2015

Improvement of surgery duration estimation using statistical methods and analysis of scheduling policies using discrete event simulation

Alexandra Blake Olsen
Iowa State University

Improvement of surgery duration estimation using statistical methods and analysis of scheduling policies using discrete event simulation

by

Alexandra Blake Olsen

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Industrial Engineering

Program of Study Committee:
Guiping Hu, Major Professor
Lizhi Wang
Frank Montabon

Iowa State University

Ames, Iowa

2015

DEDICATION

This paper is dedicated to all of the people who have and will benefit from the United States health care system and their family and friends.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my major professor, Dr. Guiping Hu, for all of her guidance throughout this process and for her willingness to learn about the health care industry alongside me. I appreciate her encouragement to pursue my interests and calming demeanor. Thank you, as well, to Lincoln Banwart for all of his work as an undergraduate research assistant. All the work he did was extremely helpful and saved me a significant amount of time. I would also like to thank the rest of my committee, Dr. Frank Montabon and Dr. Lizhi Wang, for their help and guidance. I appreciate all of the help Dr. Max Morris gave me in the statistical methods and Dr. Douglas Gemmill in Arena modeling. Without their help, the models would not have been as effective. Thank you, as well, to the rest of the faculty and staff in the IMSE department for creating an environment where it is fun to learn and students are encouraged to follow their passions.

I would like to thank those at UnityPoint Health - Des Moines, including Val Boelman, Vanessa Calderon, Brandi Vennink, Tony Nurre, and Lacey Andrews, for their willingness to work with me. I appreciate all the time they took to teach me about their processes, provide data, answer questions, and provide feedback. Without their knowledge and willingness to share that knowledge, this project would not have been possible.

I would also like to thank my peers and friends in the IMSE program for their support and making my experience at Iowa State University an enjoyable one. Lastly, I would like to thank my parents and fiancé, Kevin, for their unconditional love and support.

ABSTRACT

The United States health care system currently faces many challenges, with the most notable one being rising costs. In an effort to decrease those costs, health providers are aiming to improve efficiency in their operations. A primary source of revenue for hospitals and some clinics is the surgery department, making it a key department for improvement in efficiency. Surgery schedules drive the department and affect the operations of many other departments. The most significant challenge to creating an efficient surgery schedule is estimating surgery durations and scheduling cases in a manner that will minimize the time a surgery is off schedule and maximize utilization of resources. To identify ways to better estimate surgery durations, an analysis of the surgery scheduling process at UnityPoint Health - Des Moines, in Des Moines, Iowa was completed. Estimated surgery durations were compared to actual durations using a *t* test. Multiple linear regression models were created for the most common surgeries including the input variables of age of the patient, anesthesiologist, operating room (OR), number of residents, and day of the week. To find optimal scheduling policies, simulation models were created, each representing a series of surgery cases in one operating room during one day. Four scheduling policies were investigated: shortest estimated time first, longest estimated time first, most common surgery first, and adding an extra twenty minutes to each case in the existing order. The performance of the policies was compared to those of the existing schedule.

Using the historical data from a one-year period at UnityPoint Health - Des Moines, the estimated surgery durations for the top four surgeries by count and top surgeons were found to be statistically different in 75% of the data sets. After creating multiple linear regression models for each of the top four surgeries and surgeons performing those surgeries, the β values for each variable were compared across models. Age was found to have a minimal impact on surgery

duration in all models. The binary variable indicating residents present, was found to have minimal impact as well. For the rest of the variables, consistencies were difficult to assess, making multiple linear regression an unideal method for identifying the impact of the variables investigated.

On the other hand, the simulation model proved to be useful in identifying useful scheduling policies. Eight series based on real series were modeled individually. Each model was validated against reality, with 75% of durations simulated in the models not being statistically different than reality. Each of the four scheduling policies was modeled for each series and the average minutes off schedule and idle time between cases were compared across models. Adding an extra twenty minutes to each case in the existing order resulted in the lowest minutes off schedule, but significantly increased the idle time between cases. Most common surgery first did not have a consistent impact on the performance indicators. Longest estimated time first did not improve the performance indicators in the majority of the cases. Shortest estimated time first resulted in the best performance for minutes off schedule and idle time between cases in combination; therefore, we recommend this policy is employed when the scheduling process allows.

CHAPTER 1: GENERAL INTRODUCTION

The United States health care industry currently faces numerous challenges. In an era of changes, the system faces rising costs and increased demand for its services. This is a result of an aging population (Gupta & Denton, 2008) and the newly enforced Affordable Care Act. The Affordable Care Act, signed into law on March 23, 2010, is a federal statue aimed to improve the U.S. health care system. The goal is to make services accessible to all Americans, increase the benefit from insurance and use of the services, and reduce overall spending. One way the act works to achieve this goal is by expanding Medicaid, the federal-state funded program, to provide financial aid for medical services to low-income citizens (Davidson, 2013). As a whole, the Affordable Care Act is expected to increase the number of insured citizens by 33 million (Askin & Moore, 2012). This creates an increased demand for health services and places a larger strain on the providers.

In the United States, the healthcare industry accounts for 17.9% of the GDP as total healthcare spending exceeded $2.6 trillion in 2010 (Askin & Moore, 2012; Davidson, 2013). "The amount of public money needed to finance health care, which currently stands at 45% of all health care expenditures, is expected to double by 2050 (Gupta & Denton, 2008). In 2009, $130 billion dollars was spent due to inefficiently delivered services (Smith & Institute of Medicine (U.S.), 2013). These costs, paired with an increase in health care demand, motivate the industry to improve system operations. As the industry changes, providers are facing increasing competition (Cardoen et al., 2009). For example, statistics on costs of surgeries are available to the public, making it possible for people to shop around if they are not satisfied (Centers for Medicare and Medicaid Services, 2014). Furthermore, patients are increasingly traveling abroad for advanced and less expensive medical procedures (Burge, 2014).

As medical technology becomes more advanced and scientific discoveries are made, providing health care can become more complex, compromising the efficiency of the service (Smith & Institute of Medicine (U.S.), 2013). Electronic medical records and other technology software tools are becoming more common in the health industry. This new focus is shown in legislative acts, as the American Recovery and Reinvestment Act of 2009 included $19 billion in electronic medical record funding (Burge, 2014). With all of this new technology and capabilities, there is an abundance of potential ways to improve efficiency at the provider's fingertips. There are still bounds that need to be made to make information easily accessible to both providers and patients. This will help to ensure the best care possible is given in a cost-efficient way (Alper, Sanders, Saunders, & Institute of Medicine (U.S.), 2013).

The rising health care costs impose a burden on both patients and medical providers. Insurers do not cover provider costs, making it difficult for hospitals to make ends meet, thus, motivating providers to improve operation efficiency (Davidson, 2013). Inefficient practices are estimated to make up 13-20% of hospital's costs (Graban, 2009). Additionally, due to inefficient scheduling and administrative duties, nurses spend about 30% of their time directly taking care of patients (Smith & Institute of Medicine (U.S.), 2013). In an effort to improve the blaring inefficiencies, health care scheduling has been a strong research focus for the past several years. This extends from nurse and staff scheduling, to appointment scheduling in the hospital setting, outpatient clinic setting, and diagnostic imaging setting.

Scheduling within outpatient clinics, also called ambulatory centers, is a large area of focus in the literature, as many factors that go into patient scheduling are included in that setting. Many studies have suggested approaches to account for patient no-shows. Chakraborty et al. (2010) suggest using an algorithm to schedule times slots which account for the probability of the patient not showing up. Muthuraman and Lawley (2008) use a myopic scheduling model paired with an overbooking approach to account for patient no-shows in an outpatient setting. Many ambulatory care centers specialize in diagnostic resources. Patients schedule appointments to use various diagnostic machines such as CT scanners and MRI machines. With limited machines and time, scheduling is important in this sector. Patrick et al. (2008) looked at minimizing patient wait time while accounting for varying priority levels using approximate dynamic programming. Schutz et al. (2012) also uses heuristic approximate dynamic programming to solve the scheduling problem.

Physician lateness and interruptions can occur in both outpatient appointments and surgeries. Klassen et al. (2013) investigated the effect on outpatient scheduling and identifies the trend of physician interruptions as a "dome" effect, meaning the number of interruptions peak in the middle of the day, and are minimal at the beginning and end of the day. Simulation optimization was used to optimize time slot sizes based on physician lateness and size of the clinic. Luo et al. (2012) presents a set of differential equations that account for interruptions in diagnostic resourcing in a hospital in order to create a schedule that meets various objective functions.

The majority of health care costs come from hospital care. In 2012, 31% of health care expenditures went to hospital care, with ambulatory care accounting for 20% of these costs (Askin & Moore, 2012). At the heart of hospital care is the operating room (OR). Because surgery scheduling affects expensive resources within the post anesthesia care unit (PACU), intensive care unit (ICU), and operating rooms (ORs), surgeries alone account for more than 40% of a hospital's total revenues (Erdogan, 2010). The schedule of an OR affects many other departments of the hospital, such as nurse schedules and inpatient bed capacities (Cardoen et al., 2009).

Scheduling surgeries is a complicated process. Variability among surgery, anesthesia, and cleanup durations largely affect the OR schedule. A surgery schedule must work to minimize both indirect and direct delays. Indirect delays are the time between requesting an appointment and the time of the appointment. Direct delays are the difference between appointment time and the time the patient sees the service provider (Gupta & Denton, 2008). In some regions, waiting lists for surgeries can be problematic (Hall, 2006). Being able to meet demand by having high utilization of the ORs is important, and these ORs must also be suited to handle different urgency levels of procedures. Elective surgeries have the lowest urgency level; they are scheduled months in advance. On the other hand, emergent surgeries are those in which the patient has a minimal period of time they can wait to be admitted and start surgery. Hospitals are assigned various trauma levels, which designates the resources available in the facility for trauma and the number of patients admitted yearly (American Trauma Society, 2015). Hospitals with a high trauma certification levels must be able to accommodate emergencies, while also accommodating elective surgeries (Erdogan, 2010).

Surgery schedulers must also take into account varying priorities in the schedule. The OR managers want to keep the surgeons satisfied and have high utilization of the rooms. The surgeons do not want downtime between surgeries and prefer to be scheduled at times that work around their clinic schedules. Nurses want evenly dispersed patient flow so there are not peaks where are they are busier (Pease, 2013). Patients want minimal indirect and direct delays. These varying priorities are difficult to balance when scheduling surgeries, as they often conflict.

To confirm and further identify improvement opportunities, the elective surgery scheduling process at UnityPoint Health - Des Moines in Des Moines, Iowa was evaluated. Elective surgery scheduling is divided into two processes: surgery scheduling and add-on scheduling. At other hospitals it can also be referred to as advance scheduling and allocation scheduling (Cardoen et al., 2010). Surgery scheduling is the process by which the surgery case is requested, an estimate of duration is given, and the case is tentatively scheduled for the date requested, but a time slot is not assigned. Add-on scheduling is the process of assigning each surgery case to a time slot; the process is completed the day before the day of surgery. The standard procedure for accommodating emergency cases is to keep at least one operating room open at all times, and the add-on schedulers take care of scheduling in this occurrence. Elective surgeries were the primary focus in this since outpatient surgeries make up the highest percentage of surgeries in hospitals (Dexter, 2012). This number increases as medical technology advances, allowing minimally invasive procedures and shorter healing times for patients (Jackson, 2002).

The hospital operates using a Master Surgical Schedule (MSS), which allocates certain surgeons or groups of surgeons, blocks during the week. A block is a space of time in a given OR on a given day of the week. The schedule is cyclical, and is repeated every week, with the exception of certain groups who trade off every other week. The blocks occupy an OR for the entire day, as it is difficult to plan an exact start in the middle of the day following other surgeries. MSS processes are common and effective in the hospital setting (van Oostrum, Bredenhoff, & Hans, 2010). They take a high degree of coordination, and contain definite opportunity for optimization and improvement. The coordination and planning of the MSS procedure was not our focus for the hospital analyzed. The surgery scheduling process within the existing block schedule was the primary focus.

