The Effect of Surgeon-Scrub Nurse Collaboration on Predicting Operative Times

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE SCHOOL
OF UNIVERSITY OF MINNESOTA
BY

Yuanyuan Feng

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

Adviser: Sandra J. Potthoff, Ph.D.

September 2013

Acknowledgements

First and foremost, I would like to offer my sincere gratitude to my advisor, Dr. Sandra J. Potthoff, for her valuable mentorship and help throughout my master's research. I would like to thank Dr. Gyorgy Simon for his expertise in statistics and data mining. I'm deeply grateful for his encourages and helps when I encountered problems in my research. I would like to thank Dr. David Pieczkiewicz for his great support in the course of my master's and valuable feedback for my work at all times. I am indebted to my family members and friends who have constantly supported me all through my life.

Abstract

Background: Operating room managers need to construct the surgery schedule for the next day by synthesizing information on estimated surgery duration, staff information, and surgeons' information. The purpose of this study is to assist operating room managers' decision making one day before the surgery by developing the predictive model for operative times taking into account the staff information.

Methods: 10,960 cases in a health system in Middle West are analyzed. The outcomes are the mean absolute errors of the predictions and the correlation between the predicted operative time and the observed durations, and the predictors include surgeon-scrub nurse pair IDs, individual surgeon IDs and individual scrub nurse IDs. Lasso regression modeling on the logarithm of the operative time is performed.

Results: The unexplained variation of the residuals of the model, which only includes log scheduled duration and procedure type, can be further explained by the surgeon-scrub nurse collaboration frequency. Besides, the model that include surgeon-scrub nurse pair IDs, surgeon IDs and scrub nurse IDs can reduce the mean absolute errors by 8.47 minutes, compared with the scheduled procedure duration.

Conclusion: The more surgeons and scrub nurses collaborate, the less time a surgery will take. Including surgeon-scrub nurse pairs in the predictive model, the prediction errors can be reduced.

Table of Contents

List of Tables

List of Figures

CHAPTER 1. INTRODUCTION

1.1 Background and Context

The operating rooms (ORs) are one of the most critical and expensive resources within a hospital, with up to 42 percent of hospital revenues generated (HCFMA, 2005) and about 10 percent of hospital costs consumed (Viapiano and Ward, 2000). Operative time, which is defined as the time from incision to closure of the surgical wound, is closely related to operating room efficiency – the ratio of the output to input (Dexter et al., 2007). According to Macario (2011), every operative minute costs $15 to $20 for a basic surgical procedure and charges $62 on average.

Accurate predictions of operative times are critical to operating room management. A longer than predicted operative time results in a late start for the following surgery and, potentially, for the rest of the surgeries in that day's schedule and leads to over-utilized OR time – extra demand on staff time. A shorter than scheduled operative time results in holes on the schedule, which represent under-utilized OR time – staff are paid without working. Both over-utilized and under-utilized OR time increase the OR input, leading to inefficient use of OR time.

Most efforts in predicting operative times uses characteristics of elective cases, which can be well planned and usually scheduled more than two days in advance, and adopts regression models (Dexter et al. 2013). For example, the complexity of procedures, surgeons' estimations, and patients' characteristics are often applied as explanatory variables in the models. The estimations can be used by operating room managers in short term scheduling, when they make decisions on assigning procedures to ORs and the number of ORs needed, usually about one week before the surgery day.

1.2 Problem Statement

Operating room managers make final changes on the surgery schedule and assign staff to each procedure one day before surgery. This process is described by May et al. (2011) as the very short term scheduling phase. In very short term scheduling phase, information about estimated surgery durations, scheduled staff, and participating surgeons is synthesized by operating room managers to construct the next day's surgery schedule to achieve operating room efficiency.

There are two main tasks for operating room managers in very short term scheduling – moving cases around the ORs to reduce the number of operating rooms to be opened the next day and assigning staff to each procedures to decrease the over-utilized and under-utilized OR time. Without accurate estimations of procedure durations, moving cases to reduce the number of ORs will prolong staff working time – increasing over-utilized time. Without information about surgical team collaboration, assigning staff would be imprecise and influence the whole schedule.

Many research studies have found that statistical models to predict procedure duration improve duration estimation accuracy by moderate amounts. (Strum, 2000;

Eijkemans, 2010; Stepaniak, 2010). However, these models fail to take into account the information of surgeries gained one day before surgery, specifically the impact of surgical team collaboration on procedure surgical time.

1.3 Statement of Purpose

Operating room staffs consist of anesthetists, circulating nurses, scrub nurses, and surgical technicians. Staff assignment in very short term scheduling mainly focuses on assigning circulating nurses and scrub nurses into different procedures. The main responsibilities of circulating nurses are to check instruments pre- and post-surgery, to handle specimens, to monitor equipment during the surgery, and to respond to comfort of patients. The main duties of scrub nurses are to select and handle instruments and supplies used for the operation, to pass them to surgeons and to check all items both pre- and post-surgery. In conclusion, circulating nurses have indirect contact with surgeons and patients during the surgery, while scrub nurses have intensive communicate with surgeons and have direct contact with patients.

Previous study has quantified the relationship between surgical team familiarity and operative times. Xu et al. (2013) used generalized estimating equation regression modeling to analyze the effect of attending and assisting surgeons' familiarity on operative time. The surgical team familiarity was defined by Xu et al. (2003) as the cumulative team collaboration frequency. They found that the surgical team familiarity accounted for an additional reduction of 16 minutes after 10 prior collaborations.

The purpose of this study is to identify the relationship between procedure operative time and surgeon-scrub nurse collaboration and to build a prediction model taking into account the surgeon-scrub nurse pairs, which the operating room managers will create based on the staff information gained one day before the surgery. This study will focus on the surgeon-scrub nurse collaboration, because the interactions between the surgeon and the scrub nurse are more numerous and are likely to have a larger impact on surgical procedure operative time than the interaction among other professionals in the surgical team (Undre, 2006). This study will contribute to the existing surgical team familiarity research by exploring the effect of the collaboration of two individual professionals in the surgical team on the operative time. This study is exploratory in nature and seeks primarily to open discussion and opportunities for future research.

1.4 Research Questions

The central research questions formulated for this study were:

1. Is surgeon-scrub nurse collaboration frequency correlated with the operative times?
2. Can staff information contribute to the accuracy of the operative time predictions?

CHAPTER 2. LITERATURE REVIEW

2.1 Predicting Operative Times

Operative time prediction, due to its inherent uncertainty and unpredictable variability, had been of interest for at least 35 years (Strum et al., 2000A). Researches on operative time prediction can be divided into two parts.

The first part of researches focuses on identifying significant factors in the predictive model development. Wright et al. (1996) conducted both retrospective and prospective studies in analyzing surgeons' performance in predicting the duration of procedures. They found that surgeons provided more accurate time estimates than did the scheduling software adopted in their institution, although surgeons were often prone to underestimate the duration of surgery. Further, they developed regression models for predicting operative times by including, as the predictors, the surgeons' estimates, estimations provided by the scheduling software and other factors collected. Their model provided small improvement in predicting surgery durations. Strum et al. (2000a) analyzed procedure types' variability, individual surgeons' differences and types of anesthesia and found that surgeon was the most important factor of variability in surgical procedure times for longer surgeries. Instead of analyzing all CPT codes, they examined the three most numerous CPT codes within each of the 20 categories. They used a main-effects linear model, which included procedure types, surgeons' factors, and types of anesthesia, in comparing their effects. But they didn't analyze the interaction effects between variables. Stepaniak et al. (2010) built an ANOVA model that included five surgeon factors – gender, age, work rate, team composition and start time of the day – in estimating operative times. The most significant factors were found to be team composition, experience and time of the day. Eijkemans et al. (2010) used the different procedure types, surgeons' estimates, surgical team characteristics, and patient characteristics as predictors to predict operative times in a general surgery department. They found that characteristics of the operation and the team had the largest predictive performance, whereas patient characteristics had a modest but significant effect on operative times and that the surgeons' estimates had a substantial contribution to the prediction. Their model increased shorter-than-predicted operative time by 2.8 min per case and reduced longer-than-predicted operative time by 6.6 min per case, compared with the surgeons' estimates based on the historical data.

