Analytical Approaches to Surgical Unit Management

Mohammad Mehdi Ghotboddini

Doctor of Philosophy

Desautels Faculty of Management

McGill University,
Montreal, Quebec
December 2016

A thesis submitted to McGill University
in partial fulfillment of the requirements of the degree of
Doctor of Philosophy

DEDICATION

To my parents …

ACKNOWLEDGMENTS

First and foremost, I would like to thank Allah Almighty for his infinite blessings.

I would like to express my sincere gratitude and special appreciation to my advisor Prof. Vedat Verter for the continuous support of my Ph.D. study, for providing me with all the necessary facilities for the research, for his motivation and patience, for his positive outlook and confidence in my ability, for his wonderful personality, and for all he taught me from the heart and mind. You have been a tremendous mentor for me. Thanks a lot for everything.

I would like to especially thank other members of my Ph.D. committee, Dr. Lawrence Rosenberg and Prof. Raf Jans for their encouragement and insightful comments and suggestions. I would like to sincerely thank administrators, surgeons, and nurses in the surgical unit at the Montreal Jewish General Hospital for supporting me to address the operational challenges in the surgical unit and to collect real data for my Ph.D. thesis.

Last but not least, I would like to warmly thank my family. Words cannot express how grateful I am to my precious parents, sister, brother, and brother in law for supporting me spiritually throughout all years of my study and my life in general. I would like to express appreciation to my beloved wife, who accompanied me to finish my Ph.D. I would also like to thank my friends and officemates in Montreal for the support they've lent me over all these years.

CONTRIBUTIONS OF AUTHORS

Ghotboddini, Mohammad Mehdi, Vedat Verter, Lawrence Rosenberg, 2016. "Integrated Surgical Case Mix Management" Working Paper.

The student (first co-author) reviewed the literature, designed the modeling framework, formulated the problem, applied the solution methodology, developed the computer programs and simulation models, extended the model, and interpreted and visualized the final results. The second co-author actively contributed in designing the study framework and defining the problem, analyzing the results as well as editing the manuscript. The third co-author helped to better understand the surgical unit environment, collect the data, and verify the end results. During the whole process the first author was the main person responsible for conducting the research using guidance, comments and feedback from co-authors.

ABSTRACT

Surgical unit costs a high percentage of the hospital budget and represents one of the largest segments of the healthcare budget. Surgical unit provides direct care to diverse surgical patients through pre-surgical, operative, post-operative, and recovery steps. Through delivering the surgical services, surgical units generally face series of operational challenges and lag behind in efficiency. These challenges are amplified by conflict of interests among different stakeholders. To address these issues and deliver the best service to the surgical patients, governments or the health insurers encourage surgical units to be more cost effective. In this respect, Operations Research techniques can be applied to help hospital managers to better utilize their resources. The aim of this thesis is to provide an integrated framework for making effective decisions on the hospital surgical case-mix problem (CMP). This research is inspired by the managerial challenges at the surgical unit of the Montreal Jewish General Hospital (JGH).

The significance of this research is to mathematically model and simulate the CMP with the Master Surgical Scheduling and the Advanced Scheduling problems in an Integrated Surgical Case Mix (ISCM) model. This empowers hospital managers to enhance surgical unit efficiency by integrating surgical case mix plan, allocating Operating Room (OR) blocks to surgical divisions and surgeons, and assigning elective surgical cases to the operating rooms, at the strategic, tactical, and operational levels, respectively. The ISCM model is developed under various reimbursement mechanisms (e.g., activity based funding, and global budget) and bed configuration policies (e.g., Semi-pooled, and pooled). The usefulness of ISCM model is boosted by incorporating emergency and off-service patients into the model. A Surgical Ward Design

(SWD) simulation model is developed to i. validate the ISCM results under various scenarios, ii. explore the impact of different OR schedules on the surgical unit patient flow, and iii. simulate different surgical unit bed configurations.

From technical perspective, a stochastic integer model is developed which limits the probability of the downstream bed shortage through a Chance-Constrained programming approach. Moreover, it controls the risk of high bed shortage cost in a Conditional Value at Risk framework. The linear form of the stochastic model is approximated and calibrated with the full-scale data on 72 surgical procedures, 40 surgeons, and 7 specialties in JGH. Then, the sample average approximation method is presented to solve the ISCM model.

The results demonstrate that the stochastic ISCM model outperforms deterministic ISCM model in terms of bed shortage level and OR utilization. The activity based funding policy and the global budget with incentive policy result in the similar surgical case mix. Semi-pooled bed configuration increases the daily bed occupancy variance and the number of required beds versus the pooled bed configuration. Also, off-service patient admission is recommended mostly on Fridays, Saturdays, and Sundays for at most two patients per day. It is observed that the stochastic ISCM optimal results are quite robust to a range of bed shortage cost and OR idle/over-time cost.

RÉSUMÉ

L'unité chirurgicale coûte un pourcentage élevé du budget de l'hôpital et représente l'un des plus gros segments du budget des soins de santé. L'unité chirurgicale fournit des soins directs à divers patients chirurgicaux à travers des étapes pré-chirurgicales, opératoires, post-opératoires et de récupération. Grâce à la prestation des services chirurgicaux, les unités chirurgicales font généralement face à des séries de défis opérationnels et sont en retard sur l'efficacité. Ces défis sont amplifiés par les conflits d'intérêts entre les différentes parties prenantes. Pour répondre à ces problèmes et offrir le meilleur service aux patients chirurgicaux, les gouvernements ou les assureurs de santé encourager les unités chirurgicales pour être plus rentable. À cet égard, les techniques de recherche opérationnelle peuvent être appliquées pour aider les gestionnaires de l'hôpital à mieux utiliser leurs ressources. Le but de cette thèse est de fournir un cadre intégré pour prendre des décisions efficaces sur le cas hôpital malade chirurgical problème (CMP). Cette recherche s'inspire des défis managériaux de l'unité chirurgicale de l'Hôpital général juif de Montréal (JGH).

L'importance de cette recherche est de modéliser et de simuler mathématiquement le problème de CMP avec les programmes Master Surgical Scheduling et Advanced Scheduling dans un modèle intégré de cas de chirurgie (ISCM). Cela permet aux gestionnaires hospitaliers d'améliorer l'efficacité de l'unité chirurgicale en intégrant le plan de répartition des cas chirurgicaux, en attribuant le bloc opératoire aux divisions chirurgicales et chirurgienset en attribuant des cas opératoires aux salles d'opération (OR) aux niveaux stratégique, tactique et opérationnel respectivement. Le modèle ISCM est développé sous divers mécanismes de remboursement (p. Ex., Financement basé sur l'activité et budget global) et les politiques de

configuration des lits (p. Ex., Semi-regroupées et regroupées). L'utilité du modèle ISCM est renforcée par l'intégration des patients d'urgence et hors service dans le modèle. Un modèle de simulation de conception de salle chirurgicale (SWD) est développé pour i. Valider les résultats du ISCM sur divers scénarios, ii. Explorer l'impact de différents calendriers OR sur le débit du patient de l'unité chirurgicale, et iii. Simuler différentes configurations de lits d'unités chirurgicales.

Du point de vue technique, un modèle stochastique d'entiers est développé qui limite la probabilité de la pénurie de lit en aval par une approche de programmation Contrainte de Chance. De plus, il contrôle le risque d'un coût élevé de la pénurie de lits dans un cadre de valeur conditionnelle à risque. La forme linéaire du modèle stochastique est approximée et calibrée avec les données à grande échelle sur 72 interventions chirurgicales, 20 chirurgiens et 7 spécialités en JGH. Ensuite, une méthode d'approximation moyenne d'échantillon modifiée est présentée pour résoudre le modèle ISCM.

Les résultats démontrent que le modèle ISCM stochastique surperforme le modèle ISCM déterministe en termes de niveau de pénurie de lit et d'utilisation de OR. La politique de financement axée sur les activités et le budget global, avec une politique d'incitation, donnent lieu à un mélange de cas chirurgical similaire. La configuration de lits semi-groupés augmente la variance d'occupation journalière du lit et le nombre de lits requis par rapport à la configuration de lits groupés. En outre, l'admission hors service des patients est recommandée surtout le vendredi, le samedi et le dimanche pour au plus deux patients par jour. Il est observé que les résultats optimaux ISCM stochastiques sont assez robustes à une gamme de coût de pénurie de lit et de coût d'OR / inoccupation.

Contents

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Chapter 1

Introduction, Motivation, and Outline

1.1. Introduction

The annual budget of the healthcare system has been the center of much debate in most industrialized countries. Healthcare accounts for about 9% of the GDP[1] among OECD[2] countries (Barua & Esmail, 2013). Achieving more value out of that is the main concern for policy makers, who usually question the efficiency of the healthcare system, every time there is a demand to increase the healthcare budget. It is important for policy makers to see if they can increase the efficiency of the current system before investing more on establishing new resources. This tension presents a clear incentive for researchers to study how to improve the efficiency of the current healthcare system.

Hospitals play an important role in the healthcare system. Hospitals are publicly funded in Canada and are largely staffed by surgeons, nurses, and so on. From an economic perspective, a noticeable proportion of the annual healthcare system budget goes to hospitals. This rate was 30% in Canada in 2012, which means that 3.5% of Canada's GDP is dedicated to hospitals annually (Information, 2012). Within hospitals, surgical units are the center of attention. It is observed that 30% of the total hospital costs and 60% of its revenue derived from surgical units (Jackson, 2002). This level of expenditure demonstrates that surgical unit performance significantly affects the healthcare system costs as well as efficiency.

[1] Gross Domestic Product
[2] Organisation for Economic Cooperation and Development

Surgical unit efficiency will be even more evident because of the imbalance between surgical service demands and surgical unit resources (e.g. surgical beds, operating rooms, surgeons and nurses). On one hand, Canada has been ranked 32nd among the OECD countries with respect to the average medical bed ratio (2.80 beds per 1000 population) while the OECD average is 4.96 beds. On the other hand, the average length of stay (LOS) in Canadian hospitals is 7.7 days, which is among the top 10 OECD's longest LOS (the OECD average LOS is 7.2 days) (OECD, 2011, 2013). Therefore, from a medical perspective, patients face huge waiting times for surgical services (e.g. in Canada waiting time is on average 18.2 weeks, from GP to specialist to treatment), which might deteriorate their health as well as causing dissatisfaction (Barua & Esmail, 2013).

This Ph.D. dissertation systematically approaches surgical unit bed management. It focuses on the current managerial issues to reduce cost of surgical case planning and scheduling by decreasing the surgical bed shortage rate. This research is done in collaboration with the Montreal Jewish General Hospital (JGH). This cooperation enabled me to meet and interview several surgeons, nurses, and patients as well as hospital managers to gather a solid understanding of the pressing problems in this area. Besides, the JGH provided me with the surgical unit data to validate our analysis. This study could potentially align the best interests of hospital administrators, surgeons, and patients, and hence provides an opportunity for stakeholders to buy-in to the developed solutions.

1.2. Outline

The rest of this study is organized as follows: in the remainder of this chapter, the structure of the surgical unit patient flow will be described; and then a summary of the current operational challenges in the JGH surgical unit will be provided. Chapter 2 will provide a comprehensive literature review, which gives an overview on the surgical unit key operational concerns and solutions. The articles are classified on the basis of the key problems in the surgical unit context. Chapter 2 mainly focuses on the managerial insights of these studies; however, it also points to the key mathematical models and solution techniques, which show the application of Operations Research techniques in the surgical unit management literature. The chapter concludes by highlighting emerging new avenues for future research on the surgical unit operational issues.

Chapter 3 will study a proposed re-configuration of surgical ward beds under various bed management policies, classified as dedicated policy, pooled policy, and semi-pooled policy. A simulation model is developed to design the JGH's surgical ward. The simulation results show the sensitivity of surgical ward configuration to a range of service levels, which is defined as an index for surgical bed shortage rate. The chapter is concluded by providing the managerial insights of the study.

Chapter 4 will develop an optimization model to improve the hospital's surgical case-mix. Building a strategic decision-making model without considering the operational feasibility to implement the strategic decisions makes the results unrealistic and impractical. So, a stochastic model is developed to integrate the operational, tactical, and strategic decisions in the hospital.

The model is calibrated with the full-scale data on 72 surgical procedures, 40 surgeons, and 7 specialties in JGH. The chance-constraint approach is applied to model bed shortage in the surgical unit and apply the sample average approximation method to solve the model.

Chapter 5 will extend the proposed Integrated Surgical Case-Mix model. The effect of various governmental reimbursement policies on the hospital's surgical case-mix is studied. We analyze the impact of surgical bed configuration on the surgical unit's bed occupancy and patient flow. This chapter studies the impact of emergency patients as well as off-service patients on the optimal case-mix in JGH. These extensions provide more realistic results and meaningful insights for various stakeholders such as policy makers, hospital managers, and surgeons.

Chapter 6 will present the conclusion and possible future study. New avenues are developed to extend our research on surgical unit management.

1.3. Contributions of This Thesis

The significance of this thesis is to address the surgical case mix problem integrated with the Master Surgical Scheduling and the Advanced Scheduling problem. These problems are respectively classified into strategic, tactical, and operational decisions and their interaction is explored in this thesis. For the first time in the literature, the impact of different funding policies on the Surgical Case-Mix problem is explored. The probability of the downstream bed shortage is modeled through a Chance-Constrained programming approach and the risk of high bed shortage cost is controlled within a Conditional Value at Risk framework. Moreover, the sample average approximation method is presented to solve the ISCM model. The usefulness of the ISCM model is enhanced by incorporating emergency, off-service patients, as well as elective patients into the model. Furthermore, a surgical ward simulation model is developed to explore the impact of JGH OR schedules on the surgical unit patient flow under different surgical unit bed configurations. Findings of this research were used as valuable input for JGH to develop and implement solid plans to improve the efficiency of its surgical unit.

1.4. Surgical Unit Process

This section will describe the surgical service process to better understand the relevant problems in the literature. However, there is not a unique surgical service process for all hospitals even in the same health authority. In general, there are more than 1000 surgical procedures grouped under more or less 20 medical divisions. Each division is a formally organized unit providing similar practices of the hospital's medical staff such as orthopedic, cardiac, gynecology, urology, colorectal, general surgery, dental, and so on. While general hospitals usually provide services for many different types of ailments, specialized hospitals focus on a limited number of the aforementioned treatments.

There are two main surgical patient types: elective and non-elective patients. Elective or preplanned patients are categorized as inpatients who have to stay overnight at the hospital after the surgery, and outpatients who are discharged home on the day of surgery. Non-elective patients refer to emergency patients and urgent cases. Urgent patients are distinct from emergency cases since they can wait for the service for a short period of time, while emergency patients have to be served as soon as possible (Cardoen, Demeulemeester, and Beliën, 2010)

Elective Patients are admitted at the hospital usually on the day of surgery or one day before. Each surgeon should sequence her patients on the basis of the time allocated by the surgical specialty chief based on the operating room (OR) block schedule. Each OR block refers to 4 to 5 hours of an operating room working time in a day. The hospital might consider prescheduled OR blocks for Non-elective patients to stop cancelling elective cases to serve emergency patients.

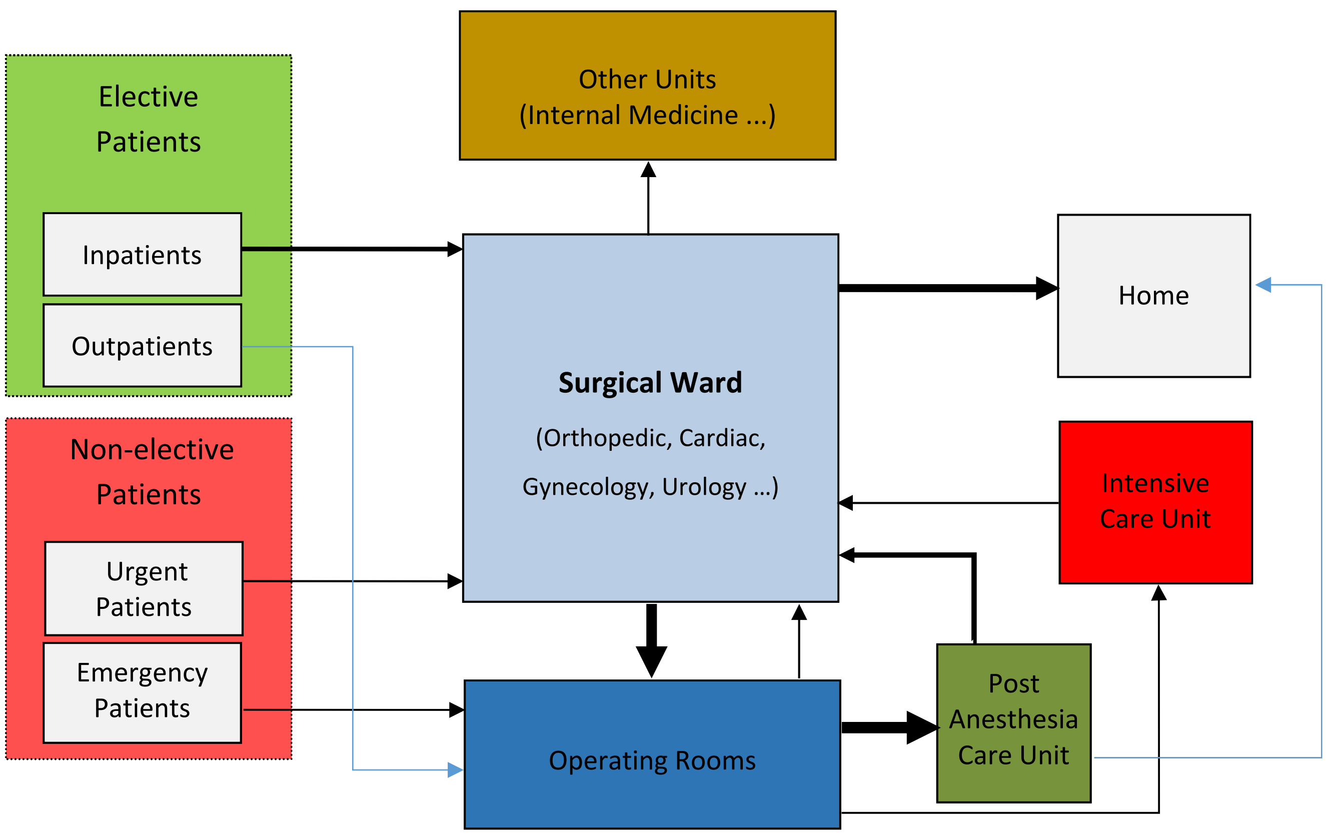


Figure 1-1: Surgical Unit Patient Flow

After the surgery, patients are usually transferred to the Post Anesthesia Care Unit (PACU) for a couple of hours. At this level, outpatients are discharged home if the patient's stable health condition is confirmed by the surgeon. However, a portion of inpatients, urgent, and emergency patients needs to stay a couple of days in the Intensive Care Unit (ICU) after the surgery. The rest are directly sent to the surgical ward for recovery. Patients will be sent to the other units if they need further medical care, but usually they will be discharged home if their health condition is stable enough. The patient's length of stay (LOS) varies either in the ICU or in the surgical ward depending her health status, yet for similar procedures it is usually within a typical range. Figure 1-1 illustrates the schematic patient flow diagram in the surgical unit.

1.5. Case Study: Montreal Jewish General Hospital

Montreal Jewish General Hospital (JGH) is an acute care teaching hospital with 637 beds. The hospital's surgical unit contains about 15 divisions, which annually serve more than 6000 patients. Surgical case cancellation is one of the main managerial concerns in the hospital. JGH cancelled more than 20 elective surgeries in June 2011, mainly due to bed unavailability in the surgical ward.

When the OR runs late, the last patients are most likely to be cancelled since the hospital does not have budget to pay anesthetists and nurses working overtime. This is not a big concern at the JGH since the surgeons can accurately predict the length of their operation using an Operating Room Information System Software (OPERA[3]). OPERA uses the last 10 similar procedures to predict the OR required time for the next operation. Moreover, a case cancellation may happen because of the unavailability of the lab results from other consultants. Also, when a higher priority case comes up, an elective patient's surgery can be cancelled. Despite all these possibilities, the surgical unit chief and the specialty chiefs at the JGH believe that the surgical bed shortage is the key bottleneck in the surgical unit patient flow.

[3] JGH implemented a state-of-the-art information system tool in 2002. OPERA supports decision-makers by OR planning and scheduling guidelines.

To support this statement, it is worth noting that two out of 13 operating rooms at the JGH are closed because of over utilized surgical ward beds. A highly congested surgical unit not only results in lots of cancellation and rescheduling problems (bringing huge financial costs to the hospital), but it also causes severe deterioration of the patient health. Hence, the JGH surgical unit seeks to gain more control through the use of its resources.

The hospital’s annual budget is the other important concern for the JGH’s managers. Governmental reimbursement policies as well as changing technology and population demographics entice the JGH managers to reassess the current surgical case-mix. In other words, the JGH surgical unit wants to gain more control on the costs and revenues of providing services by complying with the recent government reimbursement policy. As a publicly funded hospital the JGH has to serve a certain number of patients on an annual basis, known as RAMQ[4] base volume, and get reimbursed for the provided services by the government. RAMQ base volume is predefined for each surgical procedure and the hospital will be penalized if it does not meet this threshold at the end of each fiscal year. The managers try to achieve cost containment and profitability by practicing some more niche surgical services instead of the current services. This goal cannot be attained without considering the availability and integration of vital surgical unit’s resources such as the operating rooms and the surgical ward beds as well as the potential number of patients for each procedure.

[4] The Régie de l'assurance maladie du Québec (RAMQ) is the government health insurance board in the province of Quebec, Canada.

Chapter 2

Structured Review of Surgical Unit Management Problems

2.1. Introduction

This chapter provides a comprehensive review of the essential operational issues in a surgical unit. Many articles in this area focus on the surgical unit planning to efficiently utilize the unit's main resources such as operating rooms, intensive care units and main ward beds. May et al. presented a comprehensive literature review on Surgical Scheduling problems (May, Spangler, Strum, & Vargas, 2011). More than 110 manuscripts were classified under the following six labels: capacity planning, process reengineering, surgical service portfolio, procedure duration, schedule construction, and schedule execution. They also categorized the literature based on the time sensitivity of the problems. A time depended framework was developed to review all types of decisions. For example, capacity planning decisions for building a hospital should be made between 12 to 60 months before a surgery while emergency case admission decision must be taken right before the surgery. However, a typical approach is to classify the articles into the strategic, tactical, and operational problem categories, which refers to long, medium, and short term decisions. Yet May et al.'s approach is more precise. The authors noted that the surgical service portfolio area contains unexplored problems for future studies. Guerriero & Guido presented a detailed review of the scientific articles on surgical unit management (Guerriero & Guido, 2011). They classified the literature according to the hierarchical decision levels in this field. Samudra et al. presented the most recent literature review in the operating room planning and scheduling area (Samudra et al., 2016). They classified the manuscripts regarding decision type, research methodology, patient type, performance measure, and so on. These frameworks are useful but a broad review of the literature is not the scope of this chapter. This chapter develops a detailed, extensive literature review on the key challenges faced by hospital managers to support, plan, and improve the surgical unit efficiency and effectiveness. The proposed framework strives to classify the literature into 5 main domains: The first domain tackles the very initial stage of the surgical unit planning, known as Case-mix planning, although the literature on this area is relatively scarce. The second domain refers to those studies, which essentially focus on the operating rooms (ORs) at the tactical and operational level of decision-making hierarchy. How to schedule the OR blocks, sequence the patients, manage the OR idle/overtime, and many other problems are reviewed in this subsection. The third domain highlights the manuscripts on the ICU/ PACU problems. How to manage a congested ICU, sequence patients in the PACU, discharge/admit the patient in the ICU, and several other issues are reviewed under this domain. The fourth domain reviews how the care-providers (i.e. surgeons, nurses, and anaesthetists) could improve the surgical system's efficiency. The fifth domain reviews the manuscripts which study the hospital bed management issue. However, the main focus of this dissertation is on the studies that integrate the surgical bed availability into the OR scheduling models.

2.2. Case-Mix Oriented Approach

Surgical Case-Mix problem determines the types and quantities of surgical procedures to be performed in the hospital. From the early 70s researchers started to study the Surgical Case-Mix problem by developing simple mathematical models which maximizes hospital's profits or performed cases (Guerriero & Guido, 2011). Parsons et al. Parsons, Howard, Barker, and Peterson (1992) introduced the CMP as a systematic approach to improve quality and control costs. Robbins & Tuntiwongpiboon (Robbins W, 1989) developed one of the very first models on the CMP for diagnosis related groups (DRGs), a standard practice to reimburse hospitals by classifying the procedures with the same level of the expected usage of resources. Their linear model considers three capacity constraints: total hours of diagnostic services, total hours of nursing care, and dollar amount of pharmaceuticals. Moreover, the model's fourth constraint on the expected demand for each DRG guarantees that the governments' minimum expected healthcare need is satisfied. The model tries to optimally allocate the resources to the DRGs to maximize the hospital benefit. The authors also developed a sensitivity analysis to evaluate changes in the relevant costs and marginal benefits for each DRG.

However, a novel approach to the CMP was presented by Blake & Carter (Blake & Carter, 2002). Their model tackled the CMP from the hospital managers and surgeons perspective. They developed a goal programming approach to reset the type and volume of the surgeon's performed procedures, while the hospital faces an 18% budget cut. The problem was defined in the Canadian healthcare environment in which hospitals are under the global budget mechanism and the surgeons are paid on a fee-for-service basis. It was assumed that the surgeons are profit satisfier rather than profit maximizer, so they strive to maintain a minimum level of income for each year. Also it was assumed that the hospital managers want to guarantee that a minimum amount of the hospital resources (i.e. OR time) would be allocated to each surgeon through the *Block Mix* alterations. The objective of the model is to minimize the total weighted penalties when the surgeons desired revenues and the hospitals expected level of costs are not satisfied. Authors extended this model in (Blake & Carter, 2003) by comparing a set of funding policies to the hospitals and surgeons. Global budget and rate-based funding were assumed as the hospital's reimbursement method, while the surgeons are funded under either fee-for-services, or fixed salary structure. They concluded that the combination of the global budget policy and salaried surgeons increases the risk of under-servicing comparing to the fee-for-services policy. Also it was noted that under the fee-for-services policy, the hospital's reimbursement effect on the case-mix is negligible.

2.3. Operating Room-Oriented Approach

Operating Rooms (OR) are known as the most significant resource in the surgical unit, due to the huge investments and the key equipment. Hence, the OR has been a long-studied component of a surgical unit, and there is a growing amount of research addressing it in the literature. Kim et al. explored the sources of inefficiencies in the OR, measured by wasted OR time. They showed that more than 30% of the time, the preoperative and postoperative units bring inefficiency into the ORs. However, 65% of the OR-wasted time is due to surgeon unavailability, nurse shortage, prolonged turnover time, and anesthetist shortage (S.-C. Kim, Horowitz, Young, & Buckley, 1999). Table 2-1 shows the detailed results.

To optimize OR utilization, some studies applied scheduling models with different objective functions: minimizing OR idle/overtime, the risk of surgical case cancelation, patients waiting time, overcapacity cost, staffing cost, and the expected bed shortage (Beliën & Demeulemeester, 2007; Creemers, Beliën, & Lambrecht, 2012; Doulabi, Rousseau, & Pesant, 2016; Fugener, Hans, Kolisch, Kortbeek, & Vanberkel, 2014; Hans, Wullink, Van Houdenhoven, & Kazemier, 2008; Marques, Captivo, & Pato, 2015; Testi, Tanfani, & Torre, 2007).

Table 2-1: Causes of inefficiency at operating room 1

Units	Shared of OR wasted time		Causes
Preoperative	17%		Unprepared patients
Operating Room	65%	including	
		10%	Surgeon unavailability
		30%	Nurse shortage
		10%	Anesthetist shortage
		15%	Prolonged turnover time
Postoperative (such as ICU, PACU)	15%		Congested PACU
Transport	3%		Peak number of patient

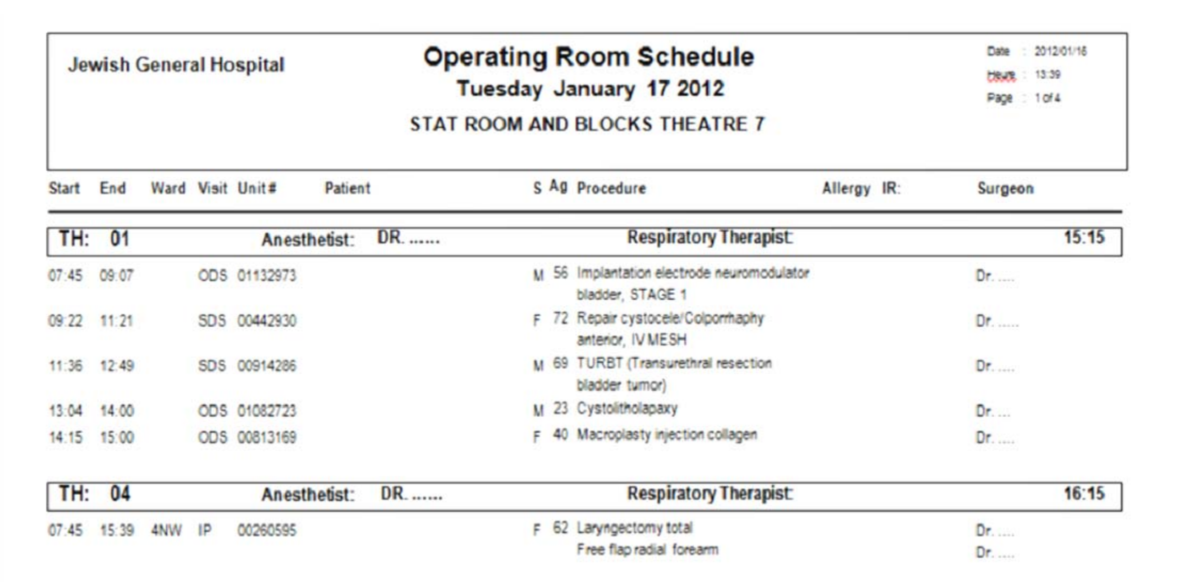
The OR scheduling problems are often split into the Master Surgical Scheduling Problem and the Patient Sequencing problem. Master Surgical Scheduling Problem refers to distributing operating room blocks among various surgeons. Patient Sequencing problem addresses patients scheduling with respect to the Master Surgical Schedule (Lee & Yih, 2014). Table 2-2 and figure 2-1 illustrate schematic OR-block schedule and patient schedule at the JGH surgical unit.

A comprehensive literature review on the OR scheduling type of problems is presented by Cardoen et al. (Cardoen et al., 2010). More than 120 manuscripts are classified on the basis of various clustering criteria. First, the reviewed articles were clustered based on their problem setting and technical structure (i.e. elective patients, emergency patients, performance measures, decision definition, research methodology, uncertainty and application of research).

Each of these articles targets one or couple of performance measures to evaluate the contribution of the developed models.

Table 2-2: Schematic OR Block Schedule

OR #	Shift	Monday	Tuesday	Wednesday	Thursday	Friday
OR 1	AM	Cardiac	Orthopedics	Cardiac	Orthopedics	Gynecology
	PM	E.N.T.	Orthopedics	Cardiac	Orthopedics	Emergency
OR 2	AM	Colo-rectal	General	Emergency	Emergency	Colo-rectal
	PM	Colo-rectal	General	Urology	General	Colo-rectal
OR 3	AM	E.N.T.	Orthopedics	Cardiac	Cardiac	Gynecology
	PM	E.N.T.	Emergency	Cardiac	E.N.T.	Emergency
OR 4	AM	Emergency	Gynecology	Emergency	Urology	E.N.T.
	PM	Emergency	Gynecology	Emergency	Urology	Orthopedics
OR 5	AM	General	General	Urology	Colo-rectal	General
	PM	Urology	Emergency	Urology	Colo-rectal	Colo-rectal



Jewish General Hospital

Operating Room Schedule
Tuesday January 17 2012
STAT ROOM AND BLOCKS THEATRE 7

Date : 2012/01/16
Heure : 13:39
Page : 1 of 4

Start	End	Ward	Visit	Unit#	Patient	S	Ag	Procedure	Allergy	IR:	Surgeon
TH: 01					**Anesthetist: DR. ......**			**Respiratory Therapist:**			**15:15**
07:45	09:07		ODS	01132973		M	56	Implantation electrode neuromodulator bladder, STAGE 1			Dr. ....
09:22	11:21		SDS	00442930		F	72	Repair cystocele/Colporrhaphy anterior, IV MESH			Dr. .....
11:36	12:49		SDS	00914286		M	69	TURBT (Transurethral resection bladder tumor)			Dr. ....
13:04	14:00		ODS	01082723		M	23	Cystolitholapaxy			Dr. ...
14:15	15:00		ODS	00813169		F	40	Macroplasty injection collagen			Dr. ....
TH: 04					**Anesthetist: DR. ......**			**Respiratory Therapist:**			**16:15**
07:45	15:39	4NW	IP	00260595		F	62	Laryngectomy total Free flap radial forearm			Dr. .... Dr. ....

Figure 2-1: A Sample Patients Schedule within OR blocks at the JGH Surgical Unit

Cardoen et al. categorized these measures as: waiting time, throughput, utilization, patient deferral, preference, leveling, makespan and general cost objective. Among these objectives the authors mostly focused on the waiting time and utilization. Second, the authors clustered the literature in three groups: studies at the patient level, surgeon level, or administrative level. The results show that most of the developed models are at the patient level and there are avenues to conduct future research at surgeon and administrative levels. Third, they categorized the literature on the basis of the solution technics. Although mathematical programming approaches have been widely applied to the OR scheduling problems, lots of studies developed Mote-Carlo and discrete event simulation models. Stochasticity in patient arrival rate and surgery duration are mainly applied into the simulation models. Cardoen et al. concluded their study by noting that although most of the reviewed studies used real data for testing the developed models, none of them provide a detailed implementation of the results. In the other study, Guerriero & Guido (Guerriero & Guido, 2011) presented a comprehensive literature review on the applications of operations research's techniques in surgical planning and scheduling processes. However, this section will go through the detail of the important studies rather than clustering the articles within different frameworks.

Lamiri et al. Lamiri, Xie, Dolgui, and Grimaud (2008) presented a stochastic mathematical model for operating room planning with elective and emergency patients. Combination of miscellaneous scheduling costs (e.g. over-utilization costs of OR) of these two kinds of patients is minimized in the objective function and a developed Monte-Carlo optimization technique is used to evaluate the stochastic model. The model does not assign the patients into their surgeons' OR blocks. In other words, they assumed that all patients could be assigned to all OR block at any time. Also the problem did not determine the sequence of the patients in an OR block. Moreover, the authors did not consider the distribution function for some of the stochastic parameters (e.g. OR time for each procedure) while the estimated amount of the other stochastic parameter implemented in the model (e.g. required OR time for emergency patients).

Denton et al. also studied allocating surgeries to the OR blocks in a multi-OR environment, when the block duration is stochastic (B. T. Denton, Miller, Balasubramanian, & Huschka, 2010). However, the model did not address the patient sequencing problem. They described the trade-off between opening a new OR, which brings a fixed set-up cost, and continuing to run the current open ORs, which impose overtime cost to the model. The OR allocation problem was formulated as a deterministic, stochastic, and robust model, in which total scheduling costs are minimized. In the deterministic model all surgical case duration is assumed to be known and constant, while the stochastic model considers the expected overtime in the formulation. Yet in the robust approach they considered an upper bound and a lower bound for the duration of surgical cases. The decision is to determine the total number of open ORs and the assignment of surgical cases into those ORs. The stochastic model is defined as a two-stage decision problem, which in the first stage the model finds the total number of open operating rooms and their assigned surgical cases, and in the second stage the model resolves how to distribute OR overtime. To solve the model they adapt the integer L-Shaped method. Also a heuristic method is applied to find a near optimal solution for the robust model. A comparison among all solution methods and models is done using numerical examples generated out of real data from a large healthcare provider.

The patient sequencing problem was modeled by Denton et al., in which the total weighted expected costs of surgeons waiting, OR idling time, and overtime is minimized (B. Denton, Viapiano, & Vogl, 2007). They applied sample average approximation method to solve the model with a real set of data. Gul et al. studied the operating room scheduling when the surgical demand and operating time is random (Gul, Denton, & Fowler, 2015). Their proposed mixed-integer model minimized the expected cost of cancelation, patient waiting time, and operating room overtime. Batun et al. studied the effect of parallel surgeries (i.e. multiple surgeries operated by the same surgeon at the same time) on patient scheduling (Batun, Denton, Huschka, & Schaefer, 2011). On the basis of their assumption each surgery could be split into 3 phases, Preincision, Incision, and Postincision. The surgeon only needs to be present in the critical phase of the surgery (i.e. Incision), so the last phase of the first surgery could be at the same time as the first part of the second surgery. Hence, the model tries to find the optimal patients' schedule which minimizes the total costs of surgeons' idle time and the ORs overtime when surgery duration is stochastic.

Although most hospitals are publicly funded, few articles focus on the specific characteristics of such environment. These hospitals usually receive a fixed annual budget and must serve a huge surgical demand. So the problem is a kind of cost containment rather than making more profit. Vijayakumar et al. presented the patient sequencing problem in which the total number of ORs is fixed and the hospital managers aim to maximize the total number of served patients with respect to a given priority (Vijayakumar, Parikh, Scott, Barnes, & Gallimore, 2013). All parameters in the model are deterministic, including surgical cases duration. All patients are sorted on the basis of their priorities; in the case that two patients have the same priority the patient sequencing rule applies. Three sequencing rules are compared in the study which contains: (i) Shortest surgery time, (ii) longest surgery time, (iii) random. They developed a heuristic method, denoted as First Fit Decreasing algorithm (FFD), to solve the model. FFD assigns the top patient in the priority list to the first open OR with respect to the surgeon's availability. The model is run by historical data and the result is compared with a publicly funded hospital, which follows first in first served discipline, in the Midwest United States on the basis of two efficiency measures: OR utilization and number of unscheduled surgical cases. Their proposed model increases the OR utilization by 20% and reduces 20% of OR working days.

Many manuscripts in the literature have focused on the operational decisions on OR scheduling, while some articles have presented the OR utilization from a strategic perspective. For example, Lovejoy & Li. focused on a decision making problem about building a new OR, or extending the working hours of the current ORs (Lovejoy & Li, 2002). The trade-off between these strategies for increasing OR capacity is investigated on the basis of three performance criteria and stakeholders' perspectives: 1. hospital profit form hospital managers perspective, 2. waiting time from patients' perspective, and 3. operation's starting time from surgeons perspective. The model tries to find an "efficient frontier" of the best solutions by maximizing the hospital profit with respect to the other two criteria as in the constraints. It is illustrated that the hospital profit is increasing in patient's waiting time and is decreasing in operation's starting time reliability. The optimal result suggests expanding the existing OR working time instead of building a new OR, which costs of 6 million dollars. These results are very dependent on the estimation of OR idle/overtime costs. . Olivares et al. assumed that hospital administrations rationally schedule the OR blocks (Olivares, Terwiesch, & Cassorla, 2008). In other words, administrative decision implicitly reflects the balance between the costs of OR idle capacity and OR overtime working. Since the final OR schedule (i.e. administrative decision) is available as an input data, the authors tried to obtain the unobservable cost function behind that decision. Two econometric methods were applied to solve the problem. The first approach is a regression analysis which tries to find some dependent variables to predict the final decision. In the second approach, denoted as structural estimation, a decision model was built with respect to the managerial concerns in a real surgical unit environment. The results show that hospital managers mostly care about OR idle time rather than delays and OR overtime working.

