Managing Variability to Improve Quality, Capacity and Cost in the Perioperative Process at Massachusetts General Hospital

By

Devon J. Price

Bachelor of Science in Mechanical Engineering, University of California, Berkeley, 2005

Submitted to the MIT Sloan School of Management and the Engineering Systems Division in Partial Fulfillment of the Requirements for the Degrees of

Master of Business Administration
AND
Master of Science in Engineering Systems

In conjunction with the Leaders for Global Operations Program at the Massachusetts Institute of Technology

June 2011

© 2011 Massachusetts Institute of Technology. All rights reserved.
This page has been intentionally left blank
Managing Variability to Improve Quality, Capacity and Cost in the Perioperative Process at Massachusetts General Hospital

By

Devon J. Price

Submitted to the MIT Sloan School of Management and the Engineering Systems Division on May 6, 2011 in Partial Fulfillment of the Requirements for the Degrees of Master of Business Administration and Master of Science in Engineering Systems

Abstract

The widely held assumption is that to improve access and quality of health care, we need to spend more. In fact, that is not necessarily true. The results of this project, performed at Massachusetts General Hospital (MGH), demonstrate that more sophisticated management of health care processes will lead to greater capacity and higher quality at lower cost.

This work includes system-level analysis of surgical patient flow and reveals several opportunities for performance improvement. The results show that management of variability, both intrinsic to and generated by the perioperative department, will result in lower patient wait times, less crowding, and ultimately higher throughput for surgical patients throughout the hospital.

The solution developed here is an “open block” scheduling policy for the operating rooms at MGH. It was designed with the aid of a discrete event simulation model, which was used to refine the policy and predict the impact of the change. By more effectively characterizing and managing the stochastic demand of non-elective surgical cases, this policy will dramatically reduce delays and open capacity for higher case volume. Specifically, it will reduce the number of non-elective surgical patients exceeding maximum recommended wait time from 30% to 2%; it will free up an average of seven inpatient beds per day; and it will lay the foundation for increased operating room utilization – by up to the equivalent of five operating rooms.

Indeed, this is merely one example demonstrating that by focusing our efforts on creative healthcare system design and management, we can meet the needs of society and spend less doing so.

Thesis Supervisor: Retsef Levi
Title: J. Spencer Standish (1945) Prof. of Management, Assoc. Prof. of Operations Mgmt

Thesis Supervisor: David Simchi-Levi
Title: Professor of Civil and Environmental Engineering and Engineering Systems Div.

1 The perioperative department is responsible for managing and performing surgical procedures, which includes the following major steps: pre-anesthesia health assessments, operating room scheduling and intake, anesthesia and surgery, post anesthesia recovery.
This page has been intentionally left blank
Acknowledgements
This thesis is the culmination of a large amount of work by many people. First, I need to thank my wife Jennifer for her support through this journey. In addition to the adventures it has provided, it has also required challenges and sacrifices. Thank you for everything you’ve done to help make these last two years possible.

I would like to thank the MGH team for the opportunity to work with them on this project. Thanks to Sue Moss\(^2\) and Peter Dunn\(^3\) for championing the drive for operational excellence in the perioperative department. Their vision and ability to create change in an environment as nuanced and complex as MGH is an inspiration. Thanks to Bethany Daily\(^4\). Her leadership, organizational skills and her insight have been instrumental in keeping us all headed in the right direction. And thanks to Marina Maziya\(^5\) for everything she has done to help access and decode the operations of MGH through its universe of data. Her expertise has been invaluable.

And of course, I must acknowledge the outstanding team I have had the pleasure of working with at MIT. Retsef Levi\(^6\), with his inspired passion, relentless energy, dogged persistence, and his ability to bring a group together has been a cornerstone of this effort. Tim Carnes\(^7\) has been exceptional to work with. The hours we’ve spent trading ideas and testing solution strategies have doubled the output of this project. I will be hard-pressed to find comparable teammates in the future. I would also like to sincerely thank David Simchi-Levi\(^8\) and Steven Spear\(^9\) for their wisdom and encouragement. Their experience and perspective have shaped this project dramatically.

\(+2\) Executive Director of the Department of Anesthesia, Critical Care and Pain Medicine
\(+3\) Executive Medical Director of Operating Rooms and Exec. Vice Chair of the Department of Critical Care and Pain Medicine
\(+4\) Administrative Director, Perioperative Strategic/Business Initiatives
\(+5\) OR Administration
\(+6\) J. Spencer Standish (1945) Prof. of Management, Assoc. Prof. of Operations Management
\(+7\) Post Doctorate Fellow, Operations Research Group, MIT
\(+8\) Professor of Civil and Environmental Engineering and the Engineering Systems Division
\(+9\) Senior Lecturer, Engineering Systems Division, MIT; Senior Fellow, Institute for Healthcare Improvement
This page has been intentionally left blank
Table of Contents

Abstract ....................................................................................................................................... 3
Acknowledgements .................................................................................................................. 4
List of Figures ............................................................................................................................9
List of Tables ...........................................................................................................................10
1 Introduction ..................................................................................................................... 11
  1.1 Background.................................................................................................................... 11
  1.2 Massachusetts General Hospital ................................................................................. 11
  1.3 Project Overview.............................................................................................................. 12
  1.4 Random vs. Non-Random Variability ........................................................................... 13
  1.5 Project Structure and Summary of Results ............................................................. 14
2 Characterization of the Perioperative Process.................................................17
  2.1 Introduction.......................................................................................................................... 17
  2.2 Methods .................................................................................................................................. 18
  2.2.1 Data Collection and Analysis.......................................................................................... 18
  2.3 Results .................................................................................................................................... 18
  2.3.1 System Analysis: Inpatient Beds as The Principal Bottleneck ............................. 18
  2.3.2 The Impact of Surgical Scheduling on the Inpatient Bed Census....................... 20
    2.3.2.1 Elective Cases............................................................................................................ 20
    2.3.2.2 Non-Elective Cases ................................................................................................. 22
    2.3.2.2.1 Non-Elective Surgical Delays............................................................................... 22
    2.3.2.2.2 Wait Times and Subsequent Hospital Length of Stay ...................................... 24
    2.3.2.2.3 Causes of Non-Elective Surgical Delays.............................................................. 25
  2.3.3 The Impact of Last-Minute Cases in the OR Schedule ........................................... 25
  2.4 Discussion.............................................................................................................................. 26
3 Design of An Open Block Policy For Non-Elective Surgical Cases...............29
  3.1 Introduction.......................................................................................................................... 29
  3.2 Methods .................................................................................................................................. 29
  3.2.1 Data analysis and simulation....................................................................................... 29
    3.2.1.1 Open block demand .............................................................................................. 30
    3.2.1.1.1 Case pooling ....................................................................................................... 31
    3.2.1.2 Variability in Case Arrivals and Durations............................................................. 33
    3.2.1.3 Open Block Simulation.......................................................................................... 33
  3.3 Results .................................................................................................................................... 34
  3.4 Discussion.............................................................................................................................. 36
4 Appendix ........................................................................................................................... 39
  4.1 Probability Distribution Functions: Service Groups 1 & 2............................................. 39
    4.1.1 Quantity of Weekday Waitlist Bookings................................................................. 39
    4.1.2 Timing of Weekday Waitlist Arrivals ....................................................................... 39
    4.1.3 Duration of Surgical Procedures .............................................................................. 40
4.1.4 Operating Room Turnover Times

5 Works Cited
List of Figures

FIGURE 1: THE PERIOPERATIVE PROCESS IS CENTRAL TO THE MGH PATIENT VALUE STREAM. OVER 40% OF ALL ADMISSIONS VISIT THE OPERATING ROOMS. PATIENTS CAN ENTER THE PERIOPERATIVE PROCESS FROM HOME OR ANY OTHER AREA OF THE HOSPITAL. .................................................................................................................. 13

