1000 simulation replications for the MU model and the MUBB model took around 50 seconds and 172 seconds respectively to execute with a 2.30 GHz Intel Core i5 processor.

3.6 Model Verification

Verification of the models was done based on the comparison of the outcomes of different scenarios to each other, to historical data and to results expected by clinical expert, while meeting with the hospital clinicians and data managers monthly to ensure that the model design and behavior make sense clinically and operationally. Sensitivity analysis for the output variable was also performed to verify the results of this model.

CHAPTER 4

QUANTITATIVE RESULTS

4.1 Baseline Measures and Model Verification

The figure below presents the historical average number of bed requests and patient discharges, the simulated average Poisson arrivals and census-dependent discharges for the single-stream MU model and for the two-stream MUBB model over 1000 replications by hour of day for Wednesdays. Simulated mean arrivals closely matches the historical mean arrivals through the day. However, simulated mean discharges seem to underestimate historical mean discharges between the hours of 2pm and 8pm.

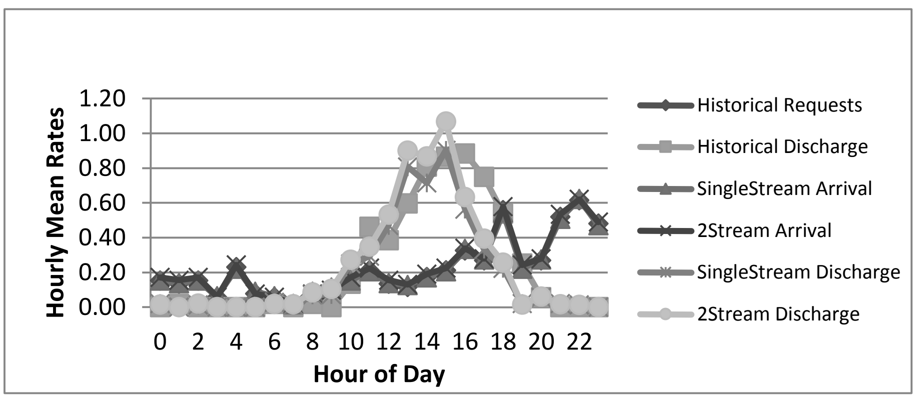


Figure 7 Wednesday Historical vs. Simulated Arrivals and Discharges

For the purpose of demonstrating the behaviors of this prototype tool, some input parameters remained constant, while data analysis and comparisons below are conducted. The initial patient census, number of blocked beds and queues were set to 28, 4 and 0 respectively, unless it is otherwise specified. Note that the start hour of the simulation is always 6am, a time before the 8am bed planning meetings among hospital clinicians. For instance, “Now” represents 6am on a day of week, 8^{th} and 32th hour into the future are 2pm. Since the daily mean arrivals for the medicine unit is approximately 5 patients, the hourly arrival or discharge rates would be less than 1. To better capture the unit performance under stress, it makes sense to look at all inputs and outputs of the model at the 75 percentile rather than simply reporting the mean values. In addition, based on historical data and clinicians' expert knowledge, Tuesday is determined to be the most “stressful” day with the highest queues. Hence, we will look at how these models predict Tuesday bed queue starting at 6am on a Sunday. Firstly, let's take a look at the behaviors of the single-stream MU model, and then we will look at the two-stream MUBB model for comparison.

4.1.1 Single-Stream MU Model Behavior

The two figures below show predicted number of bed requests and patient discharges at the 75 percentile by start day of week. The time horizon is 48 hours. If our simulation starts at 6am on a Monday, then at the 22th hour will be 4am on a Tuesday. Among 1000 Tuesday replications, 25% or 250 replications will see at least 1 bed request, while the other 75% Tuesdays will see less than 1 bed request at 4am on a Tuesday. To interpret it in terms of a 52-week year, this medicine unit will get at least one bed request at 4am on 13 Tuesdays or perhaps in a quarter of the year.

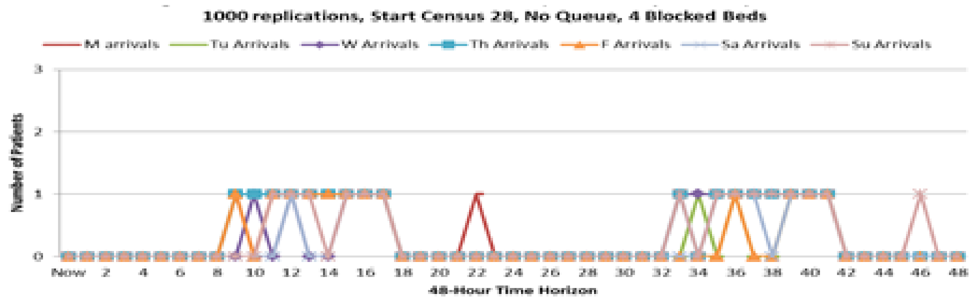


Figure 8 Single-Stream MU Baseline Arrivals at 75 Percentile by Start Day

To illustrate the number of simulated discharges at 75 percentile, the MU model predicts 2 discharges per hour for Thursdays between the 8th hour and the 10th hour. That is between 2 pm and 4pm on Thursdays. To interpret it in terms of a 52-week year, this medicine unit will complete at least 2 patient discharges per hour from 2pm to 4pm on 13 Thursdays or perhaps in a quarter of the year. That is at least a total of 6 discharges during peak 2-4pm discharge hours.

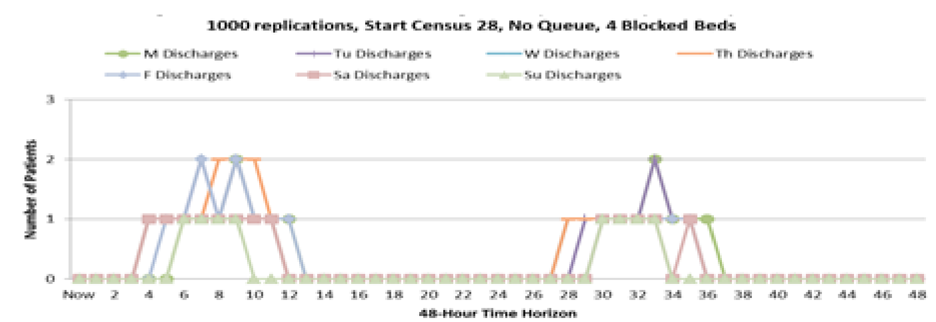


Figure 9 Single-Stream MU Baseline Discharges at 75 Percentile by Start Day

In the simple MU model, patient occupancy solely considers the number of patients in a unit. The figure below shows that most of the time patient occupancy or census is below 33. That is having one open bed. Though starting at 6am on Sunday, among 25% of the Sunday-Tuesday replications, the medicine unit reaches full capacity at the 41th hour or at 11pm on Monday. It remains full from 11pm Monday to 6am, implying that any morning bed request will be in queue until a patient discharge or transfer is completed.

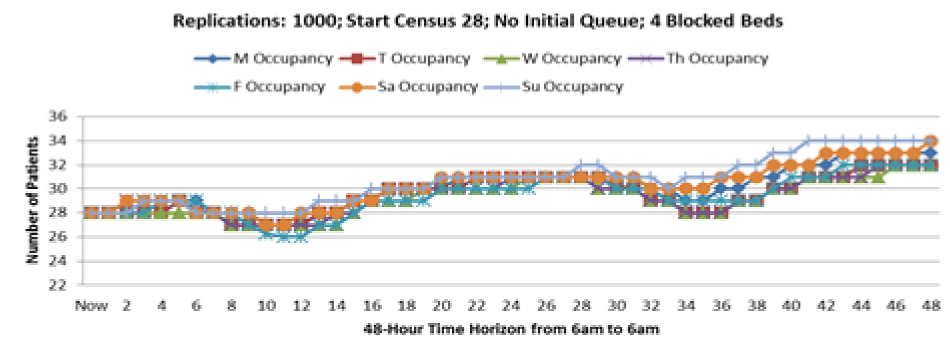


Figure 10 Single-Stream MU Baseline Occupancy by Start Day

The 48^{th}-hour queue prediction results are presented in the figure below. At 75 percentile, the 6am bed queue by day of week starting on Sunday is 0, 0, 1, 0, 0, 0, and 0 respectively. At 95 percentile, the 6am bed queue by day of week starting on Sunday is 9, 6, 8, 8, 8, 9, and 10 respectively. Among 7 days of the week, Tuesday sees the highest number of simulated queues. This simulated phenomenon reflects the seasonality is commonly seen in real life. Thus it can serve as the baseline measures in this thesis study.