The surgery scheduling process is detailed in Figure 1. A surgery can be requested days or months in advance, by one of three methods: electronically via the scheduling software, by phone, or by fax. The requests come from the doctors' offices. When the request for surgery is received, it is printed and verification is sent to the doctor's office. The patient and case information will then be recorded and the surgery will tentatively be scheduled on the day requested. A part of the case information recorded is estimated surgery duration. When recording this, the scheduler compares the estimate given by the doctor and compares it to the estimate given by the software. The software averages the last ten procedures of that type performed by that surgeon. If there are not enough data points available that meet the criteria, the average of the last ten procedures of that type in the health system, regardless of surgeon is given. If that is not available, a national average for that procedure type is given.

If the estimate given by the doctor is significantly different than the one given by the software, the team leader will investigate the reason and make a final decision on the estimate recorded. Once the final estimate is determined, a fixed amount of time is added for cleanup depending on the type of surgery. If the case is an ear, nose, or throat (ENT) procedure, 5 or 10 minutes is added to the estimated time to account for cleanup. If the case is a heart procedure or one involving a robot tool, 30 minutes is added for cleanup. For the rest, which is the majority of cases, 20 minutes are added for cleanup. Once all of the information is entered into the surgery scheduling software, the process ends and the case remains on the schedule for the day requested until it is assigned a time the day before the day of surgery.

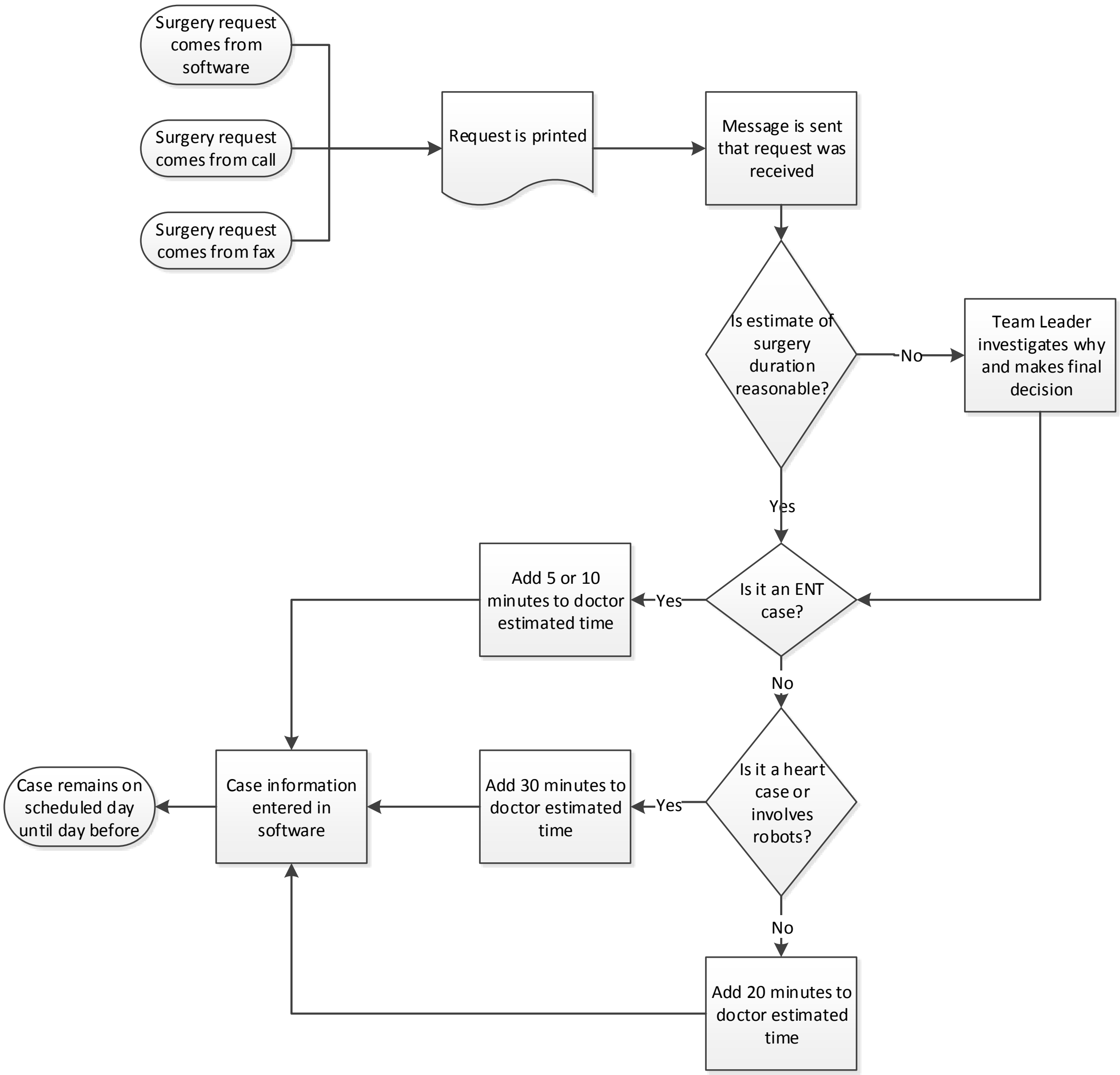

Figure 1: Surgery Scheduling Process

Forty-eight hours before the scheduled day of surgery, the cases scheduled for that day are allocated into time slots. This process, known as the add-on scheduling process, is shown in Figure 2. A case needing to be allocated a time slot is identified. Surgeons with a block are given preference for their times over surgeons without a block. If they request a certain time or order for that case or cases, the schedulers accommodate. If there is extra time in the day after all of the cases within blocks are scheduled, the unused blocks are released, and other elective cases requested for that day are added. Currently, it is rare for an elective surgery to not get scheduled on the day requested.

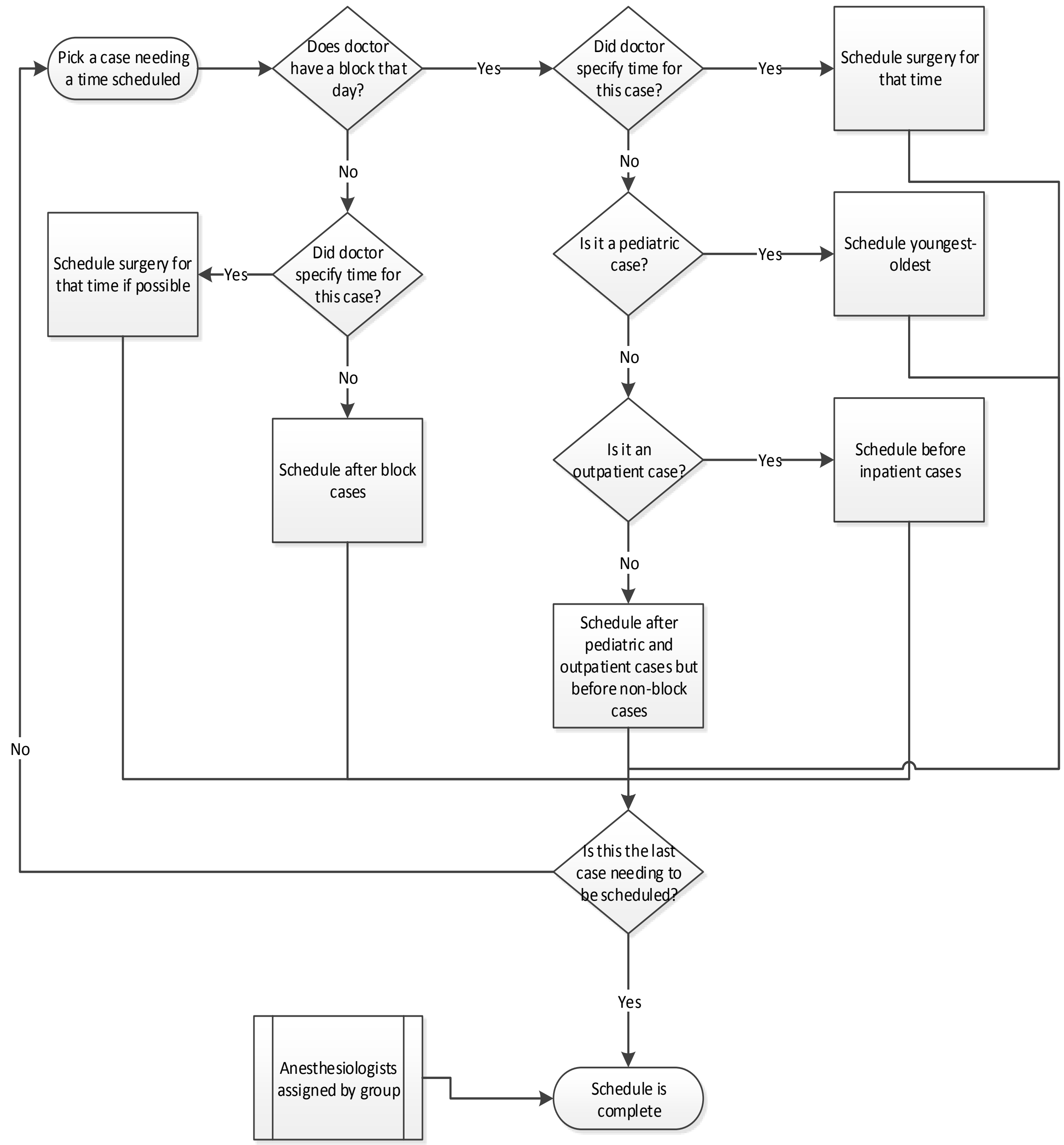

Figure 2: Add-on Scheduling Process

If the block surgeon or group of surgeons did not specify an order, the order is determined based on additional criteria. If there are multiple pediatric surgeries, they are scheduled youngest to oldest. Outpatient cases are scheduled before inpatient cases. For non-block cases, the surgeons can request a time, but understand that the time is not guaranteed, as surgeons with a block take precedence. Other considerations taken by the schedulers are medical instruments that are limited, as they must ensure cases are scheduled in a way that does not surpass the number of instruments available. Schedulers also try to group together cases with similar procedure types, age of the patients, and equipment needed. Anesthesiologists are scheduled by an external group. There are a limited number of anesthesiologists available, meaning the schedule must accommodate this as well.

Scheduling is finalized in the afternoon of the day before and the patients are notified of the time of their surgery. On the day of, a separate scheduling team is in charge of making changes if necessary. If a surgery goes significantly longer than expected, the succeeding case can be moved to another room if all persons necessary for the case are available. If the preceding surgery is being done by the same surgeon, waiting is the only option.

The biggest challenge of scheduling surgeries is correctly estimating surgery durations. The use of historical data to estimate surgery duration has been explored in the literature, with mixed results. In some of the literature, the accuracy of surgeon estimates has been compared to methods using historical data (Joustra et al., 2013; Laskin et al., 2012). Sample means have also been explored as a proper estimate of surgery duration (Broka et al., 2003; Dexter et al., 1999; Spangler et al., 2004; Zhou et al., 1999). Additionally, regression models have been found useful in the past to predict durations (Boyle et al., 2008; Eijkemans et al., 2010; Wrigh et al., 1996).

This paper aims to further investigate methods such as these by comparing surgeon estimates to reality using *t* tests and by creating linear regression models to identify key factors in duration.