FIGURE 2: CENSUS STATISTICS FOR EACH AREA OF THE HOSPITAL IN THE TYPICAL SURGICAL PATIENT’S PATH. AT MIDDAY IN THE MIDDLE OF THE WEEK, OVERCAPACITY CONDITIONS ARE LIKELY TO EXIST IN THE SURGICAL INPATIENT BEDS. THIS PROPAGATES BACK THROUGH THE PACU AND OR TO CAUSE CROWDING AND DELAYS. 19

FIGURE 3: THERE IS OFTEN A DELAY BETWEEN WHEN A PATIENT IS READY TO LEAVE THE PACU AND WHEN THE PATIENT IS DISCHARGED TO AN INPATIENT BED. THESE DELAYS COINCIDE WITH PERIODS OF OVERCAPACITY ON THE FLOOR: IN THE MIDDLE OF THE DAY AND MORE FREQUENTLY IN THE MIDDLE OF THE WEEK. THIS IS CLEARLY UNDESIRABLE FROM THE PATIENT PERSPECTIVE, BUT ALSO FROM A BUSINESS PERSPECTIVE AS IT IS MORE EXPENSIVE TO STAFF THE PACU THAN AN INPATIENT BED. ANOTHER SOURCE OF DELAY CAN INCLUDE A SHIFT CHANGE AS REPRESENTED BY THE REDUCTION IN PATIENTS LEAVING THE PACU AT 19:00. .......... 20

FIGURE 4: AVERAGE NUMBER OF CASES PERFORMED EACH WEEKDAY. THE AVERAGE OF NON-ELECTIVE AND SCHEDULED ELECTIVE CASES PERFORMED IS CONSISTENT ACROSS THE WEEK WITH THE EXCEPTION OF THURSDAY. AVERAGE OF TOTAL CASES PERFORMED EACH DAY IS INCLUDED FOR REFERENCE. .............. 21

FIGURE 5: SEGMENTATION OF PATIENTS IN THE AVERAGE MIDNIGHT INPATIENT POPULATION. SCHEDULED ELECTIVE CASES WITH A LENGTH OF STAY OF LESS THAN ONE WEEK ACCOUNT FOR THE RISE AND FALL OF THE CENSUS OVER THE COURSE OF THE WEEK. .................................................................................................................. 22

FIGURE 6: NUMBER OF NON-ELECTIVE WAIT LIST CASES EXCEEDING THE MAXIMUM RECOMMENDED WAIT TIME. 30% OF ALL WAIT LIST CASES IN THE YEAR BETWEEN JULY 1ST 2009 AND JUNE 30TH 2010 EXCEEDED THE MAXIMUM RECOMMENDED WAIT TIME. ........................................................................................................ 23

FIGURE 7: NUMBER OF NON-URGENT WAITLIST CASES PERFORMED WITHIN A GIVEN NUMBER OF DAYS AFTER BOOKING. ................................................................................................................................. 24

FIGURE 8: CAUSES OF DELAYS OF URGENT WAIT LIST CASES. ................................................................................................................................. 25


FIGURE 10: AVERAGE DAILY WAITLIST CASE VOLUME BY SERVICE. THE ONLY SERVICES WITH VOLUME SUFFICIENT TO JUSTIFY DEDICATED BLOCKS ARE ORTHOPAEDIC AND EMERG/URG SURGERY. .................................................................................. 31

FIGURE 11: AVERAGE DAILY WAITLIST CASE VOLUME BY OR GROUP. GROUPS 1 AND 2 HAVE SUFFICIENT VOLUME TO JUSTIFY AT LEAST ONE DEDICATED OPEN BLOCK. ................................................................. 32

FIGURE 12: REPRESENTATION OF OPEN BLOCK SIMULATION .................................................................................... 34

FIGURE 13: NUMBER AND PERCENTAGE OF WAITLIST CASES EXCEEDING MAXIMUM RECOMMENDED WAIT TIME PRIOR TO SURGERY; ACTUAL DATA IS COMPARED WITH SIMULATED RESULTS. ........................................................................... 35

FIGURE 14: AVERAGE OR-HOUR UTILIZATION PER DAY - COMPARISON BETWEEN ACTUAL AND WITH SIMULATED OPEN BLOCKS. ........................................................................................................ 36
List of Tables

TABLE 1: EFFECT OF WAITLIST CASE WAIT TIME ON SUBSEQUENT HOSPITAL LENGTH OF STAY. MEANS COMPARED USING STUDENT'S T.................................24
TABLE 2: SERVICE GROUPINGS FOR POOLED WAITLIST OPEN BLOCKS. CASES FROM SERVICES IN THE SAME GROUP CAN BE ACCOMMODATED IN THE SAME OPERATING ROOMS AND WITH THE SAME ANESTHESIA/RN STAFF.......................31
TABLE 3: TOTAL NUMBER OF WAITLIST CASES BOOKED BETWEEN JULY 1ST 2009 AND JUNE 30TH 2010. GROUP 1 AND 2 CASES REPRESENT 80% OF ALL WAITLIST CASES.................................................................32
TABLE 4: PERCENTAGE OF GROUP 1 AND 2 WAITLIST CASES EXCEEDING MAXIMUM RECOMMENDED WAIT FOR THOSE BOOKED ON WEEKDAYS AND ON WEEKENDS OR HOLIDAYS.................................................................33
TABLE 5: COMPARISON OF INPATIENT BED-DAYS CONSUMED BY URGENT AND NON-URGENT WAITLIST PATIENTS WITH AND WITHOUT OPEN BLOCKS..............................................................................35
TABLE 6: PROBABILITY DISTRIBUTION FUNCTIONS REPRESENTING QUANTITY OF BOOKINGS ON WEEKDAYS FOR EMERGENT, URGENT AND NON-URGENT WAITLIST CASES IN SERVICE GROUPS 1 AND 2.................................................................39
TABLE 7: DISTRIBUTIONS OF BOOKING TIMES ON WEEKDAYS FOR EMERGENT, URGENT AND NON-URGENT WAITLIST CASES IN SERVICE GROUPS 1 AND 2..................................................................................39
TABLE 8: PROBABILITY DISTRIBUTION FUNCTIONS REPRESENTING PROCEDURE DURATIONS FOR EMERGENT, URGENT AND NON-URGENT CASES IN GROUPS 1 AND 2..................................................................................40
TABLE 9: PROBABILITY DISTRIBUTION FUNCTIONS REPRESENTING LENGTH OF OPERATING ROOM TURNOVER TIME BETWEEN SURGERIES FOR GROUP 1 AND 2 WAITLIST CASES.................................................................40
1 Introduction

1.1 Background

For decades, hospitals across the country have faced challenges to providing consistent, sufficient patient access. Crowding, delays and unpredictable schedules can prevent patients from receiving timely care and lead to difficult working conditions for hospital staff (Andrulis, et al. 1991), (U.S. Government Accountability Office 2009). There are two basic approaches to solving this problem, 1) add capacity, and 2) use existing capacity more efficiently. Adding capacity is the most straightforward solution and historically it has been the option most frequently pursued. In the face of rising healthcare costs however, required capital is difficult to come by, the quantity of healthcare resources remains flat, and we are forced to endure long wait times and insufficient care. Therefore, the urgency to use existing capacity more efficiently through innovative operations management is increasing. Of course there are several challenges to doing so.