Dexter et al. (2008) conducted a systematic review of general thoracic surgery and identified that CPT codes, surgical teams, and types of anesthesia were the most important factors when predicting operative times. They suggested that information that can be gained one day before the surgery should be used in predicting operative times and that the surgeons' performance during a surgery could be adjusted by anesthesia. Cassera et al. (2009) conducted a retrospective study on laparoscopic procedures to analyze the effects of a surgical team size on the operative time. The surgical teams in their study consist of surgeons, surgeon assistants, anesthesiologists, nurses and other technicians and observers. They found that when procedure complexity and patient condition were held constant, adding one individual to a team predicted a 15.4 minutes increase in the operative time. Through their observation in ORs, they found that high turnover and short-term involvement of nurses could prolong the surgery and their findings confirmed their model's results. Zhen et al. (2012) conducted a retrospective study on general surgeries to analyze the influence of surgical team size on general surgeries' operative time. Their model indicated that when procedure complexity and patient condition were constant, adding 1 team member predicted a 7-minute increase in procedure duration. Gillespie et al. (2013) conducted a prospective observational study to describe factors that contributed to deviation from expected duration of operations. The factors they included were communication failures, intraoperative interruptions, team familiarity, unplanned surgery and prebriefings. They found that the most significant factor is the number of communication failures. Through their observation in the operating rooms, they found that the insufficient and inaccurate communication between surgeons and nurses were the most numerous communication failures in the surgical team. Elbardissi et al. (2012) used multivariate generalized estimating equation regression model to quantitatively evaluate the effects of both individual surgeon experience and the cumulative experience of attending surgeon-cardiothoracic fellow collaborations in isolated coronary artery bypass graft procedures. They found that the influence of attending-fellow pair experience far exceeded the influence of surgical experience with beta estimates for attending-fellow pair experience nearly three times that of attending surgeon experience. Xu et al. (2013) conducted a longitudinal study on bilateral reduction mammoplasty procedure to quantify the association between surgical team familiarity and operative time. They found that in addition to individual surgeon experience, team familiarity contributed to reductions in operative time, for example, surgical team familiarity accounted for an additional reduction of 16 minutes after 10 prior collaborations.

The second part of operative time researches focuses on the fitness of known distributions, especially the normal distribution and lognormal distribution for developing the predictive models. The lognormal distribution is one whose logarithms are normally distributed. Strum et al. (2000b) proposed that operative times, which take on values from zero to infinity and consist of procedures that take much longer than average times, fitted well in lognormal distribution. They tested the goodness of fit of both lognormal and normal models to their data and found that the two-parameter lognormal model was statistically superior to the normal model for modeling one and two component procedures (Strum et al. 2000B and 2003). May et al. (2000) added the parameter – location parameter to a lognormal distribution and suggested that when the skewness of the data is greater than 0.35, use the three-parameter lognormal with the location parameter, otherwise, use the two-parameter model. Spangler et al. (2004) estimated the location parameter of the lognormal distribution and found that using the best order statistic instead of the median might lead to higher goodness-of-fit.

2.2 Surgical Team Collaboration

Previous studies have suggested qualitative and behavioral methods can be applied to assess surgical team performance. Leach et al. (2009) conducted a qualitative study in assessing factors that influence the team performance. They found that effective interpersonal skills, effective communication processes, respect for one another, knowledge of their role and of the procedure, and the ability to anticipate events are contributing factors to effective team functioning. Two years later, Leach et al. (2011) conducted another qualitative study to observe and interview individual surgical team members to analyze the factors affecting the role and role interactions of members of the surgical team. This study suggested that team members who have worked with other surgical team members create a set of expectations about how they will carry out their role responsibilities. Their behaviors will be influenced by their expectations. So, the more times the team work together, the more they will know about each other and their behavior will be adjusted according to their experience.

Healey et al. (2006) described the surgeon-scrub nurse interaction in detail – "The surgeon requires instruments and swabs and the nurse provides and confirms them. Both members of this unit need to monitor each other's states and their process in order to coordinate effectively. The nurse must anticipate the surgeon's requirements and the surgeon must appreciate that the nurse depends on others to work effectively. The surgeon primarily controls their work. However, the nurse may need to shift control temporarily if an issue or a problem arises that compromises their performance or the patient's safety." From this description, we can see that the surgeon-scrub nurse interaction requires efficient communication, coordination, cooperation, leadership and monitoring – high cumulative surgeon-scrub nurse experience.

Undre et al. (2006) interviewed surgical team member from the four specialists groups (surgeons, anesthetists, operating department practitioners, and nurses) to assess the perceptions of their teamwork. They found that all professionals agreed that the surgeon-scrub nurse collaboration was weighted most among communication between other professionals and that the quality of teamwork was acceptable to the team members and could be further improved.

2.3 Conclusion

The research studies on operative time predictions identified many factors affect the operative times, which can be divided into two parts – information that can be gained at least two days before the surgery and information that can be gained one day before the surgery, which includes surgical team information. The limited number of studies that analyzed the effect of surgical team collaboration on the operative times only quantified the effects of attending surgeons- assisting surgeons collaboration.

According to our reviews in qualitative studies on surgical team collaboration, we found that the surgeon and scrub nurse collaboration is one of the most numerous and important during a surgery. Further, the studies conducted by Undre et al. (2006) and Gillespie et al. (2013) revealed that there were potential problems in the collaboration between surgeons and scrub nurses.

In this study, we analyzed the effect of surgeon and scrub nurse collaboration on operative times. We used the scheduled procedure durations in controlling the surgeons' estimations, surgeons' factors, and patients' factors, because our scheduled procedure durations were estimated based on the information gained at least two days before the surgery. Besides, in our analysis, we controlled the procedure types using primary procedure IDs. Based on the previous studies in operative time prediction, our model serves as a further attempt in improving prediction accuracy using staff information that can be gained one day before the surgery.

CHAPTER 3. METHODOLOGY

3.1 Data Description

The original data are electronic time stamped data requested from the Surgery Department of a middle size health system in Midwest. The data set contains 32,061 records for 11,259 cases. The dataset recorded one primary surgeon and multiple scrub nurses per case. The data is organized by the primary key of staff ID and staff type having a total of 24 attributes.

Since in the original dataset unique surgical procedures are broken into multiple records, the first task is to summarize each unique case into a single record.. In order to improve the computational performance, the original dataset is transformed into a sparse matrix that contains all the explanatory variables and a separate (dense) matrix that contains the actual procedure duration having the each scrub nurse ID, primary surgeon ID and procedure type ID as separate columns. Next, for each surgeon-scrub nurse pair, we created a unique ID and transformed the original data set into a sparse design matrix that has the unique surgeon-nurse pairs as columns. The cleaned data has 11,259 unique cases performed in the Surgery Department from May 11, 2011 to Oct 31, 2012. Out of the 11,259 cases, 10,960 had at least one scrub-nurse participating; the remaining cases (without the participation of a nurse) are discarded. We also discarded one case with missing procedure end time and 26 cases with missing the staff information.