A simulation was also performed to analyze scheduling rules and scenarios with an objective to minimize time off schedule and idle time of the rooms. Real cases were simulated and the simulations were validated against reality. Then, different scheduling rules were tested. Scheduling rules have been a strong topic in the literature within the manufacturing setting for many years (Vollmann, 2004). They have been recently explored within the health care industry, particularly surgery (Harper, 2002; Sciomache et al., 2005; Testi et al., 2007). Simulation is becoming more prevalent in the health care industry, both with general patient scheduling, nurse scheduling, and surgery scheduling (Everett, 2002; M'Hallah & Al-Roomi, 2014; Stiglic & Kokol, 2005; Tanfani & Testi, 2010). Surgery scheduling is an area of substantial opportunity for improvement. The impact of the results can be great due to the economic value of surgery procedures for the hospital and the current situation the health care industry faces. This paper presents additional findings on surgery scheduling processes, using real data from UnityPoint Health - Des Moines.

This thesis is structured as follows: In Chapter 2, an analysis of estimated durations versus actual durations, and several multiple linear regression models for the top surgery procedures are introduced; Chapter 3 introduces an Arena simulation model that is used to compare surgery case scenarios and analyze the impact on time off schedule and idle time; Chapter 4 summarizes the results, provides a conclusion, and details opportunities for future work.

CHAPTER 2: STATISTICAL METHODS FOR SURGERY DURATION ESTIMATION

Abstract

Surgery is a primary source of revenue in a hospital, and scheduling of surgery largely drives that revenue. A key challenge in creating schedules that minimize the amount of waiting time for patients and maximize the utilization of the operating rooms is accurately estimating surgery durations. Using data from a large Midwestern hospital, surgery duration estimations were compared to actual durations in a one-year period for the top surgeries. Statistically, a significant difference between actual and estimated durations was proven. With the goal of decreasing the difference between the estimated and actual durations, multiple linear regression models were created for the most common surgeries and used to analyze the impact various characteristics of surgery cases have on the duration. Due to the high variability of the data, the regression method was not found particularly useful in identifying strong correlations in the input characteristics.

2.1 Introduction

For hospitals, surgery is a vital source of revenue. It is estimated that 40% of a hospital's costs and 68% of its revenue come from surgery (Jackson, 2002). As health care costs rise in the United States (Askin & Moore, 2012; Davidson, 2013; Gupta & Denton, 2008), improving the efficiency of surgery scheduling has become an area of focus for both providers and researchers. Technological capabilities within surgery scheduling have been improving as well, utilizing an array of software programs to record surgery data and use it as a tool to schedule future surgeries.

When scheduling surgery, hospitals must take into account the needs of patients and surgeons, as well as their own budgets. This can be a significant challenge, as these needs include patients' requests, surgeons' schedules, and hospital resource availability. According to the Centers for Medicare and Medicaid Services (2014), data on the costs of surgery to the patient are readily available to the public. This allows patients to compare costs across hospitals in the region, and gives them the knowledge that can drive their choice to have surgery at another location if they are not satisfied with prior experiences. Surgeons can also send their services and patients elsewhere if they are not satisfied themselves with the services provided by the hospital, such as the way the surgeries are scheduled or the way their requests are handled. A hospital must meet its own financial needs as well, ensuring that the necessary utilization of their resources is met. These priorities are a challenge to balance, as they often conflict.

The core component of an effective surgery schedule is accurate surgery duration estimations. If a surgery goes longer than expected, patients have to wait and doctors and nurses may have to work overtime. Conversely, if a surgery takes less time than expected, the utilization of the operating room goes down, compromising potential revenue and services. The schedule of an operating room (OR) affects many other departments of the hospital, such as nurse schedules and inpatient bed capacities (Cardoen et al., 2009).

As technology is becoming more prevalent in administrative duties of hospitals (Burge, 2014), historical data is becoming more readily available. However, there are challenges in using historical data to make decisions. Surgeons are able to assess the individual patient's case and the complexity of it in ways a computer system cannot assess from historical data alone. For example, a case may take an average of 20 minutes based on historical data, but a particular patient could have a condition causing it to take longer. The computer system would not be aware of the condition based on historical data like the surgeon would. Another challenge with historical data is that more recent durations may be more accurate because of the introduction of advanced technology or repetition of the surgeons (Laskin et al., 2012); therefore, using data from years prior may not result in an accurate estimation. Also, there is always a chance that the data was not entered completely or accurately. Even if it is entered accurately, the use of that data is not always understood due to its relatively new development and implementation. These challenges complicate the use of historical data to make accurate surgery duration estimations.

Surgery schedulers determine the final surgery schedule and are key to an effective scheduling process. They are the first interaction with surgeons and the surgeons' offices, and how they interact can determine the satisfaction of the surgeons (Mathias, 2011). They also make final decisions on surgery duration estimations and the schedule based on surgeon input and historical data. Considerations about both indirect and direct patient delays in the scheduling process are required of surgery schedulers. Indirect delays are between the time the patient makes a request for the appointment/surgery and when it is scheduled, while direct delays are between when the appointment/surgery is scheduled and when it actually occurs (Gupta & Denton, 2008). This study only focused on the direct delays, as they are currently the primary concern for the hospital analyzed. With surgery durations as the key to an effective schedule, accurate estimation of surgery duration is highly important. This has been a focus within literature for the past several years (Dexter et al., 1999; Eijkemans et al., 2010).

Use of sample means and medians from historical data to estimate surgery durations have been studied extensively in the literature. Dexter et al. (1999) found the sample means to be a more reasonable method to predict the duration of a series of cases and turnovers than linear programming because some cases do not occur as often. Taking it a step forward, Spangler et al. (2004) found that using a second-order regression model to estimate the location parameter of a lognormal distribution for surgery duration was superior to using sample means. In a method similar to using sample means, Broka et al. (2003) found median duration times from historical data individualized by surgeon in a one year period to be useful in accurate prediction of future durations. On the contrary, Zhou et al. (1999) found surgeon and procedure type to be critical and most indicative of duration, but found historical data of average durations to not be helpful in predicting future durations.

Several studies looked at the accuracy of surgeon predictions. Laskin et al. (2012) found oral and maxillofacial surgeons to overestimate duration times more than underestimate. Larsson (2012) compared the calculated average value (CAV) for surgery duration to the surgeon estimated time (SET) for the top four surgeries. It was found that the SET system was better at identifying long cases, but the CAV system gave better estimations overall. Joustra et al. (2013) focused on getting optimal prediction of surgery durations using data mining analysis methods and historical data. Using statistical methods, specifically the Burr distribution, was found to be more accurate than surgeon estimates. In addition to this, the surgeon’s estimation was found vital to the estimation when used in combination with the statistical methods.

Surgery duration estimates made by surgeons have also been compared to duration estimates resulting from regression models. Eijkemans et al. (2010) compared surgeon predictions to a lognormal regression model using several predictors, including the surgeon prediction, level of the team, and patient characteristics like age and gender. The regression model was found to be more predicative than the surgeon estimates alone, and surgical team characteristics were more indicative of duration than patient characteristics. Boyle et al. (2008) looked at five different statistical methods and found linear regression with eleven variables to most accurately predict emergency room admissions. Wright et al. (1996) found surgeons to be more accurate than regression. The factors accounted for in their model are gender, bilateral, rank of resident assistant, essential diagnostics, case difficulty, and if historical data was used in estimation by surgeon.

Other facets of surgery durations have been analyzed as well. Koenig et al. (2011) analyzed the prediction of anesthesia times using a spreadsheet-based system and the Kruskal-Wallis test. It was observed that obesity physical status did not affect anesthesia induction times. Dexter et al. (2001) calculated an upper prediction bound using the mean and standard deviation of the natural logarithms of the durations of cases with the same procedure and surgeon. It was then used to calculate a delay time between cases performed by different surgeons. The 90% upper prediction value was found to provide the best estimate by overestimating the time needed and decreasing overtime. Other ways to improve scheduling other than investigating duration estimations were examined by Dexter et al. (2009), in which ways to estimate remaining times of surgeries using instant messaging while the surgery is in progress were investigated. Variation in durations is inevitable; therefore, methods to both estimate more accurately and improve efficiency when surgeries go overtime are worthwhile to investigate.

In this paper, we compare surgery estimations to actual surgery times using data for the top four procedure types at a large Midwestern hospital. The actual durations are then used to create a multiple linear regression model for each procedure type with five input variables, and trends of the impact of each variable is compared. Results of this comparison prove multiple linear regression methods to be an inadequate method of gaining comprehensive knowledge on the impact of each input variable. These findings can provide insight for a variety of hospitals nationally.

This paper is structured as follows. In section two, the general methods are introduced. In section three, the case study using UnityPoint Health - Des Moines in Des Moines, IA is presented, with details on the comparison of estimated and actual durations, identification of the top procedure types, linear regression models for each procedure type, and analysis of results. Section four is a conclusion.

2.2 Methods

To compare surgery duration estimates recorded by surgery schedulers to reality, a *t* test was conducted. The top four procedure types in the data set were identified and split. This reduced the sample size significantly, making it more appropriate for a *t* test and allowing comparison of estimation accuracy for each procedure type against other procedure types. For each of the top four surgeries, a *t* test was completed, comparing the estimated time versus the actual time. Data were broken down further into the top surgeons within each procedure type and a *t* test was completed for each. Isolating the surgeons within each procedure type allows identification of surgeons who are more accurate than others.

In addition to the comparison of actual and estimated durations, multiple linear regression models were created using JMP statistical software for each of the data sets identified in the *t* test process, with the aim to identify key variables affecting surgery duration. Independent variables can be added based on the data available. In this case, age of the patient, anesthesiologist, operating room (OR), number of residents, and day of the week were the independent variables. The estimates and p values for each variable were then compared across all models to identify consistencies in the positive or negative impact of each variable on the procedure duration, the dependent variable.

2.3 Case Study

The surgery scheduling process at UnityPoint Health - Des Moines, a level one trauma center and teaching hospital in Des Moines, Iowa, was evaluated and data from September 2013-2014 were used. The data include 13,874 cases.

2.3.1 Comparison of Estimated and Actual Surgery Durations

The biggest challenge UnityPoint Health's surgery schedulers face is estimating surgery durations (Lacey Andrews, personal communication, August 19, 2014). Current practice is to ask the surgeons to estimate the time the case is going to take and then add a fixed amount of time for cleanup. This is consistent with Broka et al. (2003). These estimations are not limited to sets of a certain number, but they tend to be rounded to the next five minute increment. The estimate given by the surgeon is compared to the average for the last ten surgeries for that procedure type and surgeon, given in the electronic medical record database. If the surgeon's estimate is significantly different, the specialty team leader will look into the reason why and make a decision on the final estimate.

In the one year period, the surgery cases started late 80% of the time and started exactly on time 1.4% of the time. To further assess the magnitude of this problem, *t* tests were used to compare estimated duration versus actual times of surgeries. The top four surgeries by count were evaluated, both for all of the surgeons collectively and each of the top surgeons separately. We defined the top surgeons as those that performed at least 10% of the total number of cases done with that procedure. The procedures evaluated were Laparoscopic Cholecystectomy (Lap Chol), Esophagogastrodueodenoscopy (Esophag), Cystoscopy (Cyst), and Insert Port Vascular Access (IPVA). Lapraroscopic Appendectomy (Lap App) was also a top surgery, but was not evaluated because 64% of the cases are emergency procedures, and therefore were not scheduled in the surgery scheduling process. These five procedure types make up 9.3% of the total surgeries performed in that span of time. Because the estimates were based on historical data, the surgeons did not know the estimates were being analyzed, eliminating a source of bias. For our purposes, we eliminated data points that were not complete and assumed the rest of the data was recorded accurately. The emergency surgeries were also eliminated, as they were not estimated using the same process as the elective surgeries. Twenty minutes were subtracted from each estimated time to account for the cleanup time the schedulers added in. The actual time does not include the cleanup time. Significant outliers, defined as those greater than six times the standard deviation (6σ) were removed. The results of the *t* test are shown in Table 1.