The challenge of successfully implementing system-wide operations improvements in hospitals can be traced to two primary causes: process complexity and organizational design. Hospitals are highly complex systems, coordinating thousands of activities to resolve a large quantity and diverse mix of specialized patient needs. As such, it is difficult to identify policy changes that benefit patients of a large set of services without sub-optimizing the experience of others. In addition, hospital organizations are typically designed around functions. This is ideal for developing deep expertise in solving specific, localized patient problems, but there is comparatively little institutional expertise in collaborating across functions to improve the system as a whole. As a result, functional silos are prominent, information flow is challenged, and system wide-change is slow.

1.2 Massachusetts General Hospital

As one of the largest hospitals in the US, Massachusetts General Hospital (MGH) is no exception. MGH operates over 900 inpatient beds and accommodates over 80,000 patient visits per year. It is also home to 78 medical services, 52 operating rooms and performs over 34,000 surgical cases per year. This high volume of patients with a high-mix of medical conditions requires dynamic processes and creates extremely complex and variable demand on resources.

In addition to the complexity introduced by the patient mix and organizational functions, MGH has a four-tiered mission. The mission goes beyond patient care to include education, research and community service. These often competing agendas are layered on top of the priorities of the individual functions, and can make institutional change even more difficult. Operating policy changes cannot diminish the hospital’s ability to deliver in any of these dimensions.
Recently, two factors beyond patient crowding and reluctance to add capacity have heightened the urgency to improve efficiency of the healthcare delivery process at MGH: 1) in 2009 spending growth at MGH exceeded revenue growth and 2) healthcare legislation in Massachusetts is driving a shift in payment method from fee-for-service to a bundled and global payment system, which will tie hospital financial performance more closely to efficient management of resources. MGH has partnered with MIT Sloan and the Leaders for Global Operations program to address this challenge through improved operations strategies that enhance efficiency and patient flow.

1.3 Project Overview
The goal of this project is to characterize patient flow in the perioperative process in order to reveal practicable system design changes that will increase quality, access and patient throughput without added capital or operating costs. The approach involves close analysis of patient flow data, and development of a discrete event simulation model to support solution design.

The solution developed focuses on management and reduction of system variability. MGH achieves high levels of utilization (~85%) relative to the average hospital in the U.S. (66% to 67%) (Organization for Economic Cooperation and Development 2010). One doesn’t need to look any further than queuing theory to understand that maintaining this level of utilization while reducing wait time and improving patient access will require improved management and reduction of system variability.

We use the perioperative process as a starting point for system improvement for several reasons. First, surgical cases are one of the primary revenue generators for MGH. Efficiency gains in the perioperative process have the largest financial impact on the hospital. Second, over 40% of all hospital admissions go through the operating room. Improvements in other areas of the hospital can be easily diminished or negated by an inefficient perioperative process. Finally, 85% of the surgical cases are elective and scheduled explicitly, providing a level of control over demand on resources not possible in other parts of the hospital.

Figure 1 is an illustration of the MGH patient value stream. It shows the major areas of the hospital a patient might visit during his or her stay. They include the Emergency Department (ED), the Intensive Care Unit (ICU), the Operating Rooms (OR), the Post-Anesthesia Care Unit (PACU), in-patient beds, the Pre-Admissions Testing Area (PATA) and other ambulatory clinics. The Perioperative department consists of areas in the hospital a surgical patient will visit during a surgical procedure and recovery from anesthesia. Patients can enter the perioperative department from, and be discharged to, almost any other part of the hospital. Optimal system design changes will only be found after close consideration of these interactions.
Figure 1: The perioperative process is central to the MGH patient value stream. Over 40% of all admissions visit the operating rooms. Patients can enter the perioperative process from home or any other area of the hospital.

1.4 Random vs. Non-Random Variability
Given the immense variety of patient needs, and the value placed on a practitioner’s ability to adapt and find creative ways to solve problems, variability is accepted and even embraced in a hospital setting. Indeed, unexpected patient needs are often cited as the cause of delays for subsequent patients, and it’s through heroic efforts that these problems are resolved. Because a large amount of variability is unavoidable and flexibility is needed to accommodate it, finding ways to reduce variability is difficult and gaining buy-in for policy that may limit flexibility, even more so.

When characterizing demand variability and developing solutions, it is useful to make a distinction between random and non-random variability. Random variability is inherent to the system and uncontrollable. The timing of emergent case arrivals and the nature of the medical conditions arriving patients might have are two examples. While uncontrollable, these random causes can be predicted based on historical observation and accommodated with flexible processes and buffer capacity. Non-random causes of variability on the other hand, can be controlled and eliminated to increase effective capacity without adding resources. These are typically the result of non-standard process, resulting in differences in the ways that similar work gets done. This is the reason a person with the same conditions can come to the hospital a second time and have a very different experience than he or she did the first time.
The distinction between random and non-random variability has been a framework central to industrial quality improvement since the development of statistical process control by Walter A. Shewhart in the 1930s (Shewhart 1931). W. Edwards Deming popularized the concept among manufacturing firms in Japan in the 1950s and in the United States in the 1980s (Deming 1982). Recently, principals of variability management have been applied to hospital operations improvement as well. (Litvak, Optimizing Patient Flow by Managing its Variability 2005)

### 1.5 Summary of Results

This project was conducted in two phases. The first phase focused on characterizing patient flow in the perioperative process with emphasis on distinguishing between sources of random and non-random variability. The second phase focused on developing practical changes to operating policy that would reduce variability and lay the foundation for increased quality, access and patient throughput. This document is organized according to the same framework. Chapter 2 describes the data analysis and reveals several system-level causes of non-random variability that can be addressed to reduce crowding and delays without added cost. Chapter 3 describes the detailed development of a solution for one of the highest priority opportunities identified. This involved further system characterization and simulation.

The data analysis presented in Chapter 2 includes several important findings. First, inpatient beds\(^{10}\) are the principal bottleneck in surgical patient flow. Lack of available beds frequently leads to delays and surgical cancellations. We also find that surgical scheduling practices contribute to inpatient bed crowding by creating uneven demand on bed resources. This uneven demand results in over utilization in the middle of the week, and underutilization at the beginning and end of the week. It may be somewhat surprising that this fluctuation is not the result of varying numbers of inpatient cases performed each day; rather, it is the result of uneven scheduling of inpatients with a post-surgical length of stay of less than one week.

Second, 30% of non-elective surgical patients are forced to wait longer than the maximum recommended time for surgery. In addition, patients exceeding the maximum wait time are more likely to have a longer stay in the hospital after surgery. This is bad news for the non-elective patient; but it also exacerbates the inpatient bed resource crunch, leading to more crowding and delays for other patients. Excessive wait times are most commonly caused by lack of timely operating room or staff availability. This too can be traced to suboptimal OR scheduling practices.

---

\(^{10}\) An operational inpatient bed includes the physical bed itself and the staff required to care for the patient occupying the bed.
Current OR scheduling practice seeks to accommodate non-elective patients, arriving randomly on the day of surgery, among previously scheduled patients. This limits scheduled OR utilization, contributes to the shuffling of 28% of the OR schedule on the day of surgery, and makes it difficult to marshal all resources required for a non-elective case in the required amount of time.