One of the studies which addressed the Master Surgical Scheduling problem is presented by Day et al. (Day, Garfinkel, & Thompson, 2012). They developed an OR block scheduling model by considering both the hospital administrator and surgeons' perspectives simultaneously. The proposed integrated block scheduling (IBS) model breaks the OR scheduling problem into three phases. Prior to these phases, the potential set of OR block schedules are defined for each surgeon, denoted as Dr. i's package. Each potential OR schedule is valued from surgeon and hospital perspectives, denoted as hospitals benefit and Dr. i's benefit. The first phase of the IBS model selects at most one block schedule from each surgeon's package. To this goal, an integer model was developed, in which the hospital and surgeons benefit are combined in a single objective function and the model tries to maximize it. The authors considered several parameters including: surgeons profit, hospital profit, capacity costs, inconvenience cost, lost demand, to present a more realistic model. However, one might challenge the presented objective function, which is neither from the hospital administrative perspective nor from a surgeon's perspective. By combining the two benefit terms into one objective function, the model tries to maximize the total revenue which might not be the same as the optimal equilibrium when the two stakeholders maximizes their own objectives independently.

In the second phase of the proposed IBS model, surgical cases are scheduled into the surgeons' OR blocks. An Arena simulation model was designed to evaluate the surgical case scheduling into the surgical blocks. As in the third phase, the operating rooms were dedicated to the surgeons, who have to share the OR blocks. As a result, most of surgeons with high volume of cases were assigned a full OR block while rest of the surgeons had to share an OR block. The model is recommended to the hospitals with OR utilization about 65% and to the hospitals with a large number of low volume surgeons.

In other study, Agnetis et al. presented an integer model which addresses the surgical unit master problem and the patient sequencing problem simultaneously (Agnetis et al., 2014). The objective function considers the patients' prioritization score and minimizes the overall waiting time. Agnetis et al. presented a decomposition approach to address Master Surgical Scheduling and Patient Sequencing problem based on surgery duration, waiting time and priority class of the operations (Agnetis et al., 2014). Also in one of the most recent studies, Visintin et al. presented a mixed integer programming model to capture the impact of flexible operating room, surgical team, and surgical unit on the Master Surgical Scheduling problem (Visintin, Cappanera, & Banditori, 2016). It is concluded that the flexibility of one of these resources yields significant benefits when the other two are not flexible. Scheduling a flexible operating room for elective and emergency patients is also addressed by Ferrand et al. (Ferrand, Magazine, & Rao, 2014)

2.4. PACU/ ICU-Oriented Approach

Although ORs are presented as the most important component of surgical unit, many researchers have focused on the function of the other key segments of the surgical unit such as: Intensive Care Unit (ICU) and Post Anesthesia Care Unit (PACU). After a surgical procedure, patients are transferred to the PACU to recover from anesthesia. Ideally, there are 1.5 PACU beds for every OR bed, which are equipped with an airway maintenance kit, monitors, and skilled nurses. Within the PACU the patients receive oxygen therapy, blood pressure recording, pain therapy, and so on. Once the PACU discharge criteria have been met, the patients will be transferred to the ICU, main ward, or will be discharged home. The ICU has limited curtail resources and the admission and discharge policies in the ICU directly effect on the surgical unit performance. This effect will be even more stressed in a highly congested ICU. In a crowded ICU and PACU surgical case cancellation and rescheduling increases, which affects the patient's readmission rate, waiting time and dissatisfaction increases.

The impact of surgical case sequencing on the PACU staffing was studied by Marcon & Dexter (Marcon & Dexter, 2006). Seven sequencing policies for surgical case scheduling are compared considering the percentage of days with at least one day delay in PACU admission. The rules are listed as:

- Random sequencing,
- Longest case first,
- Shortest case first,
- Johnson rule,

According to this rule, all cases are listed based on their actual OR time or PACU time, then the shortest time in the list is selected. If the time is a PACU time the procedure will be scheduled as late as possible and if it is the OR time the case will be scheduled as early as possible.

- Half increase in OR time and half decrease in OR time,

  According to this rule, the cases are sorted based on the OR times. The sequencing starts with the case correspond to the shortest OR time, then the case with the third shortest OR time, then fifth one, etc. The sequencing will end by the case with second shortest OR time.

- Half decrease in OR time and half increase in OR time,

  This sequencing rule starts with the case corresponding the longest OR time and it will continue respecting the same pattern as previous rule.

- Mixed OR time,

  According to this rule, the sequencing starts with the case corresponding the shortest OR time, then the case with the Longest OR time, and so on.

A discrete event simulation was developed to evaluate the effect of these sequencing policies on surgical unit performance. The impact of these policies on the OR utilization, PACU completion time, delays in PACU admission, and PACU staffing was presented in the study.

Price et al. studied the effect of surgical scheduling on the PACU's bed occupancy level (Price et al., 2011). To this goal, an integer mathematical model was developed which considered average patient LOS in the PACU and the ICU. The model's objective function minimizes the differences between the expected number of patient transferred into the PACU and the expected number of patients discharged from the PACU to the ICU or to the main ward. Kim et al. studied the process of transferring patient from the OR to the ICU (S. H. Kim, Chan, Olivares, & Escobar, 2015). Using a large data set of hospitalized patients, they quantified the cost of several outcomes for denied ICU admission, and then they evaluated the performance of various admission strategies.

In general, when a new patient arrives into a fully occupied ICU, surgeons/nurses have to either deny admitting the patient or prematurely discharge the current patients from the ICU. Premature discharge or demand-driven discharge refers to moving the patients to the main surgical ward, while they need to spend more time in the ICU. Dobson et al. studied the effect of ICU's patient bumping on the ICU performance (Dobson, Lee, & Pinker, 2010). They assumed that the patient with minimum remaining LOS will be discharged, who can be even the recent patient. With respect to this policy, they evaluated the ICU performance by measuring two indices: 1. the probability of bumping an ICU patient, and 2. the expected number of days remaining for a bumped patient. Each scenario is defined under a specific arrival pattern and ICU capacity. Each scenario contains both the scheduled and unscheduled patients.

From the methodological perspective, the authors developed a Markov chain model. Set of the remaining days of stay for the current ICU patients denoted as the state of the system. The state of the system would be updated at the beginning of each day, when new patients arrive. For all new arrivals random Length of Stays (LOS) are generated and some ICU patients are bumped in the case of ICU bed shortage. Since the state of the system is too large in their Markov chain model, they applied an aggregate-disaggregate method to figure out a stationary solution of the model. Then the set of states is redefined to reduce its size, computational complexity, and storage capacity. It is proved that their method dominates the Gauss-Seidel method based on computer storage capacity and time.

The study provides a numerical experiment to evaluate the model. The authors considered three arrival patterns for scheduled patients: a three-day surgical schedule, a five-day surgical schedule, and a seven-day surgical schedule. For each arrival pattern different scenarios on the basis of the number of scheduled and unscheduled patients is evaluated. The results show a significant decrease in patient bumping when the ICU capacity increases. Also, it is illustrated that the three-day schedule pattern has less bumping rate compared to the other two patterns. And finally the results reveal that the bumping rate is increasing in proportion of unscheduled patients.

However, the assumed policy for premature discharge from the ICU is challenged by Chan et al. (Chan et al., 2011). In general, premature discharge brings some clinical costs such as risk of physiological deterioration and higher mortality risk and some healthcare system cost such as an additional load on hospital resources. A comparison among a new family of discharge policies was presented by the authors, using mortality and readmission rate indices. It is worthy to note that surgical performance measures and surgeons monitoring issues are subject of huge debates in the literature which was studied by Treasure et al. (Treasure, Valencia, Sherlaw-Johnson, & Gallivan, 2002). Furthermore, Dey et al. developed some other performance measures for ICU operations (Dey, Hariharan, & Clegg, 2006).

Chan's model finds the optimal premature discharge policy by minimizing the total costs. It is assumed that all new patients have to admit to ICU and there is no cancellation or rescheduling due to ICU bed shortage; in the case that all beds are occupied one of the proposed premature discharge policies would be applied. In other words, the model would discharge the patient based on the policy with the least expected cost.

From the methodological perspective, they presented a dynamic programing model to find the best discharge policy. The state of the system was identified by the type and number of the current ICU patients as well as the arriving patients. The model deals with a huge state space so to find a robust discharge policy a greedy policy was proposed, which ignores the effect of future arrivals and patient's LOS. With this myopic policy, patients who are in the least cost class in each state will be discharged. Patients would be categorized in various cost classes regarding different criteria such as: mortality risk, readmission risk, and lowest remaining LOS. The readmitted patients will face higher mortality risk and longer LOS. Yet the authors categorized patients based on a "readmission load ratio" which is the difference of two fractions; probability of readmission over LOS after and before occurring premature discharge. The model is empirically validated using the data on more than 5000 patients. It is concluded that the best proposed policy would protect the patients with high mortality risk from a premature discharge, while it wisely prioritizes the patients with low mortality risk.

ICU utilization is the other highlighted issue in the literature. Kim et al. analyzed the admission and discharge policies in a public hospital (S.-C. Kim et al., 1999). They proposed that the ICU patients potentially arrive from four different sources: surgical ward, accidents and emergency, operating room (emergency), operating room (Elective). The authors assumed a Poisson distribution for patients' arrival and discharge rate and simulate their model using XCELL software. Several queueing measures (i.e. ICU bed utilization, average number of patients in system, average number of patients in queue, average time in system, average time in queue, probability of all beds empty, and probability of arriving patients waiting) are calculated to evaluate the operating characteristics of the ICU in a steady state. It is concluded that the surgical case cancellations are due to an inappropriate patient scheduling rather than the ICU bed shortage.

In a similar study, Ridge et al. applied a queueing theory and a simulation model for the ICU bed configuration (Ridge, Jones, Nielsen, & Shahani, 1998). Various types of emergency and elective patients (with different LOS and arrival rate) were considered as an arrival sources in the simulation model. They implemented the priority index in the model which is affected by the patient's type (i.e. emergency, elective) and the number of her previous case referral/cancellation. It was assumed that the pre-planned patients are deferred for some period of time, if the number of free beds is below a minimum level. As the result, the total number of required beds in the ICU with respect to the various patient arrival patterns was presented.

Min & Yih implemented ICU capacity constraints into an OR block scheduling model (Min & Yih, 2010). The three parameters of the model, including surgery duration, LOS in ICU, and elective patient arrival rate were the basis for generating the potential scenarios. The set of scenarios is finite, and the probability distribution of them is discrete. Due to the large size of scenarios, a sample average approximation technique was applied to solve the stochastic model. Min & Yih assumed that the number of emergency patients is stochastic, so the available OR blocks for elective patients is not fixed. The key decision variable of the model denoted by x_{ib} and y_{ib} , defined as follows.

$$x_{ib} = \begin{cases} 1 \text{ if a petient i} \in \text{I is assigned to a surgical block b} \in B \\ 0 \text{ otherwise} \end{cases}$$

$$y_{it} = \begin{cases} 1 \text{ if a patien i} \in \text{I occupies a SICU bed at day t} \in \text{T} \\ 0 \text{ otherwise} \end{cases}$$

The objective function of the model includes the expected total cost of OR overtime working and the cost of patient assignment to OR blocks. The model considers a priority score for each patient, which depends on her urgency status. With respect to this score, a waiting cost is assigned to each patient. Hence, to minimize the total cost the model expedites scheduling of the patients with higher priority. The objective function is subject to several constraints such as availability of the OR block and ICU beds. The results show that when the available ICU beds increases the OR utilization and the average number of scheduled patients increase as well, while the total cost of OR overtime and patient assignment decrease.

Furthermore, Litvak et al. studied the effect of pooling the hospitals' ICU beds (i.e. when the regional hospitals share their ICU beds) on the surgical case cancellation rate (Litvak, van Rijsbergen, Boucherie, & van Houdenhoven, 2008). A simulation model was developed to show that hospitals cooperation on sharing the ICU beds, minimizes the total cancellation rate (i.e. emergency patient admission maximizes) while the efficiency increases.

2.5. Ward Beds-Oriented Approach

Beside the aforementioned issues, to systematically approach to surgical unit management one must consider the availability of surgical ward's bed. Many studies have addressed the hospital bed management, yet the integration of downstream bed availability in the OR scheduling problem has been yet to be fully explored. However, to better differentiate among these topics, some manuscripts on both the surgical bed management and the hospital bed management are highlighted.

Queueing techniques, simulation methods, and mathematical programming have been widely used by researchers to develop an admission/discharge policy for various types of patients. Meng et al. developed a robust optimization model on the daily bed capacity planning of elective patients in a public hospital (Meng et al., 2015). The model considers both emergency and elective cases and enforced quotas for elective patients to diminish downstream bed shortage.

Adan & Vissers presented an admission planning model to optimize the patient mix. The model considers the required surgical resources (i.e. ICU beds, nursing staff, OR beds) to find the optimal mix of admitted patients on each day (Adan & Vissers, 2002; Vissers, Adan, & Bekkers, 2005).

Gorunescu et al. developed a queueing model to manage bed occupancy in a geriatric department (Gorunescu, McClean, Millard, & Correspondence, 2002). They presented an M/PH/c queueing model, which considers a Poisson arrival distribution and phase-type service distribution with a limited number of beds. The objective function maximizes the hospital's benefits by calculating an efficient number of beds given a specific probability of lost patients.

Cochran & Bharti developed a queueing network and a comprehensive simulation to study the patient flow in an obstetrics hospital. With this goal they implemented a very detailed patient flow chart into the simulation model (Cochran & Bharti, 2006). Furthermore, a time-dependent (i.e. weekend, weekday, peak hours) patient arrival pattern was proposed on the basis of the hospital's historical data for a 15 year period. This enabled them to evaluate the utilization level of all the components of the obstetrics hospital. Then an integer mathematical model was developed to minimize the mean absolute deviation of unit utilization levels. Results show that a 15% increase in the number of beds results in a 38% increase in the total number of served patients.

However, simulation is not the only method to address the hospital bed management problem. Ayvaz & Huh (Ayvaz & Huh, 2010) applied a dynamic programming approach to this problem. The model allocates a fixed capacity of the hospital beds to the emergency and elective patients assuming backlogged elective patients and lost emergency ones in the case of fully occupied beds. The authors discussed various approaches from the field of Inventory Systems and Single Resource Allocation. They proved that the safety capacity of beds for emergency patients is decreasing in the number of backlogged elective patients. Also, Best et al. developed a system dynamic model to find the optimal bed capacity and bed configuration that maximize a hospital's utility function (Best, Sandikci, Eisenstein, & Meltzer, 2015). They study the impact of forming a large unit to pool demand or forming specialized units to focus on niche medical services.

Over the last decade, one emerging research effort evident is to integrate surgical ward's bed availability into the OR scheduling problems. Helm and Van Oyen applied mix integer programming approach to model the entire hospital as a coordinated system to optimally schedule elective patients along with emergency cases (Helm & Van Oyen, 2014). The other study on integrated surgical unit modeling was presented by Chow et al. (Chow, Puterman, Salehirad, Huang, & Atkins, 2011). They focused on the OR efficiency and considered downstream bed utilization to schedule the surgeon blocks and patient types. They developed a mixed integer optimization model, denoted by Surgical Schedule Optimizer (SSO), and applied a Monte-Carlo simulation, denoted by Bed Utilization Simulator (BUS), to reduce the peak of bed occupancy. SSO helps surgical planners to obtain surgical block schedule directly or through surgical scheduling guidelines derived from it. BUS considers unplanned patients to test the schedule, obtained from SSO, surgical planners analyze the result and re-adjust the schedule, and it iterates until reaching an optimum surgical block schedule. In fact, BUS predicts the daily bed occupancy for each downstream surgical unit, and makes planners capable of seeing the effect of surgical block schedule. Hence, they reschedule the OR block if the downstream bed occupancy is not in a desired range. Fig. 2-2 is a copy of the proposed surgical scheduling diagram, illustrated in their study.

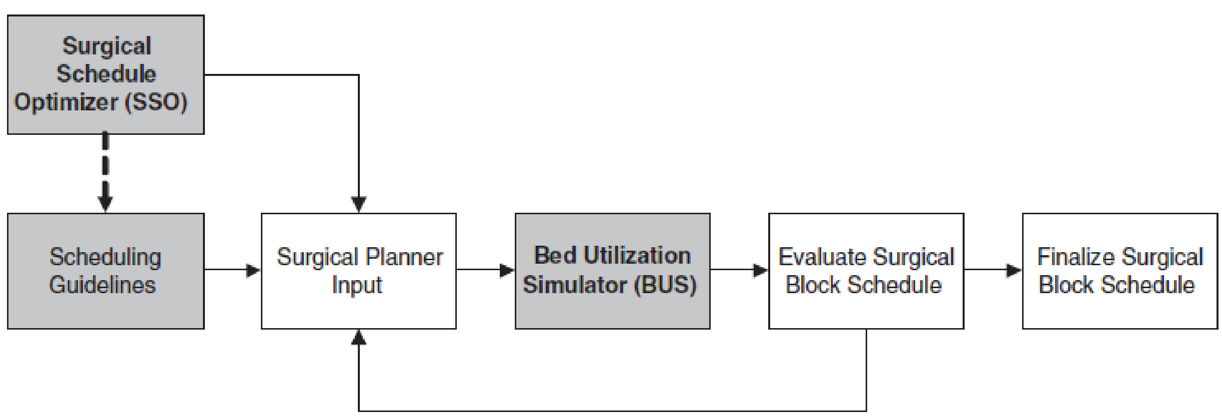


Figure 2-2: Proposed Framework for OR Scheduling Presented by Chaw et al.(2011)

The SSO's mathematical model is briefly review in this chapter to better differentiate it with the proposed model in chapter 4. The SSO minimizes the maximum number of occupied beds within the planning period for each unit. In other words, balancing the maximum number of required beds across various units is the only objective of this model.

Table 2-3: Notation and Decision Variables

Notation

b	*surgical blocks*
u	*wards*
i	weekdays (1, . . ., 5)
d	surgeons
w	weeks of the surgical schedule
p	patient types
j	days of the surgical schedule (1, . . ., 7.w)
$B(d)$	blocks associated with surgeon d
Bed_b^{pujiw}	expected number of bed-nights/days used by one patient of type *p* in ward *u* on day *j* due to surgical block *b* scheduled on day *i* of week *w*
$NumOR_b$	OR-days required for each surgical block *b*
$ORperDay^{iw}$	OR-days available on day *i* of week *w*
${ORperDaySurgeon_d}^{iw}$	OR-days available on day *i* of week *w* for surgeon *d*
$WeekBlock_b^w$	number of blocks *b* available in week *w*
$TotalBlock_b$	total number of blocks *b* in the surgical schedule
$NumCases_b^p$	number of cases for each patient type *p* in surgical block *b*

Decision variables

X_b^{iw}	1 if block *b* is scheduled on day *i* of week w, 0 otherwise
MD_u	maximum number of beds in use in ward *u* over the scheduling period

SSO Model

Objective Function	$Min\ \sum_u MD_u$	
S.t.		
Daily OR-Day Capacity	$\sum_b X_b^{iw}.NumOR_b \leq ORperDay^{iw}$	$\forall\ i,w$
Daily OR-Day Capacity per Surgeon	$\sum_{b\in B(d)} X_b^{iw}.NumOR_b \leq ORperDay\mathrm{Surgeon}_d{}^{iw}$	$\forall\ i,w,d$
Weekly Surgical Block Capacity	$\sum_b X_b^{iw} \leq \mathrm{WeekBlock}_b^w$	$\forall w,b$
Surgical Block Balance	$\sum_w \sum_i X_b^{iw} = TotalBlock_b$	$\forall\ b$
Maximum Bed Utilization Across the Scheduling Period in Each Ward	$\sum_w \sum_i \sum_b \sum_p X_b^{iw}.Bed_b^{pujiw} \leq MD_u$	$\forall\ j,u$

The SSO considers the average LOS of the elective cases to estimate the bed occupancy level of the units. However, the Bed Utilization Simulator was developed to evaluate the derived surgical block schedule. To this goal, the authors assumed that the surgical unit is uncapacitated, so no surgical case would be cancelled due to the bed shortage.

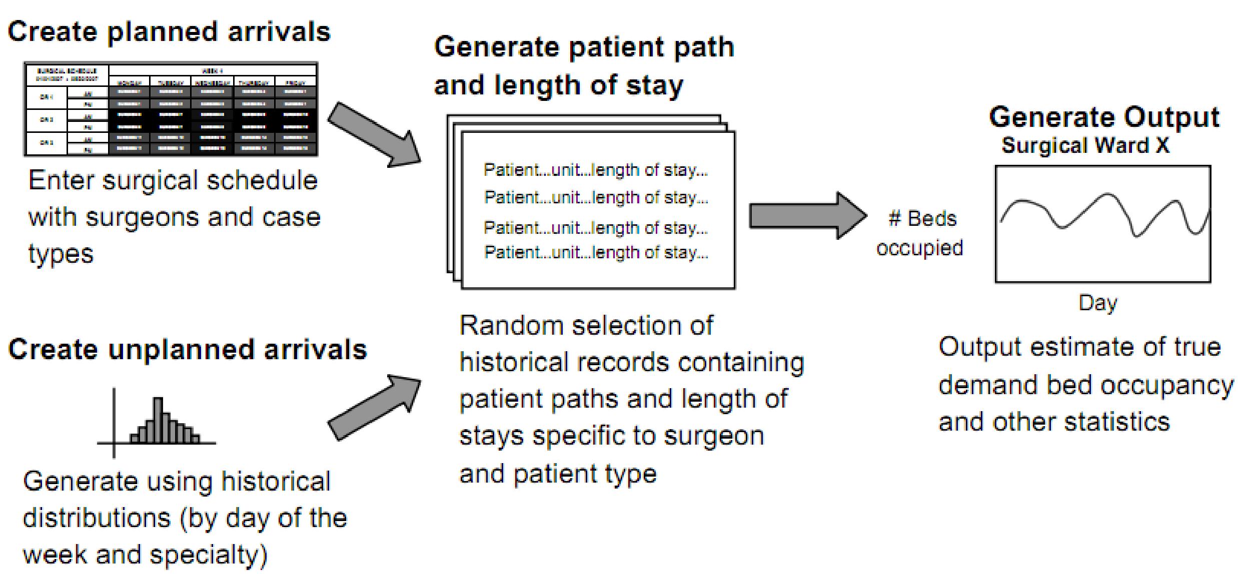


Figure 2-3: Image of the Proposed Framework for BUS model Presented by Chaw et al. (2011)

The BUS uses empirical LOS of the emergency patients as well as the elective patients, to calculate the total bed occupancies through all units. The empirical analysis also shows that the emergency patients occupy 54% of the surgical unit beds, and this rate does not significantly change over the planning period. However, this analysis shows that the elective patient's bed occupancies fluctuates over this period. An overview of the BUS structure, which is presented in their study, is illustrated in figure 2-3.

Authors suggested that the surgeons, who has high patient volume with average LOS of 2 days, should be scheduled on Monday and Wednesday. It was also suggested to group the specialties with similar ward/OR requirement into the same OR block.

2.6. Care Providers-Oriented Approach

Surgical unit performance is not only a function of quantity of available physical resources (e.g. OR, ICU, main ward's beds) or even the planning of these resources, but also it depends on the efficiency of care providers (e.g. surgeons, anesthetists, nurses). Huckman & Pisanp studied the performance of cardiac surgeons in multiple hospitals as a case study for the developed model represented by "firm-specific performance of freelancers" (Huckman & Pisano, 2006). The model explores whether or not the surgeons' performance is a hospital-specific issue and if it is, to some degree, how it can be explained. Three assumptions are the basis of the model: 1. the surgeons are freelancers who can contract with all hospitals, 2. surgeons should collaborate with nurses and anesthetist as a team in a particular hospital, 3. risk adjusted mortality rate is the only measure of the surgeon's overall performance. The authors proposed three influential factors on a surgeon's performance: surgeon effect, hospital effect, and surgeon-hospital effect; and they analyzed the performance of the surgeons who were working in different hospitals at the same time. The model was calibrated with the empirical data from a hospital in Pennsylvania. The data represented more than 38500 surgical cases, served by 200 surgeons. A logistic regression model was also developed to evaluate how the surgeon's performance is affected by the total number of surgical cases served by her. The results show a reduction in overall mortality rate by increasing surgeon's volume on a specific procedure across all hospitals. Also the results support the hypothesis that the effect of the total number of surgical cases served by a surgeon in a specific hospital on her performance in that hospital is significantly greater than her performance in the other hospitals. It shows that surgeons' overall performance is not completely transferable across the hospitals. It is recommended to the hospital managers to hire high performance surgeons and try to keep them full time in the hospital.

KC & Staats studied the effect of surgeon's previous experiences on her performance (Diwas Singh Kc & Staats, 2012). These experiences were categorized into the focal (i.e. experience with the same task) and relevant experience. The authors showed that the focal and related subtask varieties have opposite, nonlinear effects on the operation's outcome. In a novel study by KC & Terwiesch, the effect of focus on hospital operational performance was studied (Diwas Singh KC & Terwiesch, 2011). They showed that focused hospitals provide faster services with higher level of quality and lower length of stay. In another study they illustrate the impact of workload on service time and patient safety (Diwas S Kc & Terwiesch, 2009). Furthermore, and from a nursing perspective, a study on good nursing level was done by Yankovic & Green (Yankovic & Green, 2011). A summary of the highlighted articles in the literature is reviewed in table 2-4.

Table 2-4: Summary of the Reviewed Articles

Authors (year)	*Problem (Issues to be addressed)*	*Issues to be addressed*	*Objective (Performance criteria)*	*Solution Technique*	*Decisions*	*Assumption*	*Data (Application)*
Lovejoy & Li (2002)	Capacity Expansion: Building new OR Vs. adding overtime working	Building new OR / Extending the working hours in current ORs	Maximize the profit Efficient frontier for Multi objective situation: Patient: Wait to get on schedule Surgeon: Start time reliability Administration: Hospital profit	Bulk service Queue/ Approximation/ validated with Simulation	# of cases to be scheduled per day/ Probability that a Schedule procedure is on time	Poisson arrivals/ Same distribution function on the procedures' LOS/ Overtime cost	A client hospital
Huckman & Pisano (2006)	Effective criteria on the surgeon's performance: Volume & Hospital	Surgeons' Performance correction with the volume of cases in various hospitals	The risk adjusted mortality rate	Empirical/ Logistic regression	How to increase surgeon's performance: Practical guidelines for the hospital manager/ surgeons	Only the Surgeon-Hospital effect on the surgeon's performance is considered	Pennsylvania Healthcare Council year 1994-95
Olivares et al. (2008)	Reserving OR: The real underlying cost function	OR time allocation cost function in reality/ how to reserve OR time	Estimate the cost ratio for OR overtime cost and idle time cost	Empirical/ Econometrics methods	Howe to reserve OR time for each cardiac case	The surgeons and hospital administration's decisions on OR block allocation is rational and optimal	258 cardiac surgery cases
Cardoen et al. (2009)	Literature review: OR scheduling Problems	classification	Classification on the basis of manuscripts' features (e.g. technique, decision, stochasticity)/ Review of previous studies	-	Avenues for the future studies	Papers between 1998- 2009	-
May et al. (2010)	Literature review: Surgical Scheduling problems		Time frame classification/ Context classification/ Review of previous studies	-	Avenues for the future studies	Papers before 2010	-
Dobson et al. (2010)	Effects of ICU Bumping on the performance: Patient arrival rate & schedule matter	Create a planning tool to predict performance under different arrival rates and capacity load scenarios	The probability that a patient is bumped/ The expected number of days remaining for a patient to get bumped	Markov Chain/ New technique to aggregate the system's states	Best capacity load scenario/ Minimize the ICU bumping	Random LOS at ICU/ Fixed ICU capacity/ Random arrivals/ If necessary to bump k patient, we remove the k with the least remaining days	-
Denton et al. (2010)	Robust OR block scheduling & Patient scheduling	Uncertainty of surgery duration	Min Max scheduling cost	Stochastic Integer Programming/ Robust Optimization/ Heuristic solution technique	Optimal number of open ORs/ Patients schedule	complete set of surgeries is known in advance / Opening an OR has fixed cost/ OR time for each procedure has a lower and upper bound	Real data from a large healthcare provider/ Motivated by Mayo Clinic, Rochester, MN
Min & Yih (2010)	Patient scheduling: Availability of the ICU bed matters	Uncertainty of surgery duration/ Uncertainty of LOS	Minimize the cost of patient waiting and OR overtime working	Sample average approximation	Number of Open ORs/ Allocation of OR blocks to patients	Stochastic surgery duration/ Stochastic LOS/ OR block schedule is given	-
Chow et al. (2011)	OR block scheduling & Patient scheduling: Availability of ward beds matters		Balance the ward congestion/ minimize the peak of bed occupancies	Integer programming/ Monte-Carlo simulation	Practical guidelines for the hospital manager/ Optimal OR block schedule	Average LOS/ Uncapacitated surgical unit/ A fixed number of elective patients should be scheduled in a multi-week period	The admissions/discharge/ transfer system and the OR scheduling office system: Royal Jubilee Hospital, Victoria, BC
Batun et al. (2011)	Patient sequencing: Parallel surgeries		Minimize the surgeons idle time cost and OR overtime cost	L-shaped method for stochastic mixed-integer program	Optimal patient sequence	Stochastic surgery duration/ A surgeon can operate on two cases at the same time	Mayo Clinic's Division of General Thoracic Surgery at St. Marys Hospital, Rochester, MN
Day et al. (2012)	OR block scheduling & Patient scheduling: From surgeon & hospital perspectives	OR block Scheduling and surgical case allocation to the blocks	Maximize the total monetary value to the surgeons and hospital combined	Integer Programing/ Simulation	Optimal OR block schedule/ Patients schedule	OR overtime cost is 50% higher than OR regular time cost/ Shared OR block	Real data from a large healthcare provider/ Motivated by Mayo Clinic, Rochester, MN
Vijayakumar et al. (2012)	Patient scheduling/prioritization in a public hospital	Surgical case scheduling in public hospitals based on patients priority	Maximize OR utilization/ Minimize number of unscheduled cases	Mixed Integer Program/ Heuristic solution technique	Optimal patients schedule	Number of ORs is fixed Parameters are deterministic (surgery duration, etc)	A publicly funded hospital, Midwest USA
Chan et al. (2012)	Comparing a family of demand-driven discharge strategies in the ICU	Impact of various discharge scenarios on ICU	Optimal demand-driven discharge policy with the least cost/ (mortality rate, readmission rate)	Dynamic Programming	Admission/Discharge policy from ICU	Random Geometric (memoryless)LOS at ICU for each type of patient/ specific discharge cost per patient type/ Fixed ICU capacity/ one patients arrives in a short time	Over 5000 actual ICU patient

Chapter 3

Surgical Ward Design: A Case Study at Montreal Jewish General Hospital

3.1. Problem description

The JGH surgical unit annually serves more than 6000 patients grouped into more than 400 procedures, categorized as 15 main specialties or divisions (e.g. orthopedic, cardiac, gynecology, urology, colorectal, general surgery). Although JGH does perform some emergency surgeries, it is not a trauma hospital; therefore, most of its completed surgical cases are elective cases. The JGH surgical unit has over 130 staffed beds, and runs 13 operating rooms to serve these patients. Maximum average LOS belongs to vascular specialty, which is 14.3 days. With more than 1600 operations, "cataract extraction" was ranked first in terms of volume among all surgical operations in 2013. Required operating room time and LOS in the surgical ward varies among the patients. However, the patients that undergo the same procedure tend to have similar LOS distribution.

Surgical case cancellation, because of downstream bed unavailability, is one of the main managerial concerns in the hospital. The JGH surgical unit is planning to be relocated to a new building, known as *Pavilion K*. Hospital managers want to redesign the surgical unit configuration in *Pavilion K* to better meet patient demand. Managers are interested in knowing the risk of surgical case cancellation under various surgical unit configuration policies in the new surgical unit.

On the basis of the current surgical unit configuration, a certain number of beds are prioritized to serve a single specialty (e.g. oncology) or a group of specialties (e.g. colorectal and general surgery). However, the surgical beds are not completely dedicated to a particular specialty. For example, if there is no available bed for an oncology patient, she will be admitted to the other divisions with empty beds. In such a system, nurses must be able to serve all types of patients. If the beds were dedicated to the divisions separately, nurses are able to focus on a certain type of patients with similar needs, which tends to increase the quality of care. KC and Terwiesch showed that focused hospitals have better outcomes in the delivery of care, also they have lower LOS and mortality rate (Diwas Singh KC & Terwiesch, 2011). The JGH managers are interested to find out how many beds the surgical unit needs under a complete dedicated bed configuration.

In the new surgical unit configuration, JGH wants to run a new High Acuity Unit (HAU) to exclusively serve high-risk patients. High-risk patients are those with high likelihood to develop complications after the surgery and are selected on the basis of their preoperative test results, yet the exact distribution of these patients has not been documented. HAU provides a high level of nursing and monitoring, yet does not require the subspecialists of an ICU. Therefore, selected high-risk patients get extra care in an equipped HAU which prevents the start of postoperative complications. JGH's surgeons believe that HAU does not only mitigates the patient's poor outcome, but also may decrease the patient's length of stay (LOS), since the quality of care in HAU increases. However, the distribution of this LOS-reduction has not yet been calculated. Hence, to decide whether or not to open the HAU, JGH managers want to evaluate the potential effect of HAU on the total number of required surgical beds.

3.2. Analysis Procedure

3.2.1. Simulation Model

We develop a simulation model, using Arena version 14.5 from Rockwell Automation, to capture the complexity of surgical unit patient flow. This model is useful to study the impact of LOS uncertainty on surgical unit bed configuration. Our Surgical Ward Design (SWD) simulation model is a planning tool, which also helps the hospital managers to understand the impact of operating room schedule and surgical unit bed configuration on the surgical bed shortage rate. Any OR schedule can be the input of the SWD simulation model, and the output is the number of required surgical beds at each day. Note that the total number of beds is assumed to be unlimited, so, the model can show the total number of required beds to serve surgical patients. The SWD simulation model furnishes the hospital managers with a range of possible surgical unit configurations with respect to the desired surgical service level.

Total surgical unit beds are split into two parts: the main ward beds and the HAU beds. HAU beds are shared among all specialties. However, the main ward beds are distributed among the eight specialties (i.e. those with more than 100 operations in a year) on the basis of different policies. The first policy, denoted as Pooled Policy, assumes that all surgical beds are shared among all patients from all specialties. In the second policy, denoted as Dedicated Policy, all divisions work independently and admit their own patients. The third policy, denoted as Semi-Pooled Policy, allows some divisions to share their medical beds while the rest are working independently, as discussed in more detail in section 3.3.2.

After the surgery patients may be transferred either to the main surgical ward or to the HAU. It is assumed that at most 20 percent of the patients are admitted in the HAU (i.e. at most 20% of the patients are high-risk cases). If the managers decide not to open HAU all patients would be transferred to the main ward. In this case the HAU arrival rate will be zero. The third scenario is also considered in which the HAU's arrival rate is 10% of the surgical unit's arrival rate.

A schematic illustration of a sample of possible surgical unit configurations is drawn in figure 3-1, in which S1 refers to first specialty, S2 refers to the second specialty, and so on. As mentioned before, 8 specialties are considered in this simulation model. The first raw in figure 3-1 illustrates the dedicated policy whereas the second row is pooled policy.

To find the maximum number of required surgical beds for serving all patients, the total number of surgical beds is not bounded in the SWD model. Unlimited number of beds enables

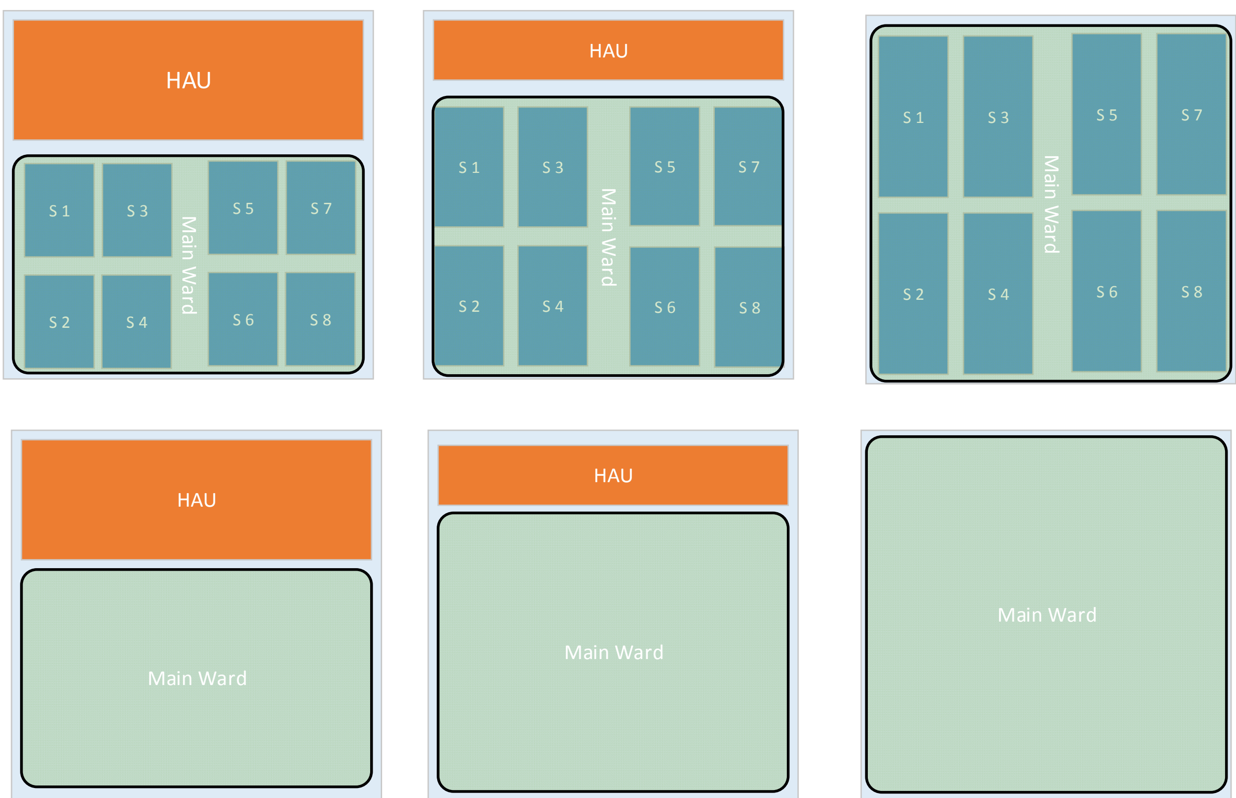


Figure 3-1: Schematic Surgical Unit Bed Configuration

us to study the sensitivity of the configuration's total required beds to the service level. The length of simulation run is set to 800 days, with a warm-up period of 100 days. The length of warm-up period is sufficient to have steady state results based on the initial results. Also, 700 days is required to achieve desired statistical precision.

3.2.2. Data Analysis

JGH provided us with the data for 1767 surgical operations (i.e. elective cases and emergency cases) in 2013. For each case, we access to the admission date, surgery date, discharge date, specialty, and the surgical procedure. Also, the average required operating room time is available for each procedure. A large amount of time was spent to clean and validate data before generating the SWD input data. Out of more than 15 specialties in the surgical unit, those

Table 3-1: JGH Data Summary

Data Summary	
Number of Data Points	639
Min Data Value	1
Max Data Value	64
Sample Mean	6.21
Sample Std Dev	8.36

Table 3-2: ARENA Dist. Fitness Result

LOS Dist. Function	Sq Error
Gamma	0.00513
Beta	0.00527
Lognormal	0.00592
Erlang	0.0137
Exponential	0.0137
Weibull	0.0205
Normal	0.0723
Triangular	0.0909
Uniform	0.108
Poisson	

Table 3-3: Fitted Gamma Distribution Info.