Table 1: *t* test **results**

Procedure	Surgeon	p value	Std dev	Mean	Lower CI	Upper CI
Cyst	All	0.00	52.32	10.16	3.21	17.12
Cyst	1	0.05	55.54	8.94	-0.02	17.90
Esophag	All	0.00	8.59	-3.77	-4.85	-2.70
Esophag	2	0.00	5.52	-5.27	-6.46	-4.08
Esophag	3	0.00	6.27	-4.58	-6.12	-3.03
Esophag	4	0.53	11.06	-0.95	-3.94	2.05
Esophag	5	0.00	6.59	-5.00	-7.46	-2.54
IPVA	All	0.00	17.51	11.70	9.31	14.10
IPVA	6	0.00	13.31	7.65	3.46	11.84
IPVA	7	0.00	10.49	7.00	3.73	10.27
IPVA	1	0.00	17.55	11.90	6.21	17.59
Lap Chol	All	0.00	26.10	15.25	11.42	19.08
Lap Chol	8	0.00	17.86	18.56	13.38	23.75
Lap Chol	9	0.06	21.88	8.00	-0.32	16.32
Lap Chol	10	0.00	17.54	13.17	7.97	18.38
Lap Chol	11	0.62	41.72	5.71	-18.38	29.80

A positive number in the mean and confidence intervals (CI) represents that the actual time took longer than the estimated time; in other words, the time was underestimated. All four of the surgery types with all of the surgeons included had statistically different values ($p<.05$) for actual time versus estimated time. Only 25% of the cases looked at were not statistically different. The trend was towards underestimating, which is consistent with Broka et al. (2003). It is worth noting that Esophag was almost always overestimated in time and the others were always underestimated in time. This portrays that completing *t* tests for each surgery type and surgeon can pinpoint procedures that are usually accurately estimated and surgeons who are better at estimating than others. The surgery schedulers often deduce this over time, but having this information could be used to train new schedulers and reduce the learning curve.

These results support the thought that estimating surgery durations is the core of the problem for surgery schedulers, and accurate estimates are the core of the entire schedule. One long surgery can throw off the rest of the day, and one short surgery can mean underutilization of the hospital's resources.

2.3.2 Impact of Variables Using Multiple Linear Regression

Factors affecting how long a surgery takes vary immensely. In order to evaluate the impact of these factors, multiple linear regression was used. The same one year period of historical data was analyzed, looking at the procedure start to procedure end time as the dependent variable, as well as the corresponding information for the independent variables. The procedure time does not include anesthesia time or cleanup time, and therefore eliminates sources of variability. A model was created for each of the top four surgeries by count and the surgeons who make up at least 10% of the cases for that procedure in the data acquired. In this case, Lap App was included and Cyst was not. Even though 64% of the Lap App cases were emergency, there is no reason to think the durations are affected by the variables looked into differently than non-emergency cases. Cyst was not used in the regression model because of its high standard deviation compared to the other procedure types (1.8-4.7 times higher).

Once again, JMP statistical software was used. The hospital schedulers and their electronic medical record software already take procedure type and surgeon into account, as is consistent with Strum et al. (2000). The additional variables included in the model were age of the patient, anesthesiologist, OR, number of residents, and day of the week. These variables were chosen because the data was available, and they realistically could have an impact on the duration. Older patients may take longer to operate on, anesthesiologists could vary in skills, different ORs could have different quality of resources, a resident could help a surgeon or take time away from a surgeon's work to teach, and the staff could perform at a different pace depending on the day of the week. Models were created for each of the top four surgeries with all surgeons included as well. The same independent variables were included with the addition of surgeon.

After each model was created, the estimate, also called the β value, of each variable was recorded, as well as the p value. To allow consistent comparison among models, each estimate was recalculated compared to a base value. The base value for each variable was chosen to be a variable that appeared in each model. A Microsoft Excel Pivot Table was then created, making it possible to compare certain the parameters for variables across all models.

2.3.3 Results

The results were fairly inconsistent, showing that multiple linear regression is not ideal for determining the impact of the variables on surgery duration. The R^2 values for the models varied, as shown in Table 2. Esophag had a high R^2 value, showing that it may be easier to estimate the duration of that procedure type. This paired with the *t* test results for Esophag is worth noting.

Table 2: R^2 values

Procedure	Surgeon	R^2
Esophag	All	0.74
	4	0.77
	1	0.60
	5	0.94
	2	0.40
Lap App	All	0.43
	9	1.00
	12	0.72
	13	0.90
Lap Chol	All	0.42
	8	0.39
	9	0.77
	10	0.77
	14	0.98
	11	0.86
IPVA	All	0.40
	6	0.63
	7	0.53
	15	0.90

Nineteen total regression models were created. Within all four top surgeries combined, there were 40 different anesthesiologists, 15 different operating rooms (there are 17 total at the hospital), 42 different surgeons, and six different days (excluding Saturday). Including the variables of age, and residents (a binary variable), there are a total of 106 variables to analyze. The maximum number of models one variable existed in was 19, the minimum was one, and the average was six. For example, Anesthesiologist 1 was included in 15 models, meaning the variable has 15 different β values. A comparison of the β values shows if there is a consistency in the impact Anesthesiologist 1 has on the surgery duration across the different procedure types.

Looking at these comparisons, there were some consistencies. The strongest one was that age did not have a large impact on duration. For each year of age, only a fraction of a minute is added to the duration, evident by the average β of -0.1. The average p value was 0.51, which is significantly higher than α=0.05, which means the impact of age is insignificant. The hospital currently only takes age into account if it is a pediatric case, by scheduling youngest to oldest. They need to accommodate small children and minimize the amount of time they have to wait for surgery. Age for older patients was found to be arbitrary; therefore, we recommend it is eliminated as a potential variable to consider for the surgery schedulers.

Residents are common in the hospital investigated, as it is a teaching hospital. Overall in the data, there was at least one resident present 35% of the time. It is highly dependent on the procedure type, as some procedures are more conducive to learning; therefore, investigating the impact of residents was found to be difficult. Residents are either used the majority of the time or not used the majority of the time, depending on the procedure type. If the value for resident is 0 a small percentage of the time, or vice versa, for a particular procedure type, it is difficult to have confidence in the estimate given. The estimates that were given varied, and the p values that were given were significantly higher than α=0.5.

For all of the other variables, it was difficult to find trends. An example of the comparison of β and p values for one of the ORs is given in Table 3. To understand the impact the variable, in this case OR 04, has on the duration, the p values and adjusted β values must be compared. For OR 04, almost all of the p values were significantly larger than α=0.05, meaning that variable is not statistically significant to the duration. Furthermore, the adjusted β values varied, which signifies that particular room had a positive impact for some surgeons and a negative impact for others. It must be noted that the amount of data points these estimates are created from vary tremendously. For example, Surgeon 2 performed Esophag in OR 04, 21 times in a year's period, but Surgeon 5 only performed a Lap App procedure in OR 04 once; therefore, the estimate for Surgeon 5 for Lap App may be less accurate than the estimate for Surgeon 2 for

Esophag because it is based on less data. These inconclusive results were common among the majority of the variables.

Table 3: OR 04 Regression Results

OR[OR 04]	p-value	Estimate (β)	Adjusted β	Std Error
Average	**0.56**	**-4.5**	**-4.05**	**18.49**
Esophag				
Surgeon 1	0.23	-3.8	-1.41	3.09
Surgeon 2	0.13	-11.8	0.53	7.59
Surgeon 3	0.83	1.8	2.60	8.07
Surgeon 4	0.59	6.2	3.34	10.77
All Surgeons	0.02	-4.2	-9.08	1.86
Lap App				
Surgeon 5	0.59	17.2	18.82	31.51
All Surgeons	0.94	-1.2	-3.61	18.02
Lap Chol				
Surgeon 6	0.52	-43.1	-43.09	60.81
All Surgeons	0.74	-4.8	-13.10	14.65
IPVA				
All Surgeons	0.98	-0.8	4.46	28.49

In the one year period of data analyzed, there are 816 different procedure codes, and often multiple codes are combined in one case. This limits the amount of data available for any one procedure type, especially when broken down even further to isolate the variables such as OR, weekday, and anesthesiologist. Interaction between variables provides an opportunity for additional impact on the duration. For example, certain surgeons may interact better with certain anesthesiologists than others. Methods used allow analysis of this interaction to be made, but this further narrows the amount of data available. Interaction of other variables, such as certain anesthesiologists in certain rooms, is not able to be analyzed using this method, unless models were made holding the room or anesthesiologist constant.

Scarcity of portions of data makes it difficult to put full confidence in all of the estimates in the models. Before making recommendations based on the estimates, the number of data points for each estimate must be investigated. If there is significant data in each model and the R^2 values are significant, the models could be used to assess if certain factors played a huge part in the duration. For example, you could see if certain procedures are done better in certain rooms. You could also see if certain anesthesiologists perform better than others. With the data from a one-year period analyzed in this paper, this was not the case for most of the models.

Although it proved to be difficult to assess many of the variables, the results did confirm that the surgeon did matter. Each individualized model was different than the overall model, and often significantly different from the others of that procedure type. When an overall model was created for each procedure type, the surgeon was included as a variable. The estimates were significantly different, as portrayed in the values for Esophag, shown in

Table **4**. This is consistent with Strum et al. (2000).

Table 4: Esophag model with all surgeons

Esophag	p-value	Estimate	Std Error
Average	**0.47**	**-1.8**	**20.17**
Surgeon 1	0.41	-12.3	14.73
Surgeon 2	0.01	31.2	11.02
Surgeon 3	<.0001	69.4	10.69
Surgeon 4	0.64	5.4	11.80
Surgeon 5	0.00	36.0	10.84
Surgeon 6	0.38	7.9	8.97
Surgeon 7	0.55	5.5	9.11
Surgeon 8	0.98	-0.4	15.07
Surgeon 9	0.94	1.2	14.93
Surgeon 10	<.0001	87.1	14.34
Surgeon 11	0.75	2.9	8.96
Surgeon 12	0.62	-7.6	15.23
Surgeon 13	0.03	-293.0	136.59
Surgeon 14	0.83	2.0	9.18
Surgeon 15	0.23	18.0	14.92
Surgeon 16	0.26	18.4	16.40

2.4 Conclusion

Estimating surgery durations is a challenge for surgery schedulers, as supported by a *t* test comparing the estimated surgery durations to actual surgery durations using data from a one year period at UnityPoint Health - Des Moines. To understand the impact of the variables that go into a surgery, multiple linear regression models were created for each of the top four procedure types. The models included variables of age of the patient, anesthesiologist, OR, number of residents, and weekday. The models proved to be ineffective in drawing significant conclusions based on the limited amount of data for certain variables. This supports the conclusion in Zhou et al. (1999).

This result can be applied to hospitals of all sizes and locations. The hospital analyzed is a level one trauma center and sees many different surgery procedures in a given amount of time. Even with the demand levels of a level one trauma center, there is not enough data for many of the procedure types to support an adequate regression model. Data sets that do have enough data have a high amount of variability in the durations, making effective comparison difficult. Therefore, this model would certainly not be useful in small hospitals with limited data, and even in large hospitals, challenges are presented. Our suggestion is to use historical data to check surgeon estimates, but not to solely predict or identify strong trends in input variables. Another suggestion is to break down the surgery duration estimation given by the surgeon into subsets, such as anesthesia, procedure, and cleanup. This will ensure that the surgeons are all using the same standards to estimate and time is not double counted, making the estimates more accurate.

Future research opportunities include trying logarithmic transformations of regression variables and trying it with more data than a one-year period. More surgeries and surgeons could be analyzed in addition to the top four analyzed in this study, especially with the *t* tests comparing estimated and actual duration times. Interactions between more variables can be added, and other variables can be investigated. Different parts of the surgery process than procedure time can be analyzed statistically as well, including anesthesia time and clean up time. Due to the nature of health services, surgery duration estimation will remain a challenge. Small steps toward understanding and predicting surgery durations can be made, and have significant impact on the success of a hospital.