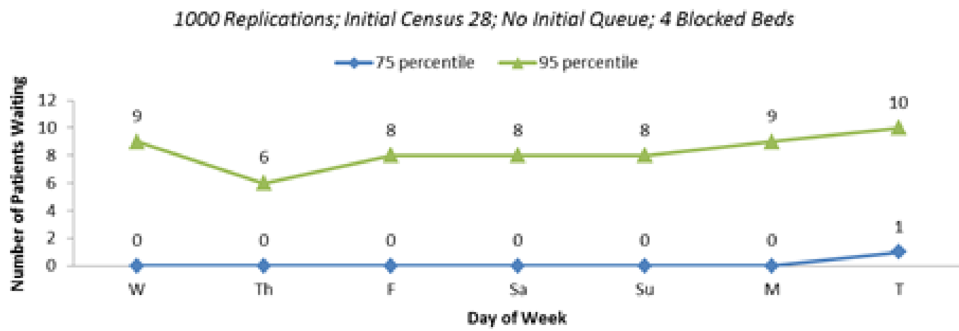


Figure 11 Single-Stream MU 48th Hour Queue Prediction by Ending Day

4.1.2 Two-Stream MUBB Model Behavior

To summarize the behavior of the MUBB model, firstly we look at and verify the simulated arrivals and discharges in the 48-hour forecast horizon. See details in the figure below for Type I arrivals by start day of the week. Type I baseline arrival pattern is similar to the total arrival pattern seen for the simple MU model.

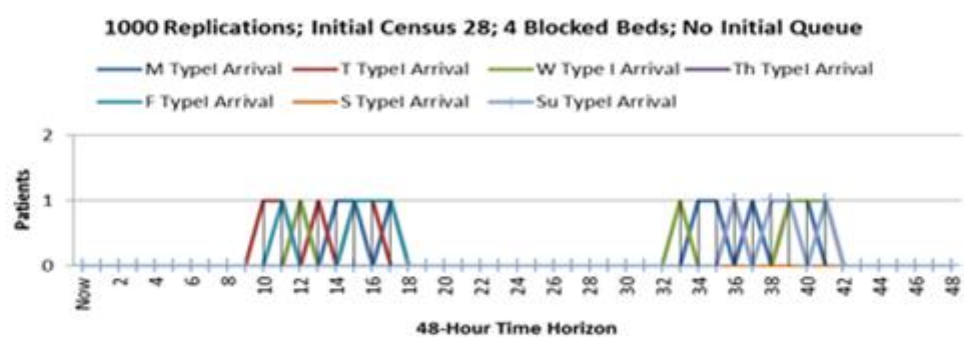


Figure 12 Two-Stream MUBB Baseline Type I Arrivals

Secondly, we look at the simulated total arrivals and discharges over 48 hours at 75 percentile by patient type and by day of week in the figure below. Assuming 1 in 3 patients are Type II patients in the stressful months of a year, starting on Monday the

MUBB model predicts that 25% of the time we see at least 10 Type I arrivals or discharges, 6 Type II arrivals and 4 Type II discharges between 6am Monday and 6am Wednesday. That is a total of 16 arrivals and 14 discharges roughly over 2 days, Monday and Tuesday. Between 6am Tuesday and 6am Thursday as well as from 6am Wednesday to 6am Friday, it is predicted that 25% of the time there will be at least 11 Type I discharges, 10 Type I arrivals, 5 Type II arrivals and 4 Type II discharges. Between 6am Thursday and 6am Saturday, it is predicted that 25% of the time there will be at least 12 Type I discharges, 11 Type I arrivals, 6 Type II arrivals and 4 Type II discharges. That is a total of 17 arrivals and 16 discharges on Thursday and Friday. Between 6am Saturday and 6am Monday, it is predicted that there will be at least 7 Type I discharges, 9 Type I arrivals, 4 Type II arrivals and 3 Type II discharges 25% of the days. That is a total of 13 arrivals and 10 discharges over the weekend. Finally, between Sunday 6am and Tuesday 6am, it is predicted that there will be at least 10 Type I arrivals, 7 Type I discharges, 5 Type II arrivals, and 3 Type II discharges. That is a total of 15 arrivals and 10 discharges over Sunday and Monday.

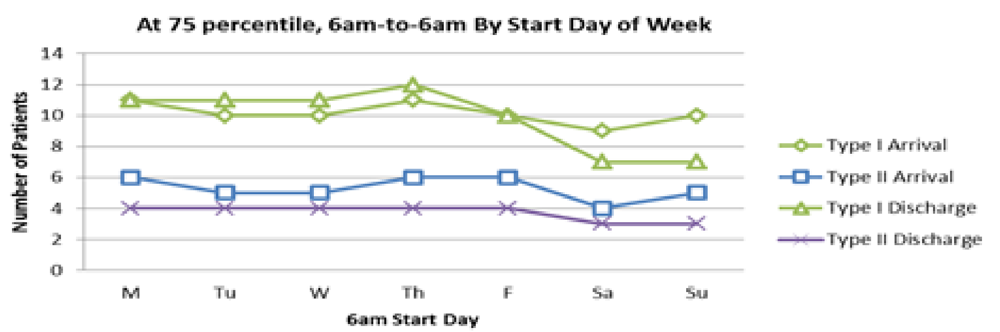


Figure 13 Two-Stream MUBB Baseline Total Arrivals, Discharges in 48 hrs

Thirdly, we look at the predicted queue size at 75 percentile at 6am in 48 hours by ending day of week and by patient type. The bed queue size is notably larger on Tuesday, Wednesday and Saturday. For instance, among 1000 replications, 25% of the time the bed queue for 6am on Tuesday is predicted to be equal or greater than 2 Type I patients and 2 Type II patients. The details are shown in figure below. Note that the percentile values are calculated separately for Type I queue and Type II patient queue. It is possible that at 6am on one Tuesday there will be 2 Type II patients waiting in the queue and no Type I patient in the queue. Type I and Type II patients may rotate staying in the bed queue. Hence, it is not appropriate to simply add the 75 percentile value of Type I queue over 1000 replications to that of the Type II queue over 1000 replications to derive the 75 percentile value of total bed queue size.

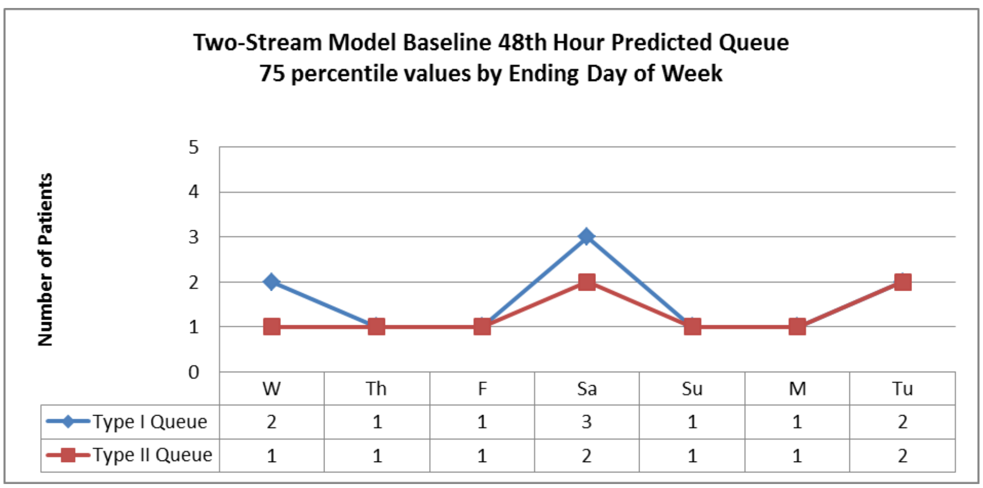


Figure 14 Two-Stream MUBB Baseline 48th Hour Predicted Queue

Finally, we look at unit bed utilization at 75 percentile in the 48-hour horizon by start day of week. Considering blocked beds, the unit bed utilization is above 90% most of the time, centered around 96%. Traditionally without considering blocked beds, the bed utilization or occupancy rate is centered at 85% and is below 90% within the 48-hour time horizon. See details in the two figures below.