Distribution Summary:	
Distribution:	Gamma
Expression:	0.5 + GAMM(7.35, 0.77)
Square Error:	0.005127**Data Summary**
Chi Square Test	
Number of intervals	19
Degrees of freedom	16
Test Statistic	89.80.115
Corresponding p-value	< 0.005

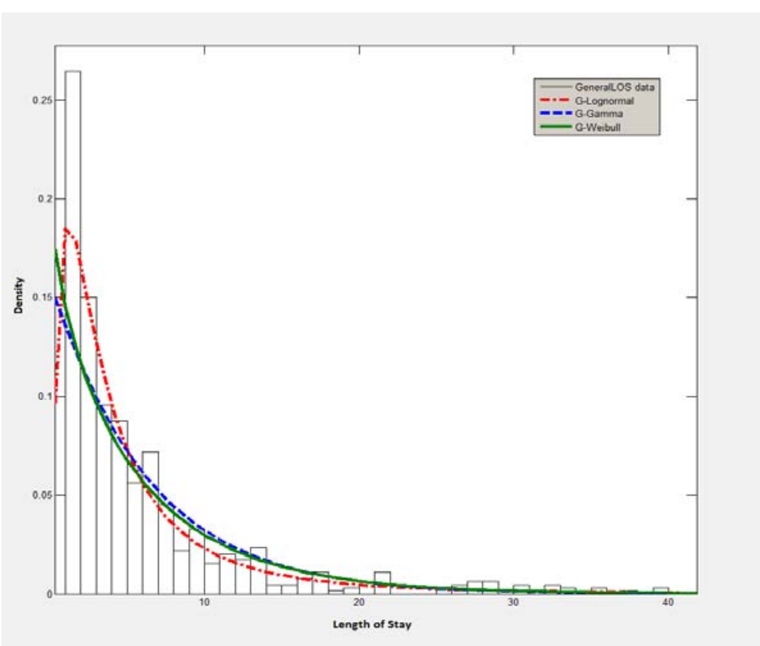


Figure 3-2: General Surgery LOS – Fitted Distribution Functions

with more than 100 operations in a year are selected in this study. For those specialities with less than 100 cases (i.e. 1% of the annual hospital's patient volume) in a year, it is not feasible to find an accurate distribution function for the patient LOS. Furthermore, specialties with less than 1% of the hospital's annual demand is not in the center of attention by the managers, since the majority of surgical beds should be dedicated to the main specialties. Hence, the model complexity is mitigated by focusing on the following specialties: Orthopedic, Vascular, Gynecology, Urology, Colorectal, General Surgery, E.N.T., and Breast Oncology. For each specialty, the best distribution function is fitted to the patient LOS, using Arena. It evaluates the goodness of fit to a function using the Chi Square Test. For example, figure 3-2 shows the 3 fitted LOS distribution functions for General surgery patients. Yet the P-value is not greater than 0.05, so the historical data is used in the SWD simulation model. Tables 3-1 to 3-3 summarize the General surgery LOS data analysis. For those specialties for which the corresponding p-value to the fitted distribution function is less than 0.05, we randomly selected the LOS from the empirical data to be assigned to an arriving patient.

Table 3-4: Operating Room Schedule for a Random Week at JGH

Specialty	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Colorectal	2	3	0	2	1	0	0
E.N.T.	0	1	5	0	4	0	0
General surgery	2	1	0	1	2	0	0
Gynecology	1	3	2	1	2	0	0
Plastics	0	0	0	1	0	0	0
Urology	0	2	1	3	1	1	0
Vascular	1	0	1	0	1	0	0
Breast Oncology	1	0	0	1	0	0	0

The JGH's operating room schedule in winter 2013 is used in our simulation model as the patient arrival rate for all specialties. This schedule is designed for periods of 6 weeks. So during a year the OR schedule is repeated 8 times. For example the OR schedule for the first week of this period is illustrated in table 3-4. As presented in the last two columns, surgeons choose not to operate on Saturdays and Sundays unless there is an emergency case.

3.3. Results

First it is assumed that the HAU has not been opened yet, and the surgical ward configuration is illustrated under pooled and dedicated policies in the following section. Second, subsection 3.3.2 comments on the effect of HAU on the pooled and dedicated policies in the main ward configuration. The semi-Pooled policy results are illustrated in subsection 3.3.3.

3.3.1. Pooled Policy vs. Dedicated Policy

Service Level Index (SLI) presents a quantitative measure of the surgical bed shortage rate. SLI shows the percentage of the staffed required bed-days in the surgical unit. As a result of SWD simulation model, the maximum number of beds occupied at any single day presents the total number of required beds to serve all surgical patients (e.g. 86 beds under pooled policy), so if the hospital managers decide to allocate 86 beds to the main surgical ward the SLI will be 100%. In the case that the managers decide to allocate a lower number of beds to the main ward, the SLI decreases as is shown in figure 3-3. This plot illustrates the required number of surgical beds under pooled policy, when there is no HAU in the surgical unit. For example, the black bar is depicted in the plot, which shows that exactly 49 beds are occupied for 34 days out of 700 days (simulation period). Also this example shows that the SLI is 42.4% if the surgical unit managers decide to assign 49 beds to the main ward. It means that with 49 beds in total, only 42.2% of the days the hospital would have a sufficient number of beds to serve its patients. As the green bar in the chart shows, configuring the main ward with 64 beds guarantees a 90% SLI, yet to increase it to 100% the surgical unit needs to staff 22 more beds. The marginal gain of each bed decreases while the SLI approaches to 100%.

Figure 3-4 illustrates the required number of surgical beds under dedicated policy. The red bar in this plot shows that 79 beds are sufficient for 90% SLI for each specialty, and 50 more beds are required to achieve 100% SLI. Service Level Index under the dedicated policy guarantees the minimum service rate by each specialty. For example, a 90% SLI here means that the service level index for each division is at least 90%, so the total surgical SLI is at least 90%. Hence, in the dedicated scenario, a more strict service level constraint is applied that increases the minimum required number of beds. However, this is part of the dedicated policy's setting and the JGH managers are willing to apply 90% SLI to each specialty when they operate separately.

In comparison, it is observed that the pooled policy results in saving 15 beds at the 90% SLI and 43 beds at the 100% SLI, which is expected as per the fundamentals of operations management. However, this was deemed useful by the project team since it estimated the extent of the savings due to pooling the beds. It is worthy to note that in both policies increasing the SLI from 90% to 100% exponentially increases the required surgical beds. However, this increase is almost linear when the LSI changes from 30% to 90%.

Figure 3-5 reports on each division's bed configuration separately. So the required number of beds for each division with respect to the desired SLI (i.e. in the range of 80% to 100%) is plotted.

It is observed that Colorectal, General and Vascular specialties consume the majority of surgical beds to serve their patients. At 80% SLI, Breast Oncology, E.N.T., Gynecology, Urology, and

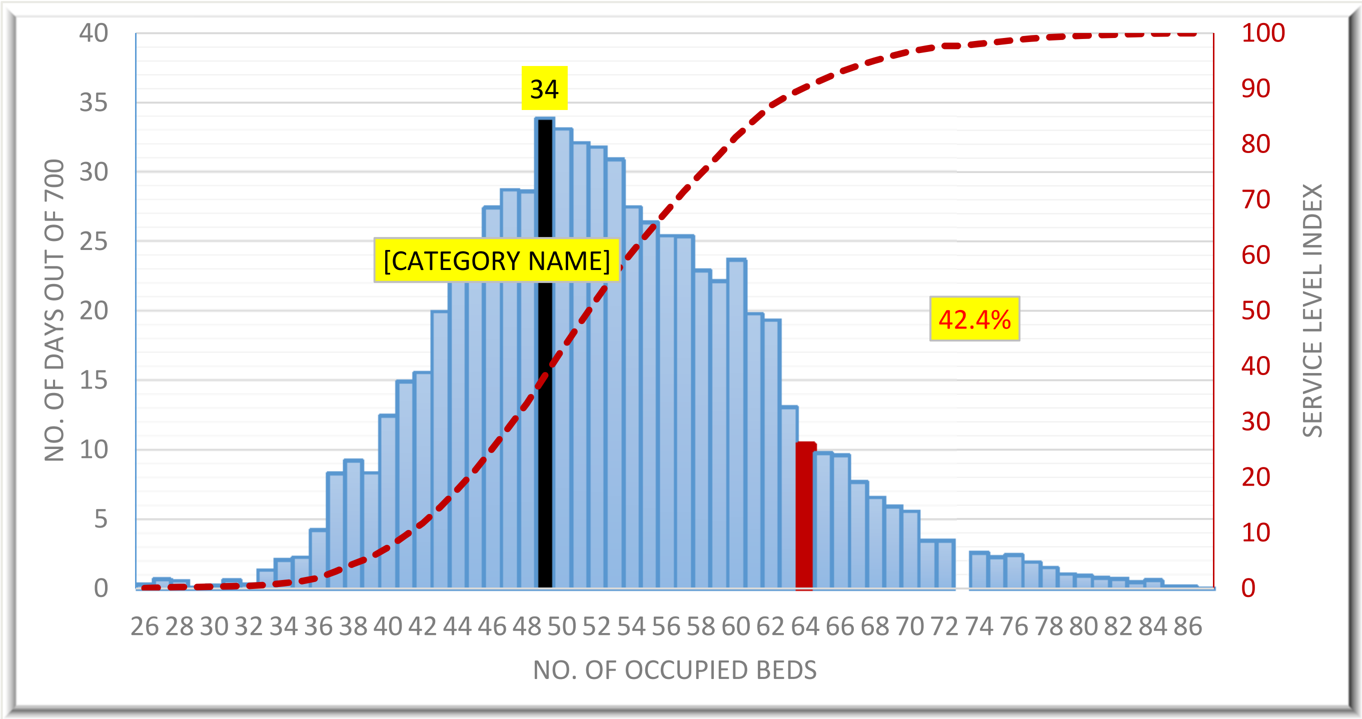


Figure 3-3: Required Bed Frequencies- Pooled Policy

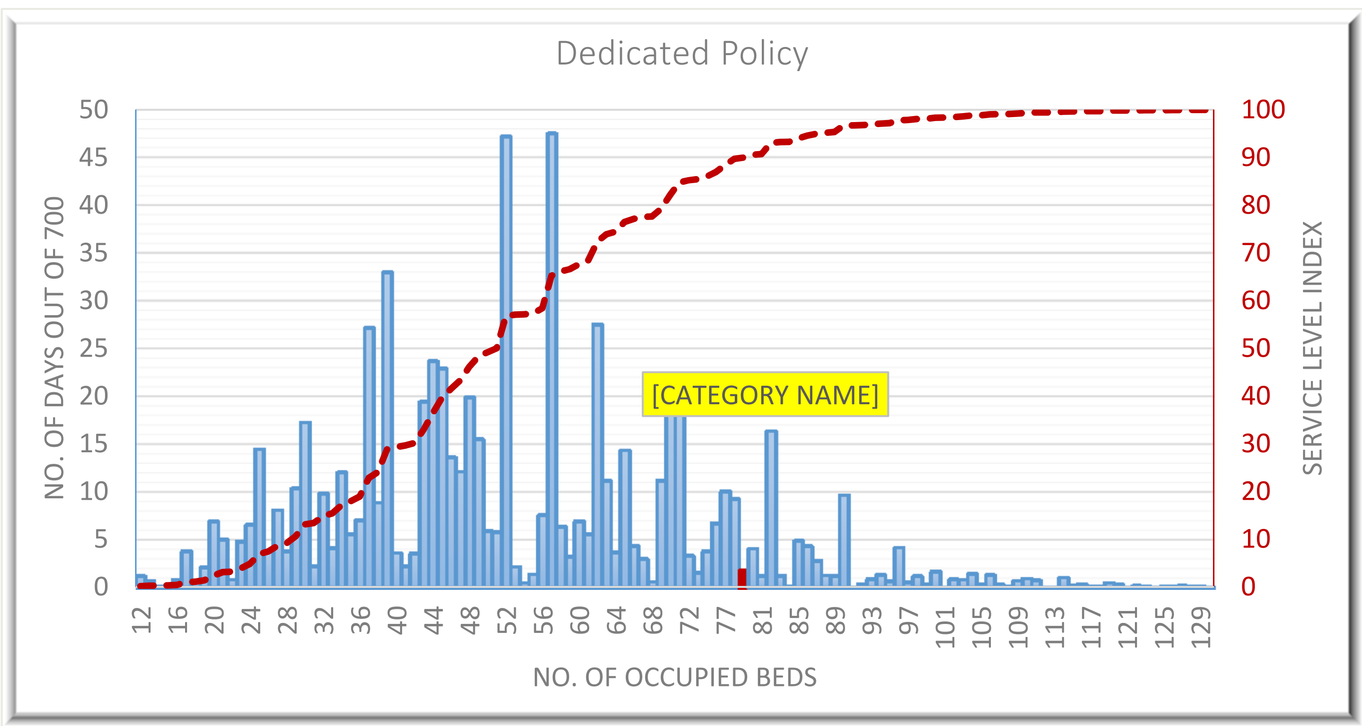


Figure 3-4: Required Bed Frequencies- Dedicated Policy

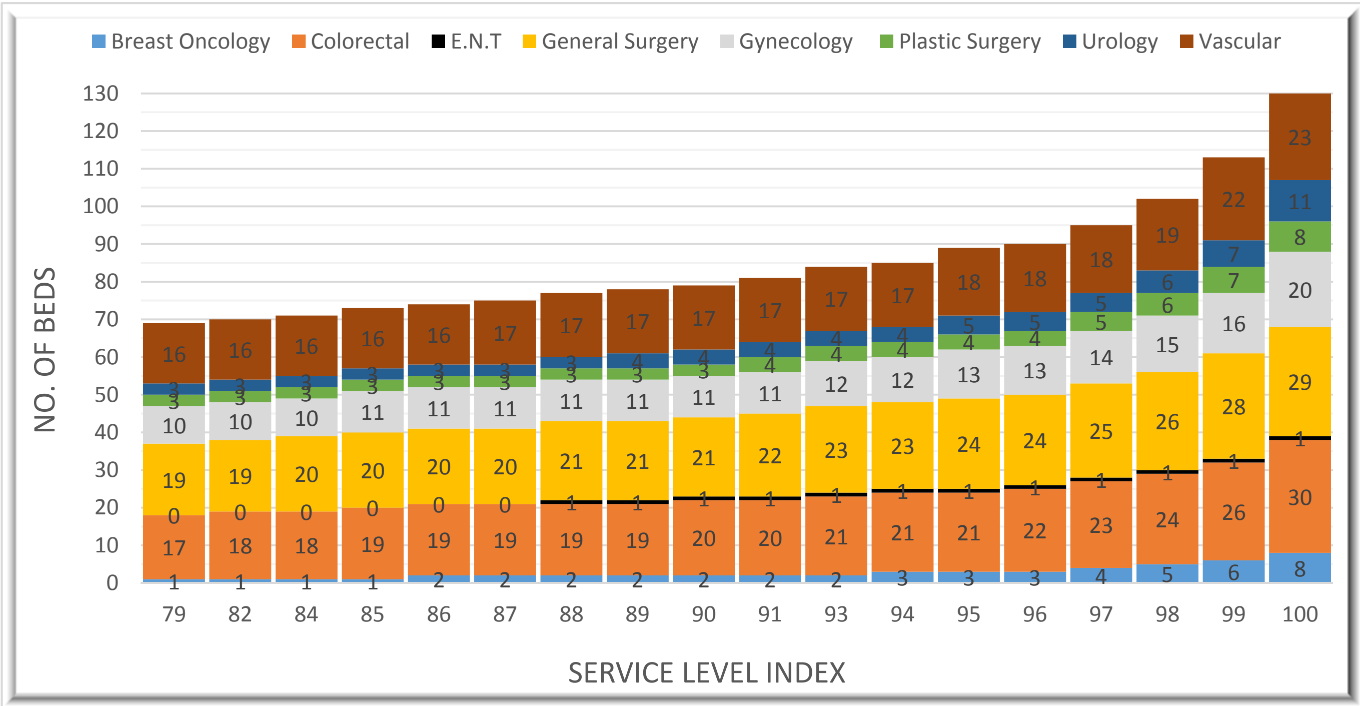


Figure 3-5: Required Beds for Each Division – Dedicated Policy

Plastic divisions require 16 beds in total, while this number is: 16 beds for Vascular, 17 beds for Colorectal, and 19 beds for General surgery divisions. For example, for E.N.T division, the surgical unit does not need any beds while keeping the SLI under 87%, since most of the patients in this specialty are discharged from JGH on the day of surgery. One should analyze the cost-effectiveness of the various procedures, to see whether or not the surgical unit resources are optimally utilized.

3.3.2. High Acuity Unit

This section will study the effect of the HAU on the surgical unit bed configuration. The HAU will be equipped with the skilled nurses as well as monitoring machines. Also, allied health professionals such as rehabilitation will be available. The patients in the HAU also have priority for diagnostic imaging and lab tests.

Hence, the JGH's surgeons believe that the high-risk patients get higher quality of care here, which also might reduce their length of stay at the surgical unit. Yet the amount of this LOS-reduction remains an unresolved matter among the project team members. In fact the LOS-reduction for each patient during her stay at the HAU is subjective and highly depended on the patients' surgical procedure, age, and some other medical pre-operational tests. To capture all potential LOS-reduction hypothesises, the SWD simulation model considers 3 LOS-reduction scenarios: 1. If the LOS-reduction is uniformly distributed in [0 , 1] days. 2. If the LOS-reduction is uniformly distributed in [0 , 2] days. 3. If the LOS-reduction is uniformly distributed in [1 , 3] days.

For each scenario the surgical unit bed configuration is evaluated under two arrival rates: 10% and 20% of the total surgical patients. Figure 3-6 illustrates the number of required surgical beds that satisfies following conditions: 1. when 20% of the surgical cases are high-risk patients, 2. these patients will save between 1 to 3 days in the HAU, and 3. the surgical beds in the main ward are organized on the basis of pooled policy. The graph covers the SLI range of 80% to 100%. For instance, if 90% SLI is depicted, the surgical unit requires 62 beds in total: 12 beds in the HAU, and 50 beds in the main ward. The total number of required beds increases to 89 beds at 100% SLI: 19 beds in the HAU, and 70 beds in the main ward. In comparison to the surgical unit bed configuration under the pooled policy, the surgical unit can save 2 beds at 90% SLI, but it needs 3 more beds to achieve 100% SLI.

Table 3-5 summarizes the SWD simulation model results for 90% SLI. Final required surgical bed depends on the HAU arrival rate, pooled/dedicated policy, and LOS-reduction interval. If JGH managers choose to open HAU and the arrival rate to this unit is 10% the surgical unit requires

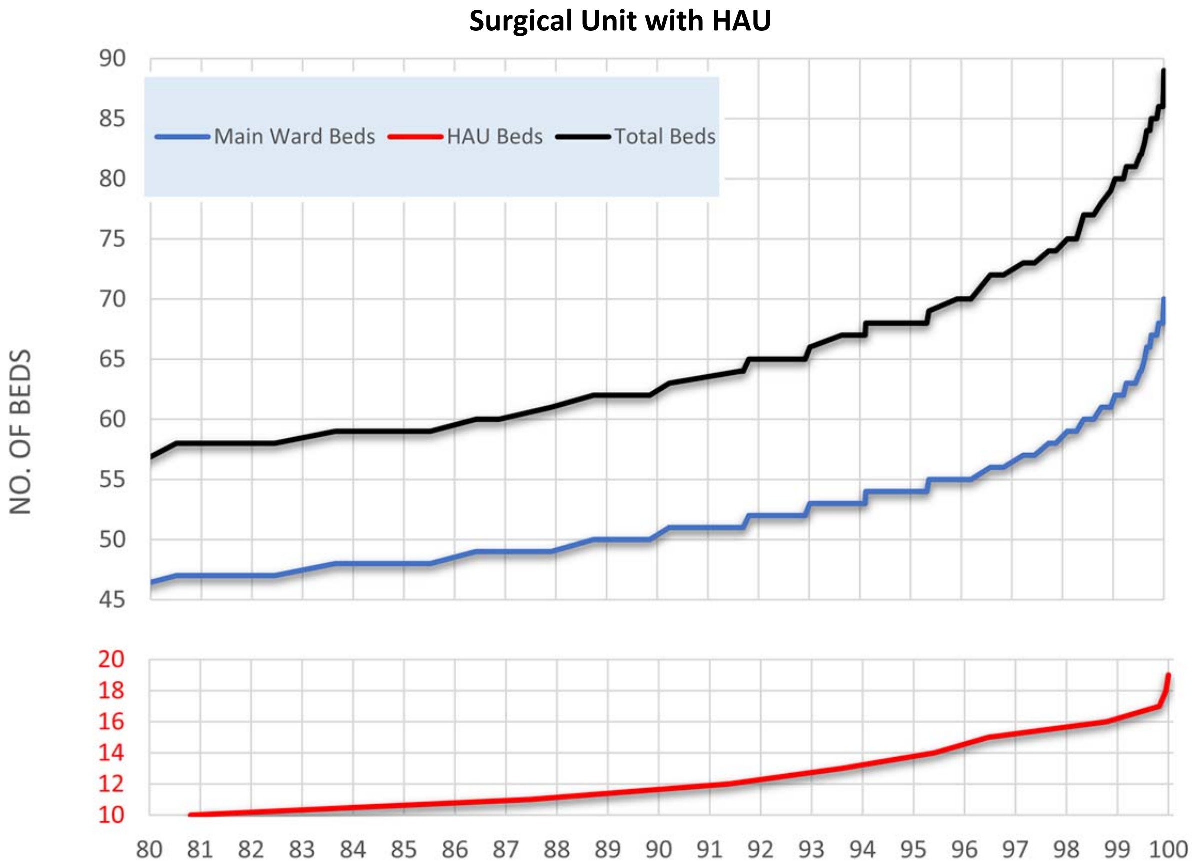


Figure 3-6: Required HAU & Main Ward Bed (Pooled Policy) – 20% Arrival Rate to HAU

between 67 to 65 beds under pooled policy, and 78 to 80 beds under dedicated policy. But if the HAU arrival rate is 20% the surgical unit requires 74 to 77 beds under dedicated policy, and 62 to 65 beds under pooled policy. Hence, given the percentage of patients going to the HAU and the configuration policy, total number of required beds is robust. Also, the surgical unit can save up to 3 HAU beds if the LOS-reduction interval increases from [0 , 1] to [1 , 3] days.

Table 3-5: Required HAU & Main Ward Bed – 90% SLI

HAU Arrival Rate	LOS-Reduction at HAU (days)	HAU Beds	Main Ward Beds		Total Surgical Unit Beds	
			Dedicated	Pooled	Dedicated + HAU	Pooled + HAU
10%	[0 , 1]	9	71	58	80	67
	[0 , 2]	8			79	66
	[1 , 3]	7			78	65
20%	[0 , 1]	15	62	50	77	65
	[0 , 2]	14			76	64
	[1 , 3]	12			74	62

3.3.3. Semi-Pooled Policy

Pooling all hospital beds is not usually feasible. Patients from different specialties might need specific post-operative care, provided by skilled nurses. In practice all nurses could not be trained to serve all types of patients. Moreover, part of surgical beds must be equipped with certain standard equipment (e.g. monitoring or respiratory devices) to serve particular patients. It is not cost-efficient to equip all beds with these kind of machines. In fact, although pooled policy gives us the minimum number of required of beds to serve all patients, regular staffed beds are not suitable to serve all patients. Yet JGH managers and surgeons do believe that certain divisions provide similar types of post-operative care to their patients. Hence, it is practical to group those divisions to share their surgical beds. To this aim, the SWD simulation model is run considering these three shared units: 1. Colorectal and general, 2. gynecology and urology, and 3. Vascular and E.N.T. The other two units work independently. It is assumed that there is no HAU under the semi-pooled policy. Table 3-6 depicts the number of required surgical beds at 90% and 100% SLI levels.

Table 3-6: Semi-Pooled Policy

SERVICE LEVEL INDEX	Shared Units			Dedicated Units		Total Surgical Unit Beds
	Colorectal & General	Gynecology & Urology	Vascular & E.N.T.	Plastic	Breast Oncology	
90%	38	13	17	4	2	74
100%	52	23	23	8	8	114

In comparison with the dedicated policy at 90% SLI, by grouping colorectal and general divisions 2 beds are saved. Grouping gynecology and urology divisions also saves 2 beds, and the third group of divisions saves only one bed. In total, under semi-pooled policy JGH surgical unit requires 74 beds at 90% SLI, while this number is 64 and 79 under pooled and dedicated policy respectively. However, the same amount of beds is required under the semi-pooled policy and dedicated policy when 20% of the patients would be admitted to the HAU.

3.4. Conclusion

This chapter has studied the surgical unit beds configuration problem. The main objective of the developed SWD simulation model was to provide a framework to evaluate the effect of various strategies in such a problem in order to make recommendations on how to utilize surgical beds efficiently. Furthermore, later on this study will use the SWD simulation model to validate the result of the optimization model. Different policies have been compared to answer all the "what if" questions raised by JGH's managers. The main outcomes are classified as follows:

1. The minimum number of required beds is 62, which occurs under pooled policy while 20% of the patients would be transferred to the HAU (assuming that the LOS-reduction is uniformly distributed in [1 , 3] days).
2. Opening the HAU does not typically increase the total number of required beds. The first underlying reason is the LOS-reduction in this unit. Second, the HAU beds are shared among all specialties which increase the HAU utilization. However, the lower bed-to-nurse ratio is expected from nurse staffing perspective.
3. Given the configuration policy (i.e. pooled or dedicated) total number of required beds is in a tight bound at the 90% service level index. In other words, the effect of uncertain parameters (i.e. the portion of high-risk patients and the potential LOS-reduction in the HAU) on the total number of required beds is small.

4. Total required surgical unit beds decrease when the LOS-reduction in the HAU increases. This reduction is more visible when a higher percentage (i.e. 20%) of the surgical patients would be admitted in the HAU.
5. The relation between SLI and number of required beds is more or less linear from 60% to 90% SLI. However, this behavior is roughly exponential form 90% SLI to 100% SLI.
6. The most important determinant of the bed - reduction is pooled policy, which significantly increases the surgical unit bed utilization. The JGH is able to save up to 19% on the total number of required beds under pooled policy in comparison with the dedicated policy.
7. JGH saves up to 5 beds in total under the semi-pooled policy in comparison with the dedicated policy at 90% SLI.

Chapter 4

Integrated Surgical Case-Mix Management

4.1. Introduction, Motivation, and Literature Review

The surgical ward is composed of several divisions. Each division is a formally organized unit providing specialized practices of the hospital's medical staff (e.g. colorectal, gynecology, orthopedic, etc). Tertiary hospitals often perform a wide range of these services. The types and quantities of surgical procedures to be performed in the hospital is the subject of the Surgical Case-Mix problem. This chapter aims to develop an integrated CMP model as a decision-making tool, which provides an opportunity for stakeholders (i.e. policy makers, hospital administrators, and surgeons) to buy-in to the developed solutions.

Surgical unit decisions are hierarchically classified into strategic, tactical, and operational decisions. The case-mix decision is a complex strategic problem that depends on several factors: the reimbursement policies enforced by the payer (i.e., the government or insurer), hospital resources (e.g. surgeons, operating rooms, equipment, and staffed beds), and the hospital's catchment area (as an indicator for patient demand) (May et al., 2011).

In the early 1970s, researchers started to study the CMP by developing simple mathematical models (Guerriero & Guido, 2011). However, the literature in this area is relatively scarce. Parsons et al. introduced the CMP as a systematic approach to improve quality and control costs (Parsons et al., 1992). Robbins developed one of the very first CMP models, which allocates the hospital resources to the diagnosis related groups (DRGs[5] - a standard practice to reimburse hospitals by classifying the procedures with the same level of the expected usage of resources) to maximize the hospital's benefit (Robbins W, 1989). Blake & Carter developed a goal programming approach to reset the type and volume of the surgeon's performed procedures (Blake & Carter, 2002). The authors assumed that the surgeons are profit satisfiers rather than profit maximizers, and there is complete cooperation between the hospital administrator and surgeons on the hospital's case-mix decision. It was not mentioned, however, why the hospital administrator should cooperate with the surgeons. The model minimizes the total weighted penalties when the surgeons' desired revenues and the hospitals' expected level of costs are not satisfied. The authors extended this model in (Blake & Carter, 2003) by comparing a set of funding policies to the hospitals and surgeons. Global budget and rate-based funding were proposed as the hospital's reimbursement method, while surgeons are funded under either a fee-for-service or fixed salary structure. Despite the fact that the proposed fixed salary structure for the surgeons might not be pragmatic, it was concluded that the combination of the global budget policy and salaried surgeons' method increases the risk of under-servicing compared to the fee-for-services policy. Since the models were presented under the assumption of budget cuts as well as complete cooperation between surgeons and administrators, the obtained solutions are rather questionable and cannot be generalized.

At the tactical level, manuscripts target the Master Surgical Scheduling Problem (MSS), which refers to distributing the operating room (OR) among various surgeons. They try to improve the surgical unit performance (e.g. measured by utilization, equity, staff costs, bed leveling), with

[5] Analogues to Canadian Case-Mix Groups's (CMG)

respect to several constraints: minimum/ maximum OR time to be allotted to each specialty (this is the CMP output as an input to the MSS), surgeon's minimum/ maximum OR time, surgeon's priority/ flexibility of the dates/times, required specialized equipment in the OR, and so on (Guerriero & Guido, 2011).

At the operational level, the dedicated OR date and time of each operation are the center of attention by authors. They divide this problem into two categories: the first is known as Advanced Scheduling, which determines the dedicated OR and date. The second category, Allocation Scheduling, is primarily concerned with the patient sequence on the day of surgery (Cardoen et al., 2010; Lee & Yih, 2014). The Advanced and Allocation Scheduling problem's (AAS) objective copes with OR utilization, OR idle/overtime cost, patient waiting time, Post-Operative bed level, and so on. The limitations with such a model include resource capacity (e.g. OR, post-operative beds, nurses), patient's priority based on the waiting list, patient health condition (i.e. emergency and urgent), and so on (Cardoen et al., 2010). A schematic mechanism on these essential decisions is illustrated in figure 4-1.

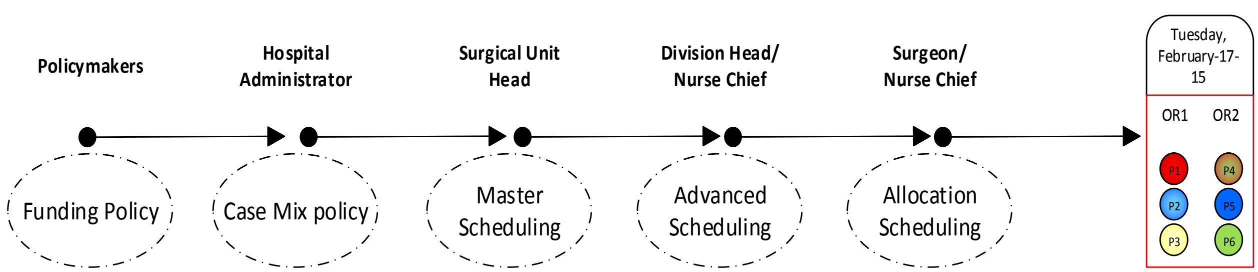


Figure 4-1: Essential Decisions that Affect the Final OR schedule

Various objective functions are studied in the structure of Master Surgical Scheduling and Advanced Scheduling models: minimizing OR idle/overtime, patients waiting time, and the expected surgical case cancellation; (Beliën & Demeulemeester, 2007; Creemers et al., 2012;

Hans et al., 2008; Testi et al., 2007) some other objectives are also highlighted in the literature. Day et al. developed a 3 stage OR block scheduling model in which a combination of the surgeons' profit and the hospital revenue was maximized simultaneously (Day et al., 2012).

In addition to the precise objective functions, the stochastic models are required to adequately capture the dynamics of the OR planning process. Denton et al. studied OR planning when the block duration is stochastic (B. T. Denton et al., 2010). They described the trade-off between opening a new OR, which brings a fixed set-up cost, and continuing to run the current open ORs, which impose overtime cost to the model. To calibrate the models with rational OR costs, Olivares et al. assumed that administrative decisions implicitly reflect the balance between the costs of OR idle time and overtime (Olivares et al., 2008). So, the authors ascertained the unobservable cost function behind the administrative decision.

Beyond the importance of the objective functions and the stochastic environment, one has to capture the combination of the required resources as a system. An emerging research effort is to integrate the availability of downstream resources (e.g. PACU, ICU, and ward's bed) into the Master Surgical Scheduling and Advanced Scheduling problems. Chow et al. focused on OR efficiency as well as downstream bed utilization to schedule the surgeon blocks and patient types (Chow et al., 2011). They developed a mixed integer model on the basis of the patients' average LOS to schedule the OR blocks. Then a simulation model was applied to evaluate the obtained OR block schedule by using the historical LOS. The simulation model calculates the peak of bed occupancy for a given OR block schedule and provides suggestions to improve it. This cycle repeats till the target level for the peak of bed occupancy is achieved. However, there is no guarantee that the optimal OR block schedule is obtained since the stochastic LOS is not part of the optimization model.

Downstream bed impact is not limited to the main ward beds. Price et al. studied the effect of surgical scheduling on the PACU's bed occupancy level (Price et al., 2011). When a new patient arrives into a fully occupied PACU/ICU, surgeons/nurses have to prematurely discharge the current patients due to downstream bed unavailability. Dobson et al. studied the effect of a specific ICU's patient bumping policy on the ICU performance (Dobson et al., 2010). However, the assumed discharging policy was challenged by Chan (Chan et al., 2011). Their model found the optimal premature discharge policy which causes the lowest mortality rate and readmission rate.

4.2. ISCM Model Approach

As illustrated in figure 4-1, the results of the CMP would be the input for the MSS problem, and the output of the MSS directly affects the AAS problem's outcomes. Despite the hierarchal classification of decisions in the surgical unit, it should be noted that, without a systematic approach to the surgical unit planning the target performance could not be achieved. In many studies the tactical and operational levels are combined in a single model or presented as hierarchical stages of the OR planning process (Chow et al., 2011; Day et al., 2012). However, the incorporation of the strategic decisions with the tactical/operational decisions has yet to be studied.

To tackle this gap, this study draws attention to an integrated approach to the case-mix problem that results in practical outcomes even from the tactical and operational perspectives. In other words, a deterministic Integrated Surgical Case-Mix (ISCM) model is developed that copes with the functionality of the derived case-mix in the following Master Surgical Scheduling and Advanced Scheduling problems. The proposed ISCM model presents a three dimensional objective function to tackle the main concerns faced by the hospital administrator: the reimbursement mechanism, the ORs' utilization, and downstream bed impact.

Since the patient LOS at the surgical unit is stochastic, the deterministic model is extended to incorporate this uncertainty. The stochastic ISCM model limits the probability of downstream bed shortage through a chance-constrained programming approach. To the best of our knowledge, for the first time in the literature, this study develops a chance-constrained programming technique for optimizing the CMP model. Note that Shylo et al. used a chance-constrained approach, based on normal approximation for surgery duration, to the OR scheduling problem (Shylo, Prokopyev, & Schaefer, 2012). Deng et al. also tackled the Advanced and Allocation Scheduling problem using a chance-constrained approach, based on an empirical probability function for surgery duration (Deng, Shen, & Denton, 2014).

In the ISCM model, the Conditional Value at Risk (CVaR) approach is also applied to control the risk of high bed shortage cost. From the modeling perspective, the interaction of chance-constraints and the proposed CVaR approach is investigated to find the impact of each one separately on the downstream bed shortage thresholds. This may enable us to reduce the size of the ISCM model for further extensions.

Furthermore, the Sample Average Approximation (SAA) technique is applied to solve the stochastic model. To this end, a linear approximation of the ISCM model is presented. The rest of the chapter is organized as follows. The next section provides a precise presentation on the mathematical model of the CMP. Section 4.4 presents the SAA solution algorithm, and the last section presents the results of the stochastic model calibrated with the full scale data. All the examples in this chapter are based on the data collected at Montreal Jewish General Hospital.

4.3. ISCM Model Definition

In this section, the Integrated Surgical Case-Mix model (ISCM) is described. A feasible set of surgical procedures, which a surgeon can schedule within an OR block is denoted as *Block Mix* (e.g. an orthopaedic surgeon can do either *two Knee Arthroplasty, Block Mix* 1 in figure 4-2, or one *Knee Arthroplasty* and one *Hip Arthroplasty* procedures, *Block Mix* 2 in figure 4-2, in an OR block).

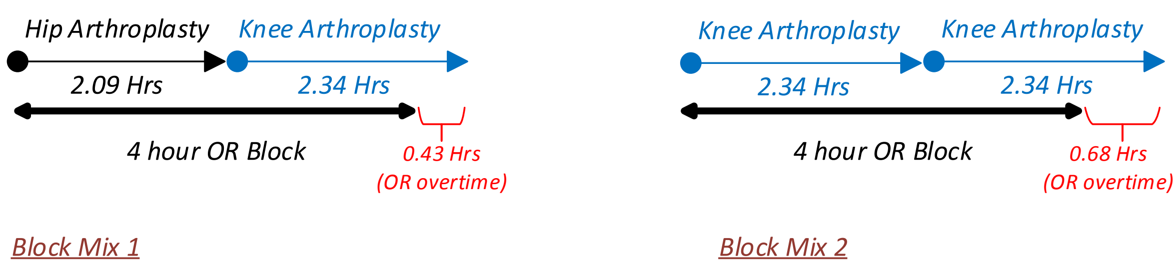


Figure 4-2: Block Mix Examples

The Block Mix is built on the basis of the average required OR time for each procedure. It is worthy to note that the *Block Mix* is not the sequence of the patients. The patients will be scheduled to the OR with respect to their surgeon's *Block Mix* and their priority in the waiting list. Also, each *Block Mix* represents the combination of the procedures while the sequence of them is not addressed in this study. Let M denote the set of all possible *Block Mixes* for all surgeons. In accordance with the current practice, it is assumed that the hospital has access to a set of S surgeons to serve a set of K procedures. Let the integer decision variable x_{smt} denote the number of OR blocks dedicated to surgeon s at day t with respect to *Block Mix* m, and the parameter N_{km} indicates the number of procedure k's cases scheduled at each *Block Mix* m. It is assumed that the surgeons do not share the same OR block and at most two OR blocks could be assigned to a surgeon each day, so $x_{smt} \in \{0,1,2\}$.

$$X_{smt} = \begin{cases} 2 & \text{if two OR blocks are assigned to surgeon s with Block Mix m in day t} \\ 1 & \text{if one OR block is assigned to surgeon s with Block Mix m in day t} \\ 0 & otherwise \end{cases}$$

4.3.1. Objective Function

The ISCM's objective function refers to the main concerns faced by the hospital administrator. The first component targets the surgical unit reimbursement. The second term represents OR utilization and the last element tackles the downstream bed impact of the case-mix decision.

4.3.1.1. Financial Component

The hospital payment mechanism is the subject of significant debates and changes in countries with publicly funded healthcare systems (Mayes, 2007). Activity Based Funding (ABF) method is one of the most common alternate methods to be contemplated by the governments of the industrialized countries (Pink, Information, McKillop, & Johnson, 2001). Governments incentivize hospitals to optimally utilize their resources by imposing ABF policy (J. Sutherland, Crump, Repin, & Hellsten, 2013).

Under the ABF approach the hospital gets reimbursed on the basis of the type and volume of cases. It is assumed that hospitals are completely flexible on the type and number of procedures they practice each year. Based on the ABF policy, surgical patients with a similar cost of hospitalization are identified under the same diagnosis-related-group (DRG). Each DRG is associated with a fixed reimbursement amount. However, this amount may not reflect costs in small hospitals (i.e. hospitals with huge fixed costs) or specialized hospitals (i.e. hospitals which serve complex surgical cases with high tech equipment). Yet, our main focus is on the tertiary hospitals. Let $\boldsymbol{R_k}$ be the reimbursement amount for procedure $\boldsymbol{k}$, so formula (4.1) shows the surgical unit reimbursement.

$$Reimbursement(ABF) = \sum_{k=1}^{K} R_k * [\sum_{t=1}^{T} \sum_{m=1}^{M} \sum_{s=1}^{S} N_{km} * x_{smt}] \quad (4.1)$$

The basic ISCM model considers ABF policy in the objective function. This approach is the basis for all of the analysis in this chapter. However, in chapter 5 other funding policies (i.e. Global Budget with Incentive, and Global Budget with Incentive and Penalty) are studied to help policymakers to discover how these policies affect hospitals' responses to better assess the trade-offs among them. (i) Under Global Budget with Incentive (GBI), this study assumes that the hospital receives a fixed budget and has to complete a certain number of surgical cases for each procedure, yet the government will reimburse any extra cases. (ii) Under Global Budget with Incentive and Penalty (GBIP), a base volume for each procedure is recommended to the managers, and the government can impose financial penalties on the hospital if it is not committed to this threshold.