CHAPTER 3: ANALYSIS OF SURGERY SCHEDULING POLICIES USING DISCRETE EVENT SIMULATION

Abstract

Surgery is a major source of revenue in a hospital. Due to the complexity of surgery cases, there is high variability in durations, making effective scheduling difficult. A simulation model was created considering a series of surgery cases in one operating room using historical data for surgery durations. The model was validated using real data as a comparison. Four scheduling policies were then evaluated: shortest estimated time first, longest estimated time first, most common surgery first, and adding an extra 20 minutes to each case in the existing order. Average minutes off schedule and average idle time between cases were compared. The policy of shortest estimated time first was found to demonstrate the best performance. This methodology can be applied to a variety of scheduling scenarios in a hospital.

3.1 Introduction

As health care costs rise tremendously in the United States (Askin & Moore, 2012; Davidson; Gupta & Denton, 2008), improving the efficiency of hospital operations has become imperative. As a vital source of revenue, surgery operations are a huge area of opportunity for efficiency improvement in a hospital setting. It is estimated that 40% of a hospital's costs and 68% revenue come from surgery (Jackson, 2002). Technology is advancing, thus aiding in the development of data collection software and systemized scheduling techniques.

When scheduling surgeries, many hospitals employ a Master Surgery Scheduling (MSS) approach. In this approach, specific surgeons or groups of surgeons are given a particular operating room (OR) on a specific day of the week to perform their surgeries. The schedule is cyclical; for example, the urology team may have OR 5 every Wednesday. There can be both centralized and decentralized MSS planning strategies. In decentralized systems, the surgeons have full control over the allocation of patients within their block time. On the other hand, in a centralized system, the OR planner makes the final decisions and typically uses data to support decisions.(van Oostrum et al., 2010) Hospitals often employ a hybrid approach, which is the case for UnityPoint Health - Des Moines, the health system analyzed in this study.

Hospitals must balance several priorities when scheduling surgeries. It is in the best interest of the hospitals to keep the surgeons satisfied, because the surgeons have the power to take their business elsewhere. Surgeons often want to be able to pick the time they perform the surgeries and prefer to not have breaks in between surgeries. Anesthesiologists are provided from an outside group, and high utilization is desired, which means they should not wait between surgeries. Surgery schedulers are also to be concerned with both indirect and direct patient delays. Indirect delays are between the time the patient makes a request for the appointment/surgery and when it is scheduled. Direct delays are between when the appointment/surgery is scheduled and when it actually occurs (Gupta & Denton, 2008). Patients anticipate entering the surgery at a certain time, and many have to fast before entering. The longer they wait, the longer they have to fast. Often, hospitals provide meals to family members of the patients when they have to wait, so the longer the wait time, the higher cost to the hospital as well (Lacey Andrews, personal communication, September 16, 2014). Rescheduling surgeries for another day is also not conducive to patient satisfaction. A scheduling system must have both high utilizations and be robust. They must be able to handle emergency situations while minimizing the number of times surgeons "cheat" the system by estimating too much or little time (van Oostrum et al., 2010). This can be a challenge when taking into account the varying priorities of those involved.

Operating room utilization can be calculated using many different methods. Surgeons or groups of surgeons who possess block schedules are held accountable to keeping a certain level of utilization during their block. This means they must ensure that they schedule enough surgeries each week within their block day and time to meet the utilization requirement, or else they will lose their block privileges. In another way, utilization can be calculated in terms of availability of operating room time and how much it is used. The interpretation adopted in this paper is that utilization is the amount of time an OR is occupied out of the working day.

Accurately estimating surgery durations is a significant challenge in the hospital setting, and also is the core of an effective surgery schedule. If a surgery goes longer than expected, patients have to wait and doctors and nurses may have to work overtime. If a surgery takes less time than expected, the utilization of the operating room goes down, compromising potential revenue and services. Scheduling of an OR affects many other departments of the hospital, such as nurse schedules and inpatient bed arrangement (Cardoen et al., 2009). The current practice is to have the surgeons provide an estimate of the surgery duration. The rationale is that the surgeon knows and understands more about that individual patient and the severity of their case than the surgery schedulers. Historical data are becoming increasingly available as electronic medical record technology advances. Many studies have been done to compare surgeon's estimates to estimates based on quantitative statistical methods using historical data (Joustra et al., 2013; Larsson, 2012; Laskin et al., 2012).

Various types of distributions used to simulate duration times have been studied in the literature. Strum et al. (2000) compared lognormal and normal distributions of surgery times. The analysis was also completed for two component procedures and the lognormal distribution model was found to perform better than the normal distribution model (Strum et al., 2003).

Spangler et al. (2004) used a simulation method to compare duration estimation models, and found the lognormal distribution with a location parameter to provide the most accurate estimation.

Simulation models have been utilized in research studies to improve surgery scheduling. Primary benefits of simulation include that it allows different scenarios and a large number of scenarios in a short time frame to be tried without making changes to the real system (Dexter et al., 1999). Stiglic and Kokol (2005) used multi-agent based simulation to optimize patient and nurse scheduling in ambulatory center. Simulation and time-series forecasting were used to identify patterns in demand. M'Hallah and Al-Roomi (2014) simulated allocating cases to rooms to enhance the under/over utilization and investigated different scheduling strategies. Everett (2002) aimed to reduce the waiting list by creating a simulation that chooses patients from the list and allocates them to time slots within the block schedule. It classifies urgency levels of the cases as urgent, semi-urgent, and routine. All of these simulation models proved to be useful in improving scheduling.

Linear programming has been used to optimize scheduling as well. Kuo et al. (2003) used a linear programming model to optimize OR time allocation among a group of surgeons in order to maximize revenue and minimize cost. Pham (2006) adopted a multi-mode blocking job shop model, formulated as a mixed integer linear program, to schedule elective and add-on cases. Tanfani and Testi (2010) applied a hybrid approach with both linear programming and simulation methods to optimize an OR schedule. They considered all sub-process from when a patient enters the system to when it is discharged. Dexter et al. (1999) found a linear programming model to be useful for minimizing labor costs, but only theoretically because many of the procedure types only occur once in several years, making the application unrealistic.

Many times integer programming models are NP-hard, making them difficult to solve and apply to a real hospital setting. Cardoen et al. (2009) created a model using heuristic algorithms based on multi-objective integer programming and branch-and-bound. The objective was to optimize operating theater schedules and the model was found to be NP-hard. Aringhieri at al. (2015) presented a two level metaheuristic algorithm that assigns both blocks to surgical specialties, as in a MSS, and also assigns patients to time slots within the blocks. This problem is also proved to be NP-hard.

Scheduling rules have been studied in the manufacturing sectors for decades (Vollmann, 2004), and have been recently applied to the health care setting. Testi et al. (2007) used a bin packing model, blocked booking method model, and simulation for a three-phase scheduling problem: determining the weekly scheduling capacity, MSS, and sequence of surgeries within the MSS. Longest Waiting Time first (LWT), Longest Processing Time first (LPT), and Shortest Processing Time first (SPT) rules for scheduling surgery cases are all evaluated. The SPT rule was found to yield the best results, with LPT resulting in the most overtime and delays. Sciomachen et al. (2005) also used simulation to compare LWT, LPT, and SPT. SPT was found to reduce the number of operations to be completed overtime by 54% when applied to a real hospital setting. LPT was found to reduce the amount of overtime. Harper (2002) simulated the following rules: First Come First Served (FCFS), LPT, SPT, and longest time first followed by shortest first (LTSC). LPT was found to be the best option, to increase the system throughput and utilization. Simulation has also been used to determine the amount of block time to provide each surgeon and which day a patient should be scheduled in order to maximize OR utilization. Dexter et al. (1999) used log-normal distributions to sample durations and simulated the following allocation rules: Next Fit, First Fit, Best Fit, Worst Fit. The term fit refers to allocating the surgery case to a time slot. We aim to investigate similar policies in a real hospital setting and additional policies.

In this paper, a simulation model was created for multiple series of surgery cases, and four different scheduling policies were evaluated: shortest estimated time first (SETF), longest estimated time first (LETF), most common surgery first, and an extra 20 minutes added to each case in the existing order. The paper is organized as follows: the methods are described in section two, including the process mapping, problem statement, model validation, and the scheduling policies evaluated. The results of the scheduling scenario comparison are presented in section three. The paper concludes with a summary and conclusions in section four.

3.2 Methods

3.2.1 Process Mapping

The health system, UnityPoint Health - Des Moines, in Des Moines, Iowa, was analyzed in this study. It includes a large hospital that is a level one trauma center and teaching hospital. The surgery scheduling process they employ and data from a one year period from September 2013-2014 were used to create the model. The hospital uses an MSS scheduling approach. Many hospitals use a similar scheduling process, making this applicable to hospitals across the country. The surgery scheduling process is as follows.

Forty-eight hours before the scheduled day of surgery, the cases scheduled for that day are allocated into time slots. This process, known as the add-on scheduling process, is shown in Figure 3. A case needing to be allocated a time slot is identified. Surgeons with a block are given preference for their times over surgeons without a block. If they request a certain time for that case or a certain order for all of their cases, the schedulers accommodate. If there is extra time in the day after all of the cases within blocks are scheduled, the unused blocks are released, and other elective cases requested for that day are added. Currently, it is rare for an elective surgery to not get scheduled on the day requested.

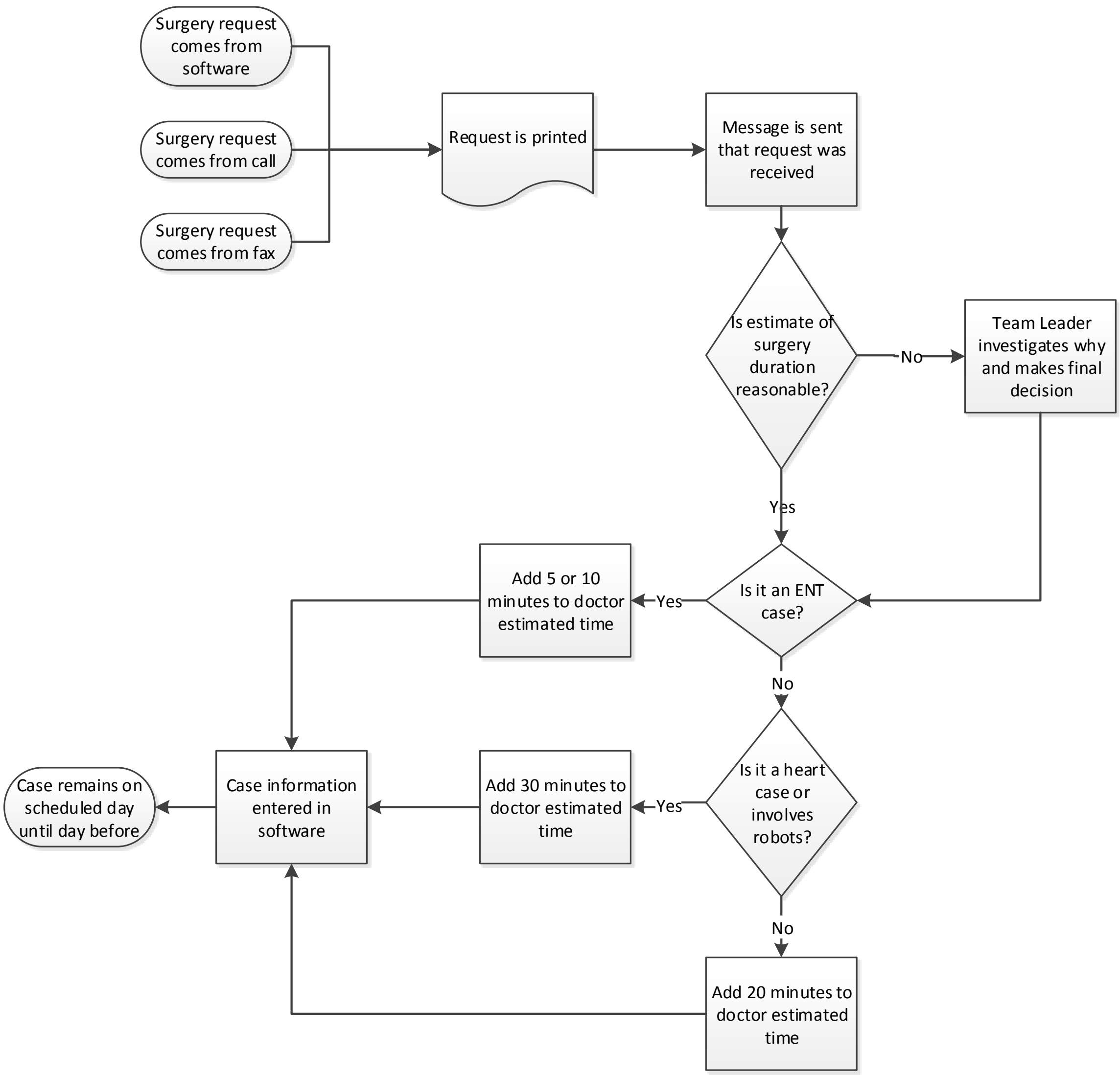

Figure 3: Add-on Scheduling Process

If the block surgeon or group of surgeons did not specify an order, the order is determined based on additional criteria. If there are multiple pediatric surgeries, they are scheduled youngest to oldest. Outpatient cases are scheduled before inpatient cases. For non-block cases, the surgeons can request a time, but understand that the time is not guaranteed, as surgeons with a block take precedence. Another consideration taken by the schedulers is the number of medical instruments available. They must ensure cases are scheduled in a way that does not surpass the number of instruments available. The schedulers also try to group together cases with similar procedure types, age of the patients, and equipment needed. The schedule must also accommodate the limited number of anesthesiologists available.