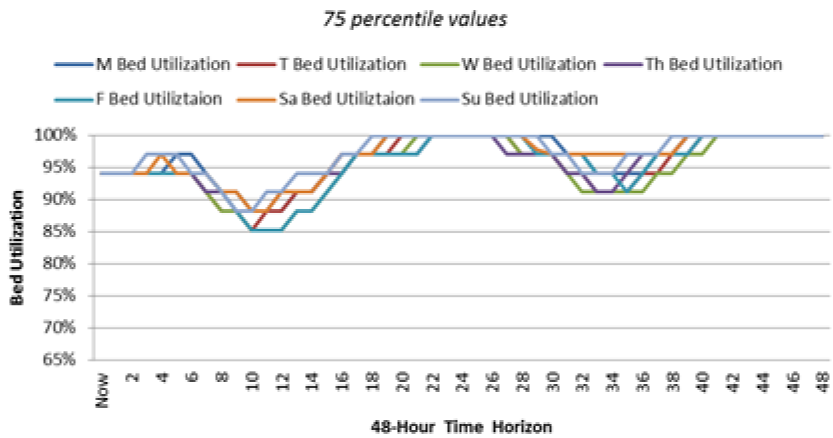


Figure 15 Two-Stream MUBB Basline Bed Utilization by Start Day

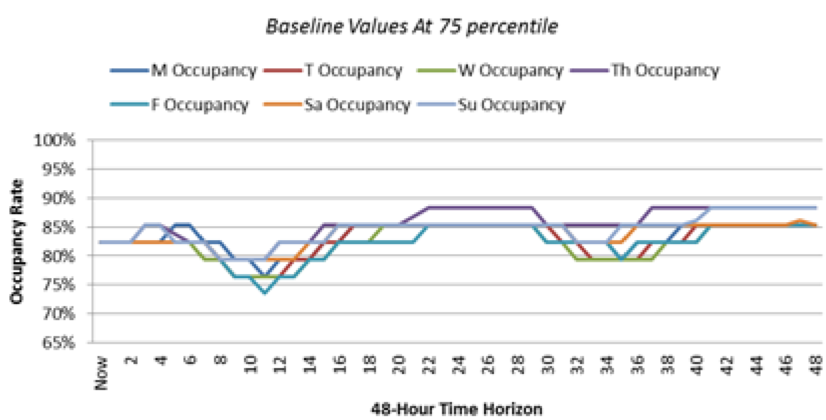


Figure 16 Two-Stream MUBB Traditional Occupancy Rate Without Considering Blocked Beds

4.2 Sensitivity Analysis

Sensitivity analysis is conducted to understand how hourly bed queue size responds to stressed arrival levels, patient type ratio and increased bed capacity, while holding controlling many input parameters such as initial patient census, initial number of blocked beds and initial queue constant as noted above. Based on the baseline measures above, Tuesday is a day when bed queue tends to be higher. It would be interesting to see how the MU model and the MUBB model predict bed queues for Tuesday. Hence, this section of the analysis sets Sunday 6am as the start time and day of the simulation in order to predict bed queue size for Tuesday morning.

4.2.1 Model Sensitivity to Arrival Levels

Baseline arrival level is identified in section 4.1. We generate two arrival levels that are lower than the baseline (60%, 80%) and two arrival levels that are higher than the baseline (120%, 140%). For the single-stream MU model, the 48th hour bed queue at 75 percentile is predicted to be 1 patient for the baseline, 120% and 140% arrival levels. The 75th percentile queue size increases sharply around 44th hour or on Tuesday early morning.

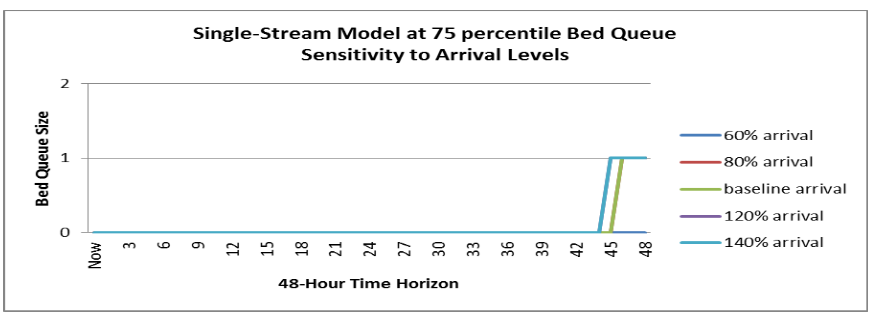


Figure 17 Single-Stream MU Bed Queue at 75 Percentile by Arrival Levels

Again for the single-stream MU model, the 48th hour bed queue at 95 percentile is predicted to be over 10 patients for the baseline, 120% and 140% arrival levels. The queue size at 95 percentile seems to rise sharply around the 34th hour or Monday early afternoon. See details below.

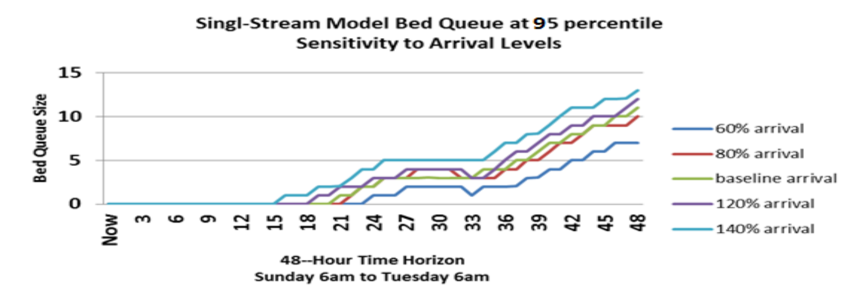


Figure 18 Single-Stream MU Bed Queue at 95 Percentile by Arrival Levels

Next, let us take a look at the two-stream MUBB model by arrival levels. The 48th hour bed queue at 75 percentile is predicted to be between 4 and 6 patients for all five arrival levels. Considering blocked beds, the 75 percentile queue in MUBB picks up around the 35th hour or noon on Monday.

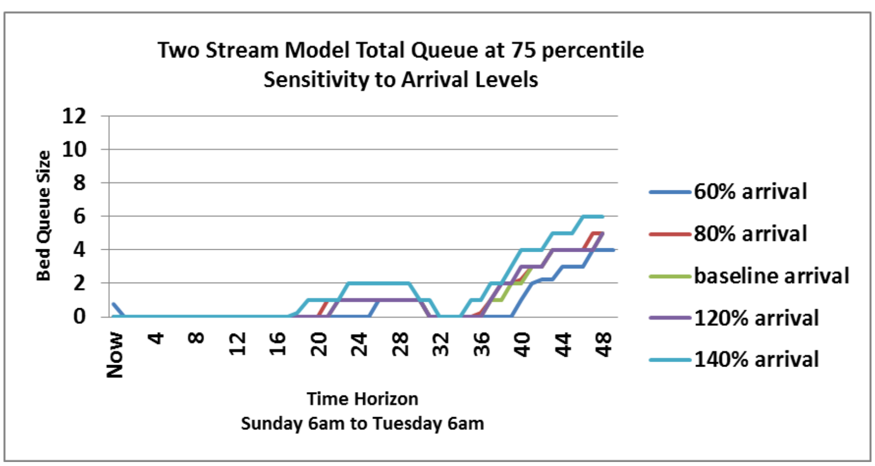


Figure 19 Two-Stream MU Total Queue at 75 Percentile by Arrival Levels

The 48th hour bed queue at 95 percentile is predicted to be between over 12 patients for all five arrival levels. The 95 percentile queue seems to grow exponentially, suggesting that 5% of the time the MUBB model is not in a steady state any more. Considering blocked beds, the 95 percentile queue in MUBB explodes around the 12th hour or 6pm on Sunday.

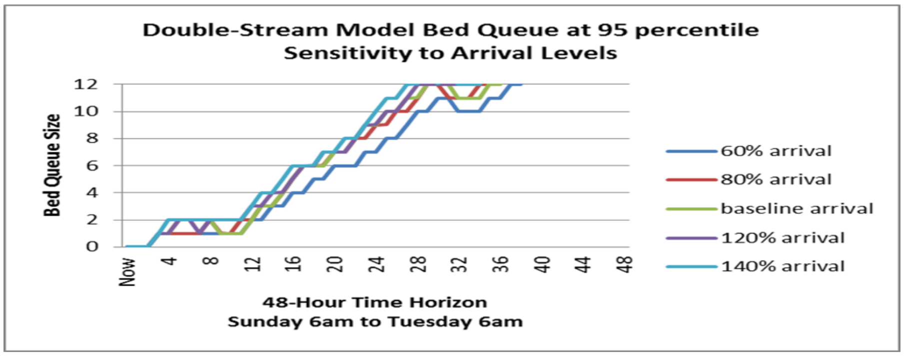


Figure 20 Two-Stream MU Bed Queue at 95 Percentile by Arrival Levels

4.2.2 Sensitivity to Patient Type Ratio

The two-stream MUBB model considers blocked beds associated with Type II patient. This section reports results concerning bed queue by patient type ratio. At baseline, It was assumed that 1 in 3 patients would be a type II patient, or 33% of all incoming patients might be a Type II patient. For the purpose of sensitivity analysis, another two patient ratios are generated to be 10% as a lower bound and 50% as an upper bound. Observing the results of MUBB simulation model, it is found that at times Type I queue is higher while Type II queue is nearly nonexistent. Vice versa, at times Type II queue is much higher than Type I queue. When Type II patient ratio is at 50%, Type II queue size at 75 percentile grows sharply to 3 patients on Tuesday morning. At 10% type ratio, the Type II queue is nearly nonexistent for the 48th hour. See figure below.