4.3.1.2. Utilization Component

Operating Room (OR) utilization is the second component of the ISCM's objective function. ORs are the most important resource at the surgical unit, and their utilization has an enormous effect on surgical unit performance. This study assumes that the OR duration is deterministic since the JGH managers and surgeons believed that the OR duration for a specific procedure is more or less predictable before the surgery. The JGH provided us with the expected OR duration for each procedure. It is assumed that a surgeon's *Block Mix* is typically defined as a regular 4 hour OR block. Each *Block Mix* represents the possible combination of surgical procedures in a 4 hour time slot. However, for us, the length of an OR block is flexible; in such a case, the hospital would face idle/overtime OR costs. So, the model tries to minimize this cost by selecting those *Block Mixes* that are close to 4 hours. Besides the OR idle/Overtime costs, each *Block Mix* imposes a fixed cost to the hospital (i.e. opening the OR, required resources that depend on the procedure type, and so on). Let M_s denote set of feasible Block Mixes for surgeon s and IO_m denote the expected cost of the OR utilization for the *Block Mix* m. The objective function considers the OR's expected costs as:

$$\sum_{s=1}^{S} \sum_{m \in M_s} \sum_{t=1}^{T} [IO_m * x_{smt}] \tag{4.2}$$

4.3.1.3. Downstream Bed Impact Component

Utilizing surgical unit resources such as operating rooms, regardless of their integration with the availability of other resources like surgical unit beds, results in cancellation and rescheduling of new elective patients or the premature discharge of current patients. Premature discharge refers to patients who require more recovery time in the surgical ward, but are discharged to make room for new patients (i.e. as a temporary solution to stop the cancellation of surgical cases). Chan et al. (Chan et al., 2011) evaluated various premature discharge policies on the basis of the mortality risk and readmission load. Eapen et al. also showed that the patients with longer LOS have significantly lower readmission rates (Eapen et al., 2013). Hence, improving surgical unit scheduling directly affects the quality of surgical service (Lucas & Pawlik, 2014).

The ISCM model tries to minimize the expected bed shortage costs, which affects the reduction of surgical case cancellations as well as premature discharges, which increase the hospital care quality and patient satisfaction. To this goal, the ISCM model accounts for the expected amount of bed shortage during the planning horizon to minimize the risk of high down stream bed shortage. This approach is similar to the Conditional Value-at-Risk (CVaR) framework, presented by Rockafellar & Stanislav (Rockafellar & Uryasev, 2000) to minimize the financial risk to portfolios. This approach will be explained later on this chapter under the stochastic ISCM model.

4.3.2. Deterministic ISCM Model

This section describes the problem settings, notations, and formulations of the deterministic Integrated Surgical Case-Mix model in which the patients' LOS is assumed to be constant. Since the LOS's stochasticity is the main source of bed shortage and case cancelation (i.e. total number of required beds is constant if the LOS is deterministic), this assumption results in an optimal case-mix which provides no downstream bed shortage. Hence, in the deterministic ISCM, the third component of the objective function, which copes with the downstream bed impact, is excluded. So, the objective function strives to maximize hospital reimbursements and minimize the OR idle/overtime costs.

Table 4-1: Set: the ISCM Model

T	:	Planning horizon	S	:	Surgeons
K	:	Procedure types	D	:	Divisions
M	:	Surgeons surgical mix	X	:	Surgeons availability/preference

Table 4-2: Notation: the ISCM Model

OR_t	:	Total available OR blocks at day t
OR_d	:	Minimum number of OR blocks reserved for division d

OR_s : Minimum number of OR blocks reserved for surgeon s

N_{km} : Number of procedure k scheduled in $Block\ Mix\ m$

B : Total number of medical beds in the ward

L_m : Maximum LOS among the procedures in $Block\ Mix\ m$

R_k : Reimbursement for procedure type k

M_s : Set of feasible $Block\ Mix$es for surgeon s

C_k : Base level for procedure k (imposed by the government)

IO_m : Expected cost of the OR utilization for $Block\ Mix\ m$

SC : Bed shortage cost at each day

η : Acceptable probability of bed shortage

V_k : Maximum number of possible procedure type k

a_{smt} : Required number of bed t days after the date that $Block\ Mix\ m$ is scheduled for surgeon s

Let a_{smt} denote the total number of required beds t days after *Block Mix m* is performed by surgeon *s*. Since the LOS for *Block Mix m'*s procedures is assumed to be deterministic, a_{smt} is also deterministic. Also, L_m denotes the latest day that at least one downstream bed is required by the procedures in *Block Mix m.* So, constraint 4.4 limits the total number of occupied beds in a day with respect to the total number of surgical ward's beds, denoted by *B.* As a resource limitation constraint, 4.5 keeps the total number of scheduled OR blocks fewer than the total available OR blocks at each day. Constraint 4.6 guarantees a minimum number of OR blocks over the planning horizon for each division. This limitation is imposed to the model, since hospital administrators might desire to keep a certain level of service for some divisions regardless of their financial impact on the objective function. This happens when the admins tend to gradually reset the current case-mix because of the managerial discretion, surgeons' contracts, and so on. Constraint 4.7 guarantees a minimum number of OR Blocks for each surgeon. In constraint 4.8, an upper bound on the total number of dedicated OR Blocks to each procedure is imposed. This upper bound depends on the hospital's catchment area and its population size (i.e. with respect to the maximum possible demand for each procedure). So, the model has to schedule at most V_k procedures of type k within the planning horizon. The last constraint considers surgeons' availability over the planning period. For example, most surgeons prefer not to work during the weekends or they might have to teach on the other weekdays. Tables 4-1 and 4-2 summarize the notations used in the deterministic ISCM model, which is also common to the stochastic version that will be introduced later.

$$\boldsymbol{Max}: \quad \sum_{k=1}^{K} R_k * [\sum_{t=1}^{T} \sum_{m=1}^{M} \sum_{s=1}^{S} N_{km} * x_{smt}] - \sum_{s=1}^{S} \sum_{m \in M_s} \sum_{t=1}^{T} [IO_m * x_{smt}] \quad (4.3)$$

S.t.

$$\sum_{s=1}^{S} \sum_{m \in M_s} \sum_{i=t-L_m}^{t} a_{sm(t-i+1)} x_{smi} \leq B \qquad \forall\, t \in T \qquad (4.4)$$

$$\sum_{s=1}^{S} \sum_{m \in M_s} x_{smt} \leq OR_t \qquad \forall\, t \in T \qquad (4.5)$$

$$\sum_{t=1}^{T} \sum_{s \in d} \sum_{m \in M_s} x_{smt} \geq OR_d \qquad \forall\, d \in D \qquad (4.6)$$

$$\sum_{t=1}^{T} \sum_{m \in M_s} x_{smt} \geq OR_s \qquad \forall\, s \in S \qquad (4.7)$$

$$\sum_{s=1}^{S} \sum_{m \in M_s} \sum_{t=1}^{T} [N_{km} * x_{smt}] \leq V_k \qquad \forall\, k \in K \qquad (4.8)$$

$$x_{smt} \in X \qquad (4.9)$$

4.3.3. Stochastic ISCM Model

To capture the intrinsic uncertainty associated with the LOS, the stochastic ISCM is developed. The goal of this model is to guarantee that the optimal case-mix would not impose a dire bed shortage on the system. A set of chance-constraints is proposed on the expected occupied beds each day. Let η denote a specified maximum probability of bed shortage over the planning horizon, $0 \leq \eta \leq 1$. Hence, the probabilistic counterpart of the (4.4) is modeled as the following joint chance-constraint:

$$P\{Y_t \geq \quad B \qquad \forall\, t \in T\} \leq \eta \tag{4.10}$$

Where P means probability. This constraint guarantees that the total occupied beds would not exceed the total number of available beds with the probability of $(1-\eta)$ through the planning horizon. In other words, the bed capacity constraint may be violated for at most η percent of the time. The chance-constraint is nonlinear and its deterministic approximation is presented in section 4.2.4.

However, the resulting case-mix may impose a high level of bed shortage, $[Y_t - B]^+$, for η percent of the time. This issue is important to hospital administrators, because surplus patients might be transferred to the other units (or other hospitals) rather than the surgical ward. So, if the gap between the number of required beds and the available beds is small, the risk of case cancellation would be negligible. With this goal in mind, the stochastic ISCM model benefits from the Conditional Value-at-Risk approach to minimize the risk of a high level of bed shortage.

Let Y be the total number of occupied beds with cumulative distribution function $\mathrm{F_Y(B)} = \mathrm{P\{Y \leq B\}}$, Sarykalin (Sarykalin, 2014) showed that the Value at Risk (VaR) is,

$$\mathrm{VaR_{1-\eta}(Y) = Min\,\{B|F_Y(B) \geq (1-\eta)\}} \tag{4.11}$$

And for a general distribution of Y, Sarykalin (Sarykalin, 2014) defined the CVaR as,

$$\mathrm{CVaR_{1-\eta}(Y) = E\big[Y|Y > VaR_{1-\eta}(Y)\big]} \tag{4.12}$$

Since the ISCM model tries to schedule the surgical cases using B beds and the chance constraint limits the probability of bed shortage to η, so the VaR in the ISCM model is B, and the $\mathrm{CVaR}_{1-\eta}(\mathrm{Y}) = \mathrm{E}[\mathrm{Y}|\mathrm{Y} > \mathrm{B}] = \mathrm{B} + \mathrm{E}[(\mathrm{Y}-\mathrm{B})|\mathrm{Y} > \mathrm{B}]$.

Let $\mathrm{E}[\mathrm{Y}-\mathrm{B}]^+$ show $\mathrm{E}[(\mathrm{Y}-\mathrm{B})|\mathrm{Y} > \mathrm{B}]$, so $\mathrm{CVaR}_{1-\eta}(\mathrm{Y}) - \mathrm{B} = \mathrm{E}[\mathrm{Y}-\mathrm{B}]^+$; and the term $\mathrm{SC} * \sum_{t=1}^{T} \mathrm{E}[\mathrm{Y}_t - \mathrm{B}]^+$ shows the expected bed shortage costs, where SC denotes the cost associated with one day bed shortage, Y_t is a stochastic variable referring to the total number of required beds on day t. This term of the ISCM's objective function is called CVaR term hereafter.

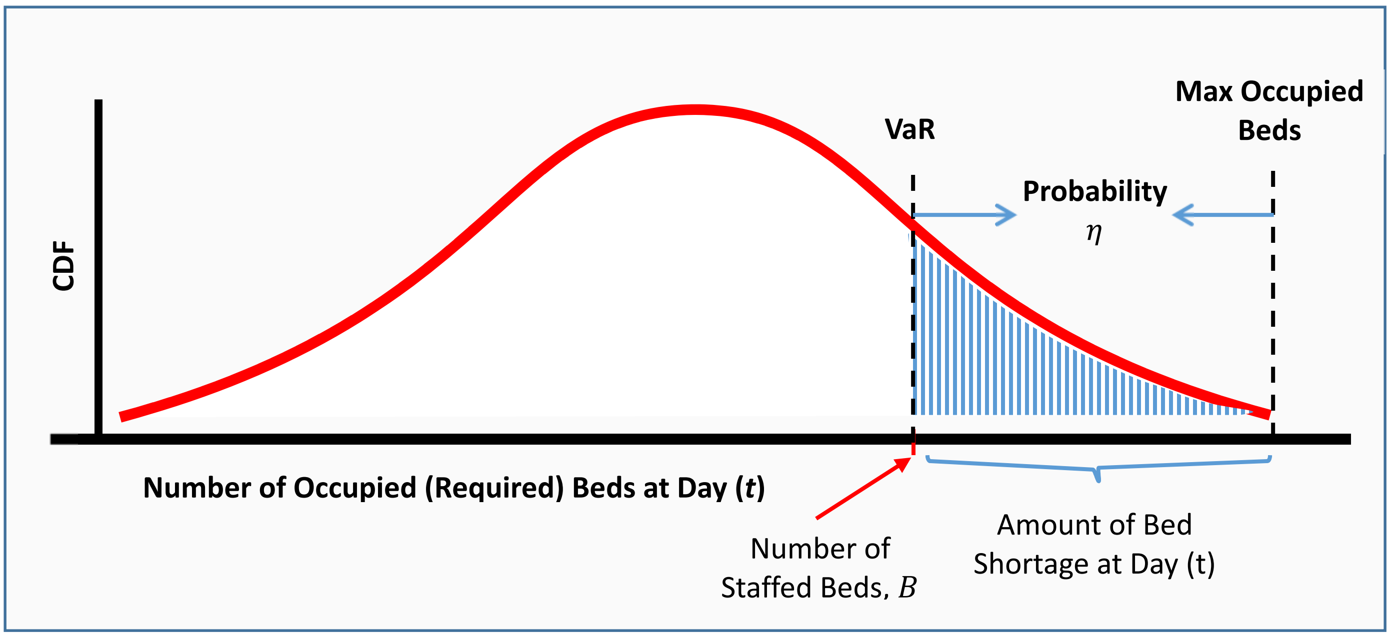


Figure 4-3: A Schematic Bed Shortage Distribution

To the best of our knowledge, the CVaR approach has not been applied at any stage of the OR planning process. A detailed definition of Y_t, in terms of x_{smt}, and the CVaR approach are presented in section 4.3.4. Figure 4-3 illustrates a schematic distribution function of bed shortage each day.

So, by considering the chance constraint and the CVaR term, the model simultaneously controls incident of bed shortage and minimizes the cost of expected bed shortage over the planning horizon. The interaction of these components is investigated later on this chapter.

The other constraints in the stochastic model are the same as the ones described in the deterministic model. So, the stochastic ISM model is,

$$\boldsymbol{Max} \quad \sum_{k=1}^{K} R_k * [\sum_{t=1}^{T}\sum_{m=1}^{M} N_{km} * x_{smt}] - \sum_{s=1}^{S}\sum_{m\in M_s}\sum_{t=1}^{T}[IO_m * x_{smt}] - SC * \sum_{t=1}^{T} E[Y_t - B]^{+} \quad (4.13)$$

S.t.

(4.5) – (4.6) – (4.7) – (4.8) – (4.9) – (4.10)

4.3.4. Approximation Model

The proposed stochastic model is not linear because of the chance-constraints and the last term of the objective function. To reformulate the model, one has to ascertain the LOS distribution function. Marazzi et al. developed a statistical analysis to find an accurate distribution function that describes the patients' LOS. They evaluated the Lognormal, Weibull, and Gamma distribution functions on a large database and concluded that the Lognormal distribution could accurately fit with the LOS distribution in the majority of the samples (Marazzi, Paccaud, Ruffieux, & Beguin, 1998).

However, as noted in section 3.2.2, several distribution functions on the JGH database are evaluated; yet, none of them could accurately fit the data. Furthermore, even if such an accurate LOS distribution function did exist for each procedure, it would be too complex to consider all these distributions in the ISCM model. Hence, the historical data is used rather than any predetermined distribution function to describe the patients' LOS.

The primary analysis of the JGH data illustrates that more than 90% of the patients' LOSs for each procedure is in the range of 3 days. As table 4-3 displays, LOS for 94% of the patients who underwent a Thyroidectomy procedure in 2013 was 1, 2, or 3 days. However, for some procedure this rage is wider (i.e. 4 days or more), so the top LOSs are selected, under each procedure, for at least 90% of the patients.

Therefore, it is assumed that the LOS for each procedure follows a discrete distribution function; and, a truncate rule is applied to find the LOS distribution based on the relative frequencies for the most common LOSs per procedure. Later on this study, the effect of this assumption on the model's outcomes will be evaluated. The LOS for a *K-type* patient is assumed to fall in $[L_{min}$, $L_{max}]$. Let $P_k(t)$ denote the probability that a *k*-type patient is discharged home exactly *t* days after the surgery. So, $\sum_{z=t+1}^{L_{max}} P_k(z)$ is the probability that a *K-type* patient stays in the hospital more than *t* days. Let $F_m(b,t)$ denote the probability that all of the *Block Mix m's* procedures require *b* beds on day *t* (i.e. *t* days after the scheduling date of the *Block Mix m*). Assume that the

Table 4-3: Sample Primary Analysis Results on the Patients' LOS

Division	**Procedure**	**LOS Distribution**										**Top 3 LOSs**
		No. of Days in the Hospital (t)	0	1	2	3	4	5	6	7	8 <=	
E.N.T.	Thyroidectomy total	No. of Patients	0	58	20	9	1	3	1	1	0	87
		Probability	0.00	0.62	0.22	0.10	0.01	0.03	0.01	0.01	0.00	0.94
Gynecology	Hysterectomy total robotic	No. of Patients	2	67	15	0	1	0	1	0	0	82
		Probability	0.02	0.78	0.17	0.00	0.01	0.00	0.01	0.00	0.00	0.98

Block Mix m contains a sequence of Q procedures, and let $W_b = \{w^1, .., w^n\}$ be the set of all $n = \binom{Q}{b}$ combinations of its procedures. Hence,

$$F_m(b,t) = \sum_{w \in W_b} \left[\left(\prod_{k \in w} \left[\sum_{z=t+1}^{L_{max}} P_k(z) \right] \right) \left(\prod_{k \in M \setminus w} \left[\sum_{z=0}^{t} P_k(z) \right] \right) \right] \quad \forall\ m \in M,\ t\ \in T \qquad (4.14)$$

Set $h_m^\zeta = \{h_{m0}^\zeta, \dots, h_{mL_{max}}^\zeta\}$ denotes a realization of the downstream bed occupation resulting the *Block Mix m* (e.g. h_{m0}^ζ is the number of required beds for *Block Mix m* at the day of surgery for ζth realization). Let Z be the finite set of all potential realizations, so, the probability of h_m^ζ is,

$$P_m^\zeta = F_m\left(h_{m0}^\zeta, 0\right) * \prod_{t=1}^{L_{max}} \left[F_m\left(h_{mt}^\zeta, t\right) | h_{m(t-1)}^\zeta \right] \qquad \forall\ \zeta \in \mathrm{Z},\ \mathrm{m}\ \in M \qquad (4.15)$$

Now, the expected total number of occupied beds at day t is defined as follows:

$$Y_t = \sum_{\zeta \in \mathrm{Z}} \sum_{s \in S} \sum_{m \in M_s} P_m^\zeta \sum_{i=t-L_{max}}^{t} h_{m(t-i+1)}^\zeta x_{smi} \qquad \forall\ t\ \in T \qquad (4.16)$$

Let E_t^ζ be a binary variable that equals one iff $Y_t^\zeta > B$. Hence, the deterministic approximation of the joint chance-constraint (4.10), $P\{Y_t \geq\ B \quad \forall\, t \in T\} \leq\ \eta$, is as follows:

$$\sum_{s \in S} \sum_{m \in M_s} \sum_{i=t-L_{max}}^{t} h_{m(t-i+1)}^\zeta x_{smi} \leq B + BigM * E_t^\zeta \qquad \forall\ \zeta \in \mathrm{Z},\ \forall\ t\ \in T \qquad (4.17)$$

$$\sum_{\zeta \in \mathrm{Z}} \left(\prod_{s \in S} \prod_{m \in M_s} P_m^\zeta \right) \sum_{t \in T} E_t^\zeta \leq\ \eta * |T| \qquad (4.18)$$

And the deterministic approximation of the CVaR term in the objective function, $\sum_{t=1}^{T} E[Y_t - B]^+$, is:

$$SC * \sum_{\zeta \in \mathrm{Z}} \sum_{t=1}^{T} \left(\prod_{s \in S} \prod_{m \in M_s} P_m^\zeta \right) * \left[\sum_{s \in S} \sum_{m \in M_s} \sum_{i=t-L_{max}}^{t} h_{m(t-i+1)}^\zeta x_{smi}\ - B \right]^+ \qquad (4.19)$$

To linearize equation (4.19), it is assumed that $\sum_{s\in S}\sum_{m\in M_s}\sum_{i=t-L_{max}}^{t} h_{m(t-i+1)}^{\zeta} x_{smi} - B = U_t^{\zeta^+} - U_t^{\zeta^-}$, while $U_t^{\zeta^+}$and $U_t^{\zeta^-}$ are non-negative variables that at most one of them can be greater than zero for scenario ζ at time t. Hence, the linearized CVaR term in the objective function is:

$$SC * \sum_{\zeta\in Z}\sum_{t=1}^{T}\left(\prod_{s\in S}\prod_{m\in M_s} P_m^{\zeta}\right) * U_t^{\zeta^+} \tag{4.20}$$

4.4. Solution Procedure

Although the deterministic approximation of the stochastic ISCM is available, size of Z increases exponentially with the dimension of the divisions, surgeons, and procedures; as a result, the expected value of (4.18) and (4.20) for a given decision X is intractable. To overcome this problem, a scenario-based method is used to optimize the model. The Sample Average Approximation (SAA) method is applied that was first presented by Verweij et al. (Verweij, Ahmed, Kleywegt, Nemhauser, & Shapiro, 2003). A set of W independent samples is randomly selected. Each sample is of size V realization of downstream bed occupancy denoted as $h_m^{\zeta_v}$, while $V \subset Z$. Then, the probability of ζ_vth realization is $\prod_{s\in S}\prod_{m\in M_s} P_m^{\zeta_v}$, and $\sum_{\zeta\in Z}\left(\prod_{s\in S}\prod_{m\in M_s} P_m^{\zeta}\right) = 1$. So, the probability of ζ_vth realization is updated as follows:

$$P^{\zeta_v} = \frac{\prod_{s\in S}\prod_{m\in M_s} P_m^{\zeta_v}}{\sum_{\zeta_v\in Z}\left(\prod_{s\in S}\prod_{m\in M_s} P_m^{\zeta_v}\right)} \tag{4.21}$$

For example, if all realizations are assumed to have the same probability, then, $P^{\zeta_v} = \frac{1}{V}$. Hence, it enables us to calculate the weighted average objective function depending on the probabilities of V realizations. Let f^W and x^Wrespectively denote the optimal objective function and the optimal solution of the stochastic ISCM model corresponding to the sample w. Then $\overline{F} = \frac{1}{W}\sum_{w \in W} f^w$ denotes the average of the optimal objective functions for W SAA problems that estimates the objective function.

To select the best solution, denoted as $x^* \in \{x^1, x^2, \dots, x^W\}$, a sample of size V' realization of downstream bed occupancy is randomly selected; and V'should be quite larger than V. Let F denote the corresponding objective function to this sample. So, x^* is the one that has the smallest objective value, that is:

$$x^* \in \arg\min\{F(x): x \in \{x^1, x^2, \dots, x^W\}\} \quad (4.22)$$

4.5. Validate the Basic Model on the Full-Scale ISCM

To achieve the main goal of this study, the ISCM model should be validated with the complete set of data. JGH provided us with the data for 1767 surgical cases (i.e. elective and emergency cases) performed as 87 surgical procedures, grouped in 7 specialties (i.e. General Surgery, Breast Oncology, Colorectal, E.N.T., Gynecology, Urology, Vascular) in 2013. For each surgical case, we have the admission date, surgery date, discharge date, specialty, and the surgical procedure type. As the first step, the data preparation process to calibrate the ISCM model is presented. And then, the SWD simulation model is extended to verify and validate the ISCM results. As the next step, the value of stochastic ISCM model and the importance of its components are studied. At the end of this section a sensitivity analysis is performed on the parameters of ISCM. But, the detailed numerical results of the ISCM model, calibrated with the full set of data, are presented in chapter 5, section 5.2.3. This enables me to better visualize the comparison among various funding policies proposed in the objective function of the ISCM model. The results are presented from strategic, tactical, and operational perspectives for each scenario.

4.5.1. Data Preparation

As illustrated in section 4.2., an increase in the number of surgical procedures exponentially increases the number of possible combinations of *Block Mixes,* which negatively impacts the solution method's effectiveness. To decrease this impact, a clustering method is developed to lessen the number of procedure types and *Block Mix* sets. For clustering categorical data, K-Mean clustering is a common approach in the literature (Costa & Cesar Jr., 2000). This approach clusters available objects into *K* number of groups on the basis of one or more criteria. It minimizes the sum of squares of distances between data and the corresponding cluster centroid. However, we are not able to fully apply this approach to our data since it is mainly interested in a certain number of clusters - *K* - regardless of homogeneity of specialties, medical features, and so on. Also, the study has to consider a limited number of procedures under each specialty while seeing multiple other clustering attributes/features. A new algorithm is proposed to group 2 or more procedures in the same cluster when i. the same surgeon can operate both of them (i.e. the same specialty); ii. they have the same reimbursement rate; iii. the difference in their average OR times is less than 30 minutes; and iv. the difference in their mean expected LOS is less than 2 days. We also consulted with the surgeons, specialists, and nurses to see if specific procedures cannot be grouped due to medical reasons. Applying our clustering algorithm resulted in 47 procedure groups as illustrated in Appendix I. The idea of clustering the procedures is also used in developing diagnosis related groups (DRG)(J. M. Sutherland & Foundation, 2011); however, DRG mostly considers medical features to cluster procedures and most of the aforementioned surgical attributes were ignored.

Although various procedures within each cluster have similar OR times and LOS distribution, yet these attributes are not exactly the same since each *Block Mix* is built on the basis of the required OR time for each cluster. So, a precise OR time must be calculated for each cluster. To this end, the weighted average OR time is calculated for all procedures within each cluster on the basis of the total number of patients historically served under these procedures. For example, "Repair hernia paraesophageal" and "Adrenalectomy" are grouped in the same cluster, while the average OR time for each procedure is 3.6 and 4.2 hours respectively, and the number of completed patients under each procedure is 6 and 9 respectively. So, the weighted average OR time assigned to the cluster is $\frac{(6*3.6+9*4.2)}{15} = 3.9\ hours$. Also, since the reimbursement rate is the same for all procedures grouped as a cluster, that rate is considered as the cluster reimbursement rate. The rest of the process to build all possible *Block Mixes* is as explained in section 4.3. Appendix II illustrates the table of all 72 *Block Mixes* for this study, which cover all feasible combinations of 47 surgical clusters that are performed by 40 surgeons in 7 divisions. To simplify calculations, it is assumed that all surgeons within a division are able to perform or operate all procedures within that division. Obviously, one can run the model with the updated surgeon expertise and preferences when the accurate data is available. Based on the given JGH data, there are 7 surgeons for the General surgery division, 3 surgeons for the Breast Oncology division, 4 surgeons for the Colorectal division, 8 surgeons for the E.N.T. division, 8 surgeons for the Gynecology division, 7 surgeons for the Urology division, and 3 surgeons for the Vascular division in the hospital.

It is assumed that the length of an OR block can be between 3.4 and 5 hours; however, the model imposes overtime cost when the *Block Mix's* OR hours is greater than 4 hours, and idle-time cost when it is less than 4 hours. It is assumed that he *Block Mixes* with the overtime OR hours impose a higher cost to model. For example, *Block Mix* 1 sequences 2 procedures *"Laparotomy exploratory"* and "*Repair hernia incisional*". The length of this *Block Mix* is 4.67 hours, which causes a 0.67 hour overtime cost to the model.

Block Mix 3 includes *"Repair hernia incisional complex"* and *"Repair hernia incisional incarcerated",* which requires 3.89 OR hours on average. This increases the objective function by 0.11 hours OR Idle-time cost, if this mix is part of the optimal solution.

Also, the LOS distribution must be investigated for each cluster. First, as explained in 4.2.4, the truncate rule is applied to find the LOS probabilities for each procedure, denoted as $\boldsymbol{P_k(t)}$. Table 4-4 illustrates a sample of these probabilities based on JGH data. It shows that for the procedure "*Cholecystectomy"* for more than 90% of

Table 4-4: Sample LOS Probabilities

Procedure (k)	**Discharge at the end of Day** (t)	**Probability** $P_k(t)$
Cholecystectomy	1	0.41
	2	0.27
	3	0.14
	4	0.10
	5	0.08
Repair hernia inguinal	1	0.35
	2	0.35
	3	0.30
Laparotomy exploratory	6	0.50
	7	0.25
	8	0.25
Dilatation and Curettage	2	0.50
	3	0.30
	4	0.20
Hysterectomy total robotic	1	0.80
	2	0.20

the patients, LOS is between 1 to 5 days, or for "*Hysterectomy total robotic"* procedure the LOSs of the patients are either 1 or 2 days. Hence, on the basis of the formula (4.14), the conditional probabilities of the total number of required beds are calculated for each *Block Mix, t* day after the surgery, $F_m(b,t)$. All these probabilities are calculated for all 72 *Block Mixes*; but, as an example, table 4-5 illustrates $F_3(b,t)$ for *Block Mix 3*. So, if the model schedule *Block Mix 3* at day t $(i.e.\ t = 1)$, it is observed that the probability of occupying 2 beds at the day of surgery ($t = 1$) is 1 and this probability is 0.872 for day $(t + 1)$. It is also shown that there will be no bed occupied at day $(t + 6)$.

Table 4-5: Distribution of Number of Required Beds for Block Mix 3, $F_3(b,t)$

Number of required beds $(\boldsymbol{b})$	Day $(\boldsymbol{t})$.						
	1	2	3	4	5	6	7
0	0.000	0.000	0.000	0.256	0.333	0.667	1.000
1	0.000	0.128	0.590	0.590	0.667	0.333	0.000
2	1.000	0.872	0.410	0.154	0.000	0.000	0.000

Table 4-6: All Possible Realizations of Required Beds for Block Mix 3

Block Mix 3: No. of Occupied Beds	Day $(\boldsymbol{t})$.						Probability of the Realization
	1	2	3	4	5	6	
Realization 1	2	2	2	2	1	1	0.051
Realization 2	2	2	2	2	1	0	0.051
Realization 3	2	2	2	2	0	0	0.051
Realization 4	2	2	2	1	1	1	0.084
Realization 5	2	2	2	1	1	0	0.084
Realization 6	2	2	2	1	0	0	0.084
Realization 7	2	2	2	0	0	0	0.105
Realization 8	2	2	1	1	1	1	0.095
Realization 9	2	2	1	1	1	0	0.095
Realization 10	2	2	1	1	0	0	0.095
Realization 11	2	2	1	0	0	0	0.118
Realization 12	2	1	1	1	1	1	0.026
Realization 13	2	1	1	1	0	0	0.026
Realization 14	2	1	1	0	0	0	0.033

To solve the problem with the Sample Average Approximation method, one needs to randomly generate different bed occupancy scenarios for all 72 *Block Mixes*. To build the scenarios, first, all possible bed occupancy realizations of each *Block Mix* are generated. Then, one realization per Block Mix is randomly selected to create a scenario. For instance, all possible realizations for *Block Mix 3* are presented in table 4-6. So, “Realization 1” could potentially be part of a scenario with the probability of 0.051. Each scenario determines the number of occupied beds at each day with respect to a *Block Mix*. A scenario has 10 columns (i.e., maximum LOS) and 2160 rows (i.e., 72 Block Mixes * 30 days – as JGH plans its elective surgical cases only on 30 week-days in a period of 6 weeks). Let R_i denotes the total number of realizations for *Block Mix i,* $(\prod_{i=1}^{72} R_i)^{30}$, shows the size of full set of scenarios.

For these seven surgical specialties, it is assumed that the hospital allocated 15 staffed beds in the surgical unit, and runs three ORs (i.e. 6 OR Block) per day. Also, it is assumed that $\eta = 0.15$ which means that for any given period time (e.g. $T = 6$ weeks), the surgical unit is allowed to face a bed shortfall for at most 15% (i.e. on average 6.3 days within 6 weeks). It is assumed that an operating room’s overtime /idle-time cost is $250 per hour and a bed shortage cost is estimated to be $400 per bed day.

To find the most appropriate number of scenarios to be fed into the model, the code is calibrated with V number of scenarios, $V \in [80,60,40,20]$. The ISCM model is coded in ILOG CPLEX Optimization Studio (Version 12.6.1.0) and run on a machine with 24 GB RAM and 16 threads working in parallel each with 2.6 GHz (2 processor) CPU. A comparison of the final results on the basis of running time, local optimality gap, number of constraints and variables is presented in table 4-7. The CPLEX is set to stop running when the local gap is less than 5%, so there is no scientific interpretation on any local gap trend from 20 to 80 scenario. Also, the objective function values are near optimal regarding to the local optimal gap. It is expected to observe that the objective function value is monotone decreasing with the number of scenarios, yet because of different local gap, the results cannot show that. Since a) the objective value does not change significantly considering the set of these scenarios, b) the running time of the code is reasonable for 20 scenarios, and c) the optimality gap is also negligible, the study uses 20 scenarios to develop the rest of the study.

Table 4-7: Comparison of ISCM results with respect to 4 set of LOS scenarios

Number of Scenarios	**20**	**40**	**60**	**80**
Objective (in $1000)	1212.98	1171.06	1162.42	1166.87
Run Time (Sec)	583	1079	2321	23567
Objective Local Gap	1.44%	4.24%	4.70%	3.78%
Constraints	15600	17274	18954	20634
Total Variables	11977	14450	16970	19490
Binary Variables	840	1680	2520	3360
Integer Variables	9455	9408	9408	9408
Other Variables	1682	3362	5042	6722

4.5.2. Extended SWD Simulation

Since the stochastic ISCM model is simplified by approximating patients' LOS distribution and a limited number of scenarios are used in the SAA algorithm to solve the model, a simulation model is deployed to validate our stochastic ISCM model's optimal results.

The basic SWD simulation model was built in chapter 3, given the current JGH OR schedule. That model considered the LOS distributions of eight specialties regardless of downstream procedures; in this chapter, the basic SWD simulation model is extended to validate the ISCM model's outputs and evaluate various OR schedules on the basis of the daily bed shortage. A schematic interface of the SWD model is illustrated in figure 4-4. The Arena model read the optimal OR block schedule from an excel file, and then it allocates staffed surgical beds to each patient based on a LOS distribution. The model discharges the patient when the LOS is completed and records the total number of occupied beds per division per day in the excel file. The run time of the model is 42 days and no warm-up period is required in the model. It is assumed that the patients, who need to stay in the surgical unit after day 42, occupy beds at the beginning of the planning period. For example, if a patient's LOS is 5 days and she is scheduled on day 40, the model assumes that she occupies one bed on days 40, 41, 42, 1, and 2.

Historical data was analyzed to find the LOS distribution for all 47 procedures within 7 divisions. The SWD simulation model is calibrated with these LOS distributions and used the optimal OR schedule from the ISCM model to find the bed occupancy histogram. The SWD simulation model is run for 50 replications. The results show that his number of replication was enough to converge to a steady state occupied beds per division and to diminish simulation error.

The SWD simulation model is able to be run for all optimal OR schedules regarding various scenarios that was discussed in section 4.4.3 and 4.4.4. However, for the sake of space and time, only one OR schedule is depicted as an example to feed and run the SWD simulation model in this chapter. From the ISCM results, it is expected to observe at most 15% bed shortage during the whole planning horizon (this is an initial setting of the model). However, figure 4-5 shows that 16 beds are required for a service level index (SLI) of 85%.

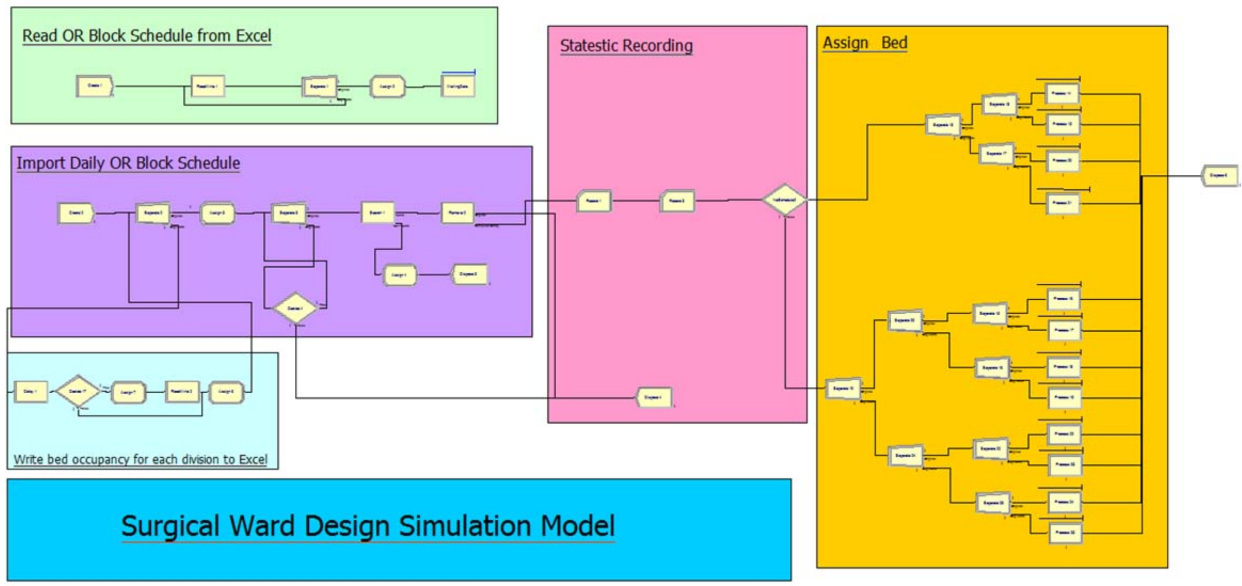


Figure 4-4: A schematic interface of the extended SWD Simulation Model

It was anticipated that more beds would be required to reach 85% SLI; since the truncate rule was applied to calculate LOS distributions for the ISCM model and it was assumed that LOS is always fewer than 10 days, although less than 5% of LOS is greater than 10 days based on JGH data. However, the SWD simulation model results prove that the ISCM model result is not far from reality for this specific OR schedule. Figure 4-5 also illustrates that the surgical unit needs at least 18 surgical beds to reach 95% SLI.

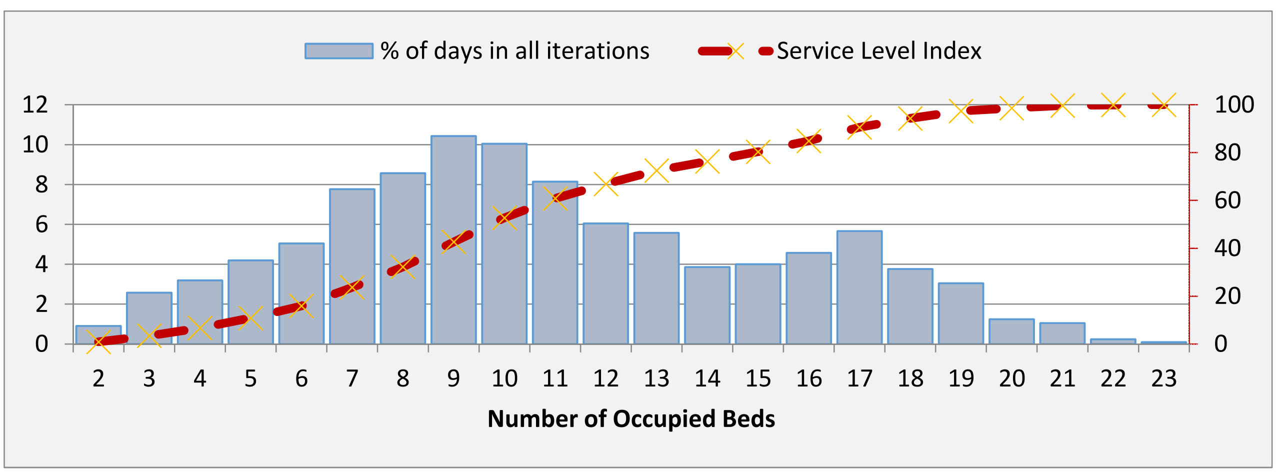


Figure 4-5:Required Bed Frequencies – As a Result of Extended SWD Simulation Model

Furthermore, Figure 4-6 illustrates further information on the required number of beds for this OR schedule among various specialties. It is observed that on average 10.9 beds are busy each day while the bed occupancy pattern is quite smooth when the 6-weeks OR schedule is repeated for 50 replications. It is also discovered that on weekends fewer beds are required to meet the demand (It is assumed that day 1 is Monday, so day 6 & 7 is weekend and so on).