The schedule is finalized in the afternoon on the day before and the patients are notified of the time of their surgery. On the day of, a separate scheduling team is in charge of making changes if necessary. If a surgery goes significantly longer than expected, the succeeding case can be moved to another room if all resources necessary for the case are available.

3.2.2 Problem Statement

Eight series of surgeries were chosen to be simulated based on the historical data of actual days and operating rooms. The series were chosen as ones that are conducive to simulation, include common procedures, and include various characteristics such as consistent surgeons, multiple surgeons, and/or multiple procedure types. Series with no procedures that had two procedure types in one case and at least ten previous cases were chosen in order to get the most accurate duration prediction possible.

Series 1 is the first Thursday of the month and has four cases completed by the same surgeon, back to back, starting two hours into the day. Series 2 is the second Wednesday of the month and has seven cases by the same surgeon scheduled back to back, starting at the beginning of the day. Series 3 is the same day as Series 2, but in a different room. It has six surgeries, with the first four done by the same surgeon, and the last two done by two other surgeons. There is a 25 minute break between the fourth and fifth surgeries. Series 4 is the second Thursday of the month and has three surgeries scheduled back to back and done by three different surgeons. Series 5 is the third Monday of the month and has six surgeries, the first three done by one surgeon, the next two done by another, and the last done by another. The last three surgeries are the same, a common surgery. Series 6 is the third Tuesday of the month in a small OR specializing in ear, nose, and throat procedures. It has six surgeries scheduled. The first surgery is completed by one surgeon, the next four are the same procedure, but done by one surgeon, and then another, and the last surgery done by yet another, for a total of four different surgeons. There is a 20 minute break between the third and fourth surgery. Series 7 is in an OR specializing in vascular procedures. It is the second Wednesday of the month and has four different surgeries and four different surgeons performing the surgeries. There is a 110 minute break between the first and second surgery, a 75 minute break between the second and third surgery, and a 15 minute break between the third and fourth. Series 8 is the fourth Wednesday of the month and is five different procedures all performed by the same surgeon back to back. The cases in all of series are done in five different operating rooms. Three of the rooms are general surgery rooms, one is a room specializing in vascular surgeries, and one room is a small room specializing in ear, nose, and throat surgeries (Lacey Andrews, personal communication, August 19, 2014). A visual representation of the eight cases simulated is shown in Figure 4. Each color represents a surgeon; therefore, cases in the same color are performed by the same surgeon.

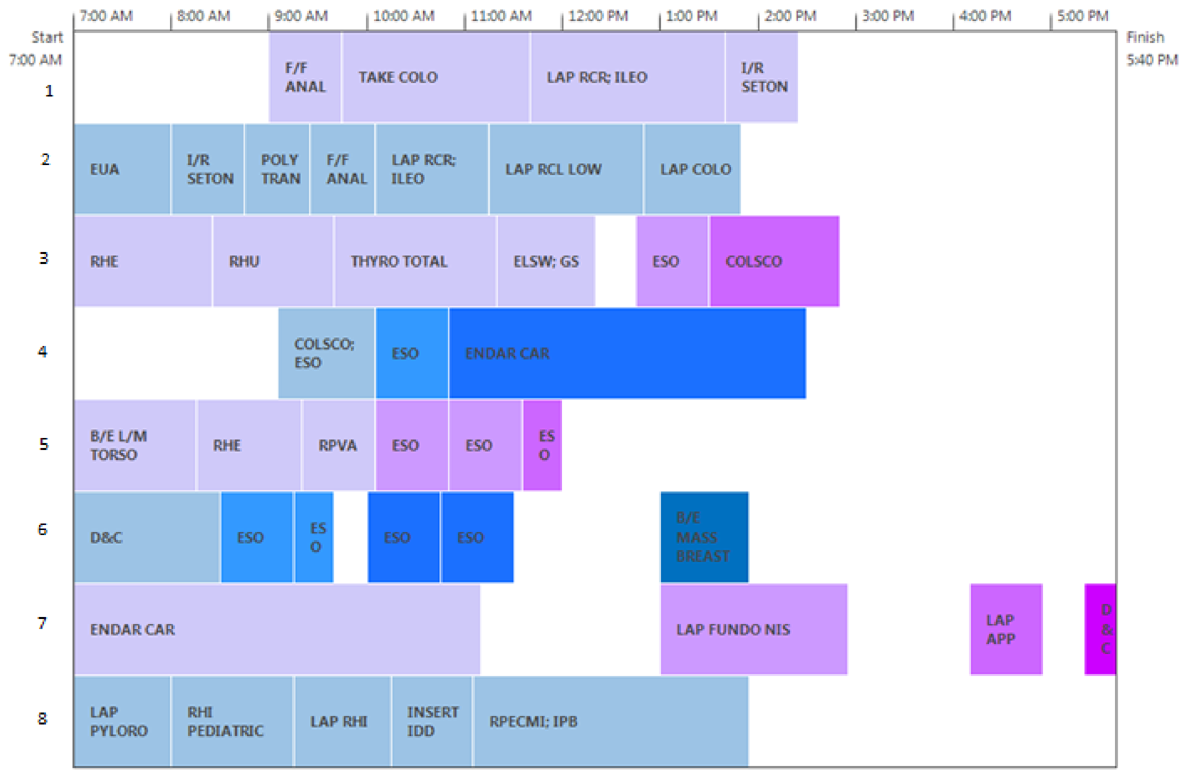

Figure 4: Timetable displaying the eight simulated cases

3.2.3 Arena Simulation

Each of the cases were simulated using Arena, and each model represented a day's schedule in one operating room. An entity represents a surgery case. In the simulation, the entity is created, and it reads the start time of the surgery from a text file. The start time is written in minutes, with time zero being 7:00 am. The model then reads the procedure type from a different text file and assigns it to that entity. The entity is delayed by a value of the surgery start time minus a value sampled from the distribution of the minutes a patient is ready before the scheduled time. A normal distribution based on the historical data for a one year period utilizing Input Analyzer was used. The distribution also includes time the patient was not ready at the scheduled time. The entity then moves to the next step as it is duplicated and the next entity, representing the next surgery case, goes through the process of being assigned a scheduled time, procedure code, and delay. The model then assigns a Ready Time to the original entity and writes it to a file. The time it takes to transfer the patient from pre-operation to the surgery room is considered arbitrary. The original entity then enters the process, or surgery. If the resource (OR) is being used by a previous case, the entity queues. The entity is processed for a duration based on the procedure type and a fixed value of 20 minutes for cleanup time is added to each duration. This value is based on the current practice of the hospital surgery schedulers. Distributions for each procedure type were estimated using the historical data and Input Analyzer. After the surgery process is done the entity leaves. Entity scheduled time, wait time, duration, and exit time are then written to a file and the entity is disposed. One thousand replications were completed and the average duration, minutes off schedule, minutes late, and start time were calculated for each surgery case using Microsoft Excel. A screenshot of the Arena simulation model with explanations is shown in Figure 5.

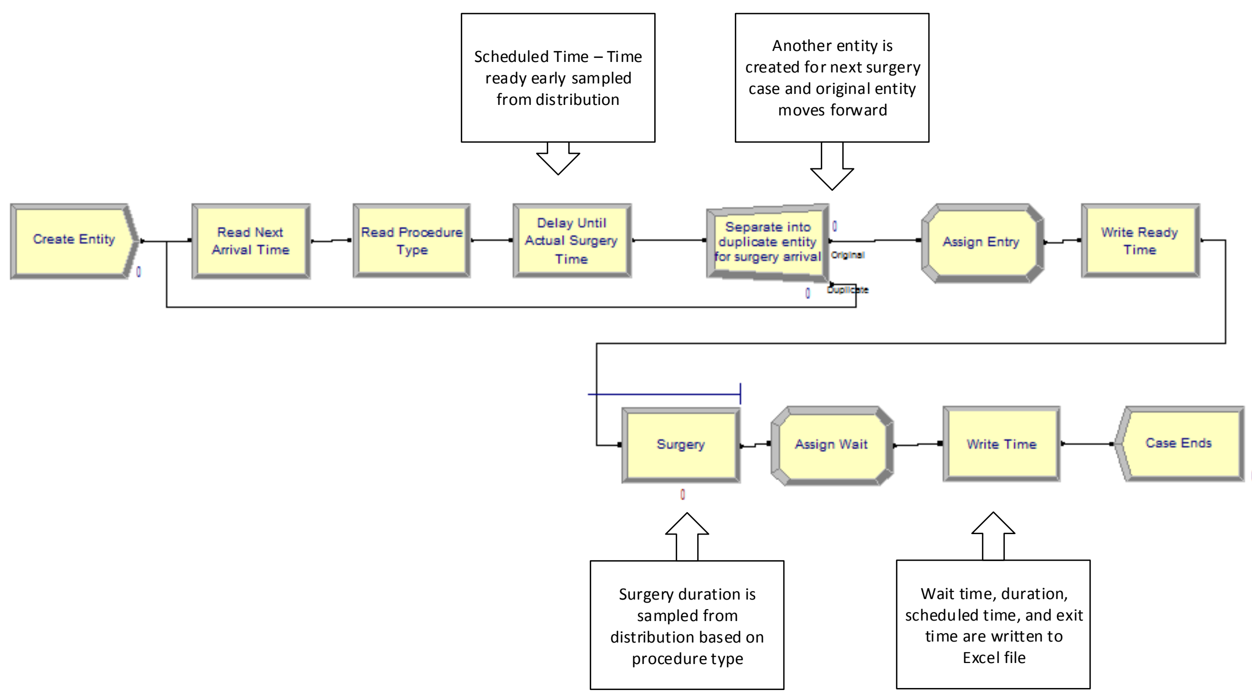

Figure 5: Snapshot of Arena model with descriptions of process

3.2.4 Model Validation

The simulation model was compared to the historical data of the actual and estimated durations to validate the model. A *t* test was done for each series of cases. In a comparison of estimated duration versus simulated duration for each series, 75% were not statistically different. The two series that were statistically different included a particular surgery that was estimated low, but the mean value was high for that procedure type. This represents the fact that the duration cannot always be estimated using historical data, as there are varying cases of severity and complexity. In a comparison of actual duration versus simulated duration, all eight cases were not statistically different, with $\alpha=0.05$, thus validating that the simulation represents reality. The simulated minutes off schedule were compared to the actual minutes off schedule as well. In 87.5% of the cases, the simulated minutes off schedule were not statistically different, with $\alpha=0.05$. When comparing the existing estimate to the simulated estimate of duration, the simulated estimate showed an improvement overall. The p value was higher for 75% of the cases. This validates that using historical data to create a distribution of durations is an accurate way to estimate durations; in many cases it is more accurate than the existing method of asking the surgeon and looking at the average of the last ten cases with the procedure and that surgeon.