Chapter 3 details the development and outlines implementation for an open block operating room scheduling policy at MGH. This policy creates dedicated operating rooms and staff resources to handle all non-elective surgeries. This will isolate elective cases performed in scheduled blocks from the variability of non-elective case arrivals. The policy is expected to reduce the number of non-elective surgical patients exceeding maximum recommended wait time from 30% to 2%. Accommodating these patients more quickly is also expected to free up an average of seven inpatient beds per day. Finally, removing these random arrivals from scheduled blocks will shift the upper bound on scheduled OR utilization to 100%, which if achieved, would be the equivalent of gaining the capacity of five operating rooms.
2 Characterization of the Perioperative Process

2.1 Introduction
According to a 2009 US Government Accounting Office report, lack of access to inpatient beds is the main factor contributing to crowding in hospitals across the country (U.S. Government Accountability Office 2009). This is reflected in the following assessment of the perioperative process at Massachusetts General Hospital. Inpatient beds are the principal bottleneck in surgical patient flow. Reducing frequency of inpatient bed overutilization should be a top priority in reducing surgical delays and improving patient throughput at MGH.

We have found two important causes of periodic inpatient bed crowding: 1) uneven scheduling\(^\text{11}\) of short-recovery elective surgical patients\(^\text{12}\) and 2) delayed waitlist patients\(^\text{13}\). The population of short-recovery elective surgical patients typically rises and falls over the course of a week. This rise and fall pushes the total average inpatient bed population to capacity in the middle of the week, while leaving inpatient beds at the beginning and the end of the week underutilized.

If not accommodated on the day of arrival, waitlist patients consume inpatient bed capacity while waiting for surgery. This is primarily an issue for non-urgent waitlist patients, who may be delayed longer than 24 hours. Furthermore, delays can cause waitlist patients to stay longer in the hospital after surgery. On average, both non-urgent and urgent waitlist cases have extended lengths of stay if they are forced to wait longer than the maximum recommended amount of time. The impact of short-recovery elective surgical patients and delayed waitlist patients on inpatient bed utilization reflect

---

\(^\text{11}\) Surgical scheduling at MGH is done using a block system. Each surgical service (eg. Orthopedic, Transplant, Cardiac, etc...) is allocated blocks of OR time spread over the course of a week. The number of blocks a service receives is determined by the Perioperative Department based primarily on historical patient volume. The services are responsible for scheduling their elective patients into the OR blocks ahead of the date of surgery.

\(^\text{12}\) These include non-ambulatory elective surgical patients with a subsequent length of stay less than seven days.

\(^\text{13}\) Non-elective surgeries, which must be performed within 24 hours of patient arrival due to severity of his or her condition, are placed on a wait list. Waitlist patients are given emergent, urgent and non-urgent designations. Emergent cases must be accommodated within 45 minutes of arrival, urgent cases within 4 hours and non-urgent cases within 24 hours. A patient must be admitted to the hospital to be placed on the waitlist. Therefore, some waitlist patients may occupy an inpatient bed prior to surgery. Waitlist cases are typically accommodated within OR blocks allocated to the services. The case will either be performed at a time that hadn’t been previously scheduled for an elective case or, if an open time cannot be found, an elective case will be delayed or cancelled.
variability in the system that can be dramatically reduced through more effective management.

Positive correlation between wait time and subsequent length of stay is not unique to MGH. Studies in 2001 and 2004 also found that longer wait times for emergency procedures lead to increased lengths of stay in the hospital and higher mortality rates (Bell and Redelmeier 2004), (Bell and Redelmeier 2001).

This chapter presents an analysis of surgical patient data that reveals the causes cited above, and discusses solutions to eliminate them.

2.2 Methods

2.2.1 Data Collection and Analysis
We used a dataset containing information on de-identified surgical patients admitted over a one-year period from July 1st 2009 to June 30th 2010. The data includes patient booking classifications and time stamps marking moves from one location in the hospital to another. This period was chosen because it includes a large sample size of 33,869 surgical cases and captures any seasonal effects on patient volume or type. Two hospital databases were used: one database that tracks patients through the perioperative process and a second database that tracks patients through the other portions of their hospital stay. Using the timestamp data, analyses were performed to highlight variability, delays and periods of crowding in the perioperative process. That information, combined with knowledge of the scheduling process, was used to test conjectures about the causes of the performance issues.

2.3 Results

2.3.1 System Analysis: Inpatient Beds as The Principal Bottleneck
As noted by others in the healthcare field, detecting the bottleneck in patient flow is not straightforward (Litvak, Optimizing Patient Flow by Managing its Variability 2005). Fluctuations in resource availability and demand throughout the perioperative process can cause the bottleneck to shift over the course of a day or week. Furthermore, a patient may be delayed multiple times and for any number of reasons over the course of his or her stay – the surgeon may be tied up with another patient in the morning and an inpatient bed may not be available in the afternoon. If we wanted to find the bottleneck by looking at where patients are most frequently delayed, it would be very difficult to obtain sufficient data. We can however deduce this information by observing occupancy trends in various areas of the hospital over time.

Figure 2 shows census statistics for each area of the hospital in a typical surgical patient’s path: Operating room, post anesthesia care unit (PACU) and surgical inpatient beds.
These charts provide strong evidence that inpatient beds are the principal bottleneck, and that periods of overutilization contribute to delays and cancellations throughout the perioperative process. When the inpatient beds reach capacity, patients are delayed in the PACU, which causes the PACU population to reach capacity. When the PACU is at capacity, patients are forced to wait in the operating rooms. When an operating room is unavailable, or there will be nowhere for a patient to go after surgery, the surgery must be delayed or cancelled.

![Graphs showing data on inpatient beds, PACU, and OR occupancy.](image)

Figure 2: Census statistics for each area of the hospital in the typical surgical patient’s path. At midday in the middle of the week, overcapacity conditions are likely to exist in the surgical inpatient beds. This propagates back through the PACU and OR to cause crowding and delays.

The chart on the right of Figure 2 shows the average surgical inpatient bed occupancy for each day of the week at midday and at midnight. Two pieces of information are immediately clear. First, beds are more highly utilized during the day than at midnight. This is caused by the fact that the discharge process lags the admission process. In other words, most of the discharge activity occurs in the afternoon, after most of the patients arriving that day have been admitted. The discharge process certainly merits a closer look and may represent an opportunity to reduce midday crowding. Second, average utilization is uneven over the course of the week. The population peaks at inpatient bed capacity of 338 on Wednesday but remains substantially below capacity Friday through Monday. Indeed, these are averages; the census can reach capacity any day of the week. However, it is more likely to reach capacity Tuesday through Thursday.

We can see the impact of inpatient bed occupancies reaching capacity in the charts in the center and on the left of Figure 2. The center chart shows the rise and fall of the typical PACU census over the course of a weekday. Peak occupancy in the middle of the day coincides with the time at which the inpatient beds are most likely to be at capacity. This suggests that patients are often delayed in the PACU because there is no inpatient bed available for them to move into. Further evidence is presented in Figure 3, which shows a delay between when an inpatient is ready to leave the PACU and when he or she is discharged to the floor. Delays peak in the middle of the day and are worst in the middle
of the week. Notice that patients to be discharged to home leave as soon as they are medically ready.

The chart on the left of Figure 2 shows the average number of patients waiting more than 30 minutes in the operating room prior to being moved to the PACU. In the middle of the day, there are typically five patients that have been waiting more than 30 minutes. Usually, at least one of them has been waiting more than an hour.

![Chart showing average number of patients waiting more than 30 minutes in the operating room prior to being moved to the PACU.](chart.png)

**Figure 3:** There is often a delay between when a patient is ready to leave the PACU and when the patient is discharged to an inpatient bed. These delays coincide with periods of overcapacity on the floor: in the middle of the day and more frequently in the middle of the week. This is clearly undesirable from the patient perspective, but also from a business perspective as it is more expensive to staff the PACU than an inpatient bed. Another source of delay can include a shift change as represented by the reduction in patients leaving the PACU at 19:00.

While there is no single cause of crowding and delays throughout the perioperative system, periods of overutilization of the inpatient beds have a substantial impact and should be addressed first. Any performance gains from improvements made upstream of the inpatient beds would be diminished or negated if the patient doesn’t have a bed to move into.