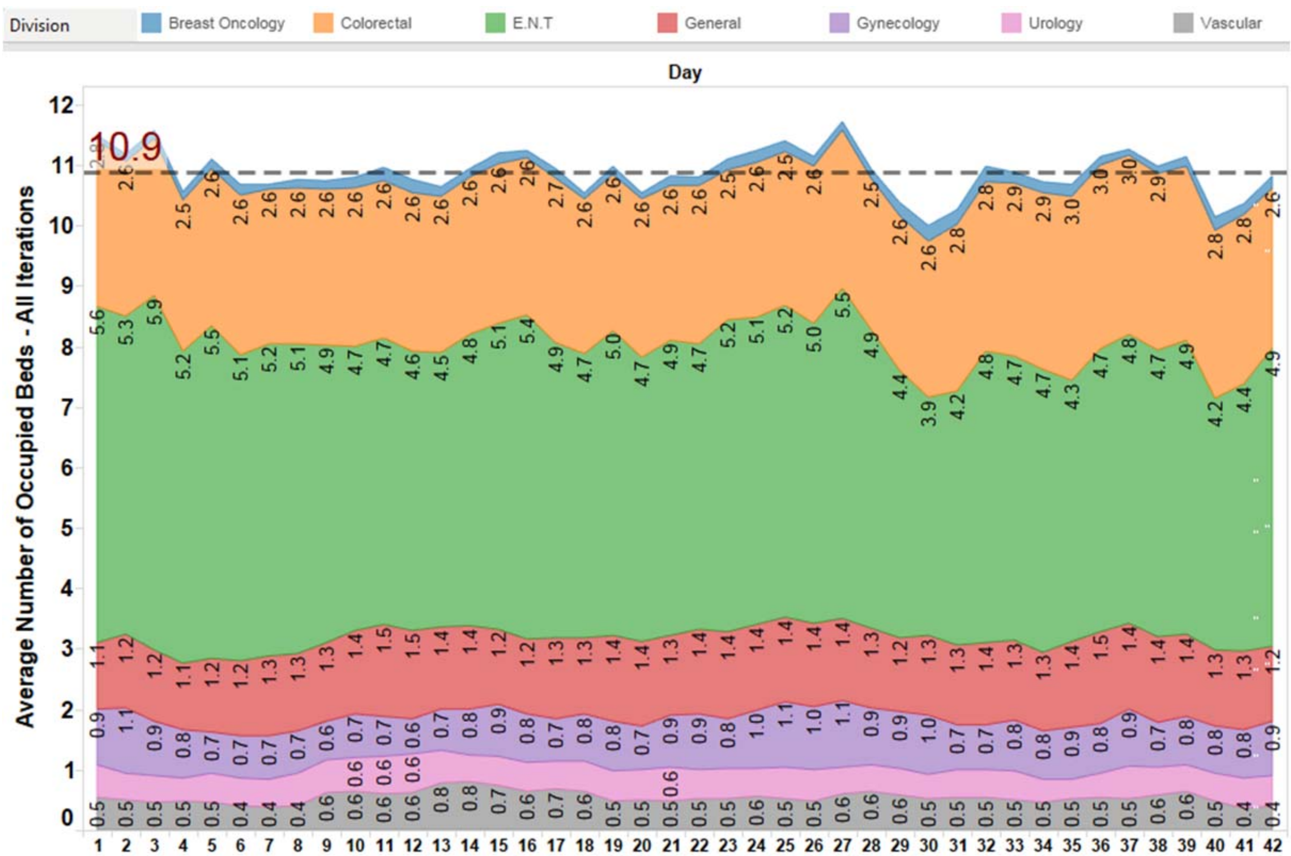


Figure 4-6: Average Occupied Surgical Beds - All Specialties

Figure 4-7 illustrates the distribution of the required number of surgical beds per day using a Whisker-Box plot. Simulation results also demonstrate that the surgical unit should expect 22.8 bed-days shortage on average over the course of 6 weeks.

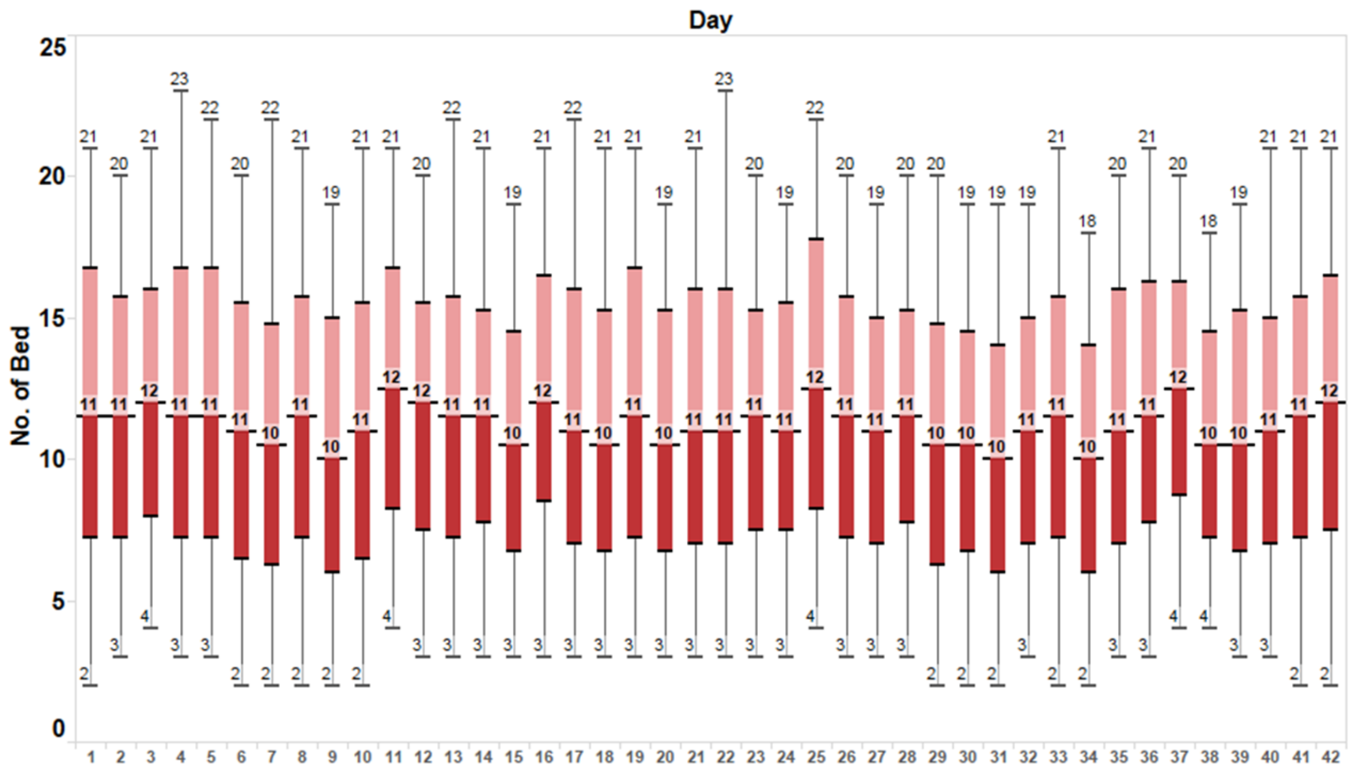


Figure 4-7: Whisker - Box Plot on the Number of Required Surgical Beds per Day

4.5.3. Deterministic Vs. Stochastic ISCM

To show the impact of the stochastic parameters on the optimal solution of the ISCM model, the model is first run using the median of all stochastic parameters (M1). Second, the model is run using the average patient LOS for each procedure (M2). Then, the results are compared with the result from the stochastic model (M3). For this comparison, the ISCM model with the ABF (i.e. Activity Based Funding Policy) objective function is chosen. First these stochastic ISCM models are solved, and then the SWD simulation model is run to evaluate the final OR schedules. Table 4-8 presents the results for both the deterministic problems and the stochastic ISCM models.

For most of surgical *Block Mixes*, the LOS data is right-skewed. So, the average of the patients'

Table 4-8: Deterministic Vs. Stochastic ISCM

ISCM Problem	Deterministic Problem (M1) (Median LOS)	Deterministic Problem (M2) (Average LOS)	Stochastic Problem (M3) (Probabilistic LOS)
Objective (in $1000)	1246.94	953.07	1212.98
Run Time (Sec)	1.9	0.84	583
Objective Local Gap	0.53%	1.66%	1.44%
No. of Assigned OR Block	178	154	179
Expected No. of Bed Shortage over 6 weeks	0	0	12
Bed Shortage Probability (SWD Simulation Results)	38.23 %	5.90%	17.4%
Constraints	1404	1404	15600
Total Variables	9583	9583	11977
Binary Variables	42	42	840
Integer Variables	9455	9455	9455
Other Variables	86	86	1682

LOSs (who will be planned) in the *Block Mixes* is greater than their medians. This negatively impacts on an accurate estimation of the required resources or the value of objective function. For example, the objective values of the M1 and the M3 are close to each other, however, the value decreases by around $250,000 for M2. This shows that the use of the average LOS might be misleading because it underestimates the objective value. Also, M2 underestimates the surgical ward workload by only using 154 OR blocks out of 180 available OR blocks. This demonstrates that JGH needs more than 15 surgical beds to be able to utilize more than 154 OR blocks, which is not true. Consequently, the SWD model shows that in reality the beds shortage under M2 is less than 5.9% since it overestimated the number of required beds by considering the average LOS.

However, when the SWD simulation model runs using the optimal OR schedule of M2, it is observed that on only 5.9% of the days the number of required surgical beds exceed the total number of available surgical beds. Figure 4-8 illustrates the distribution of required surgical beds based on the M2's optimal OR schedule. On the other hand, M1's objective value is very close to the stochastic ISCM's objective value, and it utilizes 178 OR blocks within the period of 6-week. On one hand, the running time is much less for the M2 than for M3. On the other hand, the optimal OR schedule of the M1 generates more than 38% bed shortage over the course of 6 weeks. In other words, a key difference between M1, M2 versus M3 is that the stochastic model has an accurate estimation of bed shortages, whereas the deterministic models ignore this. It shows that although M1's results seem to be appropriate for a strategic decision, one cannot validate them at the tactical or operational levels. Figure 4-9 illustrates the distribution of required surgical beds based on the M1's optimal OR schedule.

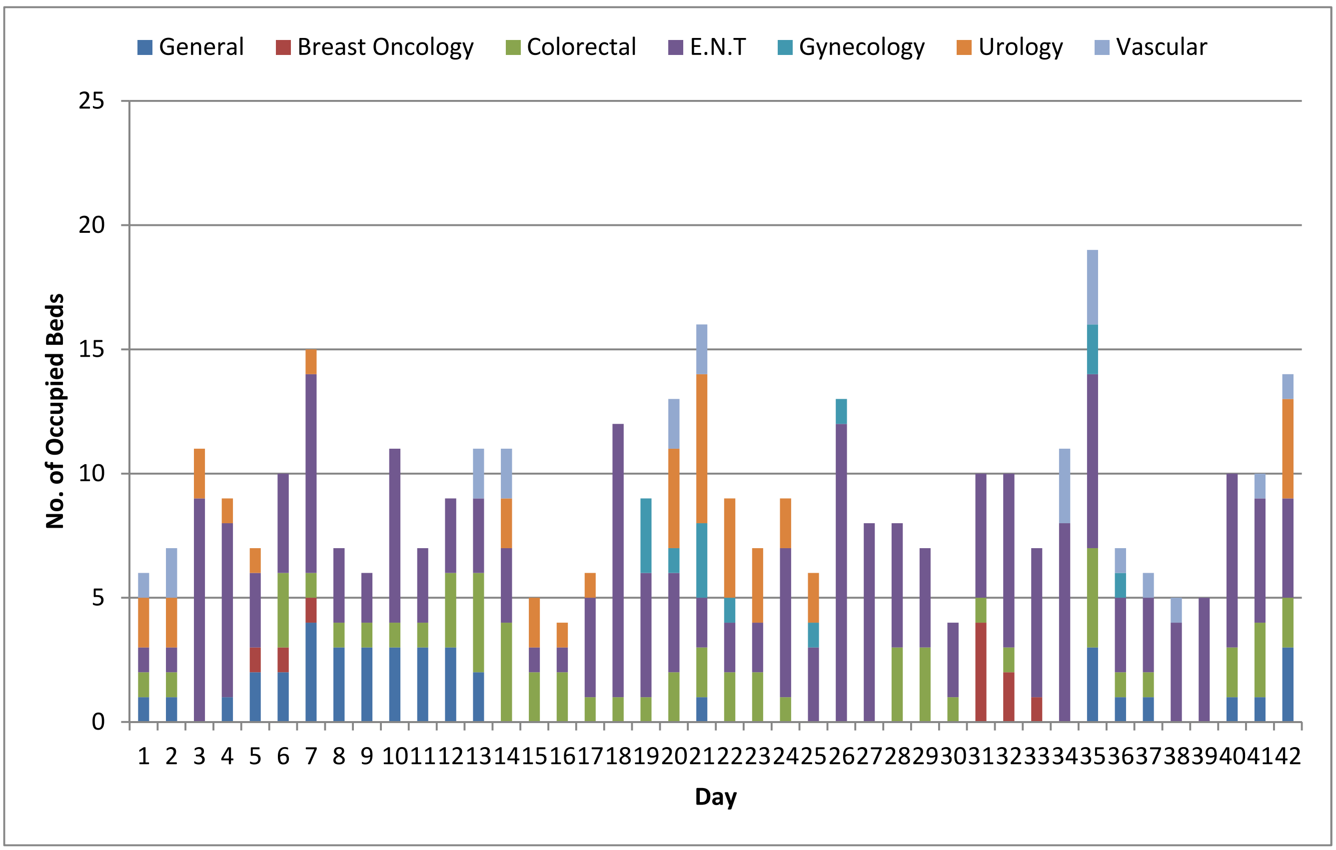


Figure 4-8: Required Surgical Beds for M2 OR Block Schedule

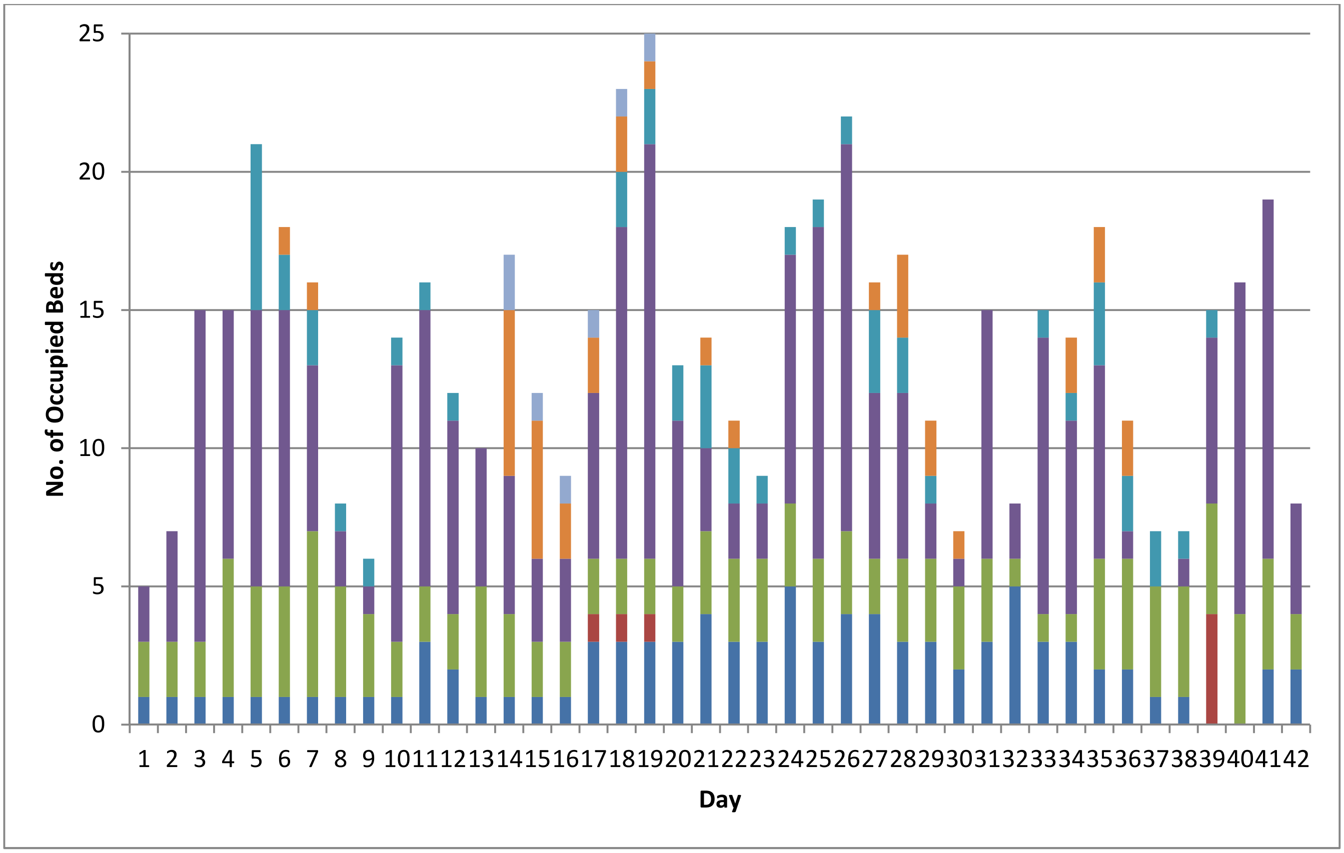


Figure 4-9: Required Surgical Beds for M1 OR Block Schedule

Table 4-9 presents dedicated OR Blocks to each division as a result of these models. It is observed that for all divisions except E.N.T. and Vascular these models dedicate the same number of OR blocks forced by the minimum number of required OR block constraints in the model. Yet, the final OR schedules of these models are completely different. The result of the SWD simulation model is illustrated based on the M3's optimal schedule in section 4.5.2. So, it is simply observed the impact of the stochastic model on the accuracy of final results. It should be noted that the "ISCM model" refers to the stochastic version of the model throughout this dissertation, unless otherwise indicated.

Table 4-9: Deterministic Vs. Stochastic ISCM – Divisions' OR Block

	Reimbursement Policy	Specialty						
		Breast Oncology	**Colorectal**	**E.N.T**	**General Surgery**	**Gynecology**	**Urology**	**Vascular**
Deterministic Problem (M1) (Median LOS)	*ABF*	3	13	106	12	16	14	13
Deterministic Problem (M2) (Average LOS)	*ABF*	3	13	77	12	16	14	19
Stochastic Problem (M3) (Probabilistic LOS)	*ABF*	3	13	105	12	16	14	16

4.5.4. CVaR and Chance-Constraint Interaction

To limit the bed shortage level as a result of a surgical mix decision, the chance-constraints are incorporated in the ISCM model. To this end, an acceptable bed shortage level (e.g. $\eta = 15\%$) is determined by the surgical unit managers, and the model is banned from exceeding this threshold. In addition, a CVaR term is embedded in the objective function of the ISCM model to control the amount of bed shortage when a surgical unit is fully occupied. This section separately studies the impact of each term on the final ISM results.

First, the CVaR term (i.e. $SC * \sum_{t=1}^{T} E[Y_t - B]^+$) is removed in the objective function to run the stochastic ISCM problem. It is assumed that there is no bed shortage cost when the model requires more than B beds. However, the chance-constraints still bound the bed shortage days at $\eta\%$. The final results show the average expected bed shortage increases from 9.05 to 12.9 bed-days within the 6-week planning period. Nevertheless, this change does not impact the optimal case-mix. Also, the SWD simulation model is run with respect to the optimal OR schedule. The results do not show a significant increase in the average required number of beds.

As explained in section 4.4.2, the SAA algorithm was applied to solve the stochastic ISCM problem and implemented an estimation of real patient LOS in the model; consequently, it is not expected that in reality the surgical unit meets 85% SLI precisely. By removing CVaR term, the SLI decreases to 82%. The CVaR term helps the stochastic ISCM model to keep the bed shortage probability around 15%. Figure 4-10 illustrates the distribution of the required number of surgical beds per day using a Whisker-Box plot when there is no CVaR term in the model. It shows that the median of occupied beds varies from 11 to 13 beds per day, while the average number of occupied beds is 10.93. The standard deviation of occupied beds increases to 4.46, which brings less smooth patient flow to the surgical unit compared to the complete stochastic ISCM. The SWD simulation results show that at most 24 beds might be required to serve all patients under the optimal OR schedule from the ISCM without a CVaR term.

Now the chance-constraint is removed and the CVaR term is kept in the ISCM to analyze how the CVaR term would resist the bed shortage increase. Final results show that this setting does not impact the optimal case-mix, but the average expected bed shortage increases to 55.3 bed-days and the average number of required beds rises to 12.04 beds. In addition, the outcomes of the SWD simulation model demonstrate that the OR schedule results in a significant decrease in the SLI from 85% to 62%. In other words, in 38% of the 6-week period the hospital would face a bed shortage. The median of occupied beds varies from 12 to 14 beds per day, which means more congestion due to the OR mix scheduling. Figure 4.11 illustrates the distribution of the required number of surgical beds per day using a Whisker-Box plot when there is no chance-constraint in the ISCM model.

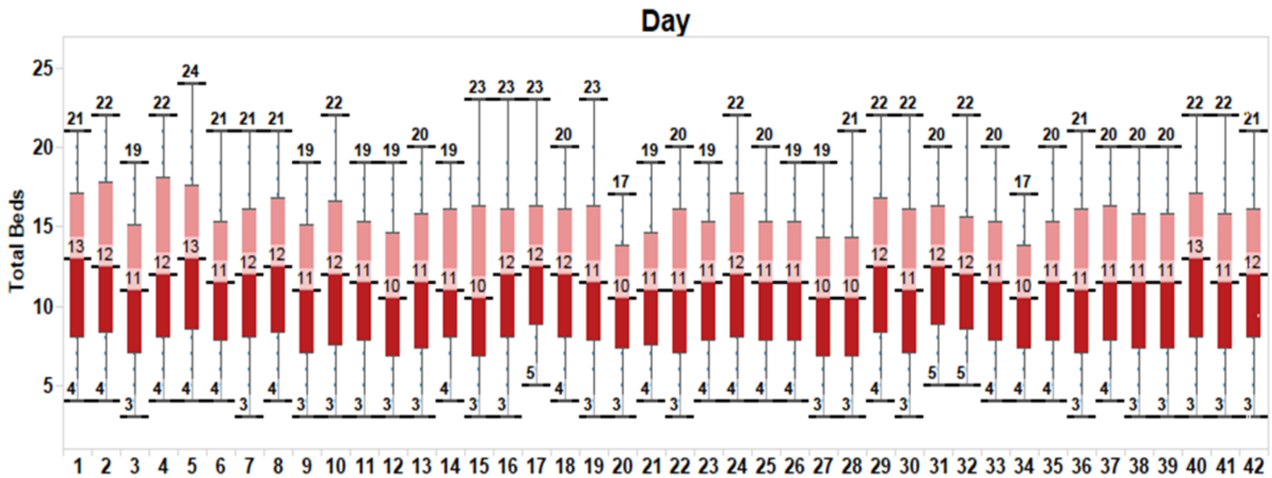


Figure 4-10: Whisker - Box Plot on the Number of Required Surgical Beds per Day – No CVaR term

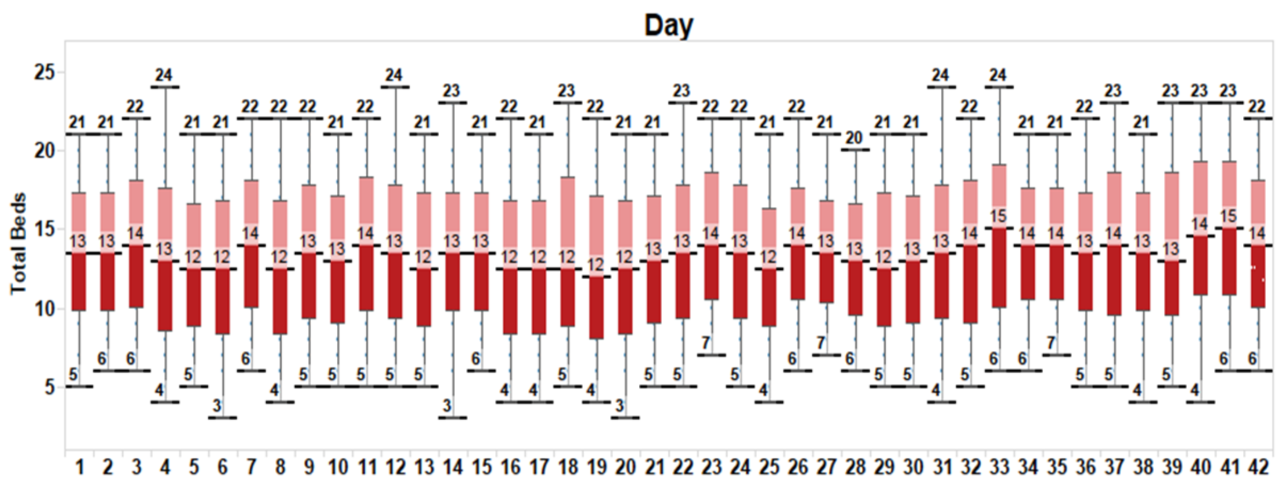


Figure 4-11: Whisker - Box Plot on the Number of Required Surgical Beds per Day – No Chance-Constraint

Furthermore, at this point a sensitivity analysis is performed on the bed shortage cost to find its impact on the average expected bed shortage when there is no chance-constraint in the ISCM model. Figure 4.12 illustrates that the average expected shortage in the number of bed negatively correlates with the bed shortage cost. This concave-up curve shows if the bed shortage cost is $2000 instead of $400 (i.e. the initial value), the CVaR term controls the bed shortage level at 12.3. The SWD simulation results also prove that SLI will be 84% for this setting.

Also a sensitivity analysis is performed on the bed shortage cost when both the CVaR term and the chance-constraint are available in the ISCM model. In figure 5.4 a concave-down curve illustrates that a bed shortage cost less than $2000 does not significantly impact the bed shortage level. In other words, the complete ISCM model is not sensitive to bed shortage cost.

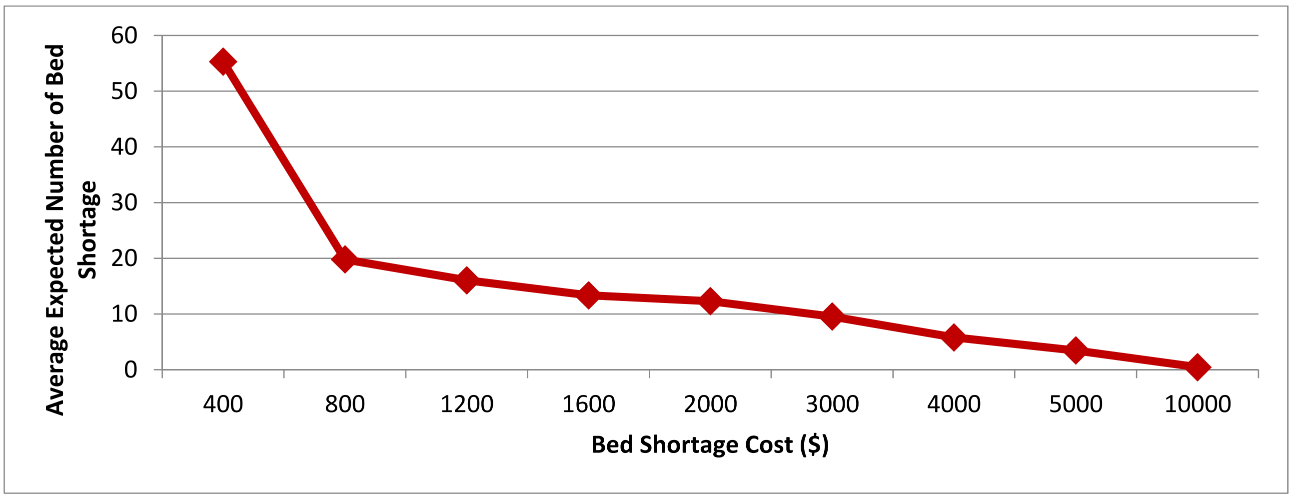


Figure 4-12: Sensitivity Analysis on the Bed Shortage Cost– No Chance-Constraint

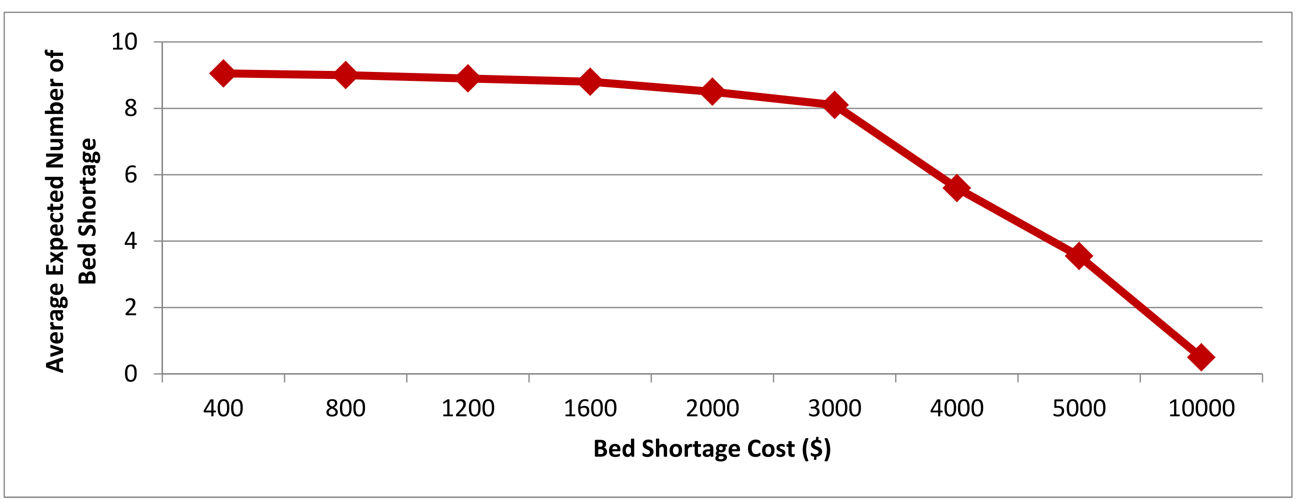


Figure 4-13: Sensitivity Analysis on the Bed Shortage Cost– Both of the CVaR and the Chance-Constraint are in the ISCM

Table 4-10: Summary of ISCM and SWD simulation model results

ISCM Setting		**Performance of the ISCM Optimal Solution**				
CVaR Term	**Chance Constraint**	Expected Bed-Shortage Days	Maximum daily Bed Shortage	Average Number of Occupied Beds	Standard Deviation of Occupied Beds	SWD Service Level Index
YES	**YES**	9.05	8	10.88	4.40	84%
NO	**YES**	12.9	10	10.93	4.46	82%
YES	**NO**	55.3	10	12.04	4.62	62%

It is concluded that the combination of the CVaR term and the chance-constraint brings the most reliability to keep the SLI at 85%, while the model is quite robust to bed shortage cost. This combination also has the least standard deviation of occupied beds, which results in a smooth patient flow in the surgical unit. Table 4-10 summarizes the final results of the stochastic ISCM and the SWD simulation model under the aforementioned settings.

4.5.5. Sensitivity Analysis of ISCM

The prevailing literature is not helpful pertaining to a reliable estimate of the OR Idle/overtime costs. JGH was unable to provide us with this information. To calibrate the ISCM model, this cost is approximated at $500 per hour. In this section a sensitivity analysis is performed on this parameter to explore how robust the final results are when the OR Over/Idle time cost changes. In this sensitivity analysis, the ABF policy is considered in the ISCM objective function. As illustrated in Figure 4-14, the hospital’s revenue decreases as OR Idle/Overtime cost increases.

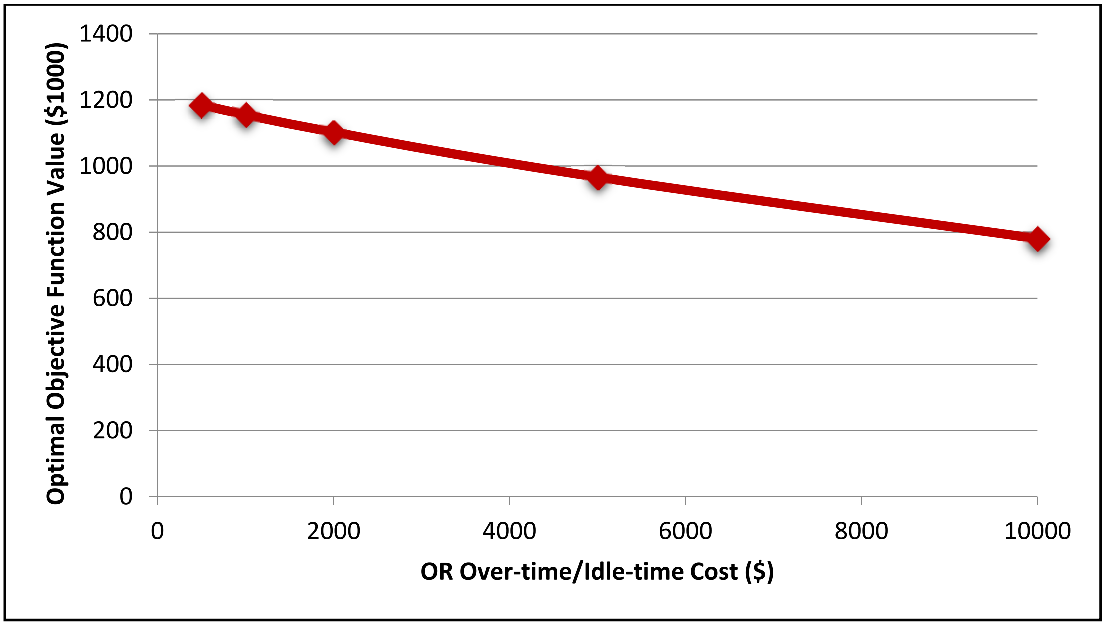


Figure 4-14: Optimal Value of the ISCM Objective Function Regarding the OR Idle/Over Time Cost

However, the optimal case-mix does not change too much when the cost is less than $2000. Table 4-11 shows a considerable reduction in the E.N.T. cases and a significant increase in Vascular procedures when the cost is $5000 or more. It is observed that the optimal number of dedicated OR blocks to surgical-mix number 72 jumps to 50. This number was 5 when the OR Idle/Overtime cost was $400. Mix number 72 contains only one Vascular procedure,"Repair AAA - aneurysm aorta abdominal". This procedure has 0.02 hours OR idle-time on average, while its reimbursement rate is $7215. However, this surgical mix was not initially part of the optimal solution due to large LOS. Also, it is observed that the optimal number of dedicated OR blocks to surgical-mix number 46 drops to 5. This number was 52 when the OR Idle/Overtime cost was $400. Mix number 46 contains "Thyroidectomy", an E.N.T. procedure with an average of 0.37 hours as Idle OR time, and $3758 as the reimbursement rate.

In general, it is concluded that the ISCM optimal results are quite robust if the OR Idle/Overtime cost is less than $2000 per hour.

Table 4-11: Optimal Surgical-Mix Regarding the OR Idle/Over Time Cost

OR OI Cost ($)	Number of Dedicated OR Blocks to each Specialty						
	Breast Oncology	Colorectal	E.N.T	General	Gynecology	Urology	Vascular
500	3	13	106	12	16	14	15
1000	3	13	107	12	16	14	14
2000	3	13	103	12	16	14	18
5000	3	13	77	12	16	14	44
10000	3	13	60	12	16	14	61

4.6. Conclusion

The Integrated Case Mix (ISCM) model was developed and evaluated in this chapter. The model simultaneously addresses i. the types and quantities of surgical procedures to be performed in the surgical unit, ii. the efficient distribution of the operating rooms among various divisions and surgeons, and iii. the optimal OR schedule. For the first time in the literature, the ISCM model developed shows how to integrate strategic, tactical, and operational decisions in the surgical unit within a single model.

To tackle the key concerns faced by the hospital administrator, the ISCM model strives to optimize the dollar amount of the hospital reimbursement, operating room utilization costs, and downstream bed shortage expenses. It also captures uncertainty in patients LOS, controls the service level, addresses surgeons' priority, and division chiefs' concerns.

The linear approximation of the ISCM model is presented and the model is calibrated and validated with the full set of real data, provided by JGH. Then, to the best of author's knowledge for the first time in the literature, the Sample Average Approximation technique was applied to solve a Surgical Case-Mix problem. Moreover, the SWD simulation model, introduced in chapter

3, was extended in this chapter to verify the ISCM results and to evaluate various OR schedules under different scenarios.

To explore the impact of LOS uncertainty on the Surgical Case-Mix problem, the deterministic and stochastic ISCM models are developed and compared using the SWD simulation model. The results demonstrate that the deterministic ISCM model, calibrated with the average LOS, overestimates the required number of beds for utilizing all available OR blocks, and underestimates the objective value. Also it is illustrated that the deterministic ISCM model, calibrated with the mean LOS, results in huge bed shortage (i.e., 38%) in the surgical unit.

From the modelling perspective, the chance-constraints limit the maximum number of days the surgical unit faces bed shortage. Moreover, the CVaR term controls the risk of high bed shortage within each day. This chapter showed that the incorporation of these terms in the ISCM model is required to meet the model's objectives (e.g., 85% SLI) and to experience a smooth patient flow. Also, the sensitivity analysis on the bed shortage cost is performed to investigate its impact on the aforementioned incorporation. The results demonstrate that when both the CVaR term and the chance-constraint are incorporated in the ISCM model, for the bed shortage cost less than $2000 the complete ISCM model is not sensitive to bed shortage cost. Furthermore, a sensitivity analysis on the OR Idle/Overtime cost (less than $2000 per hour) demonstrates that the ISCM optimal results are quite robust.

In the next chapter the ISCM model is extended to capture some external factors such as emergency and off-service patients that influence the ISCM model. Also, the ISCM model is presented under different reimbursement policies.

Chapter 5

Extensions on the Basic ISCM Model

5.1. Introduction, Motivation, and Literature Review

In chapter 4 the ISCM model was developed that integrates the surgical case mix problem with the OR block scheduling problem. The model considers all possible Block Mixes (i.e. feasible sequence of procedures within an OR block) for each surgeon to discover a realistic case-mix decision. Then the model was calibrated and run with the full set of data from JGH. And then, a sensitivity analysis was performed on the ISCM components.

Chapter 5 presents the extend ISCM model to capture its interaction with external factors: reimbursement policy, configuration policy, emergency patients, and off-service patients. To this aim, the effect of three reimbursement policies is presented on the surgical case mix problem outcomes, which helps policy makers to evaluate the consequences of the hospital funding policies on the surgical case mix.

The ISCM model considers a fully pooled bed configuration in the surgical ward. However, lots of hospitals prefer to follow semi-pooled or dedicated bed configuration policies. Less nursing training effort is required under these policies and they enable divisions' managers to easily manage their resources. To address this concern, the ISCM model is extended on the basis of the semi-pooled bed configuration policy, and the ISCM model and the SWD simulation model results are compared with the fully pooled bed configuration setting.

Although the ISCM model integrates the operational, tactical, and strategic aspects of a surgical unit case-mix problem, it does not address its cooperation with other units such as Emergency Department (ED) or Medicine. To have a comprehensive model, the study not only has to target preplanned surgical patients in the ISCM but also incorporate the emergency and medicine patients using surgical unit resources. Emergency patients might get admitted to the acute care unit through the ED when they need surgery to complete their treatment. The arrival rate of emergency patients is stochastic and makes the ISCM even more complex. This section will discuss expansion of the ISCM to capture the impact of emergency patients on the surgical unit case-mix problem and downstream bed utilization.

Occupying surgical beds might happen by medicine patients too. They might need surgery as part of their treatment, so they get transferred to the surgical unit. Also, they might get placed in the surgical unit due to bed shortage in the medicine unit. These patients are denoted as off-service patients in the literature. To the best of our knowledge, there is no study to address the impact of off-service patients on the surgical case mix problem.