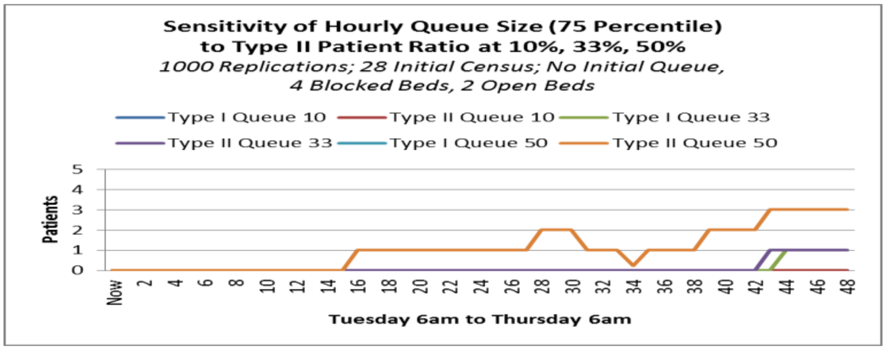


Figure 21 Sensitivity of Hourly Queue to Patient Type Ratio

Since Type I queue size and Type II queue size at 75 percentile over 1000 replications don't always add up directly, it is worthwhile to present the 75 percentile results for the total queue by patient type ratios below. Again with patient type ratio of 10%, the total bed queue at 75 percentile seems nonexistence for 6am Tuesday. At ratio 33% and 50%, the total queue size at 75 percentile is 2 or over 5 patients respectively. See details by hour of simulation below.

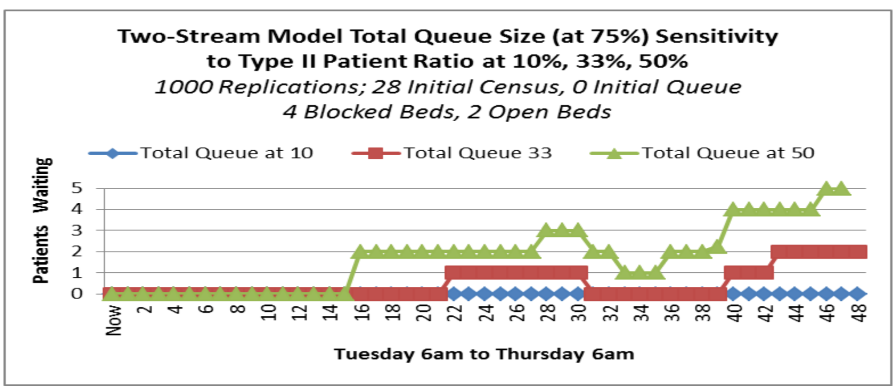


Figure 22 Two-Stream MUBB Cumulative Queue in48 hrs by Patient Type Ratio

4.3 Simply Varying Shared Bed Capacity

The simple single-stream MU model and the two-stream MUBB model respond differently to increasing shared bed capacity. Arbitrarily, starting at a unit capacity of 32 shared beds, we increase the unit capacity from 32 beds by a 2-bed interval up to 38 beds. Using the total number of patients in queue over 48 hours of simulation horizon under the 32-bed capacity scenario as baseline, we see that the single-stream MU bed queue improved 100% going from 32-bed capacity to 34-bed capacity. In other words, the single-stream MU queue is reduced to 0 as we increase unit bed capacity from 32 to 34. Given the inputs such as same arrival levels and initial parameters, the MUBB queue only improved 68% going from 32-bed capacity to 34-bed capacity and finally improved 100% or is reduced to 0 going from 34 to 36 beds. The improvement in bed queue will look different with different inputs. Controlling the inputs, the single-stream MU model behaves differently from the two-stream MUBB model under given capacity scenarios.

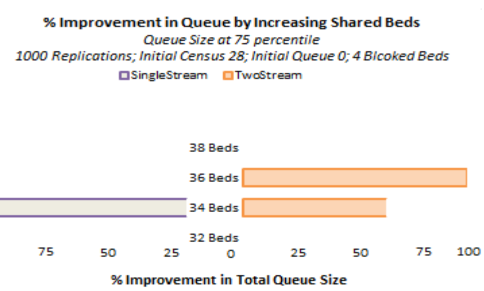


Figure 23 Percentage of Improvement in Queue by Shared Bed Capacity

4.4 Varying More Model Inputs

Besides simply varying shared bed capacity, other model inputs are varied to create the common case, better-than-common case and the stress testing scenario. The specific inputs varied include shared bed capacity, initial patient census, the number of blocked beds and the number of patients in Type I or Type II queue initially. Under all scenarios, the unit is full at start. The difference between each scenario is in the initial number of blocked beds and Type II patients in queue. According to narratives of bed placement manager at Baystate Medical Center, we set the Common Case to 29 patients, 5 blocked beds, 2 Type I patient and 1 Type II patient waiting initially. Under the Better-Than-Common case, there are 31 patients, 1 blocked bed, 2 Type I and no Type II patients in the queue initially. Under the arbitrary Stress Testing case, there are 32 patients, 2 blocked beds, 3 Type I patients and 3 Type II patients in queue initially. The shared bed capacity is increased from 34 beds to 44 beds at 1 bed per increment. The bed queue by 0, 25, 50, 75, 95 and 100 percentile and by the aforementioned scenarios is reported in the table below.

Table 3 Capacity Scenario Looking at Average Bed Queue at 1pm Thursday

Inputs : (census, blocked, Type I queue, Type II queue)

Bed Queue at 1pm Thursday

Better-Than-Common

Inputs	(31, 1, 2, 0)										
BedCapacity	34	35	36	37	38	39	40	41	42	43	44
Single-Stream Model											
0%	0	0	0	0	0	0	0	0	0	0	0
25%	0	0	0	0	0	0	0	0	0	0	0
50%	0	0	0	0	0	0	0	0	0	0	0
75%	2	1	0	0	0	0	0	0	0	0	0
95%	11	10	9	8	7	6	5	4	3	2	1
100%	45	44	43	42	41	40	39	38	37	36	35
Double-Stream Model											
0%	0	0	0	0	0	0	0	0	0	0	0
25%	0	0	0	0	0	0	0	0	0	0	0
50%	1	0	0	0	0	0	0	0	0	0	0
75%	6	5	4	3	2	1	0	0	0	0	0
95%	20	18	17	17	16	15	14	12	13	12	11
100%	70	68	71	62	61	57	59	55	63	53	52

Stress Testing

Inputs	(32, 2, 3, 3)										
BedCapacity	34	35	36	37	38	39	40	41	42	43	44
Single-Stream Model											
0%	0	0	0	0	0	0	0	0	0	0	0
25%	1	0	0	0	0	0	0	0	0	0	0
50%	4	3	2	1	0	0	0	0	0	0	0
75%	8	7	6	5	4	3	2	1	0	0	0
95%	17	16	15	14	13	12	11	10	9	8	7
100%	51	50	49	48	47	46	45	44	43	42	41
Double-Stream Model											
0%	0	0	0	0	0	0	0	0	0	0	0
25%	2	2	1	0	0	0	0	0	0	0	0
50%	6	5	4	4	3	3	2	1	0	0	0
75%	10	9	9	8	7	7	6	6	5	4	4
95%	24	23	22	21	21	20	19	19	19	17	17
100%	65	64	66	62	62	62	65	61	62	62	71

Common Case

Inputs	(29, 5, 2, 1)										
BedCapacity	34	35	36	37	38	39	40	41	42	43	44
Single-Stream Model											
0%	0	0	0	0	0	0	0	0	0	0	0
25%	0	0	0	0	0	0	0	0	0	0	0
50%	1	0	0	0	0	0	0	0	0	0	0
75%	5	4	3	2	1	0	0	0	0	0	0
95%	14	13	12	11	10	9	8	7	6	5	4
100%	48	47	46	45	44	43	42	41	40	39	38
Double-Stream Model											
0%	0	0	0	0	0	0	0	0	0	0	0
25%	0	0	0	0	0	0	0	0	0	0	0
50%	2	1	0	0	0	0	0	0	0	0	0
75%	6	6	5	4	4	3	2	1	0	0	0
95%	19	19	19	18	17	16	15	15	14	13	12
100%	61	67	73	63	61	63	67	59	59	62	56

The percentage of improvement by single-stream MU model and the two-stream MUBB model differs under the aforementioned scenarios. Specifically, the single-stream MU bed queue appears to be more sensitive to increase in shared bed capacity than the MUBB bed queue. Under the stress testing case, MU bed queue saw more than 65% improvement as shared bed capacity increases from 34 to 44. In contrast, MUBB bed queue under stress testing saw less than 50% improvement as shared beds changes from 34 to 44. See details in the figure below.