3.2 Data Analysis

3.2.1 Model Building

In the study, we built four linear models with distinct sets of predictors (Table 1). The logarithm of the actual procedure duration is the dependent variable in all four models. The actual procedure duration and schedule procedure duration are log transformed because of their right skewness (Strum et al. 2000b). Model 1 and Model 2 are developed to compare the effect of surgeon-scrub nurse collaboration on the operative times with the effect of individual surgeons and scrub nurses. Given the situation that surgeons may collaborate with different scrub nurses, which leads to new pairs, Model 3 is developed to compensate the limited records of surgeon-scrub nurse pairs in the dataset. The base model serves as the baseline for the model comparison.

Table 1. Variables in the Models

Model	Dependent Variable	Independent Variables
Base Model	Log(actual duration)	Procedure Type ID, log(scheduled duration)
Model 1	log(actual duration)	Procedure Type ID, log(scheduled duration), Surgeon-Scrub Nurse Pair ID
Model 2	log(actual duration)	Procedure Type ID, log(scheduled duration), Surgeon ID, Scrub Nurse ID
Model 3	log(actual duration)	Procedure Type ID, log(scheduled duration), Surgeon-Scrub Nurse Pair ID, Surgeon ID, Scrub Nurse ID

3.2.2 Reducing the Number of Coefficients

The four models contain a large number of covariates – 719 predictors in the base model, 1498 predictors in model 1, 824 predictors in model 2 and 1603 predictors. Thus ordinary least squares (OLS) estimates obtained by minimizing the residual squared error are unsuitable in this study – the large number of coefficients lead to over-fitting with low bias but large variance and it’s hard to interpret the predictors effects using OLS estimates (Tibshirani, 1996). This calls for reducing the number of coefficients.

The most common strategies for variable selection are subset selection methods and shrinkage methods (Figure 1).

Subset selection methods retain a subset of the variables and eliminate the rest from the model and then use least squared regression to estimate the coefficients of the retained variables. The best-subset selection is feasible for the models that have less than 40 variables. Since the models in this study have more than 700 variables, we exclude using this method. The forward stagewise selection is regarded as inefficient (Bühlmann et al., 2011), so we exclude using this method. The stepwise methods, especially forward stepwise method, are suitable for our data. However, there are two main problems with these methods. First, the stepwise methods are discrete processes of either eliminating or retaining the variable; they introduce a high variance. Secondly, they use a greedy algorithm, which builds up a model piece by piece. So they fail to consider all possible solutions. Thus, we exclude using the stepwise methods.

The shrinkage methods are more continuous and have less variability compared with subset selection. Ridge regression and LASSO impose a penalty parameter λ on the coefficients. However, ridge regression shrinks the coefficients towards zero but does not set any coefficients to zero. In certain sense, ridge regression fail to reduce the number of coefficients and it is difficult to interpret the results of ridge regression. LASSO, on the other hand, combines the benefits of both subset selection and ridge regression. It can set some coefficients equal to zero and shrink some coefficients towards zero.

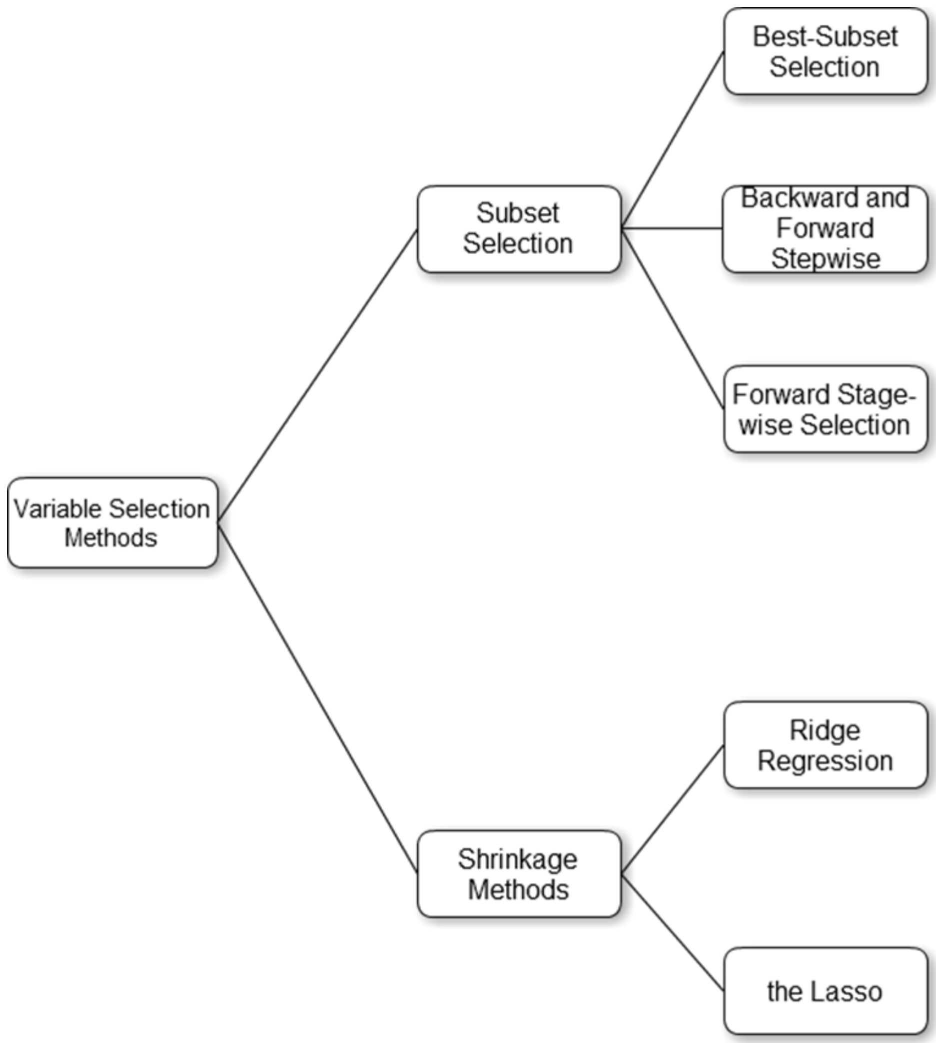


Figure 1. Variable selection methods.

LASSO estimates regression coefficients by minimizing the residual sum of squares and then applying a penalty by adding to the absolute value of the coefficients. The estimates are chosen to minimize

$$\frac{1}{2n}\sum(y_i - \alpha - \beta' x_i)^2 + \lambda\sum|\beta_j|$$

The penalty is determined by λ. If λ = 0, the lasso is the same as OLS; as λ increases, fewer the estimates are preferred. In this study, λ is chosen to get the maximum correlation between predicted value and observed value, in other words, to minimize the mean squared error (Figure 2).