### 2.3.2 The Impact of Surgical Scheduling on the Inpatient Bed Census

#### 2.3.2.1 Elective Cases

We might expect that variation in volume of inpatient elective cases scheduled each day would be the primary contributor to the rise and fall of the inpatient bed census – that the census peaks on Wednesday in other words, because more surgeries are scheduled on Wednesdays. In work published by Litvak, this has generally been the case (Litvak, Managing Unnecessary Variability in Patient Demand to Reduce Nursing Stress and Improve Patient Safety 2005). In many hospitals, simply scheduling an equal number of cases on each day of the week can substantially reduce peaks in demand on hospital resources. As shown in Figure 4, however, at MGH the average number of inpatient cases...
performed is constant across the week at 80 patients per day. The exception is Thursday, which has an average of 60. This is because fewer OR-hours are available on Thursdays due to grand rounds held in the morning. While five fewer elective cases are typically performed on Mondays and Fridays than on Tuesdays and Wednesdays, non-elective cases make up the difference.

![Average Surgical Volume by Day of the Week](image)

Figure 4: Average number of cases performed each weekday. The average of non-elective and scheduled elective cases performed is consistent across the week with the exception of Thursday. Average of total cases performed each day is included for reference.

To determine the cause of the rise and fall inpatient bed census over the course of the week, we need to look a little deeper. Figure 5 shows the segmentation of patients in the inpatient bed population. We see here that the rise and fall of the average population results almost entirely from an imbalance of inflow and outflow of patients with a length of stay less than 7 days. It is this category of patients that should be the focus of a solution to smooth the inpatient bed census. Leveling this population would free up over 11 inpatient beds on Wednesdays.
2.3.2.2 Non-Elective Cases

Waitlist cases are also a notable part of the inpatient bed population. On average, there are 12 waitlist patients occupying inpatient beds while waiting for surgery. These are all non-urgent waitlist cases. According to policy, a patient must be admitted to the hospital to be added to the waiting list to prevent it from being used as a scheduling tool to supplement the normal block scheduling process. The recommended maximum wait time for non-urgent wait list cases is 24 hours.

This leads us to another important question: how often do non-elective cases exceed the maximum recommended time on the wait list?

2.3.2.2.1 Non-Elective Surgical Delays

In the year from July 1st 2009 to June 30th 2010, 30% of all patients booked on the wait list exceeded the maximum recommended wait time between booking and the start of surgery. The maximum wait time for non-urgent patients is 24 hours, for urgent patients is 4 hours and for emergent patients is 45 minutes. Figure 6 shows the total quantity of cases that exceeded and did not exceed the wait time limits.
Figure 6: Number of non-elective wait list cases exceeding the maximum recommended wait time. 30% of all wait list cases in the year between July 1st 2009 and June 30th 2010 exceeded the maximum recommended wait time.

As discussed in the previous section, delayed non-urgent waitlist cases occupy a significant portion of inpatient bed capacity while waiting for surgery. Figure 7 shows the distribution of waiting times for these patients. The patients that wait zero days are accommodated on the day they were booked, the patients that waited one day were accommodated on the day after they were booked, and so on. While some of the cases waiting until the next day waited fewer than 24 hours, they did occupy inpatient bed capacity. Non-urgent cases occupied 4236 bed days or about 12 beds per day prior to surgery. This corresponds with the data shown in the inpatient bed population segmentation in Figure 5.
2.3.2.2.2 Wait Times and Subsequent Hospital Length of Stay

Delayed wait list cases also contribute to the inpatient bed population after surgery has been completed more than those that are not delayed. Table 1 shows the difference in hospital length of stay for wait list patients that did and did not exceed the maximum recommended wait time. On average, non-urgent cases stay one day longer and urgent cases stay two days longer if they are delayed. Of course, this represents an additional draw on inpatient bed capacity. On average, over five inpatient beds were occupied every day between July 1st 2009 and June 30th 2010 by non-elective patients who were not accommodated in the recommended amount of time.

Table 1: Effect of waitlist case wait time on subsequent hospital length of stay. Means compared using Student's t.

<table>
<thead>
<tr>
<th>Case Category</th>
<th>Post-Surgical Length of Stay (Days)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wait for Surgery Did Not Exceed Max Wait</td>
<td>Wait for Surgery Exceed Max Wait</td>
<td>Difference</td>
<td>P-Value</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>Std Dev</td>
<td>Mean</td>
<td>Std Dev</td>
</tr>
<tr>
<td>Non-Urgent</td>
<td>7.2</td>
<td>9.6</td>
<td>8.2</td>
<td>10.5</td>
</tr>
<tr>
<td>Urgent</td>
<td>6.8</td>
<td>9.9</td>
<td>8.8</td>
<td>12.2</td>
</tr>
<tr>
<td>Emergent</td>
<td>11.4</td>
<td>12.4</td>
<td>10.8</td>
<td>12.3</td>
</tr>
</tbody>
</table>

The means of length of stay for emergent cases that did and did not exceed the maximum wait time are indistinguishable based on the dataset we had available. This may be because the sample size is too small, or because the lengths of stay are truly indistinguishable between the two populations. Analysis of additional data would be required to draw a conclusion about emergent cases.
It is tempting to conclude that the length of stay is longer for patients that exceed maximum wait time because their conditions are more likely to have deteriorated during the wait. The recommended wait times are established to ensure that patients with severe conditions are accommodated in time to treat them with the highest probability of success. However, we cannot draw this conclusion unless we compare the acuity levels of each patient population. It is possible, for instance, that a larger proportion of patients that were delayed had higher levels of acuity when they were booked, and that the wait time did not impact their acuity prior to or after surgery. Until that analysis is performed, we can only observe the fact that the delayed patients have longer lengths of stay than the patients who were not delayed.

2.3.2.2.3 Causes of Non-Elective Surgical Delays
The predominant cause of delays for non-elective surgical cases is resource availability. As shown in Figure 8, an analysis of 173 urgent wait list cases that waited over four hours shows that 43% of the delays result from lack of resource (OR, surgeon or staff) availability. An additional 35% resulted from an original misclassification of the patient as non-urgent. This typically happens at night when a less experienced surgeon classifies the patient. In these cases, the status of the patient is changed from non-urgent to urgent when a more senior surgeon evaluates the patient upon arrival in the morning. By then, the patient had already waited over four hours. 11% of the patients exceeded four hours due to a delay in preparing the patient for surgery, and for the last 11%, a patient issue prevented the timely start. 89% of all urgent cases exceeding maximum wait time were the result of resource or process issues that can be addressed through policy and process changes.

![Cause Codes for Urgent Cases Exceeding 4 Hour Wait Time](image)

Figure 8: Causes of delays of Urgent wait list cases

2.3.3 The Impact of Last-Minute Cases in the OR Schedule
On average, scheduled operating room utilization jumps from 73% at 5:00pm on the day prior to surgery to 86% actual utilization on the day of surgery. 15% of the total surgical volume is scheduled at the last minute. Every day, the operating room staff makes heroic efforts to accommodate wait list cases among previously scheduled cases. However, injecting this type of variability into a system that is already highly utilized results in delays for scheduled patients and chaos for OR staff.

Further illustration of the heroics involved in accommodating surgical cases on the day of surgery can be seen in Figure 9. This shows that 28% of previously scheduled cases\textsuperscript{14} are moved in either time or location on the day of surgery. While waitlist cases are often used to fill holes in the elective schedules, if a hole does not exist, and surgeon and staff are not ready and waiting, cases must be moved and time spent gathering the required resources. This complex and reactive schedule shuffling processes drives variability and inefficiency into the system. As we've seen, lack of resource availability is a primary cause of excessive waiting times for waitlist patients.