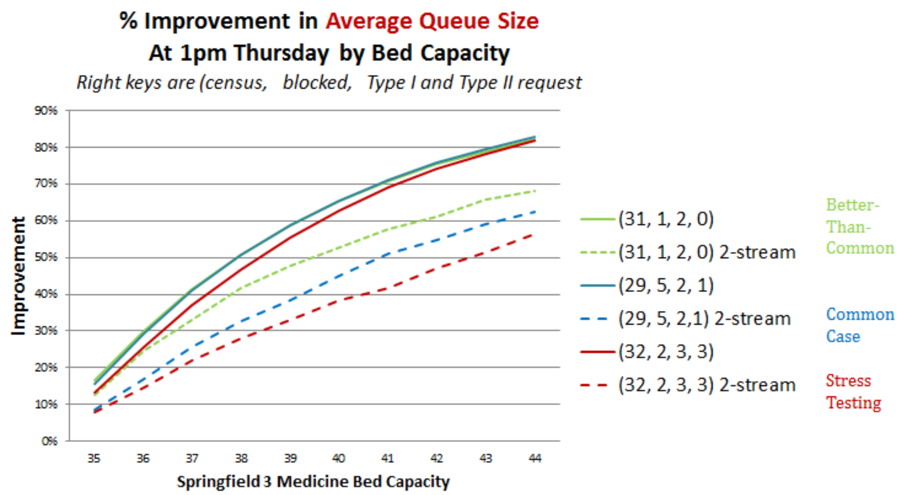


Figure 24 Percentage of Improvement in Average Queue Size at 1pm Thursday

CHAPTER 5

QUALITATIVE RESULTS

5.1 Qualitative Study Aims

Qualitative observation was employed to unravel the complex interactions between a patient and an interdisciplinary care team within acute general medicine units. The hope was to supplement the simplified Excel model and to inform future research on patient-provider information flows, especially on what will lead to process milestones that are critical for discharge planning.

5.2 Methodology in Details

Baystate Medical Center implemented interdisciplinary plan of care (IPOC) teams made up of a bedside nurse, a physician, and a case manager. A case manager usually initiates IPOC rounds between 11am and 2pm by reaching out to physicians and bedside nurses. Hence, the plan was to interview and shadow case managers on IPOC rounds. The provider-patient information flow during IPOC rounds will be simplified and categorized into a few dimensional variables, such as clinical, social, or financial. In the end, there will be a rough summary of how frequent these IPOC information criteria come up among the IPOC visits.

5.3 Results

Qualitative observation of IPOC rounds was carried out in two acute units over eight non-consecutive days, including a total of 205 patient visits made by the interdisciplinary plan of care (IPOC) teams: 131 visits in a general medicine unit specialized in acute respiratory conditions, and 74 visits in an internal medicine unit specialized in acute heart and vascular conditions. Over all, the time between each patient visit range from 2 to 6 minutes, including the time for locating clinicians, for bedside IPOC meeting, and for the occasional team discussions outside the patient room. The descriptive summary can be found in *Table 2* in the Appendix.

5.3.1 Fragmented Information and Interdisciplinary Care Team

Patient-Provider communication is multi-channel and multi-dimensional. A physician focuses on the clinical aspects of patient care such as ordering diagnostic tests, evaluating disease progression and formulating treatment plans. There are two types of physicians. A teaching physician is matched to a unit's clinical specialty and leads a team of resident or intern physicians. A hospitalist physician practices general internal medicine and does not supervise interns. A bedside nurse has a 360 degree understanding of a patient's general physical and psychological functioning level, including ability to walk, ability to carry out activities of daily living (ADL), and ability to comprehend and follow safety or medication instructions. A case manager works along social workers and interviews patients and/or their families to assess social functioning levels, order home evaluations, gather preferences regarding post-hospital discharge plans, call for financial counsel for uninsured patients and coordinate discharge placement.

Overall, many factors may characterize the discharge planning process for a particular patient, such as clinical, patient general functioning levels (physical, psychological, social), financial, as well as patient's and family's' preferences and goals. It is challenging for the IPOC teams to timely gather and navigate information from multiple provider channels to understand a patient multi-dimensionally. The results of this qualitative study motivated the workforce planning at our collaborating hospital manager to advocate for dedicated case managers or consistency of team practice.

5.3.2 Simplified IPOC Network

A case manager or a physician usually initiates IPOC rounds between 11am and 2pm, reaching out to bedside nurses and visiting each treatment rooms and to update patients about their plan of care. IPOC visits were conducted in no particular order, except for starting with the spatial-numerical order due to physical layout. Patients were assigned to nurses by spatial proximity and assigned to physicians by the complexity of their clinical conditions. In certain circumstances, a bedside nurse was called to round a hospitalist physician's patient and a teaching physician's patient at the same time. If a bedside nurse is not immediately available, the IPOC team will search for the next available bedside nurse and revisit the skipped nurses later if possible. Differences exist between the two acute medicine units, in terms of how an IPOC team is organized and the frequency of bedside nurse participation. In Acute Medicine Unit I, each case manager followed both teaching and the hospitalist teams. In Acute Medicine Unit II, in contrast each case manager was separately responsible for the teaching team and the hospitalist team each day, and case managers occasionally rotate to join the other team. Simplified organization of the internal IPOC teams is displayed in the Graph below. It doesn't account for the relationship between patient, IPOC team, and specialists external to the unit.

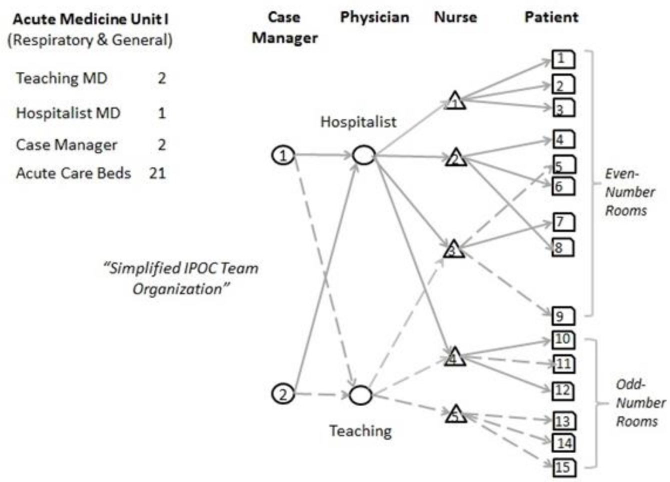


Figure 25 Graphical Representation of Simplified IPOC Network

5.3.3 IPOC Time Spent at Bedside

The IPOC team lingered longer at bedside if ......

- Physician educated a patients' about new medication / procedure
- Nurse discussed patient's functioning level and activities of daily living (ADL)
- Case manager inquired about social support or the lack thereof
- Patient/families had questions/complaints

The IPOC team spent less time by bedside if ......

- The bedside nurse was absent (about 1 in 50 chance)
- Waited for follow up with specialist
- No family was present
- The patients did not speak up for themselves

Physician skipped bedside IPOC and do hallway check in with RN and case manager if ......

- Physician discharge order was complete
- Patient was agitated, mad, unstable, going to leave against medical advice (AMA), or wanted to be left alone
- Patient was undergoing procedures or is away for testing
- Patient was cognitively impaired and had no companion (e.g. Dementia)
- Family members were frustrated and wanted private time
- Physician had visited the patients early morning, no news
- Physician was overloaded with writing discharge orders for the day

5.3.4 IPOC Information Criteria

My qualitative observation sought to understand the complex context not limited to clinical aspects that impacts the progress of a patient's care. The IPOC-patient information flow was categorized into 26 raw information variables. For details, please see the Appendix. The raw 26 IPOC variables were then summarized or assigned to six parent categories: clinical, social, discharge planning logistics, financial, level of functioning assessment, other or ancillary. Please see the table below.

Table 4 Raw IPOC Variables Classified by Greater Dimensions

Dimension	IPOC Variable
Clinical	1, 7, 8, 9, 10, 15, 16, 18, 21, 25
Social	6, 23, 24
Discharge Planning	2, 3, 5, 14, 17, 19, 26
Financial	20
Level of Functioning	4, 12, 22
Other, Ancilliary	11, 13

The figure below shows the number of outstanding information variables discussed during IPOC rounds, reflecting the patient-level variability in care and social needs. Both clinical and non-clinical factors may impact patient care progress and discharge timing. Tracking outstanding IPOC variables may be a way for future research to quantify patient discharge burden.