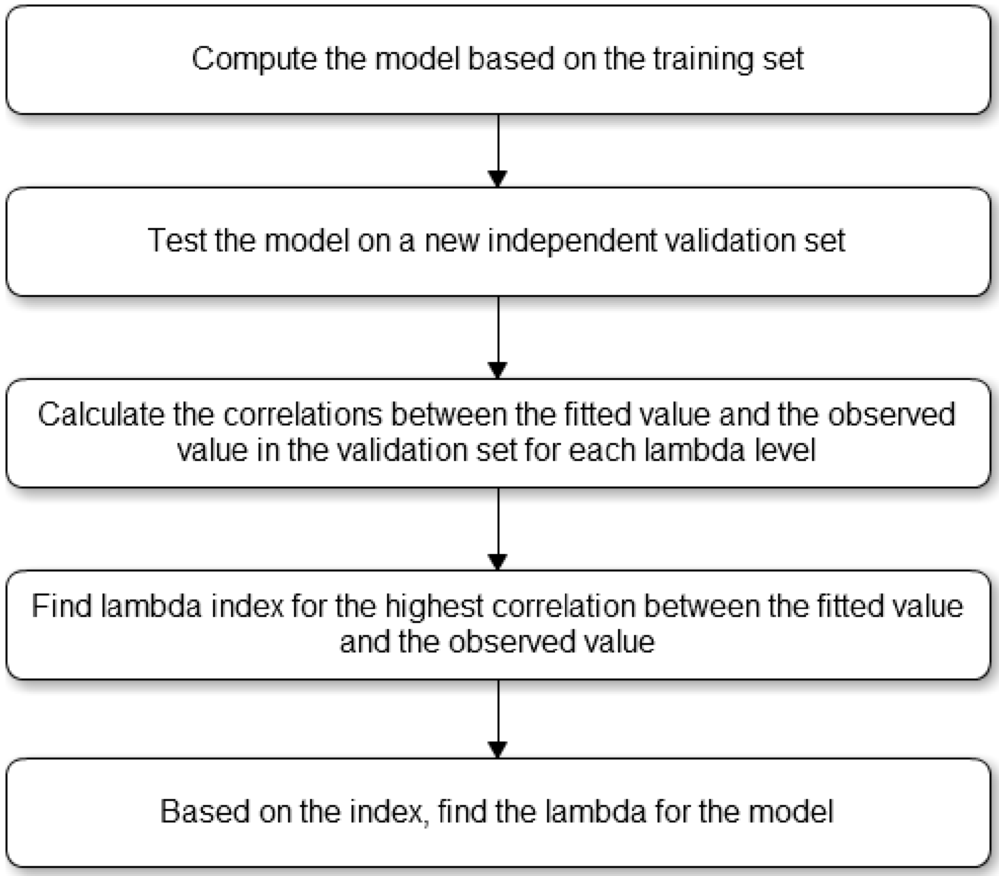


Figure 2. Choosing the penalty parameter λ.

LASSO is here chosen for four reasons. First, it can set coefficients to exactly zero, which is equivalent to excluding those variables from the model thus performing variable selection. Second, at the cost of slightly increasing the bias, LASSO can reduce the variance of the predicted values. The third reason is that correlations in the predictors are not problematic for LASSO prediction (Hebiri and Lederer, 2012). In addition, LASSO regression could analyze the effects of individual predictors on the operative time. This information is valuable for operating room managers in assigning scrub nurses to each procedure one day before the surgery.

3.2.4 Correlation between the surgeon-scrub nurse collaboration frequency and the operative time

Because the operative time is not normally distributed and couldn't meet Pearson correlation assumption, Spearman's rank correlation is used to measure the correlation between the predicted value and observed duration.

The base model is developed once and reduced the number of coefficients by LASSO, the residuals are recorded, and the mean residual for each surgeon-scrub nurse pair is computed.

Spearman's rank correlation is calculated to analyze the relationship between the surgeon-scrub nurse collaboration frequency and the mean residuals (across the surgeries where the specific surgeon-nurse pair participated). A positive correlation indicates that the more frequent surgeon and scrub nurse collaborate, the less operative time a case takes.

In order to test the significance of the correlation, permutation test is applied. The permutation test computes the null distribution for the test statistic, under the null hypothesis that the collaboration frequency is not correlated with the residual of the base model. To estimate the null distribution, 1000 samples are generated under the null hypothesis, by randomly shuffling surgeon-scrub nurse collaboration frequency. The correlations between the shuffled collaboration frequency and the mean residuals for each pair are calculated (Figure 3). Based on the null distribution of the correlations, a p-value for the correlation between the surgeon-scrub nurse collaboration frequency and the mean residuals is calculated.

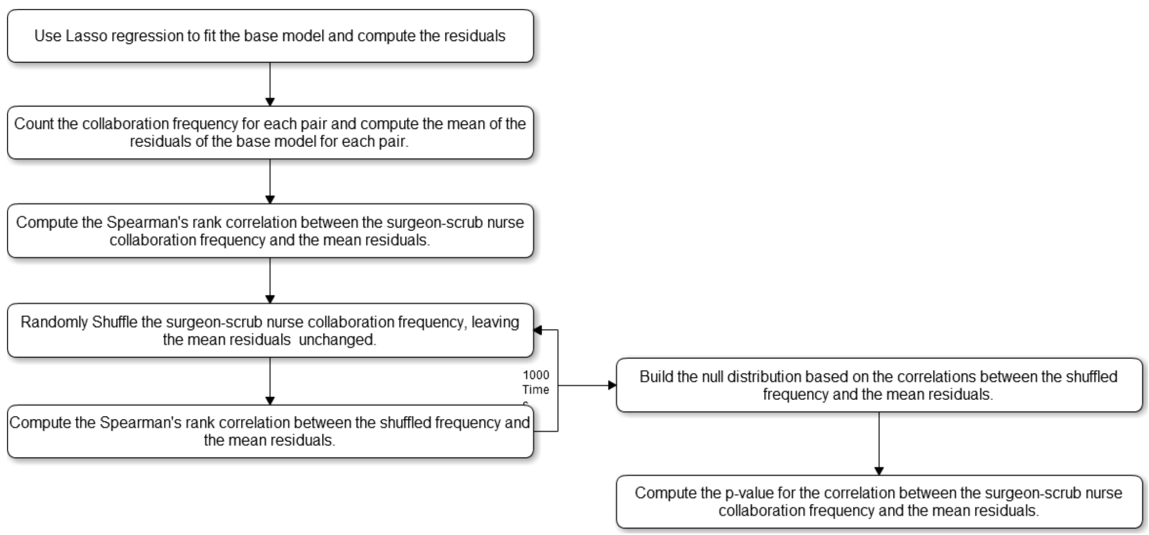


Figure 3. Permutation test.

3.2.3 Model Comparison

The four models are compared by their predictions correlations with observed operative time and the mean absolute errors. Bootstrapping resampling in which 1000 bootstrap samples are drawn with replacement within surgeon-scrub nurse pair ID is adopted in estimating the correlations and mean absolute errors (Figure 4). Wilcoxon signed rank paired test is used to compare the correlations and mean absolute errors among the four models.