In addition to the impact on patients and staff directly involved, the effects of this schedule churn ripple throughout. Bumped or moved elective patients cause other cases to be delayed and force non-prime time starts, which, due to increased overtime wages, cost the hospital more than they would have had they started on time.

2.4 Discussion
Analysis of process data for 33,869 surgical patients over a one-year period uncovered several important findings about the performance of the perioperative process.

\textsuperscript{14} These are cases that are scheduled before 5:00pm on the day prior to surgery.
First, inpatient beds are the principal bottleneck, frequently contributing to delays throughout the perioperative process, particularly in the middle of the week. Surgical patients with a post-surgery length of stay less than seven days are the primary contributor to a rise and fall of the inpatient bed census over the course of the week. Leveling this population would free up over 11 inpatient beds on Wednesdays.

Second, 30% of wait list patients exceed the recommended maximum wait time between booking and the start of surgery. Non-urgent and urgent waitlist cases that exceed the recommended maximum wait time tend to have longer lengths of stay. On average, non-urgent patients stay one day longer and urgent patients stay two days longer. These patients occupy an average of 5 inpatient beds every day that would not have been occupied if the patients had been accommodated within the maximum recommended wait time. In addition, an average of 12 beds are occupied every day by non-urgent wait list cases waiting for surgery.

The primary cause of waitlist delays is lack of resource availability. Waitlist cases arrive randomly and must be accommodated among previously scheduled elective cases. This results in a substantial amount of schedule churn in the OR. 15% of all cases are scheduled on the day of surgery, and 28% of cases that were scheduled prior to the day of surgery are either moved or bumped, or both.

Two promising potential solutions emerge from this analysis: 1) create dedicated open blocks for waitlist patients and 2) optimize the elective block schedule based on expected subsequent length of stay to level the inpatient bed population.

Creating dedicated open blocks would have two important desirable effects. First, it would ensure that sufficient resources are available to reduce the number patients exceeding maximum wait time. This would benefit the patients in critical need of timely surgery and at the same time reduce the load on inpatient bed resources, which would reduce the probability of delays for all patients. Second, it would remove the variability in the scheduled elective blocks introduced by random waitlist patient arrivals. Improving the predictability of elective blocks would provide a foundation for achieving higher levels of utilization in the future.

Optimizing the elective block schedule would be accomplished using predictions of length of stay profiles for patients of various services. With this, it may be possible to arrange the timing of surgeries for each service to balance the inflow and outflow of patients to the inpatient beds, thereby reducing overcapacity events. This idea can be extended to the PACU by changing the timing of inpatient and outpatient surgeries over the course of the day to level the census of the PACU and reduce time spent waiting in the OR.
The following chapter explores in detail the feasibility, design and impact of a dedicated open block policy for MGH. A solution for the optimization of the elective block schedule using an integer program is currently in development and further discussion will be left to a future publication.
3 Design of An Open Block Policy For Non-Elective Surgical Cases

3.1 Introduction
In chapter 2, it became clear that high priority objectives in improving quality, capacity and cost in the perioperative process at MGH include 1) minimizing the frequency of waitlist case delay due to lack of available resources and 2) managing the variability in the operating room schedule to improve predictability. Creating dedicated open blocks to accommodate waitlist patients was chosen as a solution based on its potential to simultaneously address both of these objectives.

The merit of removing waitlist cases from the scheduled population may not be initially obvious. The current scheduling process relies on waitlist patient arrivals to achieve 85% utilization. As shown in section 2.3.3 waitlist cases fill holes in the elective schedule on the day of surgery. This reliance on holes in the schedule to accommodate waitlist cases however, removes incentive to achieve 100% scheduled utilization. The cost of higher scheduled utilization is longer wait times for waitlist patients, more delayed surgeries and higher schedule churn on the day of surgery. In order to improve utilization of scheduled surgical blocks beyond 85%, waitlist cases must be accommodated in separate blocks. While this may lead to lower utilization of scheduled blocks in the short term while scheduling processes are improved, it will be impossible to achieve higher levels of utilization in the long term without it.

To arrive at a practicable open block design we need to determine how much and for how many patients, open blocks can reduce waitlist wait time while still achieving reasonable utilization of the dedicated resources. As we know from queuing theory, there is a tradeoff between waiting time and resource utilization dictated by the degree of variability in patient arrivals and processing time. Furthermore, there are limitations on which cases can be accommodated by which resources. Certainly a single surgeon does not have the specialized expertise to operate on every type of non-elective case.

The following sections explore these issues through data analysis and simulation. We find that five open blocks operating from 7:00am to 5:00pm is sufficient to accommodate 80% of all weekday waitlist cases. With this configuration, the number of cases that exceed maximum recommended wait time is reduced from 30% to 2% and open block utilization is 75%. Faster accommodation of waitlist patients also frees up an average of seven inpatient beds per day, which would further alleviate crowding in the perioperative process.

3.2 Methods

3.2.1 Data analysis and simulation
The open block design process was performed in two phases: 1) data analysis and 2) simulation. In the data analysis phase, we identified the waitlist cases that could be accommodated with common sets of resources and selected those that, on average, have sufficient volume to justify staffing open blocks. Given that 20 different services use the operating rooms at MGH, it is not initially apparent that a common set of resources would accommodate a meaningful portion of waitlist case demand. Second, we characterized the variability in arrival times, quantities and procedure durations of those cases that were deemed appropriate for open blocks. The same dataset used in our characterization of the perioperative process was used in this analysis (see section 2.2.1).

The variability analysis was used in the simulation phase. A simulation model was developed to assess the performance of open blocks and arrive at a practicable design that would minimize wait time for emergent, urgent and non-urgent waitlist patients, while maximizing resource utilization.

3.2.1.1 Open block demand
The first question to answer in designing an open block policy is whether sufficient volume of waitlist cases will exist to utilize dedicated resources at acceptable levels. The resources required to perform a surgical procedure include 1) an operating room, 2) nursing staff, 3) anesthesia staff and 4) at least one surgeon. Each of these resources is specialized to some degree. A patient will require an operating room with the proper equipment, and a surgical staff with the expertise to address his or her specific surgical needs. When a waitlist patient arrives, a surgeon from an appropriate service will be assigned depending on the patient's condition. Figure 10 shows the average daily volume of waitlist cases by service. With the exception of Orthopedic and Emerg/Urg surgery, each service can expect fewer than two cases in a 24-hour period.

Dedicating an open block to each service would not be feasible. If each service had a dedicated open block, 10 to 12 additional operating rooms would be required and utilization of those rooms would be around 20%. That level of expenditure simply is not possible. This leads us to a deeper exploration of the level of specialization required by waitlist cases. The question is: are there groups of waitlist cases from across different services with resource requirements similar enough to allow sharing of open blocks?
Figure 10: Average daily waitlist case volume by service. The only services with volume sufficient to justify dedicated blocks are Orthopaedic and Emerg/Urg surgery.

3.2.1.1 Case pooling
A closer look at the operating staff and room resource requirements of all waitlist cases reveals five groups. These are shown in Table 2. The groups are defined such that an anesthesiologist and RN team capable of handling cases from one of the services in a group are capable of handling typical cases from any of the other services in the same group. Likewise, the layout and equipment required in an operating room for one case in the group would be suitable for all other cases in the group.

Table 2: Service groupings for pooled waitlist open blocks. Case from services in the same group can be accommodated in the same operating rooms and with the same anesthesia/RN staff.