5.2. Reimbursement Policy’s Impact on the ISCM

5.2.1. Introduction

The hospital funding mechanism is the subject of significant debates and changes in most countries with publicly funded healthcare systems (Mayes, 2007). The ISCM model was developed based on the activity based funding (ABF) policy. ABF policy is based on the type and volume of procedures provided at each hospital. In addition to the ABF policy, Global Budgeting (GB) and a combination of ABF and GB methods are other common alternate methods to be contemplated by the provincial governments in Canada and other industrialized countries (Pink et al., 2001). This section will evaluate the impact of each policy on the strategy of hospitals administrators on the Surgical Case Mix problem.

Each policy has its own incentives and disincentives influencing quality, type, and volume of hospital services. Dafny studied hospital's response to the price change (Dafny, 2005). He found that the main response of the hospitals to change in reimbursement policy is coding patients to diagnosis codes (DRGs) with the largest price increases. However in the long-term, he concluded that hospitals would increase the volume of profitable surgical procedures and specialize in admissions in which they are relatively cost efficient. At the strategic level, building the ISCM model on the basis of the proposed funding policies helps the policymakers to discover how these policies affect hospitals' response to better assess the trade-offs among proposed alternatives. Also it helps hospital administrators to discover the possible utilization by contemplating case-mix reform efforts.

5.2.1.1. Global Budget with Incentive

Hospital's Global Budget (GB) refers to a fixed annual lump-sum distributed among each hospital to cover their operating costs independent of the volume or type of service provided. Global budget policy is easy to implement since it provides a stable and predictable plan for the hospitals and the government. Also, it is a suitable policy for small hospitals to cover their fixed costs.

Let GBI denote the surgical unit revenue under the global budget funding policy with incentive. Global budget is usually based on the hospital services, capacity, and historical expenditures. It does not incentivize the hospitals to utilize their unused capacity to increase the surgical case volume, resulting in lengthy waiting lists. Therefore, depending on the hospital capacity, the government defines a certain number of cases for each procedure type, C_k, and if the hospital practice more than this threshold, it will get reimbursed for the extra cases, denoted as $[\sum_{t=1}^{T}\sum_{m=1}^{M} N_{km} * x_{smt} - C_k]^+$, in addition to the hospital's GB. In other words, the government buys increased volume of surgical cases from the hospital. Hence,

$$Reimbursement(GBI) = GB + \sum_{k=1}^{K} R_k * [\sum_{t=1}^{T}\sum_{m=1}^{M}\sum_{s=1}^{S} N_{km} * x_{smt} - C_k]^+ \tag{5.1}$$

And then the objective function 4.13 must be updated to:

$$Max \tag{5.2}$$

$$GB + \sum_{k=1}^{K} R_k * [\sum_{t=1}^{T}\sum_{m=1}^{M}\sum_{s=1}^{S} N_{km} * x_{smt} - C_k]^+ - \sum_{s=1}^{S}\sum_{m\in M_s}\sum_{t=1}^{T}[IO_m * x_{smt}] - SC * \sum_{t=1}^{T} E[Y_t - B]^+$$

5.2.1.2. Global Budget with Incentive and Penalty

To decrease the waiting lists, government can impose financial penalties on the hospital if it is not committed to the C_k threshold. Note that GB is usually calculated on the basis of the procedure cost rate which is lower than R_k, hence $GB < \sum_{k=1}^{K} R_k * [\sum_{t=1}^{T}\sum_{m=1}^{M}\sum_{s=1}^{S} N_{km} * x_{smt}]$. However, at the end of the year if the number of patients with procedure k deviates from C_k the hospital will payback or get reimbursed with the rate of R_k per patient depending on whether it falls short or exceeds the threshold C_k respectively. Therefore, the financial term is defined as:

$$Reimbursement(GBIP) = GB + \sum_{k=1}^{K} R_k * [\sum_{t=1}^{T}\sum_{m=1}^{M}\sum_{s=1}^{S} N_{km} * x_{smt} - C_k] \tag{5.3}$$

Hence, the objective function of the ISCM model is:

$$Max \tag{5.4}$$

$$GB + \sum_{k=1}^{K} R_{\mathrm{k}} * [\sum_{t=1}^{T} \sum_{m=1}^{M} \sum\nolimits_{s=1}^{S} N_{km} * x_{smt} \; - \; C_{\mathrm{k}}] - \sum_{s=1}^{S} \sum_{m \in M_s} \sum_{t=1}^{T} [IO_m * x_{smt}] - SC * \sum_{t=1}^{T} E[Y_t - B]^+$$

5.2.2. Evaluating Reimbursement Policies

To compare these policies, two settings for the ISCM model are considered. First, it is assumed that the manager might face some restrictions due to the minimum number of OR Blocks dedicated to each division or surgeon. This setting is denoted as "Constrained Problem". Second, it is assumed that there is no such restriction in front of the managers. This setting is denoted as "Relaxed Problem". Under both settings, it is assumed that the total number of procedures cannot exceed a fixed number since there is a finite number of patients that require a specific surgery. This threshold is assumed to be 2 times the average completed cases for each procedure annually on the basis of historical data. Also, the same parameter that is used in different models is assumed to have the same value in all models. Table 5-1 illustrates the results of all three policies under both settings. It is observed that for both the ABF and GBIP funding policies, the ISCM model results in similar optimal solution under both the Relaxed and the constrained scenarios, which is interesting. In other words, hospital managers should make similar strategic decision on the Surgical Case Mix problem under both policies. From a mathematical point of view it is proved that these objective functions must result in the same optimal surgical case-mix:

Proof:

$$Reimbursement(GBIP) = \ GB + \sum_{k=1}^{K} R_{\mathrm{k}} * [\sum_{t=1}^{T} \sum_{m=1}^{M} \sum_{s=1}^{S} N_{km} * x_{smt} \ - \ C_{\mathrm{k}}]$$

$$= \ GB - \sum_{k=1}^{K} R_{\mathrm{k}} * C_{\mathrm{k}} + \sum_{k=1}^{K} R_{\mathrm{k}} * [\sum_{t=1}^{T} \sum_{m=1}^{M} \sum_{s=1}^{S} N_{km} * x_{smt}$$

$$= \ GB - \sum_{k=1}^{K} R_{\mathrm{k}} * C_{\mathrm{k}} + Reimbursement(ABF) \quad \square$$

And the $GB - \sum_{k=1}^{K} R_{\mathrm{k}} * C_{\mathrm{k}}$ term is a constant number. The ABF and GBIP objective functions are maximized when surgical procedures with the highest marginal benefit stand in the final solution. Under the GBIP policy, if these procedures do not satisfy the base volume threshold, the hospital is penalized but still makes more money by operating the most beneficial procedures using limited available resources. The objective function value, however, might be different for these two funding policies. Table 5-1 shows that the objective function value is $1212980 for the ABF policy while this value is “Global Budget +$10850”. If the hospital’s global budget is less than $1202130, the government saves on the total reimbursement amount to the hospital by applying the GBIP policy.

Regarding the relaxed setting, it is observed that the Colorectal Surgery, the General Surgery, the Gynecology, and the Urology are not among the most attractive divisions. E.N.T. procedures use minimum resources, yet bring the highest funding to the hospital. Also, the Vascular Surgery is ranked as the second most important division for the hospital.

The story is different for GBI funding since the managers have to plan for the base volumes regardless of their marginal benefits. So, under both constrained and relaxed settings the priority goes to those procedures listed under the hospitals’ commitment to the government. Then, the remaining resources are dedicated to the E.N.T division to optimize the objective function of the ISCM model. In general, it is obvious that the objective functions are greater in relaxed settings compared to the constrained setting.

Table 5-1 illustrates that the ISCM model results in a higher optimal objective function value under the GBIP compared to the GBI policy for the same setting. From the mathematical view, the GBIP is a more flexible policy comparing to the GBI policy, and any feasible solution under the GBI policy setting is also a feasible solution under the GBIP policy setting. Hence, the ISCM model with GBIP policy must has an optimal objective function value that is at least equal (and likely larger) than the optimal value of the ISCM model with GBI policy.

Table 5-1: Funding Policies Comparison

	Funding Policy	Objective Function Value ($1000)	***Specialty***						
			Breast Oncology	**Colorectal**	**E.N.T**	**General Surgery**	**Gynecology**	**Urology**	**Vascular**
Constrained Problem	*ABF*	1212.98	3	13	105	12	16	14	16
	GBI	Global Budget + 6.30	2	6	87	10	30	22	13
	GBIP	Global Budget + 10.85	3	13	108	16	16	14	13
Min OR Block-ISCM Constraint			2	0	20	10	30	16	12
Relaxed Problem	*ABF*	1399.80	2	0	160	0	0	0	16
	GBI	Global Budget + 7.80	3	0	96	16	32	19	14
	GBIP	Global Budget + 181.80	2	0	160	0	0	0	16

It is important to note that since a limited number of procedures under each division is studied in the ISCM (due to data unavailability), the model might result in different outcomes if the mix of procedures changes under each division or if a complete set of hospital resources such as available ORs and surgical beds are considered.

5.2.3. Strategic, Tactical, and Operational Outcomes of Stochastic ISCM

This section presents outcomes of the stochastic ISCM model based on three different settings denoted as alternative A, B, and C. The results are compared from strategic, tactical, and operational perspectives. Alternative A applies the constrained setting, described in previous section, to the ISCM model with the ABF objective function. Alternative B imposes the relaxed setting to the ISCM model with the ABF objective function. Alternative C also explains the results of ISCM with respect to the relaxed setting under the GBI funding policy. Figures 5-1, 5-2, and 5-3 illustrate OR Block allocation among all divisions. Boxes with the same colour represent the same specialty, while each box is split into several smaller boxes representing surgeons. For example, figure 5-1 shows that more than half of available OR blocks are dedicated to the E.N.T. division. Depending on the minimum number of required OR blocks for each E.N.T. surgeon, the ISM model decided to allocate 32 OR blocks to surgeon no.15 and only 1 OR Block to surgeon no. 19. It is observed that surgeon no. 15's share from available OR Blocks is greater than all the other 6 divisions. So, to have a more realistic model a maximum number of OR Blocks must be imposed to each surgeon. Also we are able to figure out if the hospital managers need to have more surgeons within each division. However, this does not change the strategic guideline to focus on the E.N.T. division. The importance of the E.N.T. division is even more significant based on figure 5-2 when there is no commitment to dedicate some OR Blocks to the other divisions.

Alternative - A

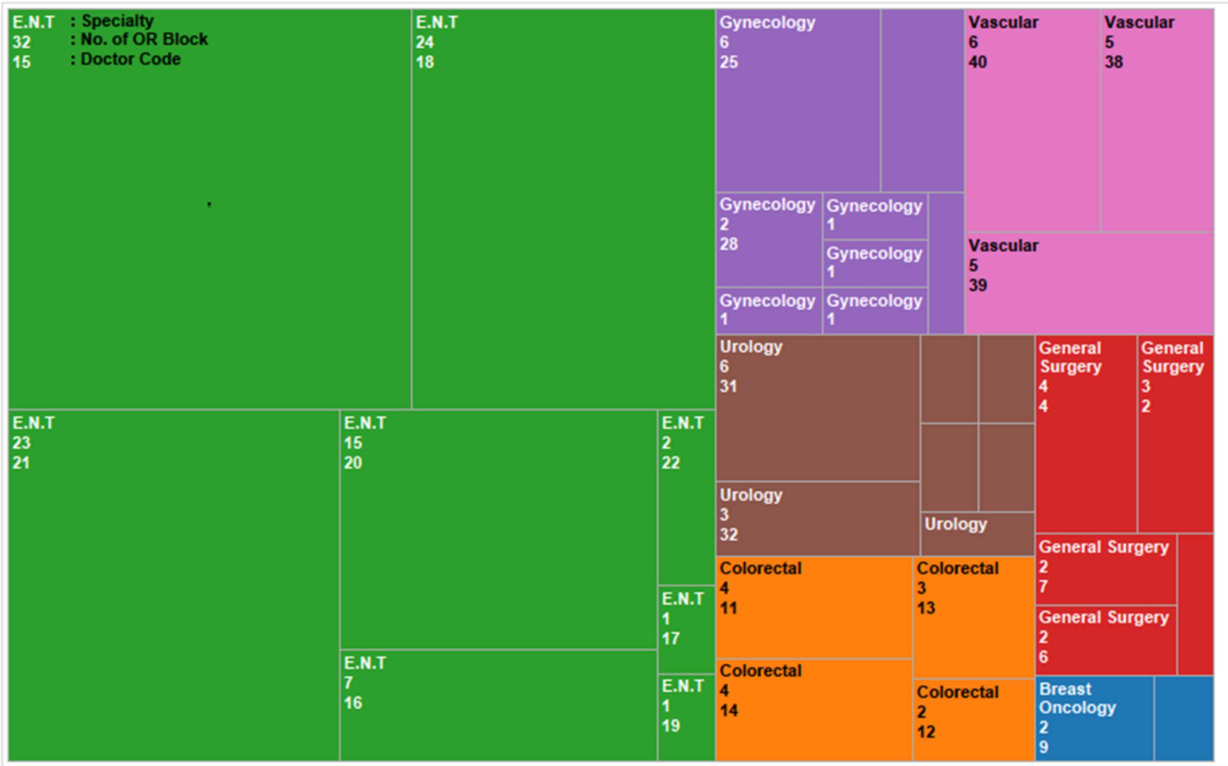


Figure 5-1: Strategic Allocation of OR Blocks to Divisions – Alternative A

Alternative - B

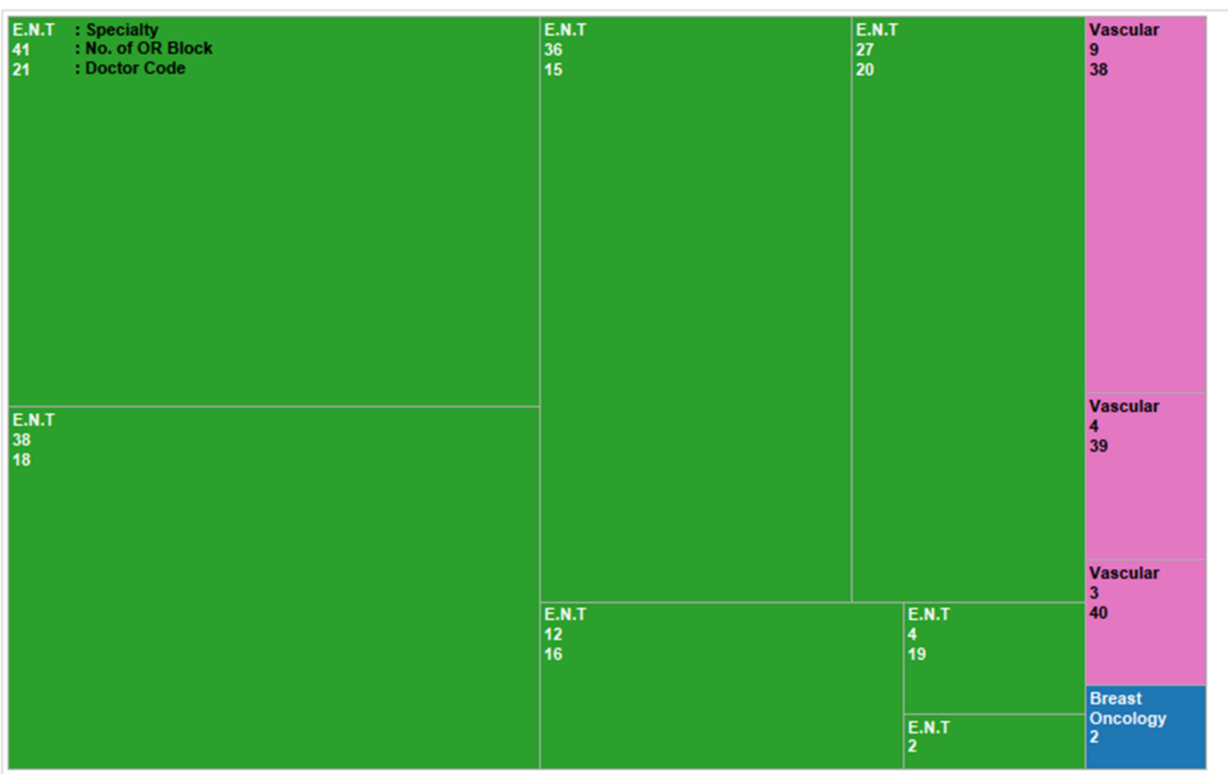


Figure 5-2: Strategic Allocation of OR Blocks to Divisions – Alternative B

Alternative - C

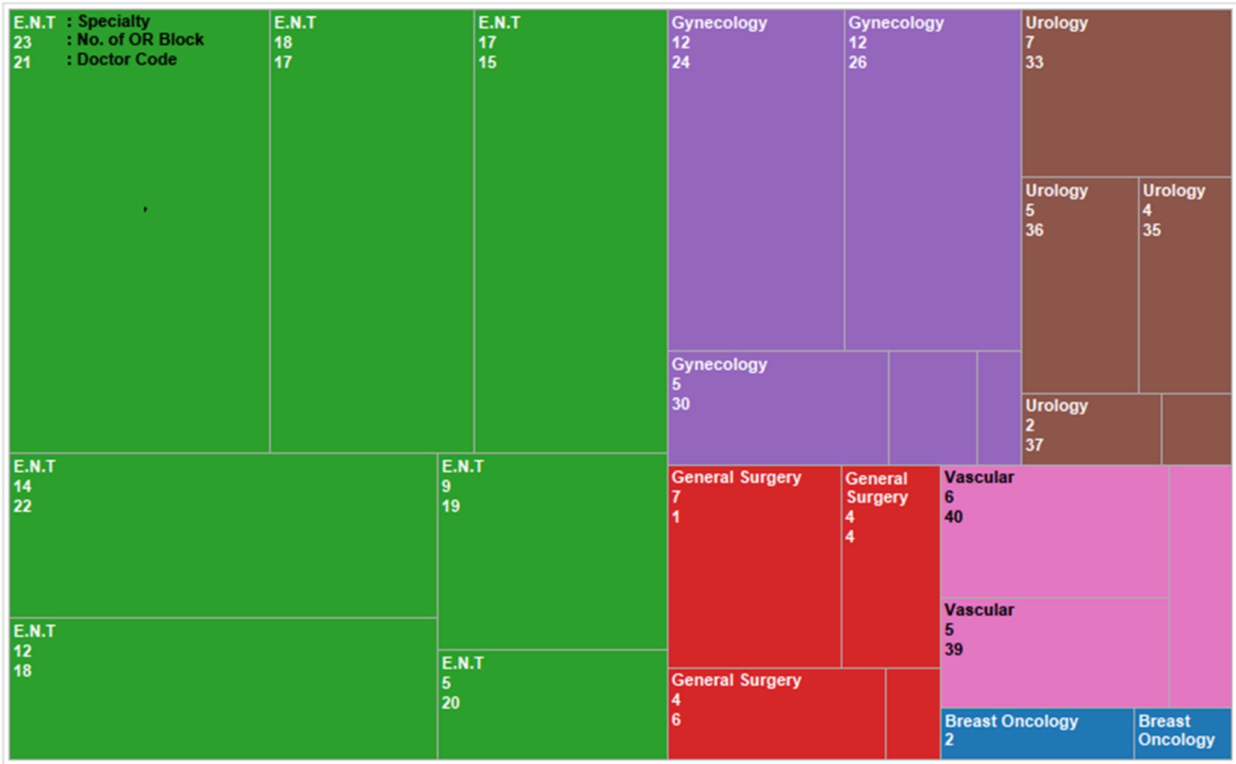


Figure 5-3: Strategic Allocation of OR Blocks to Divisions – Alternative C

Figure 5-4 illustrates the *Block Mix* schedules for all three alternatives. Out of 72 *Block Mixes*, 33 *Block Mixes* are not part of the optimal OR schedules for any of the alternatives. This means that 17 surgical procedures out of all 47 procedures are not part of the final solution. 6 of these 17 procedures are among the top 10% of procedures with the longest LOS. It was predictable that procedures such as the "Hartmann's procedure" and the "Resection sigmoid", which are considered in *Block Mix 19* and *Block Mix 20,* have an average 4.5 hours OR time and 10 days LOS, and are not the most attractive procedures for the managers. However, the "Resection abdomino-perineal with colostomy" is part of the optimal solution since the reimbursement rate is $7215, while the average OR time is 6.7 hours, and the average LOS is almost 10 days. In general, it is observed that the ISCM model tries to avoid scheduling any Block Mix with a high LOS and OR time. Figure 5-5 shows that *Block Mixes* 46 & 47 are the most attractive ones for alternatives A and B. These Block Mixes cover a group of “Thyroidectomy” procedures. The median LOS for these procedures is 1 day, and the average OR time is less than 1.4 hours. Its reimbursement rate is $3973, and there is a sufficient number of patients waiting for this service. This gives us an approximation of the marginal value of hospital resources such as ORs and surgical beds. Figure 5-6 illustrates the number of dedicated OR blocks for all procedures. And figure 5-7 presents the surgeons’ schedule for the first two weeks of the 6 week planning horizon for alternative A.

Specialty: Breast Oncology | Colorectal | E.N.T | General Surgery | Gynecology | Urology | Vascular

Mix / Alternative

	1	2	4	6	9		11	16	22	24	25	28	30	31	33	36	38	41	42		43	45			46			47			51			52	54	57		58	59		60	62	64	65		67	68	71			72		
Day	C	A	C	A	A	B	C	A	A	A	C	C	C	C	C	C	C	C	A	C	C	A	B	C	A	B	C	A	B	C	A	B	C	C	C	A	C	C	A	C	A	A	A	A	C	A	C	A	B	C	A	B	C
1							1													4			1		4	4			1				1															2					
2				1											4							1			4	2	2		3																				1				
3																1				4			1		3	2			2														1			1		1				1	
4		2								1										4			2						2		1	2		2										1		1							
5			1																									1	3					1											1						5	3	3
8			1												2							1			4	5			1						1	1	1													2			
9													1							1	1				2	3	2	3	3																	1				1			
10	1									1					3							1			2	3		1	3																2			1					
11												1								2		1			2	4			1		1		1									1		1	1				1				
12			1							1																		3	5					3	2									1	1	1						1	
15																3			1	1					5	2			2					1															2	1			
16					1										2					3					3	3	1	1	3																	1							
17	1		1					1			1				3		1					1	2		3	3			1													1											
18	1			1																4		3	1		2	3			1					1															1				
19	1	2								1																		2	6					2											1	1	2						
22	1					2												1				1			1	1		3	3									1										1		3			
23															2					2		1	2						1										3	1								2	3	1			
24																									5	5	4		1				1						1	1													
25	2								4						2					1		1	2			3			1														1		1								
26		1								1										1								2	6					2									1		1	1	2						
29															2								1		1	4	1									1	2											4		1			
30																				1	2				3	2		2	3			1			2			1								1							
31	1	2			2									3		1				1					1	3			1			1		1					1										1				
32		1													2							1				2		3	4	2										1					1	1							
33	2									1																		2	6					2	1						1		1		1								
36							1												1	3		1			3	3			3									1	1											1			
37							1					1			2										5	3	2					1														1			2				
38				1							1									1		3	2	2	1	2			1																1	1				1			
39		1													4							2	3		2				2			1										1			2								
40	2									2																		4	6						2										2								

Figure 5-4: Block Mix Schedule for Alternative A, B, and C

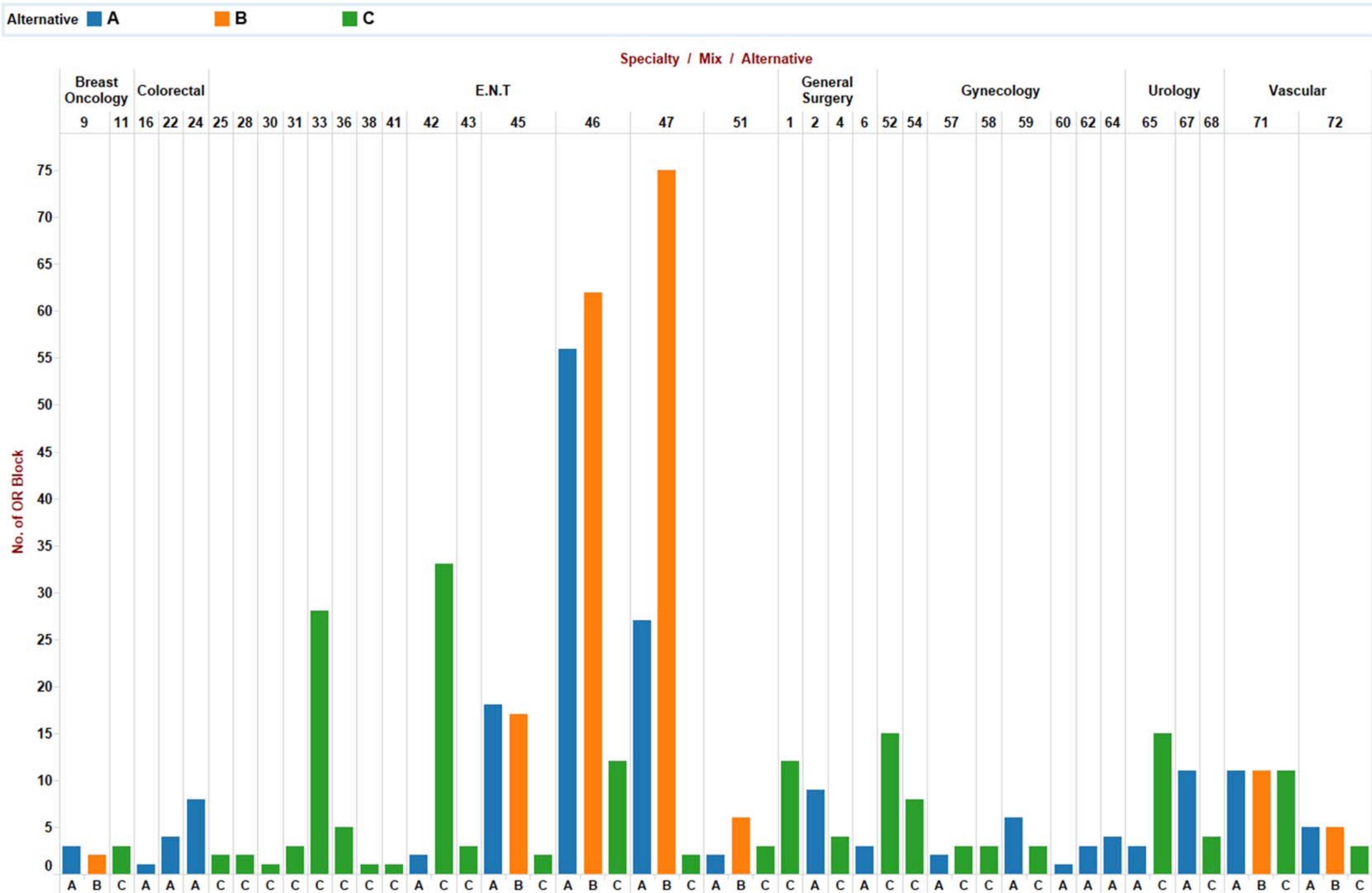


Figure 5-5: Total OR Blocks Dedicated to Each Mix for Alternative A, B, and C

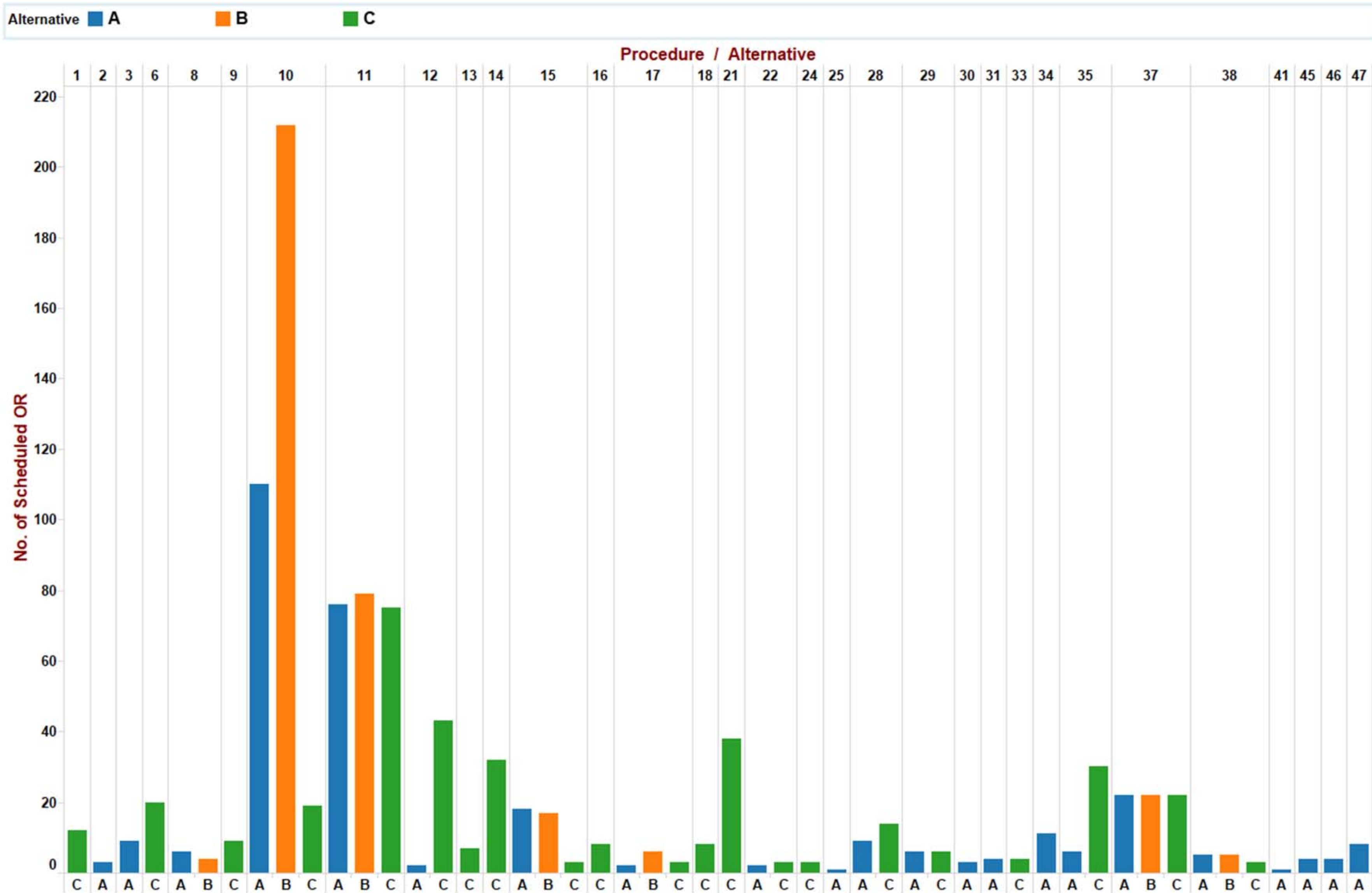


Figure 5-6: Total OR Blocks Dedicated to Each Procedure for Alternative A, B, and C

Specialty
Colorectal E.N.T General Surgery Gynecology Urology Vascular

Alternative / Mix

A

Day	Surgeon	2	6	24	45	46	47	51	57	62	64	65	67	71	72
1	15					2									
	18					2									
	40													2	
2	2		1												
	15					2									
	18					2									
	21				1										
3	15					1									
	21					2									
	25										1				
	31												1		
	40													1	
4	4	2													
	14			1											
	19							1							
	34											1			
	36												1		
5	18						1								
	38														2
	39														2
	40														1
8	15					2									
	20					1									
	21				1	1									
	24								1						
9	15						2								
	18						1								
	21					2									
	32												1		
10	13			1											
	15					1									
	18						1								
	21				1	1									
	39													1	
11	15					2									
	16				1										
	17							1							
	25									1					
	33											1			
12	14			1											
	15						1								
	18						2								
	31												1		
	35											1			

Figure 5-7: Surgeons' Schedule for the First 2 Week, Alternative A

5.3. Pooled Vs. Semi-Pooled Divisions

In the ISCM model, surgical beds are assumed to be shared among all patients from all specialties. This assumption is defined as pooled policy in chapter 3. Although many hospitals aspire to a pooled bed configuration, in reality certain divisions share their beds among themselves. This section studies the impact of a semi-pooled bed configuration policy on the ISCM problem. This approach results in a more realistic outcome from various stakeholders' perspectives.

JGH surgical beds are divided mainly into two segments. General Surgery and Colorectal divisions share surgical beds, denoted as "Group 1", and the rest of the beds are shared among other divisions, denoted as "Group 2". To apply the semi-pooled bed configuration policy on the ISCM model, the chance-constraint 4.10 is split into two chance-constraint sets. Each chance-constraint limits the probability of bed shortage for those divisions that share available beds. The total number of main ward beds, $B = 15$, is split between them with respect to each groups' demand distribution. Let $B_1 = 5$ and $B_2 = 10$ denote the total number of dedicated beds in groups 1 and 2 respectively, and Y_t^1 and Y_t^2 be the total number of beds occupied by the patients in these groups. So, the chance-constraint (4-10) is split into two chance-constraints as follows:

$$Pr\{Y_t^1 \geq \; B_1 \qquad \forall\, t \in T\} \leq \; \eta \tag{5.5}$$

$$Pr\{Y_t^2 \geq \; B_2 \qquad \forall\, t \in T\} \leq \; \eta \tag{5.6}$$

The final ISCM results demonstrate that the optimal value of the objective function does not significantly change under the semi-pooled bed configuration. Table 5-2 shows that the optimal

OR blocks dedicated to each specialty also remain the same under both bed configurations. Hence, it is necessary to know the impact of each bed configuration on surgical bed utilization.

Table 5-2: Optimal OR Block Distribution under Pooled and Semi-Pooled Bed Configurations

Bed Configuration	No. of Allocated OR Block to each Specialty						
	Breast Oncology	Colorectal	E.N.T	General	Gynecology	Urology	Vascular
Semi-Pooled	3	13	104	12	16	14	17
Pooled	3	13	105	12	16	14	16

Also, the SWD simulation model results show that the patients in groups 1 and 2 occupy 4.23 and 7.18 beds on average, respectively (Figures 5-8, 5-9). In total 11.42 beds are occupied on average, while this number is 10.88 under a pooled bed configuration. Figures 5-10 and 5-11 illustrate that the JGH needs 5 and 12 beds for groups 1 and 2 to reach 85% SLI. In other words, 17 beds are needed in total to serve 85% of the required bed-days. Also, the standard deviation of occupied beds for semi-pooled bed configuration is 4.55, while this number is 4.46 for pooled bed configuration. Further details on the distribution of the required number of surgical beds per day for groups 1 and 2 are illustrated in Figure 5-12 and 5-13 using a Whisker-Box plot. Our analysis demonstrates that, as expected, fewer beds are required under the pooled versus the semi-pooled bed configuration for the same service level. However, note that within the semi-pooled bed configuration, nurses with similar expertise are serving appropriate patients. In reality, all nurses are not trained for all specialties. So, our analysis shows that at most one more bed is required under the semi-pooled bed configuration. To compare these two bed configurations economically, one must study the cost and feasibility of training nurses to be able to serve all types of patients, and the cost and feasibility of adding a bed into the system.