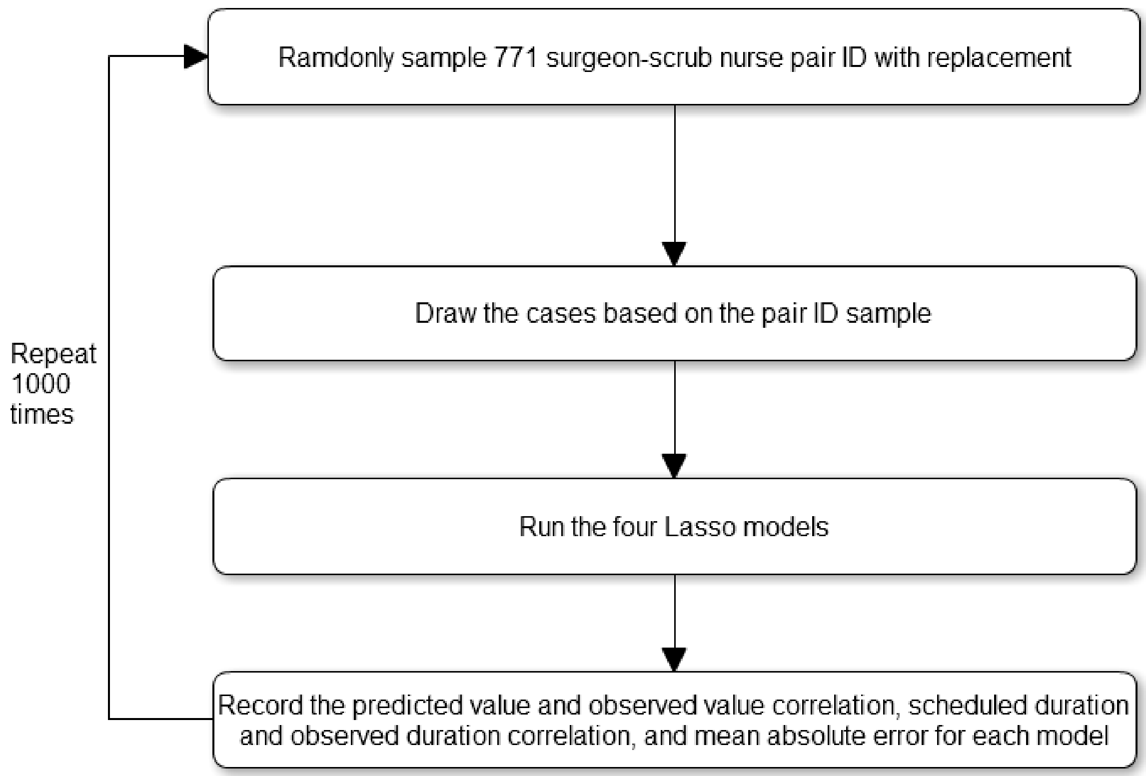
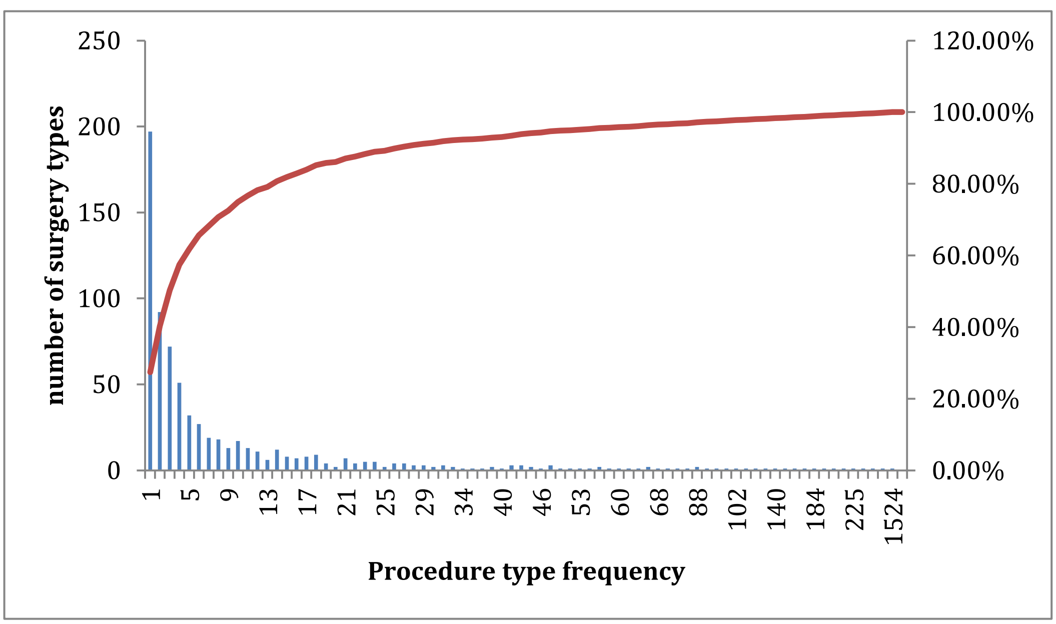


Figure 4. Bootstrapping steps

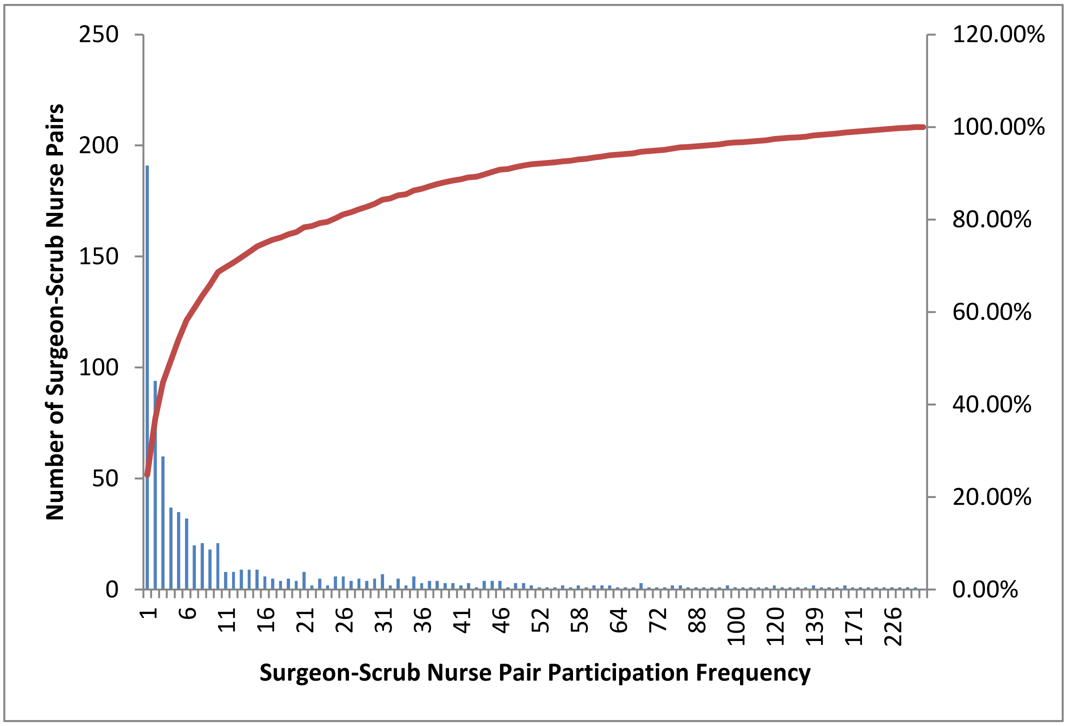
CHAPTER 4. RESULTS

4.1 Procedure Duration Characteristics

Out of 11259 surgeries, surgeon-scrub nurse collaborations are available for 10960 cases. The excluded 296 cases are bronchoscopy, MRI, ERCP and ERCP with stent. These types of surgeries often only require circulating nurses and surgical technicians. There are 74 unique primary surgeons, 31 scrub nurses, 771 pairs of surgeon-scrub nurse collaboration, and 718 procedure types. Figure 5 shows that 80% surgery types had been performed under 13 times and that 80% surgeon-scrub nurse pairs had collaborated within 25 times. Although the sample size is limited, it is sufficient to analyze the surgeon-scrub nurse pairs' effects on the operative time for this specific health system.



a. Histogram of Procedure type frequency



b. Histogram of surgeon-scrub nurse pair participation frequency.