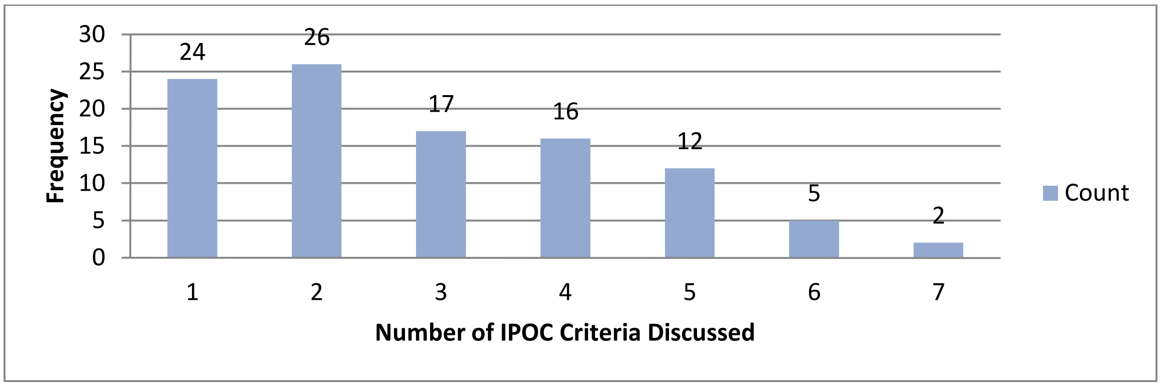


Figure 26 Number of IPOC Variables Discussed During IPOC Rounds

5.3.5 Multi-Dimensional Discharge Criteria

Case managers usually take into consideration many complex factors for patient discharge planning. And this is a process that happens in parallel to physicians' clinical work. The table below is a qualitative summary of the discharge criteria based on my interviewing case managers from the general medicine unit. Some of these factors definitely would affect patient discharge timing. For instance, hospital shuttle is only available twice a day; family pickup usually is later in the evening; new admissions to post-acute care facilities shouldn't be later than 6pm. The discharge planning and community care coordination processes are multi-dimensional and are not always perfectly integrated with the clinical milestones of patient care.

Table 5 Multi-Criteria Discharge Planning Decision Making

NON-CLINICAL DISCHARGE CRITERIA	LEVELS		DISCHARGE MILESTONES	
			HOW	WHERE
Physical Level of Functioning	High		Shuttle, Drive, Family	Home Care
	Medium		Shuttle, Ambulance	Home Care, PCP-VNA, or Rehab
	Low		Ambulance, Family	Home Care, PCP-VNA, or Rehab
			FAMILY INTERVIEW AND EVALUATION	
Social Level of Functioning	With Family Support	Challenging	Social Worker, Home Evaluation	
		Not Challenging	No Social Worker	
	No Family Support	Challenging	Social Worker, Patient Advocate, Homr Evaluation	
		Not Challenging	Patient Advocate, Home Evaluation	
			HOW	SPECIAL COUNSEL
Psychological Dimension	Agitated		AMA	Behavioral Helath, Social Worker
	Not Agitated		Normal	N/A
	Declining Mental Ability (Dementia)		Companion Required	Geriatrics, Neurology
	Normal Mental Ability		Normal	N/A
			IF POST-HOSPITAL CARE FACILITY	
Financial	Insured		Normal	
	Uninsured, Expired		Insurance applications filed and processed on weekdays	

CHAPTER 6

CONCLUSION

6.1 Introduction

This thesis study was set out to identify the requirements for building a Monte Carlo micro-simulation model to predict future patient arrivals, discharges and bed queue for an adult medicine unit, and has implemented a prototype dashboard in Excel VBA, with the hope to provide bed placement team with future bed queue information as alerts for surge capacity planning prior to bed crisis, or admission and discharge control. This study has also sought out to compare two model configurations to judge the value of considering blocked beds for hospital unit performance analysis, and to lay some groundwork for future research. A comparative analysis of the two models aforementioned reveals a difference in queue improvement as unit bed capacity is varied.

6.2 Significance Recap

All in all, it is valuable to analyze hospital downstream unit performance and to understand unit-level bed queue dynamics. Specifically, it could be worthwhile to predict hospital unit performance in the near term under stressed scenarios in order to design early alerts for activating surge capacity planning, patient admission or discharge control. Besides, it seems financially worthwhile to study the phenomenon of blocked beds associated with Type II patient.

6.3 Assumptions and Limitations

6.3.1 Simplified arrival rates

For the adult medicine unit in this study, the hospital clinicians conducted a statistical test on patient encounters by Diagnosis-Related Groups (DRGs) for fiscal year 2010, 2011, 2012 and 2013 and found no significant difference in patient demand among those years." If there would exist a sudden outbreak of contagious diseases or an extensive epidemic in the next few years, the generalized historical arrival rates used in the prototype model would be inappropriate. Collaboration between hospital clinicians and engineers is required then to re-configure the prototype model for it to yield realistic and useful insights into hospital unit performance.

6.3.2 Seasonality in Demand Patterns

According to clinicians' expert knowledge and historical data, the usual time frame for notable hospital congestion by month of year was post-Christmas winter months and by day of week was on Tuesdays. This thesis study attempted to quantify the impact of blocked beds on bed queue in a 48-hour time. The prototype tool does not consider demand seasonality by DRGs. The idea is that clinical practitioners may input system parameters such as current patient census to define the current state of the system and ask the model for 48-hour bed queue predictions under stressed scenarios. If a clinical user wishes to run the model with arrival rates varied by quarter of year, it is possible and easy to do that.

6.3.3 Limited Data on Type II Patient Discharge

Today's hospital information system has no way to automatically track the arrivals and discharges of Type II patients. Type II patients span across many Major Diagnostic Groups (MDCs), see the tale in the Appendix. The definition of Type II patient comes from grouping patients that require a private room but are assigned to a shared room. Hence, assumptions had to be made regarding the patient type ratio for bed requests and for discharges. The limitation is that the model does not automatically update or verify the patient type ratio assumption as hospital's clinical case mix changes in the future.

6.3.4 Patient Admission Prioritization

I had the opportunity to interview bed placement managers and other clinicians concerning how the bed assignment process works. The takeaway I had was that so many complex factors play into their decisions, and the clinicians cannot simply give priority to patients by Type I or Type II specified in this study. The general understanding about the bed queue rule is first come first serve, and that all patients going to adult general medicine units without telemetry requirements fall under the same acuity level. The bed placement manager says first-come-first-serve is what they do in practice. And it is acceptable and appropriate to say that every patient in the waiting line have an equal chance of getting admitted if their bed requests arrived within the same hour.

6.3.5 Simplified Discharge Simulation

My qualitative results reflect the complexity in discharge planning and hospital-to-community care coordination. Many non-clinical variables play into the timeliness of patient discharges, but little data is available about it. Hence, my models only make very crude and aggregate-level prediction about near-future patient discharges. If the hospital were to change its traditional discharge process, or if the post-acute care facilities are changing their admission time window, then the discharge timing distributions by hour of day and day of week may change accordingly. My results can only make sense or be improved, if someone updates the discharge distributions by hour and day.

6.4 Implication

The quantitative results of this study implies that if a general medicine unit only has shared beds but accepts patients who may need private rooms, it is important to consider blocked beds in unit performance modeling and capacity planning scenarios. The qualitative results suggest that patient discharge modeling shall consider various non-clinical factors and milestones that run in parallel to the clinical care progress.

In addition, traditionally bed occupancy rate only counts the actual number of patients in a unit, underestimating the actual unit bed utilization level due to blocked beds. Since many clinical practitioners and unit managers use unit bed occupancy as a key performance indicator, this measure can be misleading at times when Type II patient ratio rises significantly. Ideally a unit described above shall have all private treatment rooms, but in reality the hospital needs to carefully consider financial consequences of expanding or redesigning its existing physical facility.

The prototype model was meant to be a short-term prediction tool that can be used daily to inform bed placement managers on admission and discharge control as well as bed crisis planning. The tool was supposed to be used independent by nurses in administration; however, direct support from an in-house modeler or developer may be necessary for maintenance and updates. Last but not least, the clinicians as well as the hospital administration thought this tool would in fact serve better if it evolves to be a unit redesign evaluation tool.