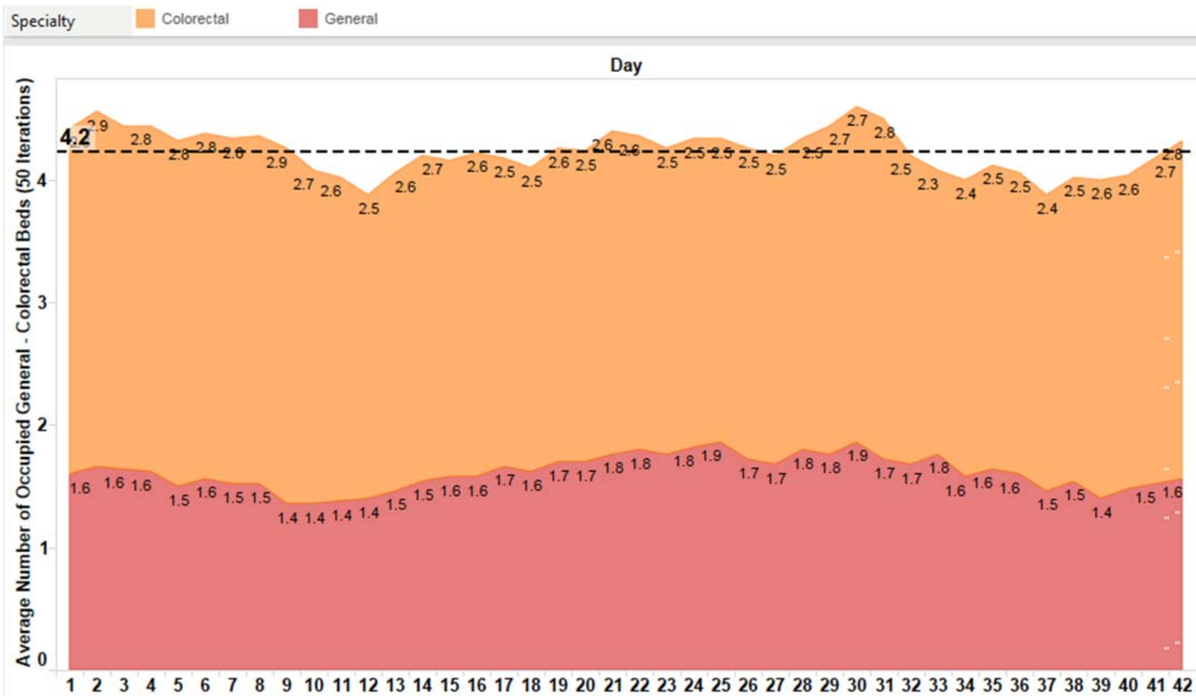


Figure 5-8: Average Occupied Surgical Beds – Group 1

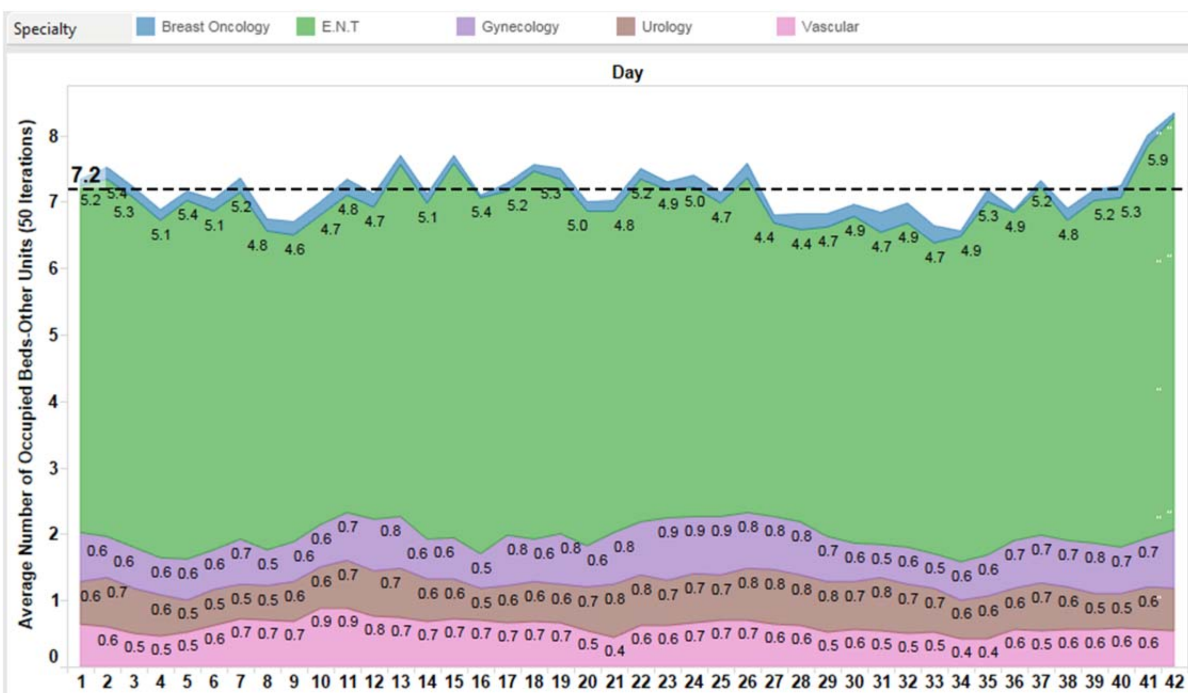


Figure 5-9: Average Occupied Surgical Beds – Group 2

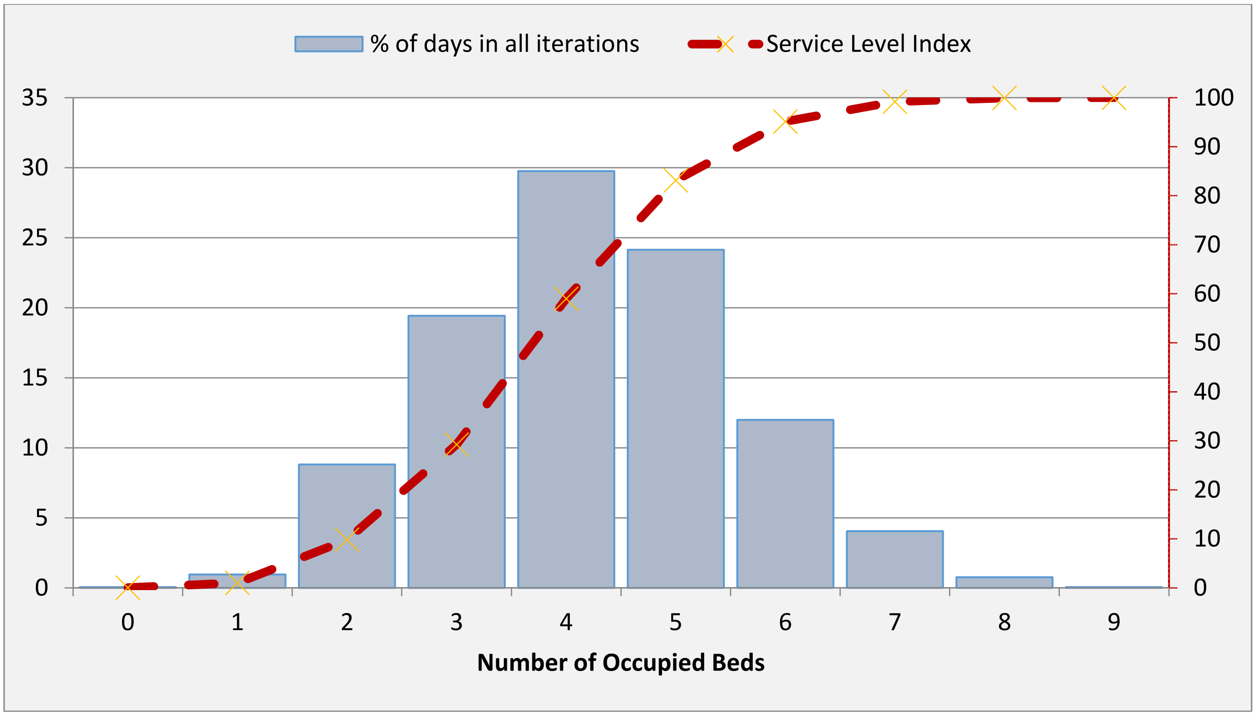


Figure 5-10: Required Bed Frequencies –Extended SWD Simulation Model for Group 1

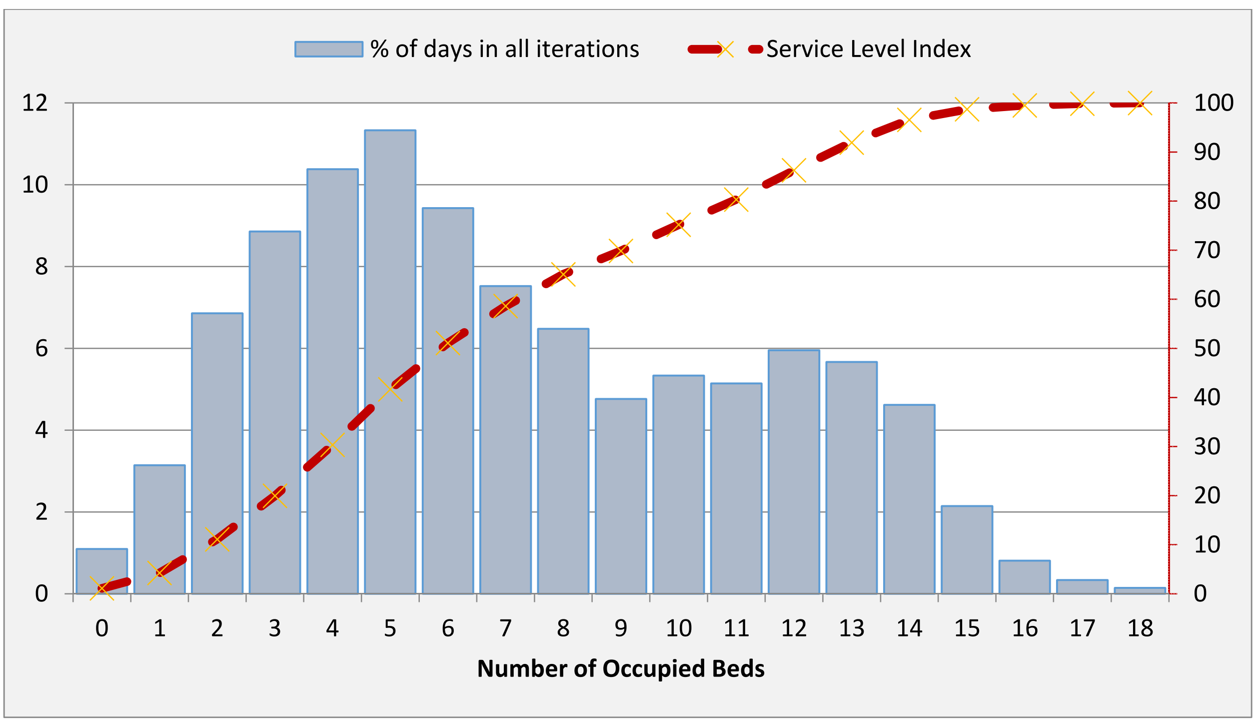


Figure 5-11: Required Bed Frequencies –Extended SWD Simulation Model for Group 2

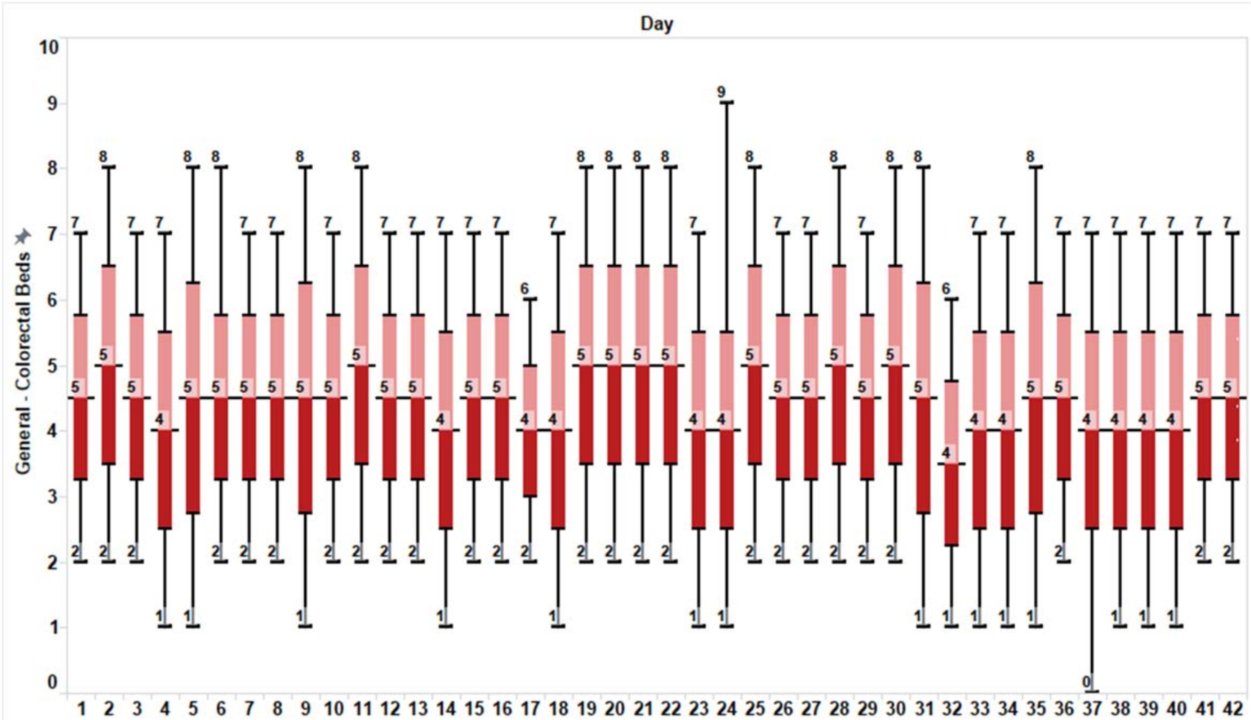


Figure 5-12: Whisker - Box Plot on the Number of Required Surgical Beds per Day – Group 1

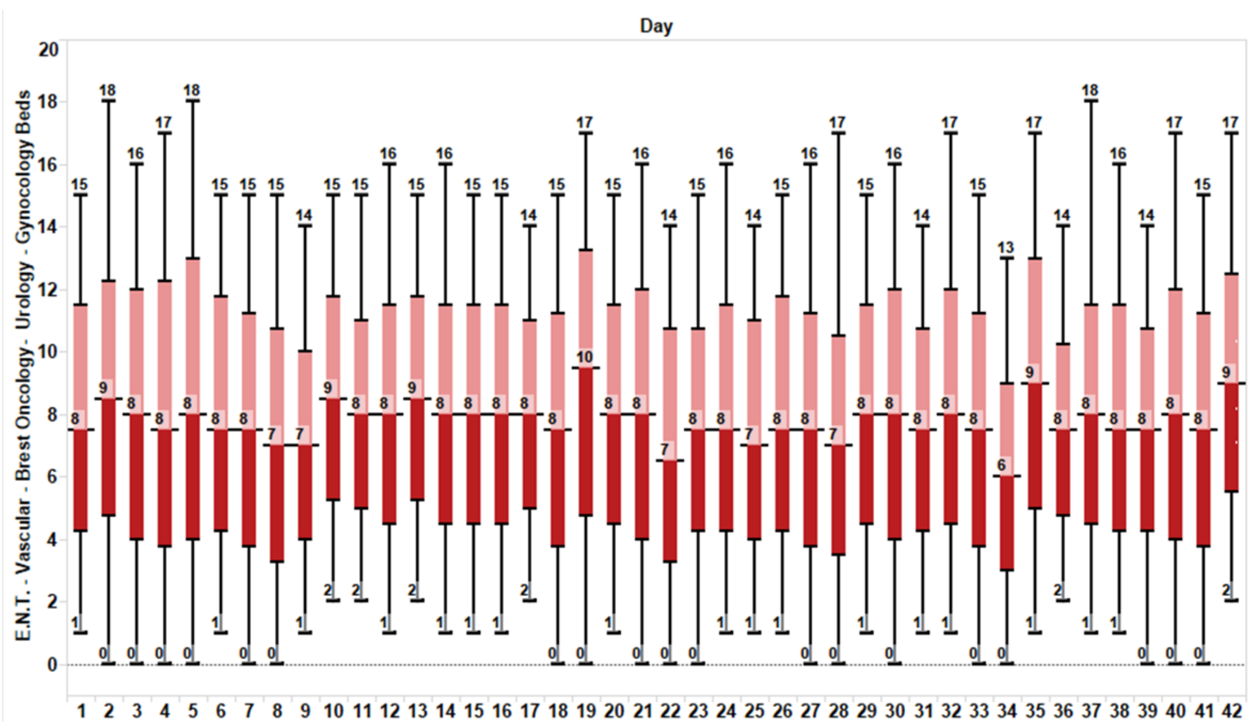


Figure 5-13: Whisker - Box Plot on the Number of Required Surgical Beds per Day – Group 2

5.4. Emergency Patients

The ISCM model tried to find the best surgical mix for the elective patients. It was assumed that the total number of beds is fixed for these patients. However, many hospitals serve emergency patients as well as elective cases. To build a realistic model, the impact of emergency patients on the hospital's surgical mix must be considered. Serving emergency patients results in uncertainty in the available inpatient beds for elective patients each day. So, total number of surgical ward beds for elective patients, denoted as B_e, has to be a stochastic variable in the ISCM model, which makes the model even more complex.

Emergency patients also affect the availability of other resources such as ORs, nurses, and surgeons. When there is no trauma service (as in our case, JGH), a significant majority of the emergency patients do not need to go to OR immediately after getting admitted to the hospital. Such hospitals often reserve a certain number of OR blocks only for emergency patients. Also, nurse ratio mostly depends on the total number of beds not the type of patient's entry, emergency or elective. So, this section only focuses on the impact of emergency patients on the surgical ward beds and elective patient scheduling.

This study only focuses on 72 main surgical procedures operated in JGH. So, only emergency cases for the same set of procedures are considered. On the basis of historical data it is observed that 34 emergency patients on average are admitted to the hospital in 6-week under these 72 procedures. Since the number of emergency patients is not large enough to study each procedure individually, all emergency patients are considered as a single group.

Figure 5-14 illustrates the arrival rate and pattern of the emergency patients. Almost all of these patients are operated during the week. On Tuesdays the least number of emergency cases transfer from the ED to the surgical unit, while the surgical unit experiences the highest rate on Wednesdays and Thursdays. Also, historical data shows that almost always less than three emergency patients are operated on per day. This histogram is used to generate 20 Arrival Scenarios for emergency patients during the planning horizon of six weeks.

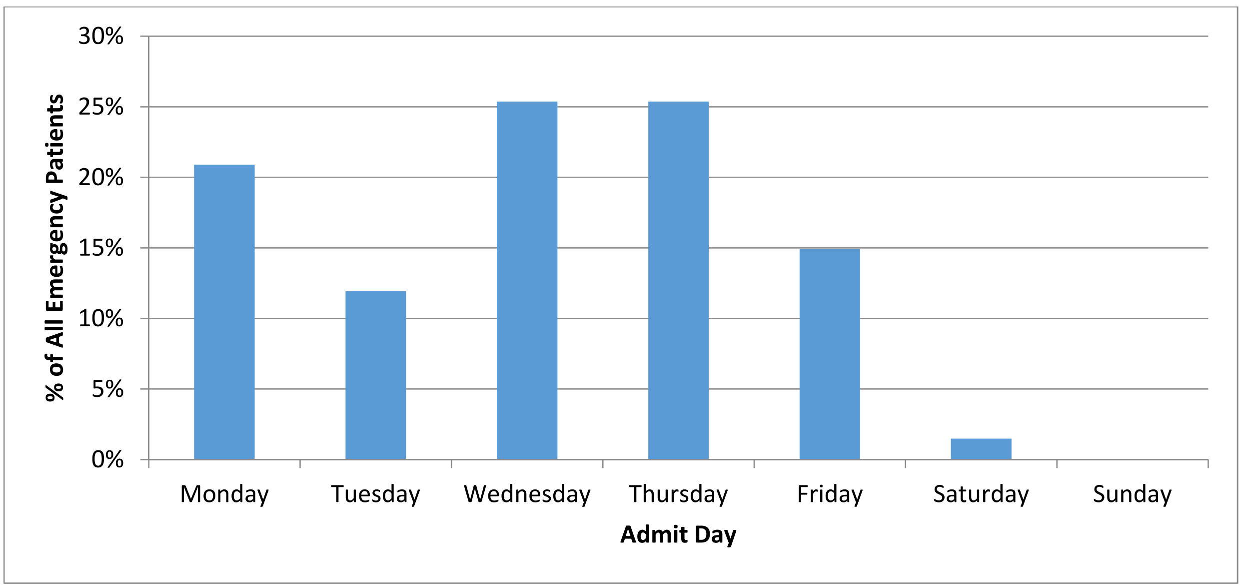


Figure 5-14: Emergency Patients Arrival Pattern

Then emergency patients LOS distribution is explored. As illustrated in Figure 5-15, less than 20% of patients have LOS greater than 15 days, so the truncate rule is applied here while it is assumed that all these 20% of emergency patients stay at the hospital for 16 days.

To include emergency patients into the ISCM model two OR *Block Mixes* are added to the model, the first one considers one emergency patient and the second one considers two emergency patients. For each emergency OR *Block Mix*, different bed occupancy scenarios are generated. To build the scenarios, first, all possible bed occupancy realizations of each emergency *Block Mix* are generated. Then, one realization per Arrival Scenario of the emergency patients is randomly selected.

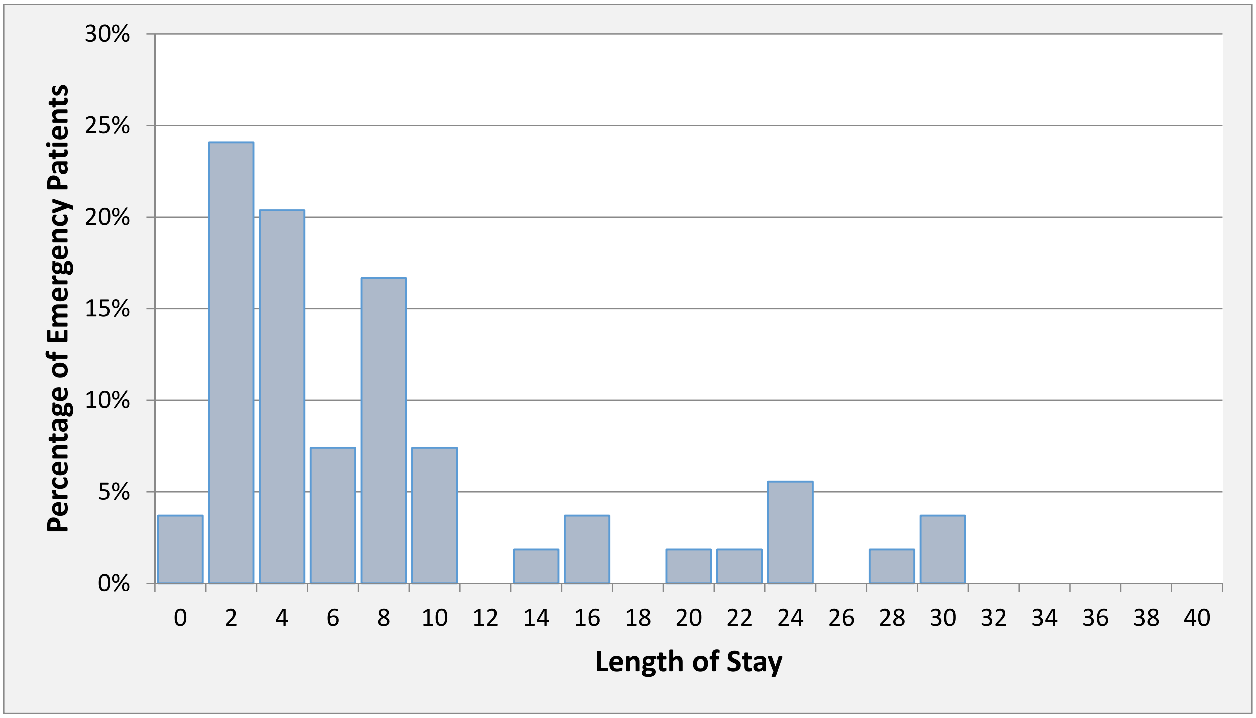


Figure 5-15: Emergency Patients Length of Stay Histogram

To run the ISCM model constraint 4.4 must be replaced with:

$$\sum_{s=1}^{S} \sum_{m \in M_s} \sum_{i=t-L_m}^{t} a_{sm(t-i+1)} x_{smi} \leq B - B_{et} \qquad \forall\, t \in T \qquad (5.7)$$

While B_{et}denotes number of occupied beds by emergency patients on day t. Now the ISCM model is run with 15 beds, shared between elective and emergency patients. Given the optimal OR schedule, the SWD simulation model is used to find its impact on the surgical bed occupancy. It is observed that the average number of occupied beds is 4.6 for emergency patients and 8.0 for elective cases. Note that, not all of the OR blocks are used when only 15 beds are dedicated to the emergency and elective cases, so the ISCM schedules fewer elective patients than the elective-only base model, hence, fewer number of beds are required to serve them. Figure 5-16 and 5-17 illustrate more details on the average occupied bed by each surgical specialty.

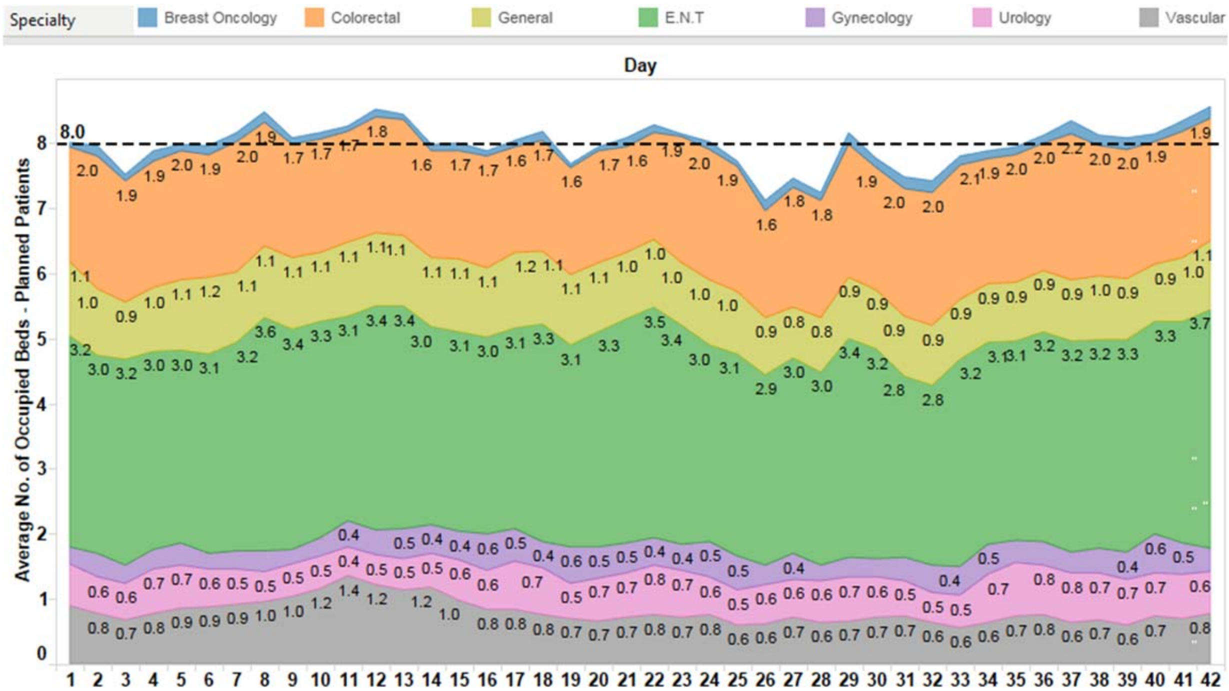


Figure 5-16: Average Occupied Surgical Beds – Elective Patients

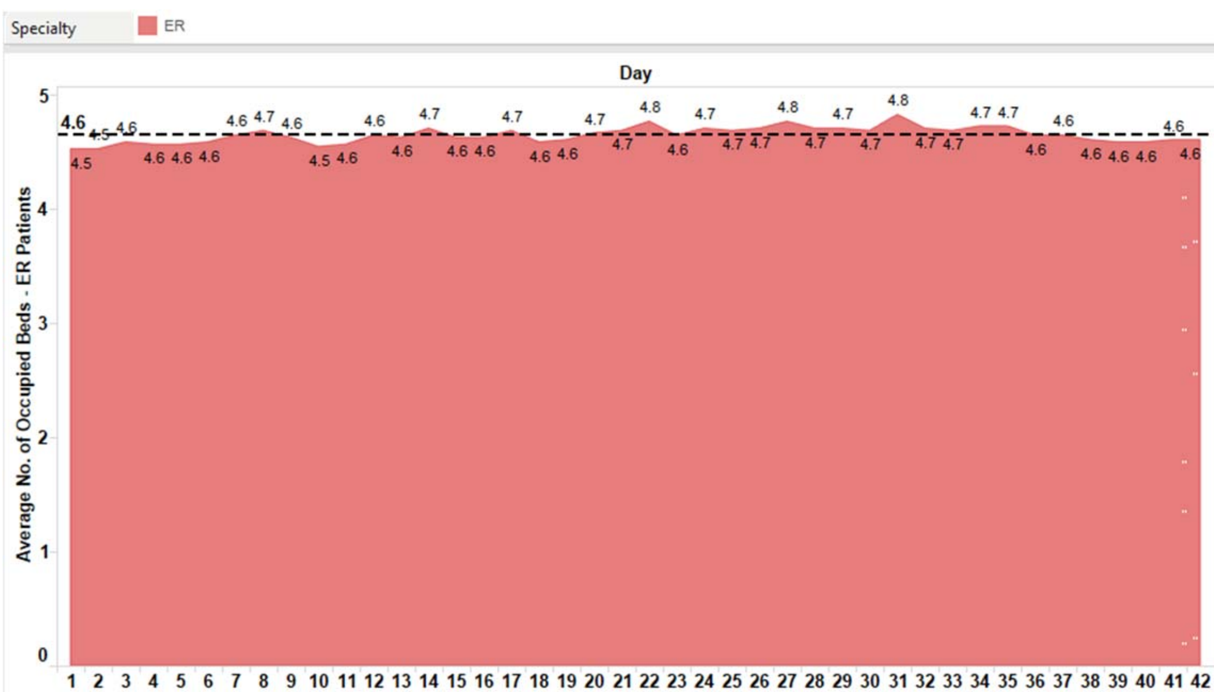


Figure 5-17: Average Occupied Surgical Beds – Emergency Patients

Furthermore, figures 5-18 and 5-19 present the SLI for elective and emergency patients. To reach 85% SLI, 11 surgical beds must be equiped for elective patients and 5 surgical beds for emergency patients. As explained in section 4.5.2, note that the truncate rule was applied to calculate LOS distributions for the ISCM model and the LOS was assumed to be always fewer than 10 days, so the SWD simulation model results shows that 16 beds are required in total instead of 15 beds.

The optimal result of the ISCM shows that only 144 out of 180 OR blocks were utilized, and the objective function decreased from $1212980 to $847800. We increase the total number of available beds, denoted as B, and rerun the ISCM to analyze the marginal value of each bed. Results show that 23 more OR blocks are utilized when B shifts from 15 to 17 beds. Yet, only 10 more OR Blocks are scheduled when B shifts from 17 to 19 beds. And finally, the surgical unit is able to use 2 more OR blocks if it opens one more bed. The results show that the marginal value of each bed is zero when there is 21 beds. Table 5-3 presents further information on the ISCM optimal results under various scenarios on B. It is observed that the dedicated OR Blocks to almost all divisions except E.N.T. and Vascular do not change when B shifts from 15 to 20 beds.

Table 5-3: Surgical Ward Bed Sensitivity Analysis – ISCM with emergency Patients

No. Of Available Beds	No. of Utilized OR Blocks	Dedicate OR Blocks							Objective Value (1000$)
		Breast Oncology	Colorectal	E.N.T	General	Gynecology	Urology	Vascular	
15	144	3	13	66	12	17	14	19	847.8
17	167	3	13	86	12	16	14	23	1030.1
19	177	3	13	100	12	16	14	19	1159.5
20	179	3	13	105	12	16	14	16	1202.0

This occurs since the model has to dedicate a certain number of OR Blocks to each specialty. But, the ISCM dedicates any unutilized OR Blocks to the most profitable Block Mixes, which are under E.N.T. and Vascular surgery, when there are some extra surgical beds available in the ward.

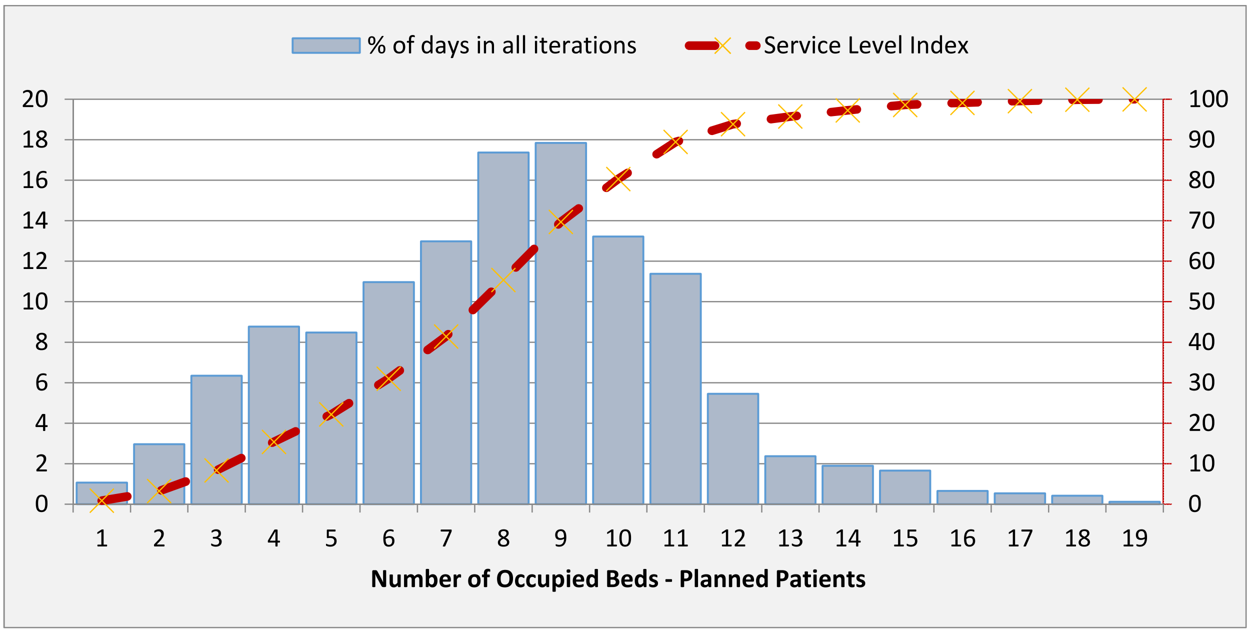


Figure 5-18: Required Bed Frequencies –Extended SWD Simulation Model for Elective Patients

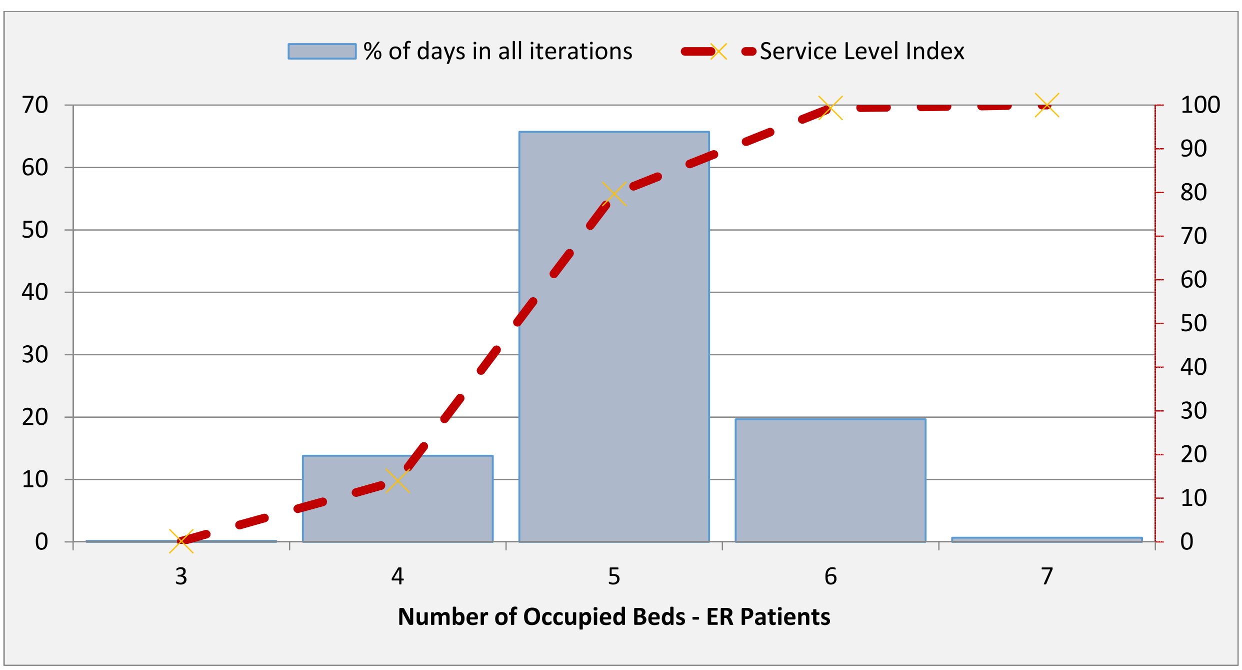


Figure 5-19: Required Bed Frequencies – Extended SWD Simulation Model for Emergency Patients

5.5. Off-Service Patients

The bed shortage problem is not limited to the surgical unit. Patients might admit or transfer to an inpatient bed regardless of their main service group (e.g. surgical, medicine). As explained in the ISCM model, it is assumed that the surgical patients might use other units' beds for $\eta\%$ (e.g. η =15%) of the planning days. However, the opposite usually happens to the surgical unit. In other words, for example the medicine unit transfers its patients to the surgical unit when there are not enough beds in that unit and some beds are free in the surgical unit. This section will study the impact of these patients on the surgical unit bed management.

It is assumed that there are always 3 non-surgical patients ready to transfer to the surgical unit if there are unoccupied beds. Let o $(o \leq 3)$ show the number of off-service patients transferred to the surgical unit, and $Y_{ot} = 1$ if o off-service patients are transferred to the surgical unit at day t. Also, b_{ot} denotes the required number of beds t days after the date that o number of off-service patients are transferred to the surgical unit. It is assumed that the surgical unit charges other units for each day per off-service patient. Let ϑ represent this rate, and ϑ is assumed to be equal to the bed shortage cost.

In addition, three off-service admission policies are considered. The first policy allows the off-service patients to stay one day at the surgical unit. The second and third policies extend this number to 2 and 3 days at most. So, let q show the maximum number of days that an off-service patient is allowed to stay in the surgical unit. To consider off-service patients in the ISCM model, constraints (5.9) and (5.10) must be added to the model and replace constraint

(4.4) with constraint (5.8). This constraint counts the occupied beds by all off-service and surgical patients.

$$\sum_{s=1}^{S}\sum_{m\in M_s}\sum_{i=t-L_m}^{t} a_{sm(t-i+1)}x_{smi} + \sum_{i=t-q}^{t}\sum_{0=0}^{O} b_{o(t-i+1)} * Y_{ot} \leq \; B \quad \forall\, t \;\in T \qquad (5.8)$$

$$\sum_{o=0}^{3} Y_{ot} \leq 1 \qquad \forall\, t \;\in T \qquad (5.9)$$

$$Y_{ot} \in \{0,1\} \qquad \forall\, t, o \;\in T \qquad (5.10)$$

Also, the objective function must be extended by adding the off-service bed charges, so:

$$Max \qquad (5.11)$$

$$\sum_{k=1}^{K} R_k * [\sum_{t=1}^{T}\sum_{m=1}^{M}\sum_{s=1}^{S} N_{km} * x_{smt}] - \sum_{s=1}^{S}\sum_{m\in M_s}\sum_{t=1}^{T}[IO_m * x_{smt}] -$$
$$SC * \sum_{t=1}^{T} E[Y_t - B]^+ + \vartheta * \sum_{i=t-q}^{t}\sum_{0=0}^{O} b_{o(t-i+1)} * Y_{ot}$$

Since the length of stay of an off-service patient is stochastic, it is assumed that each off-service patient will either discharge home or transfer to his origin unit with a daily probability of β (e.g. $\beta = 0.1$). The surgical unit admits at most 3 off-service patients per day, so, 4 Arrival Scenarios must be considered. Different bed occupancy scenarios need to be generated for each Arrival Scenario. To this aim, first, all possible bed occupancy realizations of each state are generated, and then one realization is randomly selected per Arrival Scenario of the off-service patients. Then the proposed SAA methodology is applied, to solve the ISCM model.

Final results do not show a significant change in the objective value of the ISCM model. In other words, it is observed that the off-service admission policy does not negatively impact the surgical unit reimbursement amount (as long as they can charge the sending unit an appropriate amount). Figures 5-20, 5-21, and 5-22 illustrate the volume of off-service patients, transferred to the surgical unit under 3 different admission policies during a 6-week planning horizon. With respect to the first off-service admission policy, figure 5-20 shows that the surgical unit usually admits one off-service patient each day. However, there are three days that the surgical unit beds are fully occupied by the surgical patients, so no off-service patient is admitted on those days. As the model allows off-service patients to stay more than one day in the surgical unit, the number of admitted patients decreases under the second and third policies.

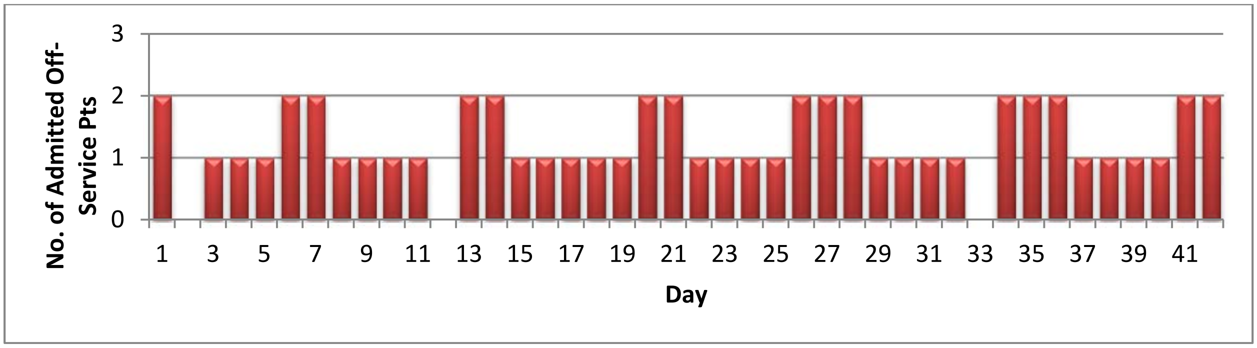


Figure 5-20: Off-Service Patients Arrival Schedule under Admission Policy No. 1

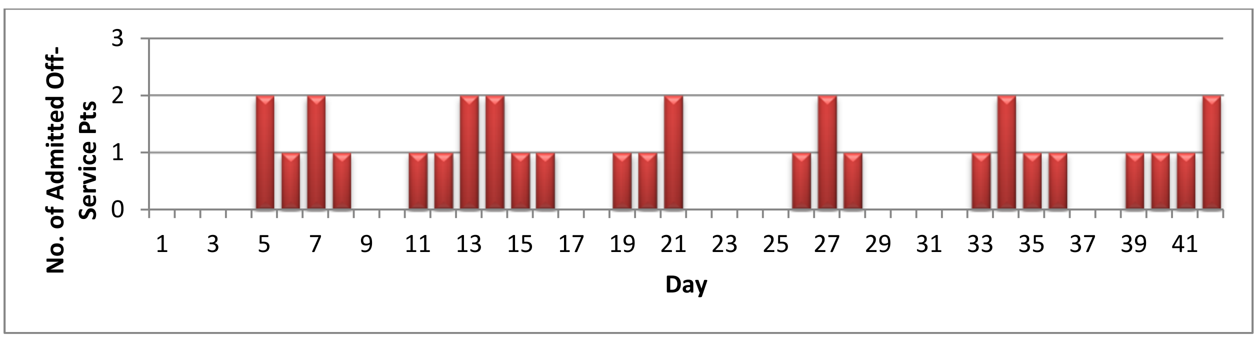


Figure 5-21: Off-Service Patients Arrival Schedule under Admission Policy No. 2

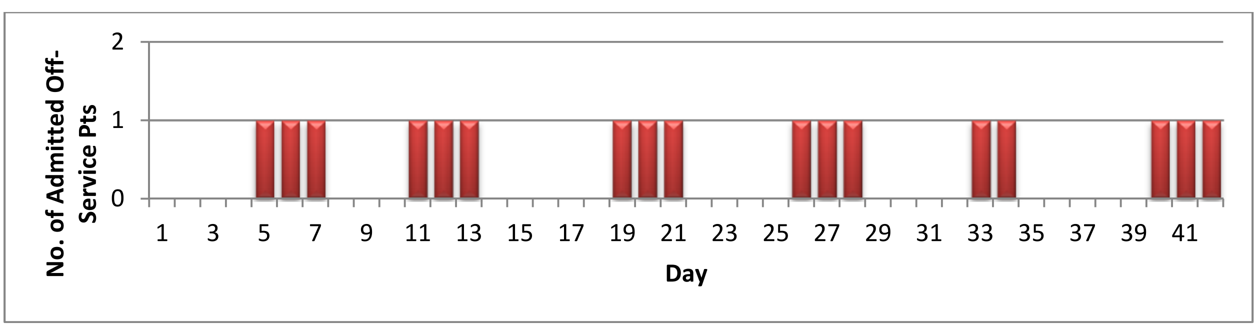


Figure 5-22: Off-Service Patients Arrival Schedule under Admission Policy No. 3

Figures 5-23, 5-24, and 5-25 illustrate the daily distribution of off-service patients' admission under all admission policies. These figures aggregate the admission schedule of off-service patients within a six-week planning horizon. For example, Figure 5-23 shows that for four Mondays in the period of six weeks the surgical unit should admit one (colored in orange) off-service patient and for the rest two Mondays it admits two (colored in red) off-service patients. It is also observed that under Admission Policy No. 1, for all six Saturdays and six Sundays in our planning horizon, two off-service patients are admitted. In all policies, admission rate is higher on the weekends rather than weekdays. On Tuesdays, Wednesdays and Thursdays the results suggest to admit no off-service patients under second and third policies, and at most one off-service patient under the first policy.