Figure 5. Variables distributions.

As Figure 6 shows, both actual duration and scheduled duration are right skewed and they are fit well in the log transformations. Thus this study uses the lognormal models in predicting the operative times.

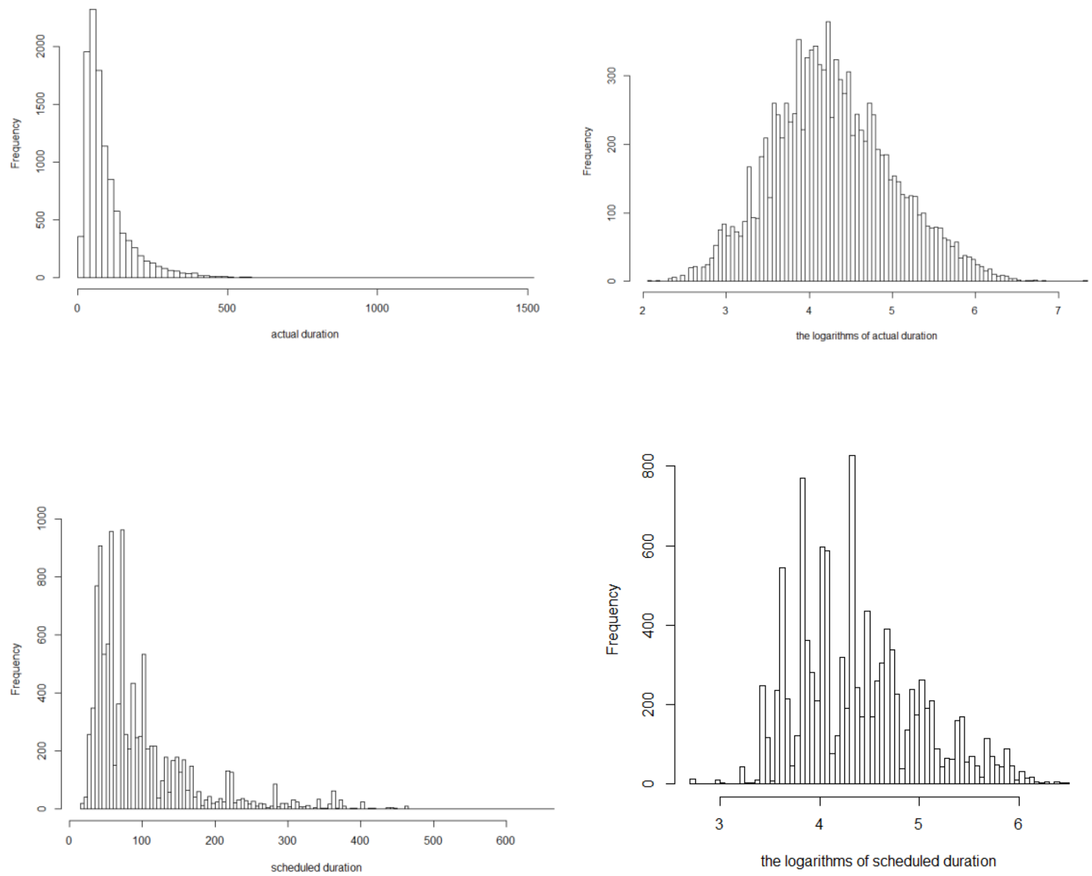


Figure 6. Histograms of surgery durations and log (surgery durations).

As Table 2 shows, the scheduled durations are prone to overestimate surgery durations that can lead to under-utilized OR time. The standard deviation for estimated error is large and increases as the scheduled case duration increases. It reflects the embedded uncertainty in prediction of surgery durations. The mean absolute error of scheduled durations is 30.79 minutes, which is the duration of some procedures such as phacoemulsification.

As Figure 7 shows, the differences between scheduled durations and actual durations varied within each procedure type. The plot indicates that there are other variables that can further explain the variation of operative times. Figure 8 shows that the differences between scheduled durations and actual durations among procedures that performed at least 50 times varied for individual primary surgeon. Each surgeon can perform different procedures for several times. The same color dots in each surgeon's category indicate that for the same procedure and the same surgeon, the differences between scheduled durations and actual durations varies. Thus, the scheduled duration, procedure type, and physician ID cannot explain the variance of the operative times well. There must be other variables that result in the variance of the operative times and we hypothesize that the surgeon-scrub nurse collaboration can improve the accuracy of operative time prediction.

Table 2. Descriptive statistics of actual and scheduled durations

	Mean (min)	**SD (min)**	**Median (min)**	**Mean Absolute Error (min)**
Actual Duration	95.33	81.90	69	30.79
Scheduled Duration	119.43	74.15	75	
Over-predicted Case (n=6952)	-32.4	27.98	-18	
Under-predicted Cases (n=4008)	37.18	48.52	22	