3.2.5 Scheduling Policy Comparison

To evaluate the impact of scheduling policies, a variety of scheduling strategies have been simulated with the eight series. The baseline order is based on the existing schedule. Three of the eight series included breaks in between one or more surgeries. To allow for consistent comparison across scheduling rules, the existing schedules for those three series were adjusted so the breaks were eliminated. This adjustment results in each of the eight series representing sample cases, not necessarily actual cases that need to be fixed. The real data was used to provide a realistic set of sample series.

Each series was tried with four different scenarios: shortest estimated time first (SETF), longest estimated time first (LETF), most common first (MC), and adding a fixed 20 minutes to each case. The rationale behind most common first is that the most common surgeries would have more historical data for durations and the duration estimate may be more accurate. Twenty minutes was chosen as the fixed amount because it was found to be the average time between the patient arriving in the OR and the procedure starting. This time is referred to as anesthesia time. The current scheduling process does not directly account for anesthesia time like it does cleanup time. For each scenario, the average minutes off schedule, average minutes late, utilization, and average idle minutes between cases were calculated. Minutes off schedule is defined as the minutes the case started before or after its scheduled time. A positive number indicates the case started after the scheduled time, and a negative number indicates it started before the scheduled time. Minutes late is defined as the minutes the case started after the scheduled time, with any cases starting earlier being classified as zero minutes late. Idle minutes between cases are defined as the number of minutes the OR was unused between sequential cases.

3.3 Results and Analysis

When comparing the average minutes off schedule for the scheduling rules for all the series, adding twenty minutes extra time resulted in the lowest average minutes off schedule 88% of the time. This means that when compared to the existing order, it resulted in the most positive difference between the existing number of minutes off schedule and the number of minutes off schedule with the scheduling rule. This can be seen in Figure 6 in the purple line.

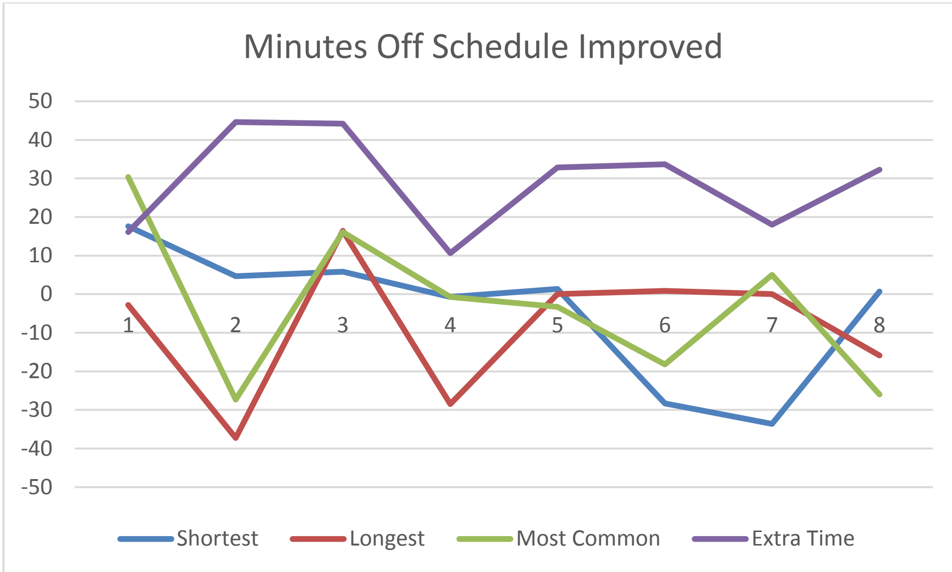

Figure 6: Scenario comparison results

This is intuitive, as allowing more time will always reduce the amount of time off schedule. In one series, putting the cases in order of most common to least common proved to be more effective in lowering the average minutes off schedule than adding extra time. In this series, the least common surgery had a difference between the estimated time and simulated time of 55 minutes. In the most common first scenario, this case would be scheduled last, eliminating the extra minutes off schedule for a succeeding case. In the current schedule, that case is second, so adding an extra twenty minutes to the case will leave the case being 35 minutes off schedule, still throwing off the rest of the cases. However, if an extra twenty minutes is added to each case, the idle time between cases increases, as shown in Figure 7. Extra time is represented in the purple line, and a negative number denotes that there was not an improvement in idle time from the existing schedule. As shown, adding an extra twenty minutes to each case never improved the idle time.

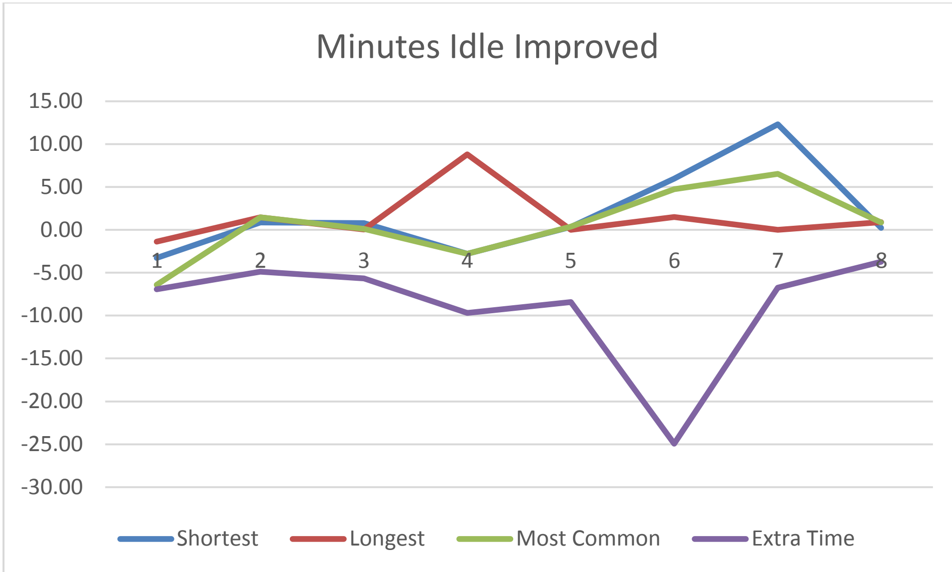

Figure 7: Idle time scenario comparison

In all eight cases, the number of idle minutes increased by an average of nine minutes when twenty minutes was added to each case. This affects utilization, and if there were four cases, and three sets of nine minutes between cases, that would be an extra 27 minutes, enough time for a small surgery case. An ideal scheduling policy would minimize the tradeoff between time off schedule and minutes of idle time, making the addition of extra time on the existing schedule not an ideal policy.

If the addition of an extra twenty minutes is eliminated, other scheduling rules emerge as the best option. It should be noted that shortest estimated time first is the best option in 50% of the cases. It was an improvement in both minutes off schedule and idle minutes in 67% of all series, making it the most ideal policy out of those investigated. There were two series in which SETF was not an improvement in minutes off schedule. These series each included a surgery case that was estimated to be one of the shortest, but simulated to go significantly over the estimation time. These cases went significantly over time in reality as well. By putting those cases at the beginning of the order, the schedule is thrown off for the rest of the day and the minutes off schedule are increased. This scenario is indicative of the fact that a single case can throw off an entire day's schedule and it is difficult to know which case that will be.

For most common first, the trend was not as clear. The policy resulted in the lowest number of minutes off schedule in two of the series, ignoring the extra time strategy. It obtained the highest number of minutes off schedule in three of the series, and somewhere in the middle for the remaining three series. Most common first had a mixed result in minutes of idle time as well; therefore, it can be said that the frequency of procedure type does not make a significant difference in the accuracy of the estimation. Longest estimated time first resulted in highest number of minutes off schedule in three of the eight series. It resulted in the lowest minutes off schedule in one series, only surpassing most common first by less than a minute. LETF only showed significant improvement in minutes idle in one series; therefore, we recommend avoiding LETF when possible.

It should be noted that it can often be difficult to schedule surgeries in whatever order is best. For example, it is more conducive to surgeon satisfaction if the cases scheduled for a surgeon to perform are consecutive. Surgeon satisfaction is a high priority for many hospitals; therefore, we recommend using the scheduling policies to schedule consecutive cases for a given surgeon, as opposed to all cases in a series together. They are more difficult to use with series in which surgeons perform a small number of cases in a row and there are multiple surgeons that do so. It is not realistic to schedule shortest to longest estimated duration for the entire series if the surgeons will have to wait in between their cases. The same goes for the other scheduling policies. Use of the policies can be highly beneficial within series that include one surgeon, which is common.

3.4 Conclusion

Surgery is a significant source of revenue for hospitals. Due to the nature of surgery, durations can vary tremendously, making it difficult to effectively schedule. A simulation model was created to study one operating room's schedule for a given day. Using data from a large Midwestern hospital, eight series of surgery cases were simulated. Five different scheduling policies were simulated within each series to identify the scheduling policy that resulted in the lowest number of minutes off schedule and the lowest number of idle minutes between surgery cases. The four policies investigated were shortest estimated time first (SETF), longest estimated time first (LETF), most common first (MC), and adding twenty minutes of extra time to each case. SETF was found to be the strategy that resulted in both the lowest minutes off schedule and minutes of idle time. This finding is consistent with research findings in the health care setting (Sciomachen, Tanfani, & Testi, 2005; Testi et al., 2007) and the manufacturing sectors (Vollmann, 2004). Results of this simulation model can be applied to many other hospitals, as scheduling processes are similar in both the United States and other countries.

Minimizing the number of minutes surgeries are off schedule will have a positive effect on patient and surgeon satisfaction, reduce the amount of reactive work the schedulers must do if a surgery goes significantly over time, and lowers costs to the hospitals for providing lunch vouchers or other accommodations. Minimizing the number of idle minutes between surgeries allows for additional surgeries to be scheduled, which brings in more revenue to the hospital and ultimately helps more patients. By following surgery scheduling policies that minimize minutes off schedule and idle time, hospitals are more effective in providing high quality services to the patients.

There are several ways this study can be advanced and further research can be done. Series with additional characteristics can be simulated to allow a more in depth comparison of scheduling policies. The current model does not include interactions between operating rooms, such as transferring a patient from one OR to another. In reality, if a surgery goes significantly longer than expected, the succeeding case may be transferred to another available OR if the patient and team are ready. This interaction complicates the schedule, but should be investigated to make the simulation model more realistic. Another opportunity for improvement is to include other steps of the surgery process, including pre and post-operation. After arriving at the hospital or surgery center, the patient goes through a pre-operation procedure. After the surgery is completed, the patient goes through a post-operation procedure. These procedures, most notably pre-operation, may have inefficiencies themselves that affect the surgery schedule. Cleanup time should be analyzed as well. This was not conducted in this study due to lack of sufficient data.

CHAPTER 4: GENERAL CONCLUSIONS

The health care industry faces many challenges, with one of the primary being rising costs. A significant opportunity to lower those costs to health care providers is to improve patient scheduling. This can apply to patient scheduling in many arenas, including inpatient, outpatient, and diagnostic imaging appointments. For large inpatient and outpatient providers, namely hospitals, surgery is a significant source of revenue; therefore, any improvement would have a large economic impact. When it comes to running an effective surgery department, accurate scheduling is extremely important. The biggest challenge in accurate scheduling is estimating surgery durations and scheduling in a manner that will minimize minutes off schedule and idle time between cases.