The surgical unit should admit 39, 23, and 17 off-service patients under first, second and third policy respectively. The ISCM model expects 54, 58.9, and 46.7 bed-days to serve these off-service patients respectively. Although the surgical unit allows off-service patients to stay 3 days under the third policy, it requires the least number of bed-days among all policies. This policy stops the model to admit off-service patients on weekdays even if there is a free bed, since the surgical unit needs that bed for future surgical patients.

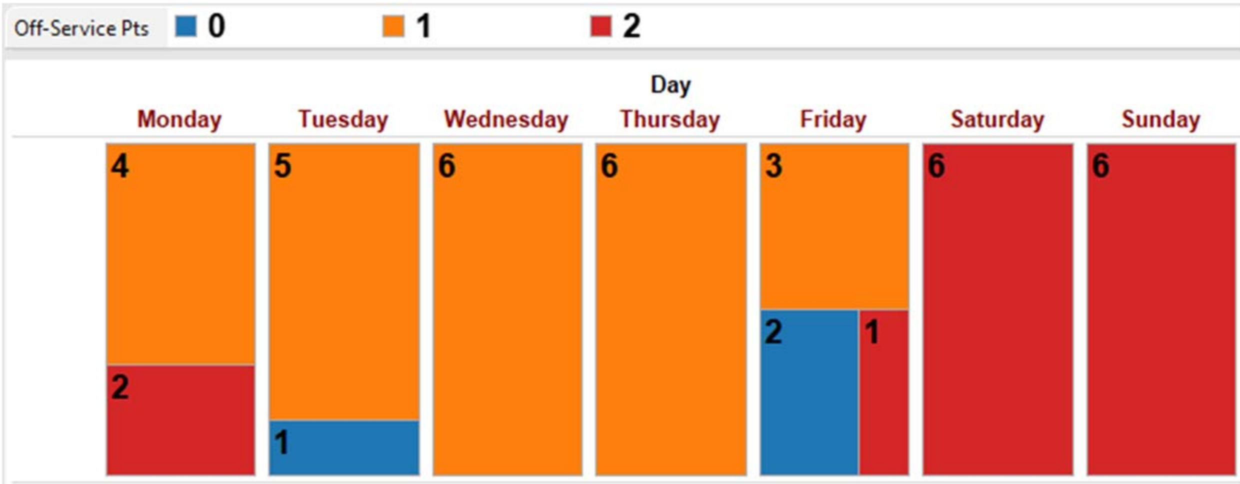


Figure 5-23: Daily Distribution of Off-Service Patients Arrival under Admission Policy No. 1

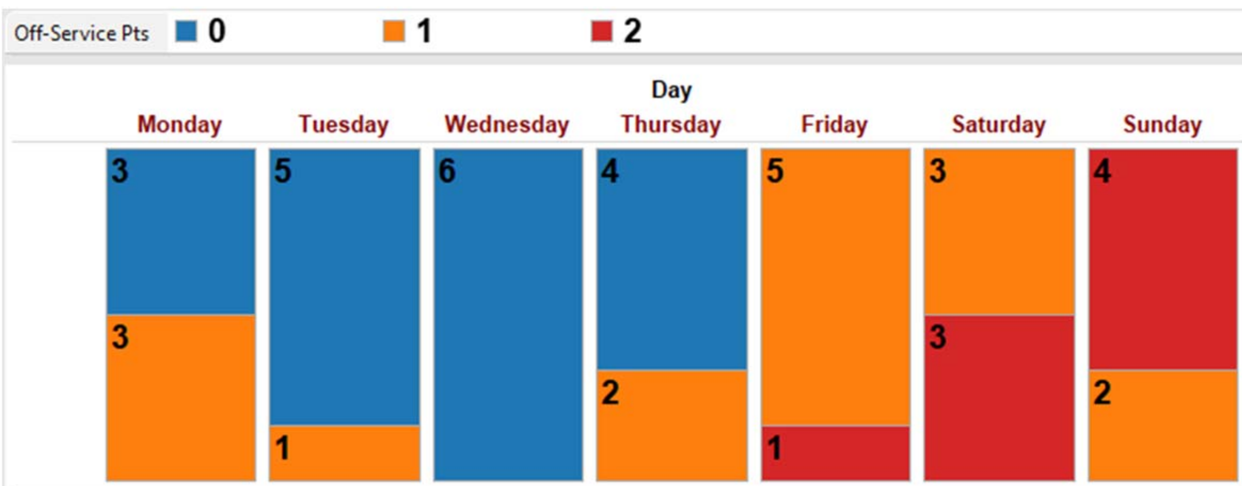


Figure 5-24: Daily Distribution of Off-Service Patients Arrival under Admission Policy No. 2

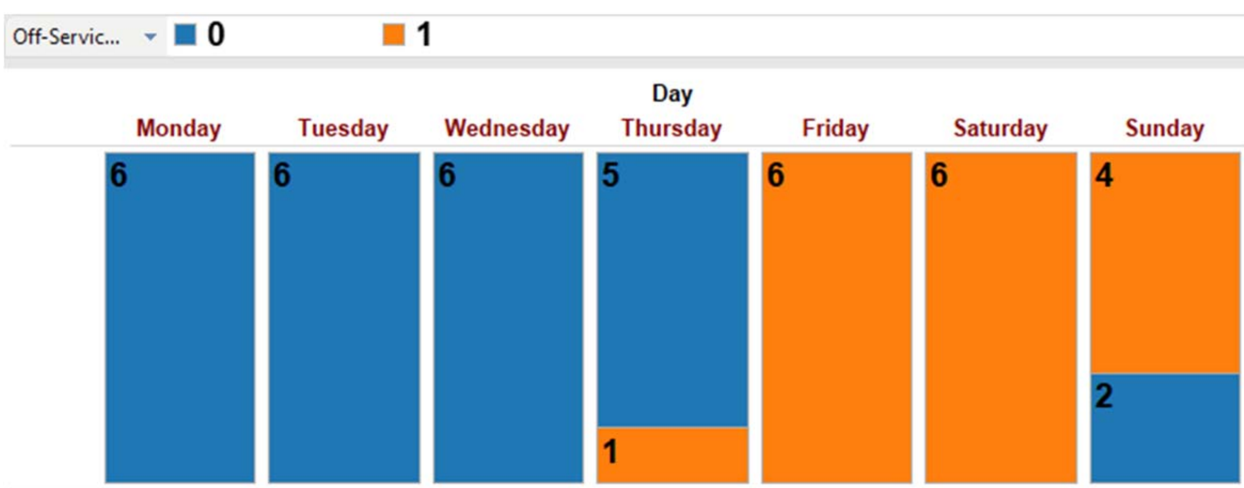


Figure 5-25: Daily Distribution of Off-Service Patients Arrival under Admission Policy No. 3

5.6. Conclusion

In this chapter the ISCM model is extended to address the impact of external factors on the surgical case mix problem. To help the policy makers to respond governmental funding policy in a long term, trade-offs analysis are performed among three funding alternatives i) activity based funding (ABF), ii) global budget with incentive (GBI), and iii) global budget with incentive and penalty (GBIP). Two scenarios were assumed, either the policy makers are completely flexible (i.e., Relaxed Problem) or they have to apply a lower bound (i.e., Constrained Problem) on the OR allocation among various divisions and surgeons. Also, it is assumed that the total demand for each surgical procedure cannot go beyond a threshold. The results demonstrate that the hospital managers would similarly respond to the ABF and GBIP policies under both scenarios. Furthermore, from the government perspective, a global budget range is found which results in the same surgical case-mix under the GBIP policy compared to the ABF policy. The ISCM results (e.g., OR Block allocation among all divisions and the *Block Mix* schedules at the strategic, technical, and operational levels) under all aforementioned settings were illustrated in this chapter. It was observed that 33 out of 72 *Block Mixes* are not part of the optimal OR schedules for any of the funding policies. This translates to 17 surgical procedures out of all 47 procedures. Yet, only 6 of these 17 procedures are among the top 10% of procedures with the longest LOS. In general, it was found that more than 50% of the OR blocks are assigned to the E.N.T. procedures under the constrained scenario. This was because they bring the highest funding to the surgical unit. The Colorectal Surgery, the General Surgery, the Gynecology, and the Urology are not very attractive divisions from a financial perspective. However, not all surgical procedures are addressed in this study since there was no access to the full data on them, and this might bias the results presented.

As the second factor, the impact of bed configuration policy on the ISCM model was evaluated. The results demonstrate that the semi-pooled bed configuration increases the daily bed occupancy variance. That also slightly increases the number of required beds for the same SLI under pooled bed configuration. However, a minor impact was observed on the value of the objective function and the optimal surgical mix. So, it is recommended to extend this study to explore all advantages and disadvantages of various bed configuration policies from managerial and medical perspectives.

The ISCM model was also extended to address emergency patients which brought uncertainty in the availability of the surgical beds. The results showed that about one third of the surgical beds are occupied by these patients who prevent all OR blocks to be utilized. Hence, it is concluded that the surgical unit has to add 5 more beds to fully utilize OR blocks and serve the elective patients at 85% SLI.

At the end, off-service patients were incorporated in the ISCM model. To smooth patient flow across all units (e.g. medical unit), different off-service patient admission policies were proposed. In all policies, the optimal results suggest a higher off-service patient admission on weekends and Fridays rather than the rest of the week since the off-service patient LOS is stochastic and they might stay more than one day in the surgical unit.

Chapter 6

Conclusion & Future Study

6.1. Conclusion

This study systematically approaches surgical unit management. The surgical case-mix problem is addressed at the strategic level while we control its impacts on the surgical unit's patient flow at the operational level. The study was initiated after lots of meetings with surgical unit managers, surgeons, and nurses at the Montreal Jewish General hospital, talking about their main operational concerns there. A comprehensive literature review was conducted on the key challenges faced by hospital managers to support, plan, and improve the surgical unit efficiency and efficacy. The literature was classified into five main domains: Case-Mix Planning problems, Operating Room oriented problems, ICU/ PACU oriented problems, bed management problems in the main ward, and care providers-oriented issues. To the best of our knowledge the main concerns of the stakeholders at JGH are not studied as an integrated model in the literature.

Chapter 3 presented the Surgical Ward Design (SWD) simulation model, a planning tool to study the impact of different bed configuration approaches, OR schedules, and actual LOS of the patients on the surgical unit occupancy rate. The Service Level Index (SLI) was defined to capture the ratio of downstream surgical beds unavailability. Using SLI, we compare pooled, semi-pooled, and dedicated bed configuration policies. It is concluded that the JGH is able to save up to 19% on the total number of required beds under the pooled policy in comparison with the dedicated policy at 90% SLI. Furthermore, the impact of the High Acuity Care Unit on the surgical unit patient flow was studied. Several scenarios were developed on high-risk patients' arrival rate to visualize the patient flow under various possible circumstances.

In chapter 4, an integrated approach to the surgical case-mix problem was developed. In the Integrated Surgical Case-Mix (ISCM) model, the operational and tactical details are imbedded in a strategic decision. In other words, the developed model copes with the functionality of the derived strategic case-mix in the following Operating Room Scheduling problem. This novel approach helps hospital managers to make strategic decisions that are indeed feasible at the operational level.

At the operational level, we defined the Operating Rooms' *Block Mix, a* feasible set of surgical procedures one specific surgeon can schedule within a 4-hour OR block. At the tactical level, we decided how to allocate divisions' resources to these predefined *Block Mix*; and at the strategic level, the ISCM decided on how to dedicate surgical unit resources to the divisions within a 6-week planning horizon. To this end, a multi-dimensional objective function was defined to tackle the reimbursement mechanism of the surgical unit, the ORs' utilization, and downstream bed utilization. The ISCM was calibrated with JGH data on 40 surgeons, and 87 surgical procedures under 7 specialties (i.e. General Surgery, Breast Oncology, Colorectal, E.N.T., Gynecology, Urology, Vascular). Figure 6-1 illustrates more detail on the main components of the ISCM model.

In chapter 5, the ISCM is extended based on three reimbursement policies: Activity-Based Funding (ABF), Global Budget with Incentive (GBI), and Global Budget with Incentive and Penalty (GBIP). It is logical that the ISCM resulted in the same surgical case-mix under ABF and GBIP. We observed that the selected surgical procedures in the colorectal surgery, the General

Surgery, the Gynecology, and the Urology are not among the most attractive ones. E.N.T. procedures bring the highest funding to the hospital and using the least portion of resources. Also, the Vascular Surgery is ranked as the second most important division of the hospital. However, under GBI funding the priority goes to those procedures listed under the hospital's commitment to the government. However, the study focused on the procedures' reimbursement rates and the cost of each procedure is not considered since JGH was not able to provide us with the data. As the main limitation of this study, that might impact on the marginal revenue of the optimal surgical case-mix. However, the ISCM model is completely capable of considering procedures' costs in future studies.

Figure 6-1: Key Components of the Integrated Surgical Case-Mix Model

The ISCM model was extended to address other key levers of the surgical unit bed shortage. We studied the impact of semi-pooled bed configuration, emergency patients, and off-service patients on the ISCM model's outcomes. The results demonstrated that the optimal value of the objective function as well as the optimal case-mix do not significantly change under semi-pooled versus pooled bed configuration. However, fewer beds were required under pooled versus semi-pooled bed configuration for the same SLI. Also, the bed shortage level and the standard deviation of occupied beds increased under semi-pooled bed configuration policy.

Furthermore, to address emergency patients, the ISCM model was developed by considering their arrival pattern, LOS distribution, and OR schedules based on JGH historical data. It was assumed that some OR blocks are reserved for the emergency patients and they just share surgical beds with elective patients. The final results confirmed that the surgical unit requires five more beds (i.e. in total 20 beds) to serve all emergency and elective cases. When the number of surgical unit beds was 15 beds in total, JGH could not utilize 36 out of 180 available OR Blocks due to the emergency patients. These 36 OR Blocks were mostly occupied by the most profitable patients from E.N.T. and Vascular divisions when emergency patients are not considered in the ISCM model.

The study was completed by considering the impact of off-service patients on the surgical unit's patient flow. Since the LOS of an off-service patient is stochastic, three admission policies were defined, which limit the maximum number of their stay in the surgical unit to 3, 2, and 1 days. The impact of these three admission policies was evaluated on the ISCM outcomes. The results show that the off-service admission policy does not have a significant negative impact on the surgical unit reimbursement amount. However, it was observed the admission of more off-service patients per day under the 1 and 2-day admission policies versus the 3-day policy. Also, the 3-day admission policy has the least amount of occupied bed-days among all policies. So, it was recommended to the surgical unit managers to take the 2-day admission policy.

From modeling perspective, the application of chance-constraint programming on the surgical case-mix problem is developed in this study for the first time to the best of our knowledge. This approach could explain and illustrate the required amount of surgical beds as well as the real probability of bed shortage. So, the stochastic ISCM model limited the probability of downstream bed shortage through a set of chance-constraints. The model was calibrated with patients' real LOS distribution pulled out from JGH historical data.

Furthermore, we implemented the concept of CVaR in the ISCM model to minimize the risk of high downstream bed shortage. It was concluded that the integration of the CVaR term (in the objective function) and the chance-constraints results in robust ISCM's outcomes regardless of the actual bed shortage cost.

Furthermore, the stochastic ISCM model was linearized and the Sample Average Approximation (SAA) technique was applied to solve the model. Then the SWD simulation model was extended (on the basis of empirical probability functions for each procedure LOS distribution and the patient flow framework for 7 specialties and 72 procedures) to validate and verify the final results of the ISCM. The SWD simulation model was fed by the optimal case-mix and OR schedule to illustrate the surgical unit's 6-week bed occupancy pattern.

6.2. Future Study

In this study, we demonstrated that the main reason for cancelation of surgical procedures is the downstream bed unavailability. We believe that neither capacity expansion nor early discharge of patients is the best approach to manage surgical bed unavailability. We propose that the underlying problem can be studied from three perspectives: First, it is related to the integration of two important decisions: a. allocation of operating room time blocks to surgical specialties and individual surgeons (at the operational level), and b. planning on the surgical case-mix (at the strategic level). This has been studied in this dissertation.

The second perspective is the poor cooperation of surgical/ medicine units with other care provider centers such as residential care sites. Surgical patients, who are assessed and approved for either residential care services or home support services, have to spend lots of days waiting for a free space. They occupy surgical beds while they do not need that service anymore. We need to investigate the impact of such collaboration on surgical unit patient flow in the future.

The third perspective, which is the incorporation of various incompatible stakeholders' (i.e. hospital administrators, and surgeons) incentives and actions in the planning process, is yet to be explored. Although the quality of care is the main goal of all stakeholders, we must consider conflicts of interest in the surgical unit planning. To be realistic, hospital managers aim to smooth patient flow in the surgical department and maximize the expected financial surplus of the hospital, as surgeons are concerned with their levels of income and effort. The key point is that the surgeons have better information on their patients' health status and illness severity than any other stakeholder. On the other hand, hospital managers decide on the surgical unit resources to make available to surgeons.

Appendix I

Cluster #	Procedure	Specialty
1	Laparotomy exploratory	General surgery
2	Whipple's operation	General surgery
3	Repair hernia paraesophageal	General surgery
	Adrenalectomy	General surgery
4	Hepatectomy	General surgery
	Hepatectomy partial	General surgery
5	Pancreatectomy subtotal	General surgery
6	Repair hernia incisional simple	General surgery
	Repair hernia incisional complex	General surgery
	Repair hernia incisional recurrent	General surgery
7	Repair hernia incisional incarcerated	General surgery
8	Mastectomy total and dissection lymph node axillary	Breast Oncology
9	Mastectomy segmental and dissection lymph node axillary	Breast Oncology
10	Thyroidectomy total with unilateral central neck dissection	E.N.T.
	Thyroidectomy total	E.N.T.
	Parotidectomy superficial	E.N.T.
	Parathyroidectomy	E.N.T.
	Completion hemithyroidectomy	E.N.T.
	Parathyroidectomy (1 gland with intra-operative PTH)	E.N.T.
11	Completion thyroidectomy	E.N.T.
	Thyroidectomy subtotal	E.N.T.
12	Tympanoplasty	E.N.T.
	Stapedectomy	E.N.T.
	Hemithyroidectomy	E.N.T.
13	Excision cyst branchial cleft	E.N.T.
	Laryngoscopy	E.N.T.
14	Transfer submandibular gland	E.N.T.
	Transfer salivary gland	E.N.T.
	ESS Wizard (endoscopic sinus surgery)	E.N.T.
	ESS complet (endoscopic sinus surgery)	E.N.T.
15	Dissection neck functional	E.N.T.
16	ESS partial (endoscopic sinus surgery)	E.N.T.
	ESS removal tumor (endoscopic sinus surgery)	E.N.T.
17	Excision wide lesion skin	E.N.T.
18	Excision gland submandibular	E.N.T.

	Septorhinoplasty partial	E.N.T.
19	Tympanomastoidectomy	E.N.T.
	Laryngectomy partial	E.N.T.
	Dissection neck radical	E.N.T.
20	Lymphadenectomy	Gynecology
	Salpingo oophorectomy robotic	Gynecology
	Hysterectomy total robotic	Gynecology
21	Myomectomy	Gynecology
22	LTH (Laparoscopic total hysterectomy)	Gynecology
	LAVH (Laparoscopic assisted vaginal hysterectomy)	Gynecology
23	Excision wide local lesion vulva	Gynecology
	Vulvectomy simple / partial	Gynecology
24	Repair cystocele/Colporrhaphy anterior	Gynecology
25	TAH (Total abdominal hysterectomy)	Gynecology
	Oophorectomy	Gynecology
	Hysterectomy vaginal total	Gynecology
	Hysterectomy supracervical	Gynecology
26	TAHBSO (Total abdominal hysterectomy with bilateral salpingo-oophorectomy)	Gynecology
27	Hysterectomy radical robotic	Gynecology
28	Salpingectomy	Gynecology
29	Salpingo-oophorectomy	Gynecology
30	Dissection lymph node pelvis	Gynecology
	Debulking robotic	Gynecology
31	Interval debulking TAHBSO	Gynecology
32	Urethroplasty stage 1	Urology
33	Urethroplasty	Urology
34	Nephrectomy partial	Urology
	Prostatectomy radical	Urology
35	TURP (Transurethral resection prostate)	Urology
	TURP (Transurethral resection prostate) Holmium laser	Urology
36	Endarterectomy carotid	Vascular
37	Angioplasty and stenting iliac artery endovascular	Vascular
	Angioplasty and stenting femoral artery endovascular	Vascular
38	Repair AAA (aneurysm aorta abdominal) endovascular	Vascular
	Repair AAA (aneurysm aorta abdominal)	Vascular
	Repair AAA (aneurysm aorta abdominal) fenestrated endovascular	Vascular
	Repair AAT (aneurysm aorta thoracic) endovascular	Vascular
39	Repair aneurysm popliteal	Vascular

40	Closure ileostomy	Colorectal
41	Colectomy abdominal total	Colorectal
42	Excision lesion rectum transanal	Colorectal
43	Hartmann's procedure	Colorectal
44	Hemicolectomy left	Colorectal
	Resection sigmoid	Colorectal
	Anterior resection (without ileostomy)	Colorectal
	Low anterior resection rectum (without ileostomy)	Colorectal
	Low anterior resection with ileostomy	Colorectal
45	Hemicolectomy right	Colorectal
46	Repair fistula in ano	Colorectal
47	Resection abdomino-perineal with colostomy	Colorectal

Appendix II

Possible Block Mixes			
Specialty: General surgery	Cluster Sequence	Weighted Average OR Time (hours)	Reimbursement Rate ($)
Mix1	1,6	4.67	3243
Mix2	3	3.93	3973
Mix3	6,7	3.89	5852
Mix4	6,6	3.44	3758
Mix5	7,7	4.34	7946
Mix6	2	11.92	7215
Mix7	4	7.78	3973
Mix8	6	5.02	1879
Breast Oncology			
Mix9	8,8	4.07	7946
Mix10	8,9	3.56	5852
Mix11	9,9,9	4.72	5637
Mix12	9,9	3.06	3758
Colorectal			
Mix13	40,40	4.43	14430
Mix14	40,42	4.54	14430
Mix15	40,46	3.17	14430
Mix16	41	4.47	7215
Mix17	42,42	4.65	14430
Mix18	42,46	3.28	14430
Mix19	43	4.65	7215
Mix20	44	4.92	7215
Mix21	45	3.14	7215
Mix22	45,46	4.22	14430

Mix23	46,46,46,46	4.08	28860
Mix24	47	6.68	7215
E.N.T.			
Mix25	16,16,16	4.43	2847
Mix26	16,15	3.66	2828
Mix27	16,16,13	4.43	3777
Mix28	16,13,13	4.43	4707
Mix29	14,14	4.01	3758
Mix30	14,15	4.22	3758
Mix31	14,18	3.67	5852
Mix32	14,12	3.93	3758
Mix33	14,11	3.56	3758
Mix34	14,10	4.01	3758
Mix35	18,15	3.88	5852
Mix36	18,12	3.60	5852
Mix37	18,10	3.74	5852
Mix38	13,13,13	4.41	5637
Mix39	13,10	3.50	3758
Mix40	13,15	3.65	3758
Mix41	12,12	3.86	3758
Mix42	12,11	3.49	3758
Mix43	12,10	4.00	3758
Mix44	12,15	4.15	3758
Mix45	11,15	3.77	3758
Mix46	11,10	3.63	3758
Mix47	10,10	4.14	3758
Mix48	10,15	4.29	3758
Mix49	15,15	4.44	3758
Mix50	19	4.07	3973

Mix51	17	3.87	3973
Gynecology			
Mix52	21,21	4.18	1898
Mix53	21,24	4.17	2313
Mix54	21,28	3.63	4922
Mix55	21,29	4.28	4922
Mix56	21,25	4.50	4922
Mix57	22,28	4.40	5852
Mix58	24,29	4.28	5337
Mix59	28,29	3.74	7946
Mix60	28,25	3.96	7946
Mix61	20	4.96	7215
Mix62	30	3.97	3973
Mix63	27	5.63	3973
Mix64	31	6.26	7215
Urology			
Mix65	35,35	3.58	3758
Mix66	35,32	3.51	9094
Mix67	34	4.78	1879
Mix68	33	5.28	1879
Vascular			
Mix69	36,37	4.89	3758
Mix70	39	3.58	7215
Mix71	37,37	4.13	3758
Mix72	38	3.98	7215

References

Adan, I., & Vissers, J. (2002). Patient mix optimisation in hospital admission planning: a case study. *International Journal of Operations & Production Management, 22*(4), 445-461.

Agnetis, A., Coppi, A., Corsini, M., Dellino, G., Meloni, C., & Pranzo, M. (2014). A decomposition approach for the combined master surgical schedule and surgical case assignment problems. *Health care management science, 17*(1), 49-59.

Ayvaz, N., & Huh, W. T. (2010). Allocation of hospital capacity to multiple types of patients. *Journal of Revenue & Pricing Management, 9*(5), 386-398.

Barua, B., & Esmail, N. (2013). Waiting Your Turn: Wait Times for Health Care in Canada, 2013 Report | Fraser Institute.

Batun, S., Denton, B. T., Huschka, T. R., & Schaefer, A. J. (2011). Operating Room Pooling and Parallel Surgery Processing Under Uncertainty. *Informs Journal on Computing, 23*(2), 220-237. doi: 10.1287/ijoc.1100.0396

Beliën, J., & Demeulemeester, E. (2007). Building cyclic master surgery schedules with leveled resulting bed occupancy. *European Journal of Operational Research, 176*(2), 1185-1204.

Best, T. J., Sandikci, B., Eisenstein, D. D., & Meltzer, D. O. (2015). Managing Hospital Inpatient Bed Capacity Through Partitioning Care into Focused Wings. *M&Som-Manufacturing & Service Operations Management, 17*(2), 157-176. doi: 10.1287/msom.2014.0516

Blake, J. T., & Carter, M. W. (2002). A goal programming approach to strategic resource allocation in acute care hospitals. *European Journal of Operational Research, 140*(3), 541-561.

Blake, J. T., & Carter, M. W. (2003). Physician and hospital funding options in a public system with decreasing resources. *Socio-Economic Planning Sciences, 37*(1), 45-68.

Cardoen, B., Demeulemeester, E., & Beliën, J. (2010). Operating room planning and scheduling: A literature review. *European Journal of Operational Research, 201*(3), 921-932.

Chan, C. W., Farias, V. F., Bambos, N., Escobar, G. J., Management, S. S. o., & Farias, V. F. (2011). Optimizing Intensive Care Unit Discharge Decisions with Patient Readmissions. *SSRN*. doi: OPEN_ACCESS_POLICY

Chow, V. S., Puterman, M. L., Salehirad, N., Huang, W., & Atkins, D. (2011). Reducing surgical ward congestion through improved surgical scheduling and uncapacitated simulation. *Production and Operations Management, 20*(3), 418-430.

Cochran, J. K., & Bharti, A. (2006). Stochastic bed balancing of an obstetrics hospital. *Health care management science, 9*(1), 31-45.

Costa, L. d. F. D., & Cesar Jr., R. M. (2000). Shape Analysis and Classification: Theory and Practice.

Creemers, S., Beliën, J., & Lambrecht, M. (2012). The optimal allocation of server time slots over different classes of patients. *European Journal of Operational Research, 219*(3), 508-521.

Dafny, L. S. (2005). How do hospitals respond to price changes? *American Economic Review, 95*(5), 1525-1547. doi: 10.1257/000282805775014236

Day, R., Garfinkel, R., & Thompson, S. (2012). Integrated block sharing: A win-win strategy for hospitals and surgeons. *Manufacturing & Service Operations Management, 14*(4), 567-583.

Deng, Y., Shen, S., & Denton, B. (2014). Chance-Constrained Surgery Planning Under Uncertain or Ambiguous Surgery Duration. *Available at SSRN 2432375*.

Denton, B., Viapiano, J., & Vogl, A. (2007). Optimization of surgery sequencing and scheduling decisions under uncertainty. *Health care management science, 10*(1), 13-24.

Denton, B. T., Miller, A. J., Balasubramanian, H. J., & Huschka, T. R. (2010). Optimal allocation of surgery blocks to operating rooms under uncertainty. *Operations Research, 58*(4-part-1), 802-816.

Dey, P. K., Hariharan, S., & Clegg, B. T. (2006). Measuring the operational performance of intensive care units using the analytic hierarchy process approach. *International Journal of Operations & Production Management, 26*(8), 849-865.

Dobson, G., Lee, H.-H., & Pinker, E. (2010). A model of ICU bumping. *Operations Research, 58*(6), 1564-1576.

Doulabi, S. H. H., Rousseau, L. M., & Pesant, G. (2016). A Constraint-Programming-Based Branch-and-Price-and-Cut Approach for Operating Room Planning and Scheduling. *Informs Journal on Computing, 28*(3), 432-448. doi: 10.1287/ijoc.2015.0686

Eapen, Z. J., Reed, S. D., Li, Y., Kociol, R. D., Armstrong, P. W., Starling, R. C., . . . Ezekowitz, J. A. (2013). Do countries or hospitals with longer hospital stays for acute heart failure have lower readmission rates? Findings from ASCEND-HF. *Circulation: Heart Failure, 6*(4), 727-732.

Ferrand, Y. B., Magazine, M. J., & Rao, U. S. (2014). Partially Flexible Operating Rooms for Elective and Emergency Surgeries. *Decision Sciences, 45*(5), 819-847. doi: 10.1111/deci.12096

Fugener, A., Hans, E. W., Kolisch, R., Kortbeek, N., & Vanberkel, P. T. (2014). Master surgery scheduling with consideration of multiple downstream units. *European Journal of Operational Research, 239*(1), 227-236. doi: 10.1016/j.ejor.2014.05.009

Gorunescu, F., McClean, S. I., Millard, P. H., & Correspondence. (2002). A queueing model for bed-occupancy management and planning of hospitals.

Guerriero, F., & Guido, R. (2011). Operational research in the management of the operating theatre: a survey. *Health care management science, 14*(1), 89-114.

Gul, S., Denton, B. T., & Fowler, J. W. (2015). A Progressive Hedging Approach for Surgery Planning Under Uncertainty. *Informs Journal on Computing, 27*(4), 755-772. doi: 10.1287/ijoc.2015.0658

Hans, E., Wullink, G., Van Houdenhoven, M., & Kazemier, G. (2008). Robust surgery loading. *European Journal of Operational Research, 185*(3), 1038-1050.

Helm, J. E., & Van Oyen, M. P. (2014). Design and Optimization Methods for Elective Hospital Admissions. *Operations Research, 62*(6), 1265-1282. doi: 10.1287/opre.2014.1317

Huckman, R. S., & Pisano, G. P. (2006). The firm specificity of individual performance: Evidence from cardiac surgery. *Management Science, 52*(4), 473-488.

Information, C. I. f. H. (2012). National Health Expenditure Trends, 1975 to 2012: Canadian Institute for Health Information (CIHI).

Jackson, R. L. (2002). The business of surgery. Managing the OR as a profit center requires more than just IT. It requires a profit-making mindset, too. *Health management technology, 23*(7), 20.

Kc, D. S., & Staats, B. R. (2012). Accumulating a portfolio of experience: The effect of focal and related experience on surgeon performance. *Manufacturing & Service Operations Management, 14*(4), 618-633.

Kc, D. S., & Terwiesch, C. (2009). Impact of workload on service time and patient safety: An econometric analysis of hospital operations. *Management Science, 55*(9), 1486-1498.

KC, D. S., & Terwiesch, C. (2011). The effects of focus on performance: Evidence from California hospitals. *Management Science, 57*(11), 1897-1912.

Kim, S.-C., Horowitz, I., Young, K. K., & Buckley, T. A. (1999). Analysis of capacity management of the intensive care unit in a hospital. *European Journal of Operational Research, 115*(1), 36-46.

Kim, S. H., Chan, C. W., Olivares, M., & Escobar, G. (2015). ICU Admission Control: An Empirical Study of Capacity Allocation and Its Implication for Patient Outcomes. *Management Science, 61*(1), 19-38. doi: 10.1287/mnsc.2014.2057

Lamiri, M., Xie, X., Dolgui, A., & Grimaud, F. (2008). A stochastic model for operating room planning with elective and emergency demand for surgery. *European Journal of Operational Research, 185*(3), 1026-1037.

Lee, S., & Yih, Y. (2014). Reducing patient-flow delays in surgical suites through determining start-times of surgical cases. *European Journal of Operational Research, 238*(2), 620-629. doi: 10.1016/j.ejor.2014.03.043

Litvak, N., van Rijsbergen, M., Boucherie, R. J., & van Houdenhoven, M. (2008). Managing the overflow of intensive care patients. *European Journal of Operational Research, 185*(3), 998-1010.

Lovejoy, W. S., & Li, Y. (2002). Hospital operating room capacity expansion. *Management Science, 48*(11), 1369-1387.

Lucas, D. J., & Pawlik, T. M. (2014). Readmission after surgery. *Advances in surgery, 48*, 185.

Marazzi, A., Paccaud, F., Ruffieux, C., & Beguin, C. (1998). Fitting the distributions of length of stay by parametric models. *Medical care, 36*(6), 915-927.

Marcon, E., & Dexter, F. (2006). Impact of surgical sequencing on post anesthesia care unit staffing. *Health care management science, 9*(1), 87-98.

Marques, I., Captivo, M. E., & Pato, M. V. (2015). A bicriteria heuristic for an elective surgery scheduling problem. *Health care management science, 18*(3), 251-266. doi: 10.1007/s10729-014-9305-z

May, J. H., Spangler, W. E., Strum, D. P., & Vargas, L. G. (2011). The surgical scheduling problem: Current research and future opportunities. *Production and Operations Management, 20*(3), 392-405.

Mayes, R. (2007). The origins, development, and passage of medicare's revolutionary prospective payment system. *Journal of the History of Medicine and Allied Sciences, 62*(1), 21-55. doi: 10.1093/jhmas/jrj038

Meng, F. W., Qi, J., Zhang, M. L., Ang, J., Chu, S., & Sim, M. (2015). A Robust Optimization Model for Managing Elective Admission in a Public Hospital. *Operations Research, 63*(6), 1452-1467. doi: 10.1287/opre.2015.1423

Min, D., & Yih, Y. (2010). Scheduling elective surgery under uncertainty and downstream capacity constraints. *European Journal of Operational Research, 206*(3), 642-652.

OECD. (2011). *Health at a Glance 2011: OECD Indicators*: OECD Publishing.

OECD. (2013). *Health at a Glance 2013: OECD Indicators*: OECD Publishing.

Olivares, M., Terwiesch, C., & Cassorla, L. (2008). Structural estimation of the newsvendor model: An application to reserving operating room time. *Management Science, 54*(1), 41-55.

Parsons, R. J., Howard, M. J., Barker, B., & Peterson, L. A. (1992). Case Mix Management: Rationale, Implementation, and Administration. *Journal of hospital marketing, 6*(1), 163-174.

Pink, G. H., Information, C. I. f. H., McKillop, I. D., & Johnson, L. M. (2001). *The financial management of acute care in Canada: a review of funding, performance monitoring and reporting practices*: Canadian Institute for Health Information.

Price, C., Golden, B., Harrington, M., Konewko, R., Wasil, E., & Herring, W. (2011). Reducing Boarding in a Post-Anesthesia Care Unit. *Production and Operations Management, 20*(3), 431-441.

Ridge, J., Jones, S., Nielsen, M., & Shahani, A. (1998). Capacity planning for intensive care units. *European Journal of Operational Research, 105*(2), 346-355.

Robbins W, T. N. (1989). Linear Programming Is a Useful Tool in Case-Mix Management. *Healthcare financial management, 43*(6), 114–116.

Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of risk, 2*, 21-42.

Samudra, M., Van Riet, C., Demeulemeester, E., Cardoen, B., Vansteenkiste, N., & Rademakers, F. E. (2016). Scheduling operating rooms: achievements, challenges and pitfalls. *Journal of Scheduling, 19*(5), 493-525. doi: 10.1007/s10951-016-0489-6

Sarykalin, S. S., Gaia (2014). Value-at-Risk vs. Conditional Value-at-Risk in Risk Management and Optimization *INFORMS Tutorials in Operations Research* (Chapter 13, pp. 270 - 294).

Shylo, O. V., Prokopyev, O. A., & Schaefer, A. J. (2012). Stochastic operating room scheduling for high-volume specialties under block booking. *Informs Journal on Computing, 25*(4), 682-692.

Sutherland, J., Crump, R. T., Repin, N., & Hellsten, E. (2013). Paying for hospital services: a hard look at the options. *CD Howe Institute Commentary, 378*.

Sutherland, J. M., & Foundation, C. H. S. R. (2011). *Hospital payment mechanisms: an overview and options for Canada*: Canadian Health Services Research Foundation.

Testi, A., Tanfani, E., & Torre, G. (2007). A three-phase approach for operating theatre schedules. *Health care management science, 10*(2), 163-172.

Treasure, T., Valencia, O., Sherlaw-Johnson, C., & Gallivan, S. (2002). Surgical performance measurement. *Health care management science, 5*(4), 243-248.

Verweij, B., Ahmed, S., Kleywegt, A. J., Nemhauser, G., & Shapiro, A. (2003). The sample average approximation method applied to stochastic routing problems: a computational study. *Computational Optimization and Applications, 24*(2-3), 289-333.

Vijayakumar, B., Parikh, P. J., Scott, R., Barnes, A., & Gallimore, J. (2013). A dual bin-packing approach to scheduling surgical cases at a publicly-funded hospital. *European Journal of Operational Research, 224*(3), 583-591.

Visintin, F., Cappanera, P., & Banditori, C. (2016). Evaluating the impact of flexible practices on the master surgical scheduling process: an empirical analysis. *Flexible Services and Manufacturing Journal, 28*(1-2), 182-205. doi: 10.1007/s10696-014-9208-9

Vissers, J., Adan, I. J., & Bekkers, J. A. (2005). Patient mix optimization in tactical cardiothoracic surgery planning: a case study. *IMA Journal of Management Mathematics, 16*(3), 281-304.

Yankovic, N., & Green, L. V. (2011). Identifying good nursing levels: a queuing approach. *Operations Research, 59*(4), 942-955.