6.5 Recommendations for Future Research

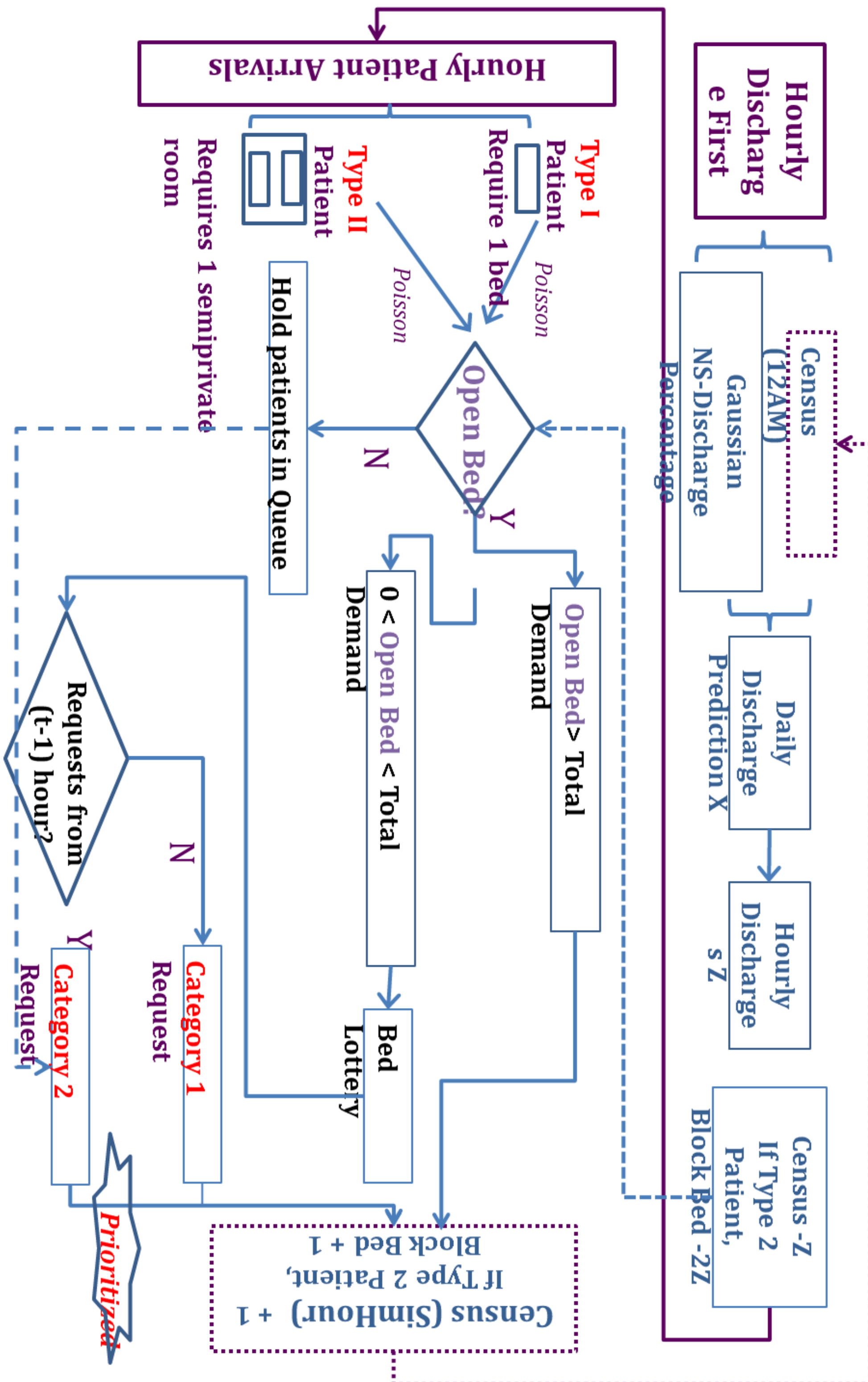
- Predicting Type II arrivals using DRGs or Major Diagnostic Categories (MDCs)
- Discharge prediction at the patient level by discharge destinations (home vs. facility)
- Studying alternative unit design (shared vs. privates) under stressed scenarios
- Modeling hospital-to-community care coordination in parallel to the clinical care
- Studying inter-unit patient transfer dynamics

6.6 Final Words

It is demonstrated that considering blocked beds is worthwhile for a general medicine unit with shared beds but accept patients who require private rooms. In summary, this thesis study identified the key inputs for building a simple unit-level bed crisis alert tool. Both the quantitative and qualitative results help lay the groundwork for more elaborated modeling in future research.

APPENDIX A

FIGURE 27 FULL EXCEL VBA MODEL



APPENDIX B

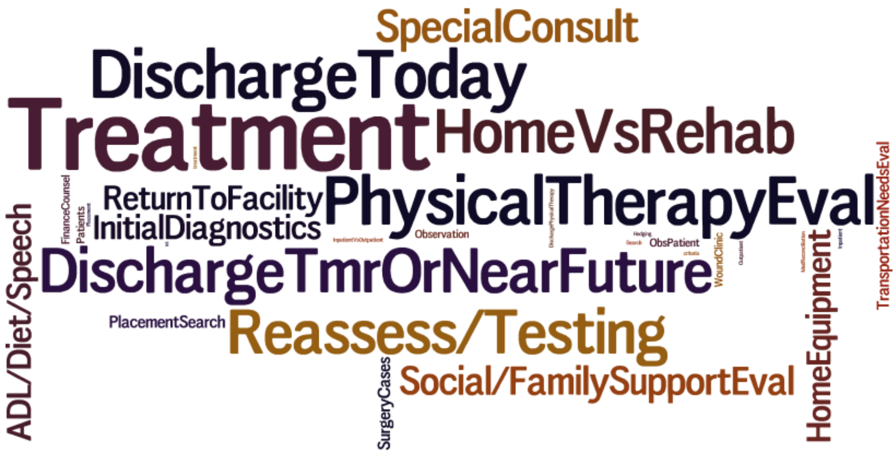
TABLE 6 MU PATIENT LOAD BY MDCS 2010_2011

General Medicine Units (GMU) Load by Major Diagnostic Groups (MDC)

MDC Code	Yearly	Q1	% Total	Q1 vs Total	Note
4			24.19%	3.63%	Respiratory
6			14.67%	0.14%	Digestive
18			8.46%	1.30%	Infectious / Parasite
11			7.79%	-0.46%	
5			6.33%	-1.00%	
10			6.21%	-0.46%	
7			5.89%	-0.77%	
9			5.85%	-1.22%	
21			4.23%	-0.29%	
8			3.19%	-0.70%	
1			3.19%	-0.77%	
16			2.70%	0.00%	Blood
20			1.78%	0.43%	Alcohol/Drug Induced Mental Disorders
3			1.64%	0.51%	Ear Nose Mouth Throat
25			1.15%	-0.04%	
23			0.86%	0.17%	HLTH STAT/Other
19			0.53%	0.09%	Mental Disorders
17			0.26%	-0.20%	
14			0.25%	-0.11%	
12			0.23%	-0.02%	
2			0.19%	-0.06%	
0			0.18%	-0.18%	
13			0.14%	0.00%	
24			0.05%	-0.05%	
22			0.02%	0.05%	Burns
Unassigned			0.02%	-0.02%	

APPENDIX C

FIGURE 28 COMMON IPOC VARIABLES IN MU



APPENDIX D

TABLE 7 TWENTY-SIX IPOC VARIABLES

Var	Description
1	Treatment
2	Discharge tomorrow or near future
3	Home vs Rehab
4	Physical Therapy Evaluation
5	Discharge today
6	Evaluating social and family support
7	Reassess with testing
8	Special Consult
9	Initial Diagnostics or Workups
10	Medication Reconciliation
11	Transportation Need Evaluation
12	Inquiry of ADLs, Diet, Speech
13	Inquiry of Home Equipment
14	Returning to Facility Confirmation
15	Surgery or Post-Surgery
16	Wound Clinic
17	Discharge Physical Therapy
18	New ED Admits
19	Placement Search and Hedging
20	Finance Counsel
21	Observation Patients
22	Fall Risk Complaint
23	Medication Noncompliance
24	Agitated Patients or potential AMA
25	Inpatient vs. Outpatient treatment criteria
26	Post-Surgery boarding, Intercare is full??

APPENDIX E

PSEUDO-CODE

<u>Single-Stream Model Census Balancing</u>

Index

t_n ≡ time index at hourly granularity; 0 to 48, where 0 denotes now

n ≡ replication number; 0 to 1000

q ≡ quarter of year

Parameter

K ≡ unit capacity

x_* ≡ beginning bed census at t_0

z_* ≡ beginning number of blocked beds

L_* ≡ beginning bed queue or backlog at t_0

$\lambda_{t,q}$ ≡ mean hourly bed request rate

$d_{t,q}$ ≡ mean hourly discharge rate

Variable

$Z_{t,q}$ ≡ total ending demand in time period t; integer

$Q_{t,q}$ ≡ *actual capacity adjusted by blocked beds*

$y_{t,q}$ ≡ number of blocked beds in time period t; binary

i ≡ patient type; 1 blocked beds required; 0 otherwise; binary

$x_{t,q}$ ≡ ending bed census in time period t and quarter q; integer

$L_{t,q}$ ≡ ending bed queue size in time period t; integer

$A_{t,q}$ ≡ total bed requests in time period t and quarter q

$D_{t,q}$ ≡ total patients discharges in time period t and quarter q

$A_{i,t,q}$ ≡ bed requests by patient type i in time period t and quarter q

$D_{i,t,q}$ ≡ discharges by patient type i in time period t and quarter q

Ending bed census and ending bed queue are represented as difference equations and calculated via a recurrence definition presented below. First set of equations do not consider blocked beds:

$Z_{0,q} = x_* + L_* + A_{0,q} - D_{0,q}$ (1) ... Initialization

$Z_{t,q} = x_{t-1,q} + L_{t-1} + A_{t,q} - D_{t,q}, \forall t$ (2) … Total Demand

If $Z_{t,q} < K$, then $x_{t,q} = Z_{t,q}$; $L_t = 0$, $\forall t$ (3) … Balancing

If $Z_{t,q} > K$, then $x_{t,q} = K$; $L_t = Z_{t,q} - K$, $\forall t$ (4) … Balancing