To investigate surgery duration estimations, the scheduling process at UnityPoint Health - Des Moines in Des Moines, Iowa was analyzed. Historical data on estimated and actual durations for the top four surgeries and the top surgeons within those surgeries were compared using a *t* test. The results show that 75% of the data sets investigated had estimates and actual durations that were not statistically different. To further investigate factors that go into how long a surgery will last, multiple linear regression models were created again for the top surgeries and surgeons. The input variables included were age of the patient, anesthesiologist, operating room (OR), number of residents, and day of the week. The β values for each input variable were compared across the models to find consistencies. Age was found to consistently have a minimal impact on the duration. For the other variables, consistencies were difficult to find, making multiple linear regression not an ideal method to find strong correlations among input variables. The correlation (R^2) varied largely as well, making multiple linear regression an invalid way to predict surgery durations as well.

In an effort to improve the surgery scheduling process within use of the current estimation process, simulation models were created, each to represent a series of cases in one operating room during one day. Eight series based on real series at UnityPoint Health - Des Moines were chosen based on how conducive the cases were to simulation and the variety of characteristics the series possessed. All simulations were validated with data on reality for those series. Four scheduling policies were investigated for each of the eight series: shortest estimated time first (SETF), longest estimated time first (LETF), most common first (MC), and adding an extra 20 minutes to each case in the existing order. SETF was found to be the best option for minimizing the minutes off schedule and minutes the OR is idle between cases. We recommend applying this policy to consecutive cases performed by a common surgeon, or cases that are all performed by different surgeons; it is not as realistic when applied to a series that includes multiple surgeons performing multiple surgeons, as surgeons prefer to perform their cases consecutively.

Results of this study and recommendations can be made to multiple types of surgery centers and even to other types of patient scheduling, such as outpatient appointment scheduling. As always, there are numerous opportunities for further research. The surgery process can be broken down into sub-processes, including pre-operation, anesthesia, procedure, cleanup, and post-operation. These sub-processes contain opportunities for improvement in efficiency. Other factors in surgery duration can also be investigated and other methods such as lognormal regression models can used. Other scheduling opportunities are also available, such as nurse and doctor scheduling, diagnostic resource scheduling, and inpatient bed scheduling. Statistical methods and simulation methods can be highly useful in improving health care efficiency. This study shows a few ways to do so, and provides recommendations based on those methods.

REFERENCES

Alper, J., Sanders, J., Saunders, R. S., & Institute of Medicine (U.S.). Roundtable on Value & Science-Driven Health Care. (2013). *Core measurement needs for better care, better health, and lower costs : counting what counts, workshop summary*.National Academies Press.

American Trauma Society. *Trauma Center Levels Explained*. Retrieved from http://www.amtrauma.org/?page=TraumaLevels

Aringhieri, R., Landa, P., Soriano, P., Tànfani, E., & Testi, A. (2015). A two level metaheuristic for the operating room scheduling and assignment problem. *Computers & Operations Research, 54*(0), 21-34.

Askin, E., & Moore, N. (2012). *The health care handbook : a clear and concise guide to the United States health care system*. St. Louis: Washington University in St. Louis.

Boyle, J., Wallis, M., Jessup, M., Crilly, J., Lind, J., Miller, P. et al. (2008). Regression forecasting of patient admission data. *Conference proceedings : 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference,* 3819-3822.

Broka, S. M., Jamart, J., & Louagie, Y. A. (2003). Scheduling of elective surgical cases within allocated block-times: can the future be drawn from the experience of the past? *Acta Chirurgica Belgica,* 103(1):90-94.

Burge, R. C. (2014). *Engineering solutions to America's healthcare challenges*. Boca Raton: Taylor & Francis.

Cardoen, B., Demeulemeester, E., & Beliën, J. (2009). Optimizing a multiple objective surgical case sequencing problem. *International Journal of Production Economics, 119*(2), 354-366.

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

Centers for Medicare and Medicaid Services. (2014). *Medicare Provider Utilization and Payment Data: Outpatient*. Retrieved from https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Provider-Charge-Data/Outpatient.html

Chakraborty, S., Muthuraman, K., & Lawley, M. (2010). Sequential clinical scheduling with patient no-shows and general service time distributions. *IIE Transactions, 42*(5), 354-366.

Davidson, S. M. (2013). *A new era in U.S. health care : critical next steps under the Affordable Care Act*. Stanford, California: Stanford Briefs, an imprint of Stanford University Press.

Dexter, F. (2012). A Brief History of Evidence-Based Operating Room Management: Then and Now. *Anesthesia and analgesia, 115*(July 2012), 10-11.

Dexter, F., Epstein, R. H., Lee, J. D., & Ledolter, J. (2009). Automatic updating of times remaining in surgical cases using bayesian analysis of historical case duration data and instant messaging; updates from anesthesia providers. *Anesthesia and analgesia, 108*(3), 929.

Dexter, F., Macario, A., Traub, R. D., Hopwood, M., & Lubarsky, D. A. (1999). An Operating Room Scheduling Strategy to Maximize the Use of Operating Room Block Time: Computer Simulation of Patient Scheduling and Survey of Patients' Preferences for Surgical Waiting Time. *Anesthesia and Analgesia, 89*, 7-20.

Dexter, F., Traub, R., & Qian, F. (1999). Comparison of Statistical Methods to Predict the Time to Complete a Series of Surgical Cases. *Journal of Clinical Monitoring and Computing, 15*(1), 45-51.

Dexter, F., Traub, R. D., & Lebowitz, P. (2001). Scheduling a delay between different surgeons: cases in the same operating room on the same day using upper prediction bounds for case durations. *Anesthesia and analgesia, 92*(4), 943.

Eijkemans, P. D. Marinus J. C., van Houdenhoven, P. D. M., Nguyen, M. S. T., Boersma, P. D. E., Steyerberg, P. D. Ewout W., & Kazemier, P. D. G. (2010). Predicting the Unpredictable: A New Prediction Model for Operating Room Times Using Individual Characteristics and the Surgeon's Estimate. *The Journal of the American Society of Anesthesiologists, 112*(1), 41-49.

Erdogan, S. A. (2010). Surgery Scheduling and Planning. *Wiley Encyclopedia of Operations Research and Management Science*.

Everett, J. E. (2002). A Decision Support Simulation Model for the Management of an Elective Surgery Waiting System. *Health Care Management Science, 5*(2), 89-95.

Graban, M. (2009). *Lean hospitals : improving quality, patient safety, and employee satisfaction*. Boca Raton: CRC Press.

Gupta, D., & Denton, B. (2008). Appointment scheduling in health care: Challenges and opportunities. *IIE Transactions, 40*(9), 800-819.

Hall, R. W. (2006). *Patient flow : reducing delay in healthcare delivery*. New York: Springer.

Harper, P. R. (2002). A Framework for Operational Modelling of Hospital Resources. *Health Care Management Science, 5*(3), 165-173.

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.

Joustra, P., Meester, R., & van Ophem, H. (2013). Can statisticians beat surgeons at the planning of operations? *Empirical Economics, 44*(3), 1697-1718.

Klassen, K. J., & Yoogalingam, R. (2013). Appointment system design with interruptions and physician lateness. *International Journal of Operations & Production Management, 33*(4), 394-414.

Koenig, T., Neumann, C., Ocker, T., Kramer, S., Spies, C., & Schuster, M. (2011). Estimating the time needed for induction of anaesthesia and its importance in balancing anaesthetists' and surgeons' waiting times around the start of surgery. *Anaesthesia, 66*(7), 556-562.

Kuo, P. C., Schroeder, R. A., Mahaffey, S., & Bollinger, R. R. (2003). Optimization of operating room allocation using linear programming techniques. *Journal of the American College of Surgeons, 197*(6), 889-895.

Larsson, A. (2012). The accuracy of surgery time estimations. *Production Planning & Control, 24*(10-11), 891-902.

Laskin, D. M., Abubaker, A. O., & Strauss, R. A. (2012). Accuracy Of Predicting The Duration Of A Surgical Operation. *Journal of Oral and Maxillofacial Surgery*, 71(2),446-7.

Luo, J., Kulkarni, V. G., & Ziya, S. (2012). Appointment Scheduling Under Patient No-Shows and Service Interruptions. *Manufacturing & Service Operations Management, 14*(4), 670-684.

Mathias, J. M. (2011). Surgical Scheduling: Taking an important role to the next level. *OR Manager, 27*(3).

Muthuraman, K., & Lawley, M. (2008). A stochastic overbooking model for outpatient clinical scheduling with no-shows. *IIE Transactions, 40*(9), 820-837.

M'Hallah, R., & Al-Roomi, A. H. (2014). The Planning and Scheduling of Operating Rooms. *Computers & Industrial Engineering, 78*, 235-248.

Patrick, J., Puterman, M. L., & Queyranne, M. (2008). Dynamic Multipriority Patient Scheduling for a Diagnostic Resource. *Operations Research, 56*(6), 1507-1525.

Pease, A. (2013). The Hospital Game: Optimizing Surgery Schedules to Save Resources, and to Save Lives. *SAS Global Forum 2013, 154*(2013), 1-14.

Pham, D.-N. (2006). Surgical Case Scheduling as a Generalized Job Shop Scheduling Problem. *European Journal of Operational Research, 185*(1), 1011-1025.

Rockwell Automation Technologies, Inc. (2007). Arena (computer software).

Schütz, H.-J., & Kolisch, R. (2012). Approximate dynamic programming for capacity allocation in the service industry. *European Journal of Operational Research, 218*(1), 239-250.

Sciomachen, A., Tanfani, E., & Testi, A. (2005). Simulation models for optimal schedules of operating theatres. *International Journal of Simulation*, *6*(12-13), 26-34.

Smith, M. D., & Institute of Medicine (U.S.). Committee on the Learning Health Care System in America. (2013). *Best care at lower cost : the path to continuously learning health care in America*. Washington, D.C.: National Academies Press.

Spangler, W., Strum, D., Vargas, L., & May, J. (2004). Estimating Procedure Times for Surgeries by Determining Location Parameters for the Lognormal Model. *Health Care Management Science, 7*(2), 97-104.

Stiglic, G., & Kokol, P. (2005). Intelligent patient and nurse scheduling in ambulatory health care centers. *Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, 5*, 5475.

Strum, D. P., May, J. H., Sampson, A. R., Vargas, L. G., & Spangler, W. E. (2003). Estimating times of surgeries with two component procedures: comparison of the lognormal and normal models. *Anesthesiology*, *98*(1), 232.

Strum, D. P., May, J. H., & Vargas, L. G. (2000). Modeling the uncertainty of surgical procedure times: comparison of log-normal and normal models. *Anesthesiology, 92*(4), 1160.

Strum, D. P., Sampson, A. R., May, J. H., & Vargas, L. G. (2000). Surgeon and type of anesthesia predict variability in surgical procedure times. *Anesthesiology*, *92*(5), 1454.

Tanfani, E., & Testi, A. (2010, February). Improving surgery department performance via simulation and optimization. *Health Care Management (WHCM), 2010 IEEE Workshop* (pp. 1-6). IEEE.

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

van Oostrum, J., Bredenhoff, E., & Hans, E. (2010). Suitability and managerial implications of a Master Surgical Scheduling approach. *Annals of Operations Research, 178*(1), 91-104.

Vollmann, T. E. (2004). *Manufacturing planning and control systems for supply chain management* (5th ed.). New York: McGraw-Hill.

Wright, I. H., Kooperberg, C., Bonar, B. A., & Bashein, G. (1996). Statistical modeling to predict elective surgery time. Comparison with a computer scheduling system and surgeon-provided estimates. *Anesthesiology*, *85*(6), 1235.

Zhou, J., Dexter, F., Macario, A., & Lubarsky, D. A. (1999). Relying solely on historical surgical times to estimate accurately future surgical times is unlikely to reduce the average length of time cases finish late. *Journal of Clinical Anesthesia*, *11*(7), 601-605.