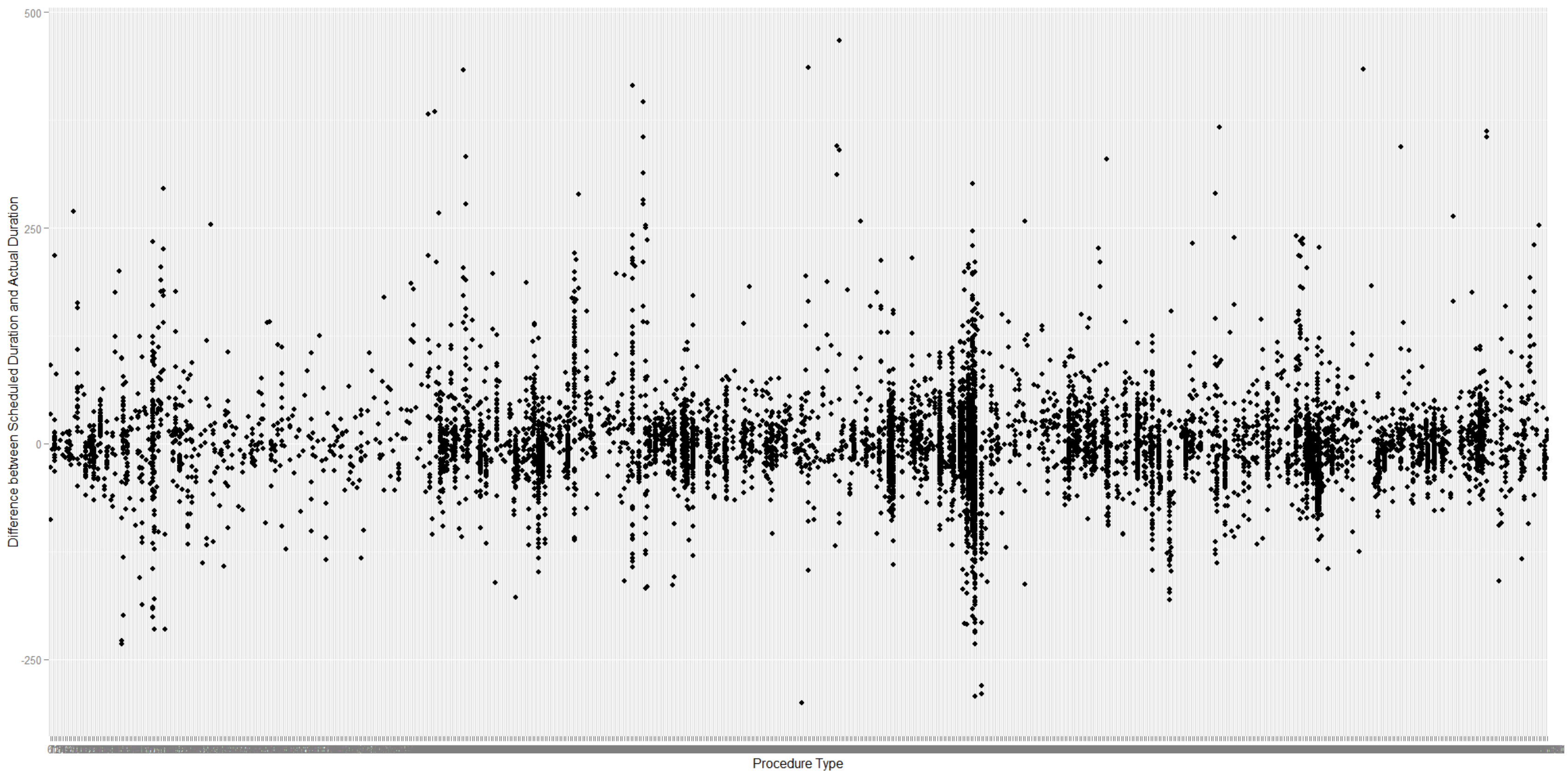


Figure 7. Differences between scheduled durations and actual durations varied within each procedure type.

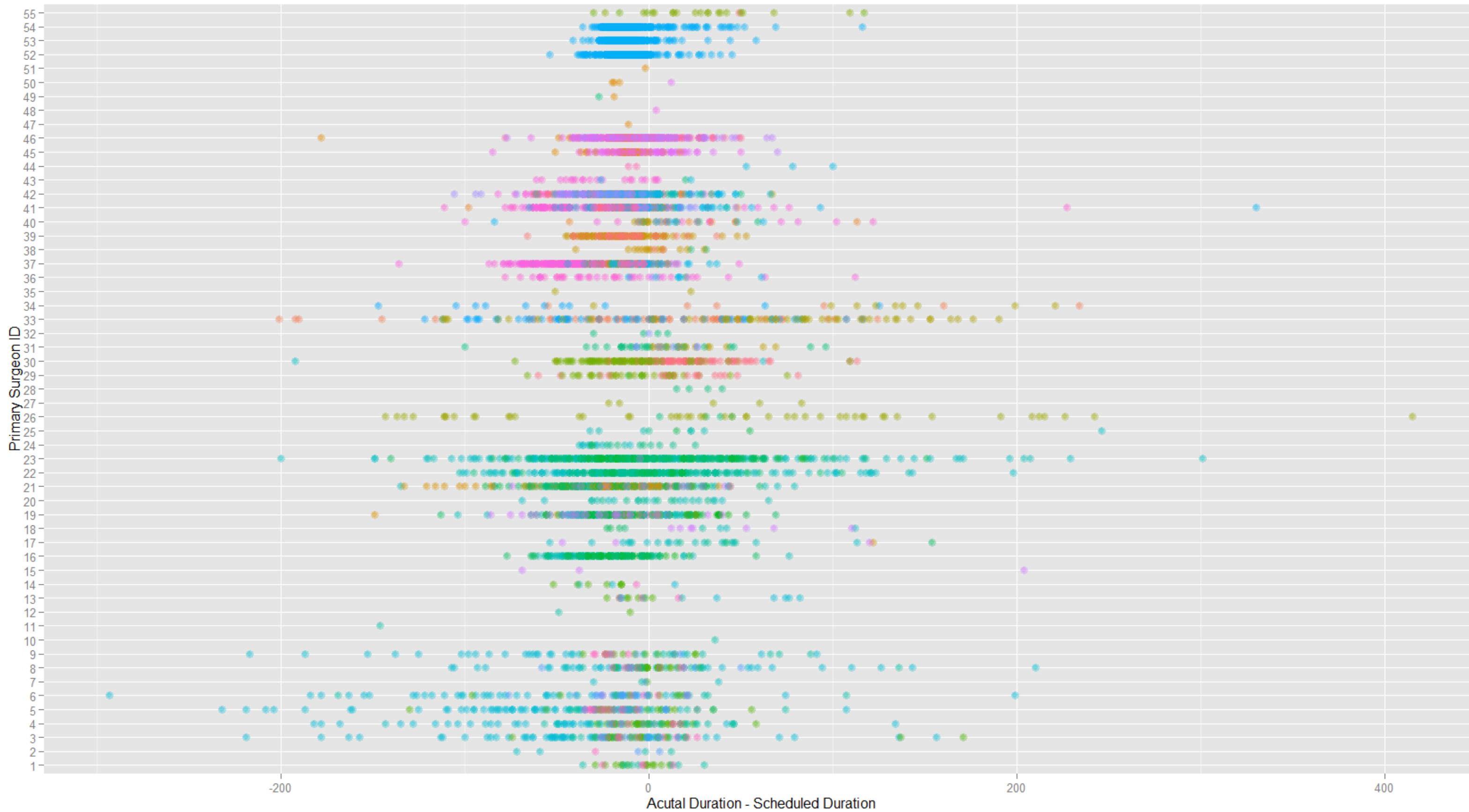


Figure 8. Differences between scheduled durations and actual durations among procedures that performed at least 50 times varied for individual primary surgeons.

4.2 The Results of Models

Figure 9 shows the regularization paths for the coefficients of the models fit the cleaned data. Each curve represents a coefficient in the model. The x-axis is a function of lambda, the regularization penalty parameter. The y-axis gives the value of the coefficient. The graph shows how the coefficients “enter the model” (become non-zero) as lambda changes. The lambda in this study is chosen based on the correlation between fitted values and the actual values in the validation data. There are 719 predictors in the base model, 1490 predictors in Model 1, 824 predictors in Model 2, and 1595 predictors in Model 3. The logarithm of chosen lambda for the base model is -6.29 and 568 predictors retained in the final base model. The logarithm of chosen lambda for Model 1 is -6.38 and 983 predictors retained in Model 1. The logarithm of chosen lambda for Model 2 is -6.84 and 657 predictors retained in Model 2. The logarithm of chosen lambda for Model 3 is -6.02 and 796 predictors retained in Model 3.