<table>
<thead>
<tr>
<th>GROUP</th>
<th>SERVICE</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>PEDIATRIC SURGERY</td>
<td>EMERG/URG SURGERY</td>
<td>ORTHOPAEDIC SURGERY</td>
<td>THORACIC SURGERY</td>
<td>BURN</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>ANGELARY SERVICES</td>
<td>GENERAL SURGERY</td>
<td>NEUROSURGERY</td>
<td>CARDIAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>ANESTHESIOLOGY</td>
<td>UROLOGY</td>
<td>VASCULAR SURGERY</td>
<td>VASCULAR SURGERY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>TRANSPLANT</td>
<td>GYNECOLOGY</td>
<td>ORAL/MAXILLOFACIAL SURGERY</td>
<td>PAIN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>PLASTIC SURGERY</td>
<td>SURGICAL ONCOLOGY</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 11 shows the average daily wait list case volume for each group. Typically, 3 to 4 cases can be performed in an operating room on a given day. Groups 1 and 2 have the highest case volume, which is clearly sufficient to justify at least one dedicated open block per group.
As shown in Table 3, cases in groups 1 and 2 represent 80% of all wait list case booked. Addressing this population will have the largest impact at minimum cost and are therefore the focus of initial design and implementation of an open block policy at MGH. Group 1 and 2 cases represent 80% of those cases booked on weekdays as well. Cases booked on weekends and holidays are handled by call teams and would therefore not be affected by open block policy.

Table 3: Total number of waitlist cases booked between July 1st 2009 and June 30th 2010. Group 1 and 2 cases represent 80% of all waitlist cases.

<table>
<thead>
<tr>
<th>Waitlist cases Booked</th>
<th>Weekday</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>1707</td>
<td>2306</td>
</tr>
<tr>
<td>Group 2</td>
<td>1995</td>
<td>2780</td>
</tr>
<tr>
<td>Other</td>
<td>929</td>
<td>1243</td>
</tr>
<tr>
<td>Grand Total</td>
<td>4631</td>
<td>6329</td>
</tr>
<tr>
<td>Groups 1 and 2 as percentage of total</td>
<td>80%</td>
<td>80%</td>
</tr>
</tbody>
</table>

It is important to note that an equal proportion of waitlist cases exceed maximum recommended wait time on the weekends and on weekdays. As shown in Table 4, this is consistent for each case type: emergent, urgent and non-urgent. While the focus of this open block solution is on cases booked on weekdays, these numbers merit development of approaches to improve access for weekend waitlist patients as well. Group 1 and 2 waitlist cases booked on weekdays represent 58% of the entire waitlist case volume and 58% of the cases that exceeded maximum recommended wait time.
Table 4: Percentage of Group 1 and 2 waitlist cases exceeding maximum recommended wait for those booked on weekdays and on weekends or holidays.

<table>
<thead>
<tr>
<th>Case Type</th>
<th>Group 1</th>
<th>Group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Weekday</td>
<td>Weekend/Holiday</td>
</tr>
<tr>
<td>Emergent</td>
<td>25%</td>
<td>20%</td>
</tr>
<tr>
<td>Non-Urgent</td>
<td>32%</td>
<td>33%</td>
</tr>
<tr>
<td>Urgent</td>
<td>16%</td>
<td>17%</td>
</tr>
</tbody>
</table>

3.2.1.2 Variability in Case Arrivals and Durations

As shown in Figure 11 the average number of cases booked in each group is nearly constant across weekdays. However, to properly design and assess the performance of dedicated open blocks we need to characterize the variability of demand within each day. With the help of Stat::Fit, a statistical software package available with ProModel, the data for 1) quantity of bookings each day, 2) duration of surgical procedures and 3) room turnover time, were fitted to theoretical probability distribution functions. The distribution function for each parameter was selected based on comparison of goodness-of-fit test results between 30 continuous and discrete theoretical distributions. The goodness-of-fit tests included the Chi-squared test, the Kolmogorov Smirnov test and the Anderson Darling test. The distribution with the relative best fit was chosen for each parameter. All distributions selected met the minimum acceptance requirement: probability of Type 1 error of less than 5%. The timing of bookings over the course of each day is also highly variable. However, this parameter could not be fit to a theoretical distribution function. Empirical probability distributions were used to characterize waitlist case arrival timing. For each parameter, separate functions were defined for emergent, urgent and non-urgent waitlist cases in each service group. The distribution functions are shown in Table 6, Table 7, Table 8 and Table 9 in the appendix, section 4.1.

3.2.1.3 Open Block Simulation

Using the data characterization described above, a discrete event simulation model was developed to determine the minimum number of open blocks required to ensure group 1 and 2 wait list cases were accommodated within the maximum wait time. Figure 12 is a representation of the simulation model.

The model assumes all waitlist cases must be accommodated in open blocks and that elective cases may not be scheduled in the open blocks. If a waitlist patient arrives and capacity is not available, the patient waits until resources become available. Resources are allocated to each patient based on priority. Emergent cases have highest priority and non-urgent cases have lowest priority. If an emergent case arrives while an urgent or non-urgent patient is waiting for a resource to become available, the emergent case moves to the front of the line and is accommodated first. Patients in the same priority category are served on a first come, first served basis.
Each type of case can arrive at any time of day with probability as represented in the distribution functions shown in section 4.1.2. If an emergent or urgent case arrives outside of open block hours, a call team will take the case. Four call teams are available during non-scheduled hours. If a non-urgent case is scheduled outside of open block hours, the case will wait until the next morning before it is accommodated. Between surgical procedures, a turnover team is required to prepare the operating room for the next patient. The turnover time varies according to the distributions defined in section 4.1.4.

To determine the optimal open block design, two parameters are adjusted between each simulation run: 1) the number of open blocks available to each group and 2) the time of day each open block is available. The results of the simulation runs are used to predict patient wait time and open block utilization given the number and timing of open block availability. The impact of waitlist wait time on inpatient bed resources is also determined for each configuration.

### 3.3 Results

After several iterations it was determined that five dedicated open blocks operating from 7:00am to 5:00pm is the best configuration. Two of the five are owned by Group 1 services and three are owned by Group 2 services. This solution was selected for relative simplicity of implementation and expected performance improvement.

With five open blocks operated according to the assumptions described in section 3.2.1.3, the number of patients exceeding max wait time would be reduced from 30% to 2%.

Figure 13 shows a comparison between the number of group 1 and 2 patients exceeding max wait time with and without open blocks.
Figure 13: Number and percentage of waitlist cases exceeding maximum recommended wait time prior to surgery; actual data is compared with simulated results.

In addition, 2514 inpatient bed-days per year, an average of seven additional beds per day, would be made available with the implementation of open blocks. Table 5 shows a comparison of inpatient bed-days consumed by urgent and non-urgent waitlist patients with and without open blocks. For non-urgent patients, the reduction in bed-days under the open block scenario is a result of two factors: 1) fewer patients occupying inpatient beds while on the wait list waiting for surgery and 2) fewer patients exceeding recommended maximum wait time, thereby reducing the average post surgical length of stay. For urgent patients, the reduction in bed-days consumed is solely a result of shorter average length of stay due to fewer patients exceeding maximum wait time. Refer to Table 5 in section 2.3.2.2.2 for details on the impact of patient wait time on subsequent hospital length of stay.