Two-Stream Model Bed Assignment and Census Balancing

Preparation

1. Create one 24 by 7 table to store historical arrivals for all patients
2. Multiply the total mean rates above by assumed type ratio for Type I, II patients
3. Store the above data in 24 by 7 tables in spreadsheet and load as arrays in Excel VBA
4. Gather user input of start hour and day from spreadsheet as index
5. Denote the index hour and day as "Now" or the 0th hour
6. Find arrival and discharge rates for the corresponding index in the historical table
7. Walk down the table by the index and gather data for now and the next 48 hour
8. Create 49 by 1 "Input" tables to store the gathered data for Poisson simulation

Simulation

1. Clear spreadsheet cells for four 1000 by 49 tables "Simulated"
2. Walk through 0th to 48th hour of parameter in "Input" table
3. If the hourly parameter in "Input" table is not 0, then simulate Poisson arrivals for 1000 replications and write to tables "Simulated"
4. For census-dependent discharge simulation, see separate peudocode
5. The tables "Simulated" are storing variables named type1request, type2request, type1discharge and type2discharge

Processing simulated changes and update patient census

1. Create several 49 by 1000 tables for variables named census, blocked, type1queue, type2queue, total queue, openbed, m (total hourly bed demand)
2. Let m = type1queue + type2queue * 2 + type1request + type2request * 2

3. Set total queue to the sum of type1queue and type2queue

4. Set openbed to unit capacity minus blocked and census

5. Gather user input of initial parameters including number of patients in the unit (census), unit bed capacity, number of blocked beds, number of patients waiting for beds (bed queue) by patient type

6. Create temporary variables for tracking the progress of lottery admission, including temp_openbed, temp_type1request, temp_type2request, temp_type1queue, temp_type2queue

7. Create one 4*200 table to store bed request indicators for lottery admission. The 200 rows store holding bed requests to be processed. The 4 columns are indicator variables named patient_type, lottery rank, request_category, isProcessed

8. For 1000 replications and 48 hours, get the number of open beds at the end of last hour, type1discharge, type2discharge

9. If simulated type1discharge is less than census minus blocked, and simulated type2discharge is less than blocked, and then subtract discharges from open bed accounting for release of blocked beds when necessary

10. If the condition in 8. Is false, and then subtract the actual number of type 1 or 2 patients present in the unit from open bed to represent the discharge process, accounting for release of blocked beds when necessary

11. Besides the variable open bed, update variable blocked after discharges

12. If there is no open beds or no bed requests after discharge, update queue and move to next hour

13. Proceed to next hour when the total bed demand m for the hour is less than the number of beds. (No queue)

14. When 0 < open bed < total demand, the lottery admission takes place.

15. Write in lottery admission table for processing bed requests, including type1queue, type2queue, type1request, type2request

16. For the variable lottery rank, generate a random number between 0 and 1 for each bed request. Bed request with higher score is favored!

17. Multiply lottery score with 0.1. For instance, a score of 0.56 will become 0.056. Then reward Category 2 request (waiting for more than an hour) with 0.5. The new score will be 0.556 and 0.056 for type 1 or for type 2 patients respectively.

18. Sort bed requests by lottery score from large to small

19. Process bed requests from the top of the list until open bed is 0; update queue and other variables for every hour and every replication

20. Check if the patient type is feasible for the number of open beds, and ignore infeasible ones

Census-Dependent Discharge Simulation

Sets

Day of Week is denoted by i , 1 ... 7
Hour of Day is denoted by j , 0 ... 23
SimHour is denoted by t , 0 ... 48 *(49 hours spread over 3 days)*
SimDay is denoted by s , 0 ... 2
Sim Replication is denoted by n , 0 ... 999
'Calendar quarter is denoted as CQuarter, 1..4

* *Hour of Day = (SimHour + StartHour) mod 24*
* *Day of Week = StartDay + rounddown((SimHour + StartHour)/24,0)*

Parameters

Start Hour of Simulation is denoted by StartHour
Start Day of Week is denoted by StartDay
Mean Historical "12am Census" is denoted as HistoricalCensus(i)
"12am Census" on StartDay is denoted by iCensus
Mean percentage of NS-Discharge by "12am Census" is denoted by dPmean(i)
Standard deviation of NS-Discharge by "12am Census" is denoted by dPStddev(i)
Empirical probability density of discharge is denoted by HourlyDischarge(j, i)
Cumulative distribution of discharge is denoted by DischargeCDF(j,i)

* *NS stands for "non-sameday"*
* HistoricalCensus, dPmean, dPStddev, HourlyDischarge are tables of historical data

Variables

Gaussian simulated daily discharge percentage P(s, n)

Simulated daily discharges is denoted by X(s,n)
Simulated hourly discharges is denoted by Z(t,n)
Random number generated for each discharge "trial" is denoted by score(X)
Simulated 12am Census for 2nd and 3rd SimDay is denoted by sCensus(s)

Initialization

SimHour = 0
SimDay = 0
StartHour = "user input"
StartDay = "user input"
iCensus = "user input" or "historical"

Procedures

1. The table of HourlyDischarge PDF is converted to DischargeCDF
2. Find iCensus, dPmean, dPStddev by Day of Week i from tables of historical data
3. Model daily NS-Discharge percentages with Gaussian distribution
    a. P(s, n) = Norminv(rnd(), dPmean, dPStddev)
    b. Question – Or keep this constant for all 1000 replications?
4. Calculate daily discharges X(s,n)
    a. For StartDay, X(s, n) = round (P(s, n) * iCensus)
    b. For following SimDay s, X(s,n) = round(P(s,n) * sCensus(s))
5. Run through X discharge trials to obtain hourly discharge distribution array Z(t, n)
    a. If X > 0, then generate a random number "score" between 0 and 1 for X trials
    b. Loop through each hour of day j:
        i. If score is greater than or equal to dischargeCDF (j, startday)
        ii. And If score is less than dischargeCDF (j+1, startday)
        iii. Then Z(j,0) += 1

Output

X, Z, sCensus *(Update Open Beds)*

BIBLIOGRAPHY

Agency for Healthcare Research and Quality. HCUPnet - Healthcare Cost and Utilization Project. January 2009. http://hcupnet.ahrq.gov/

AHA Annual Survery of Hospitals (2012 & 2013). American Hospital Association. Retrieved from < http://www.aha.org/ >

Armony, M., Israelit, S., Mandelbaum, A., Marmor, Y. N., Tseytlin, Y., & Yom-Tov, G. B. (2013). Patient Flow in Hospitals: A Data-Based Queuing-Science Perspective. Accepted to *Stochastic Systems*, New York University Stern School of Business.

Bhattacharjee, P., & Ray, P. K. (2014). Patient Flow Modeling And Performance Analysis Of Healthcare Delivery Processes In Hospitals: A Review And Reflections. *Computers & Industrial Engineering*, *78*, 299–312. Http://Doi.Org/10.1016/J.Cie.2014.04.016

Brailsford, S. C., Harper, P. R., Patel, B., & Pitt, M. (2009). An Analysis of The Academic Literature On Simulation And Modelling In Health Care. *Journal of Simulation, 3(3),* 130–140. Http://Doi.Org/10.1057/Jos.2009.10

Brailsford, S., & Vissers, J. (2011). OR In Healthcare: A European Perspective. *European Journal of Operational Research, 212(2),* 223–234. Http://Doi.Org/10.1016/J.Ejor.2010.10.026

Chan, C. W., Dong, J., & Green, L. V. Queues with Time-Varying Arrivals and Inspections with Applications to Hospital Discharge Policies.

Fildes, R., Nikolopoulos, K., Crone, S. F., & Syntetos, A A. (2008). Forecasting and Operational Research: A Review. *Journal of The Operational Research Society, 59(9),* 1150–1172. Http://Doi.Org/10.1057/Palgrave.Jors.2602597

Helm, J. E., AhmadBeygi, S., & Van Oyen, M. P. (2009, January). The flexible patient flow simulation framework. In IIE Annual Conference. Proceedings (p. 803). Institute of Industrial Engineers-Publisher.

Khurma, N., Salamati, F. & Pasek Z.J. (2013). Simulation of Patient Discharge Process and Its Improvement. *Proceedings of the 2013 Winter Simulation Conference,* 2452–2462.

Kim, S. H., Chan, C. W., Olivares, M., & Escobar, G. (2014). ICU Admission Control: An Empirical Study of Capacity Allocation and Its Implication for Patient Outcomes. Management Science, 61(1), 19-38.

Most Unscheduled Hospital Admissions Now Come Through ER. (2013 Jun 20). Retrieved from <http://newsroom.acep.org/2013-06-20-Most-Unscheduled-Hospital-Admissions-Now-Come-Through-ER>

Ozen, A., Balasubramanian, H., Samra, P., Ehresman, M., Li, H., Fairman, T., & Roche, J. (2014). The Impact of Hourly Discharge Rates and Prioritization on Timely Access to Inpatient Beds. *Proceedings of the 2014 Winter Simulation Conference.*

Xu, K., & Chan, C. W. (2014). Using Future Information to Reduce Waiting Times in the Emergency Department via Diversion.