Table 5: Comparison of inpatient bed-days consumed by urgent and non-urgent waitlist patients with and without open blocks

<table>
<thead>
<tr>
<th></th>
<th>Non-Urgent</th>
<th>Urgent</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-Surgery</td>
<td>Post-Surgery</td>
<td>Post-Surgery</td>
</tr>
<tr>
<td>Inpatient Bed-Days Consumed Per Year</td>
<td>Actual</td>
<td>2462</td>
<td>18979</td>
</tr>
<tr>
<td>Open Blocks</td>
<td>1083</td>
<td>18234</td>
<td>6046</td>
</tr>
<tr>
<td>Reduction in Annual Bed-Days Consumed w/ Open Blocks</td>
<td>1379</td>
<td>745</td>
<td>390</td>
</tr>
<tr>
<td>Additional Beds Available Per Day w/ Open Blocks (Avg)</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 14 shows the impact of the open block policy on operating room utilization. The analysis indicates that utilization of the five open blocks would be 75% and that removing these operating rooms from scheduled blocks would require a very little
increase in OR utilization. Current average utilization of 52 scheduled operating rooms is 85.8%. If the five operating rooms were set aside, the remaining 47 operating rooms could accommodate all scheduled cases with a utilization of 86.4%. This increase is smaller than expected because turnaround time for waitlist cases is higher than for scheduled elective cases. When waitlist cases are removed from the scheduled blocks, average turnaround time in those blocks is reduced.

Note that 86.4% represents required actual utilization of the 47 scheduled blocks. But without waitlist cases arriving on the day of surgery, scheduled utilization prior to the day of surgery must be 86% as well. This means that the current average scheduled utilization must improve from the current 74%. Note also however, that without waitlist cases relying on holes in the scheduled blocks, target scheduled utilization can now be raised to 100%. The predictability gained in the scheduled blocks by separating out the random arrivals of waitlist cases would help make this an achievable goal. Hence, the foundation would be laid to improve effective OR capacity. Improving from 86% to 100% scheduled utilization would be equivalent to gaining five additional operating rooms per day.

3.4 Discussion
Implementing an open block policy at Massachusetts General Hospital would have numerous benefits to patients and hospital operations. First, it would allow perioperative
services to reduce the number of wait list cases that exceed maximum recommended wait time from the current 30% to 2%. This would have a profound impact on the patient’s experience. It would also reduce pressure on inpatient bed resources, the primary bottleneck in perioperative patient flow. Freeing up an average of seven inpatient beds per day could dramatically reduce crowding and delays, and open capacity for growth. Finally, separating the stochastic demand of non-elective cases from elective blocks would dramatically improve schedule predictability, reduce the number of changes on the day of surgery (see section 2.3.3) and lay the foundation for improved OR utilization in the future.

Implementation will require several major changes in the form of revised scheduling policies and increased collaboration across services. The following are suggested guidelines for open block scheduling policy: 1) All patients booked on the waitlist must be in-house patients that medically require surgery within 24 hours, 2) all other cases must be booked into scheduled blocks. A non-urgent case may be booked into the open blocks if he or she does not medically require surgery within 24 hours only if 1) the scheduling surgeon does not own a scheduled block and the case cannot be accommodated in an unused scheduled block owned by another surgeon or 2) the case cannot be scheduled within the week and the scheduling surgeon has at least 86% utilization of his or her assigned blocks. These guidelines are meant to ensure that the open blocks are used for their intended purpose and to limit their use as a scheduling tool to supplement surgeon’s allocated blocks.

Increased collaboration across functions will also be critical. This has implications for compensation models and the culture. The compensation model must be adapted to provide incentive for surgeons to staff open blocks and perform surgeries for patients associated with other services. The culture must shift to embrace a team approach to managing variability. This will be a challenge in an environment accustomed to and reliant on skilled individuals reacting to emergencies and other forms of variability. However, managing variability will be critical to reducing cost and improving quality and access to healthcare. Separating the random demand of surgical emergencies from the predictable demand of elective cases with open blocks will be a profound first step.
This page has been intentionally left blank
4 Appendix

4.1 Probability Distribution Functions: Service Groups 1 & 2

4.1.1 Quantity of Weekday Waitlist Bookings
Table 6: Probability distribution functions representing quantity of bookings on weekdays for Emergent, Urgent and Non-Urgent waitlist cases in service groups 1 and 2

<table>
<thead>
<tr>
<th>Case Type</th>
<th>Group 1 Distribution Function</th>
<th>Expression</th>
<th>Group 2 Distribution Function</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emergent</td>
<td>Poisson</td>
<td>P(0.5)</td>
<td>Binomial</td>
<td>B(9., 7.054e-002)</td>
</tr>
<tr>
<td>Urgent</td>
<td>Poisson</td>
<td>P(2.1111)</td>
<td>Poisson</td>
<td>P(1.4285)</td>
</tr>
<tr>
<td>Non-Urgent</td>
<td>Binomial</td>
<td>B(31, 0.1333)</td>
<td>Poisson</td>
<td>P(5.8333)</td>
</tr>
</tbody>
</table>

4.1.2 Timing of Weekday Waitlist Arrivals
Table 7: Distributions of booking times on weekdays for Emergent, Urgent and Non-Urgent waitlist cases in service groups 1 and 2

![Histograms of booking times for Emergent, Urgent, and Non-Urgent cases in service groups 1 and 2.](image-url)
### 4.1.3 Duration of Surgical Procedures

Table 8: Probability distribution functions representing procedure durations for emergent, urgent and non-urgent cases in groups 1 and 2

<table>
<thead>
<tr>
<th>Case Timing</th>
<th>Case Type</th>
<th>Distribution Function</th>
<th>Expression</th>
<th>Group 1</th>
<th>Distribution Function</th>
<th>Expression</th>
<th>Group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>7:00am - 5:00pm</td>
<td>Emergent</td>
<td>Pearson6</td>
<td>57. + P6(2.1390, 6.3406, 326.15)</td>
<td>Pearson6</td>
<td>LogLogistic</td>
<td>12. + P6(4.2963, 13.623, 722.77)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Urgent</td>
<td>LogLogistic</td>
<td>25. + 94.19<em>1</em>1/(1/(1-U(0.5,0.5))-1))**1/2.63</td>
<td></td>
<td>Lognormal</td>
<td>25. + L(167.3, 147.85)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-Urgent</td>
<td>Lognormal</td>
<td>17. + L(16.04, 108.06)</td>
<td></td>
<td>LogLogistic</td>
<td>17. + 124.7*(1/(1/(1-U(0.5,0.5))-1))**1/2.353</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Emergent</td>
<td>Weibull</td>
<td>65. + W(1.2391, 115.60)</td>
<td></td>
<td>LogLogistic</td>
<td>40. + 168.7*(1/(1/(1-U(0.5,0.5))-1))**1/2.187</td>
<td></td>
</tr>
<tr>
<td>5:00pm - 7:00am</td>
<td>Urgent</td>
<td>LogLogistic</td>
<td>26. + 88.60*1/(1/(1/(1-U(0.5,0.5))-1))**1/2.850</td>
<td></td>
<td>LogLogistic</td>
<td>44. + 101.0*(1/(1/(1-U(0.5,0.5))-1))**1/2.016</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-Urgent</td>
<td>LogNormal</td>
<td>28. + L(105.16, 117.26)</td>
<td></td>
<td>LogLogistic</td>
<td>30. + 123.4*(1/(1/(1-U(0.5,0.5))-1))**1/2.571</td>
<td></td>
</tr>
</tbody>
</table>

### 4.1.4 Operating Room Turnover Times

Table 9: Probability distribution functions representing length of operating room turnover time between surgeries for group 1 and 2 waitlist cases

<table>
<thead>
<tr>
<th></th>
<th>Group 1</th>
<th>Group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution Function</td>
<td>Expression</td>
<td>Distribution Function</td>
</tr>
<tr>
<td>Beta</td>
<td>B(2.85, 1.04, 27., 60.)</td>
<td>Beta</td>
</tr>
</tbody>
</table>
Works Cited