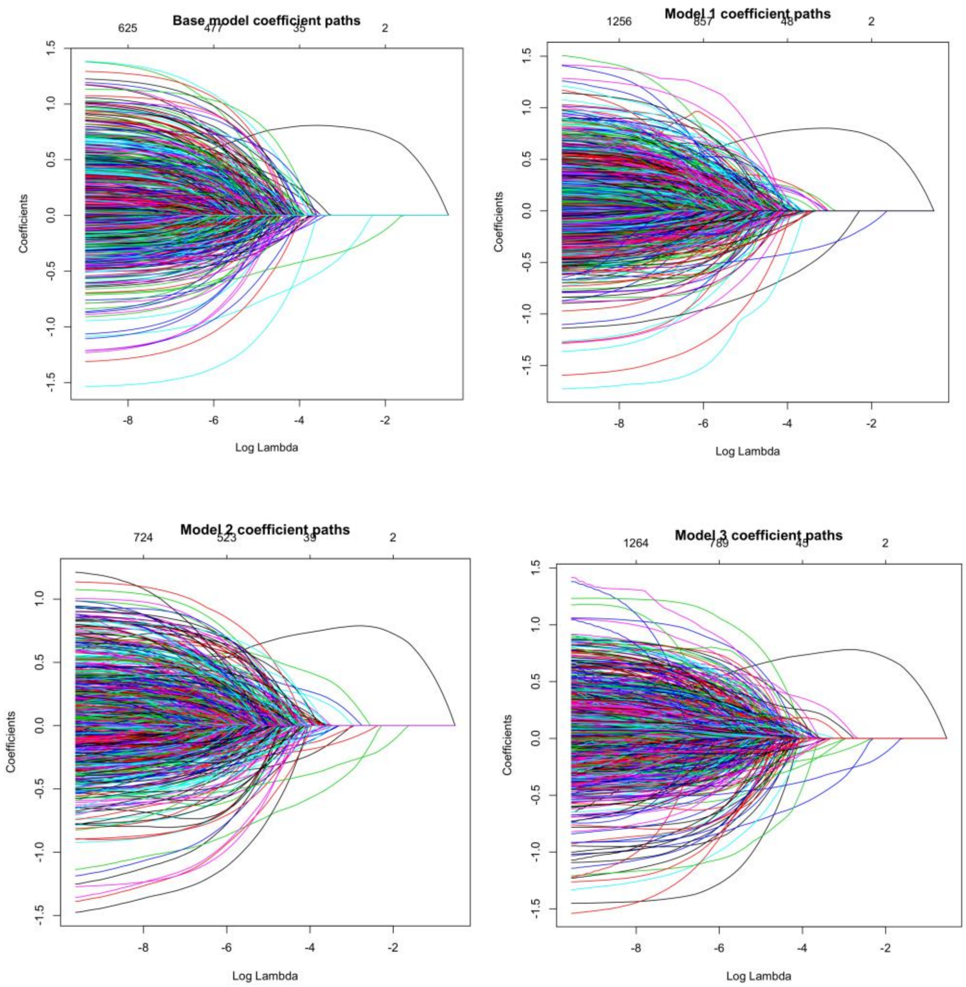


Figure 9. The regularization paths for the coefficients of the models.

After bootstrapping resampling, the mean coefficients for the retained predictors are calculated (Appendix). There are 181 surgeon-scrub nurse pairs in Model 1 that can increase or decrease the operative times by more than 20%. 89.5% of these pairs prolong the operative duration. There are 27 surgeon IDs (24) and scrub nurse IDs (3) in Models that can increase or decrease the operative times by more than 20%. There are 98 variables excluding the surgery types and the logarithm of scheduled duration in Model 3 that can increase or decrease the operative times by more than 20%. 89.8% of these variables are surgeon-scrub nurse pairs, most of which prolong the surgery. With more predictors entering the model, Lasso shrinks the coefficients of the retained variables. Thus, the coefficients in Model 3 are prone to be less than those in Model 1 and 2. The highest coefficient in Model 3 is 1.92, meaning that when this pair participates the surgery the operative time will be prolong by 1.92 times. The lowest coefficient is 0.38, meaning that when this pair participates the surgery its operative time will be decreased by 0.62 times.

4.3 Correlation between the surgeon-scrub nurse collaboration frequency and the operative time

The correlation between the surgeon-scrub nurse collaboration frequency and the mean residuals of the base model is 0.109. Figure 10 shows the frequency distribution of correlations under null hypothesis that the collaboration frequency is not correlated with the residuals of the base model. The vertical line in Figure 10 represents the location of 0.109. The p-value for the correlation between collaboration frequency and the residuals is $0.014 < 0.05$. Thus, we reject the null hypothesis and conclude that the unexplained variation in the residuals of the model, which only includes log scheduled duration and procedure type, can be further explained by the surgeon-scrub nurse collaboration frequency. Thus, we conclude that surgeon-scrub nurse collaboration frequency is significantly correlated with operative time. Further, the vertical line is located in the right tail of the distribution. We can conclude that the more surgeon and scrub nurse collaborate, the less time the surgery will take.

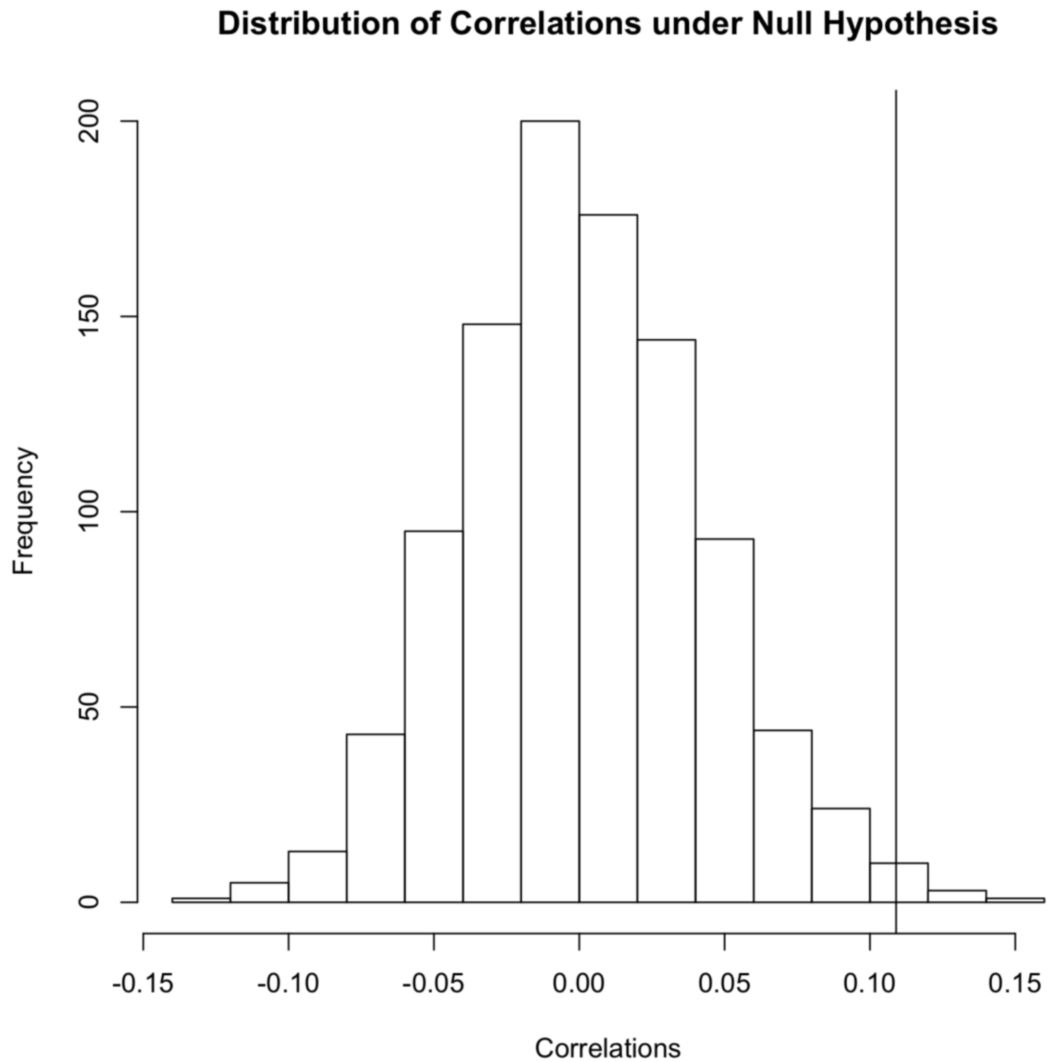


Figure 10. Frequency distribution of the correlations under null hypothesis.

4.4 Model Comparison

The four models are fitted using 1000 bootstrapped samples. The Spearman correlations between the predicted durations and observed durations are estimated for each sample. Figure 11 shows the distributions of correlations for the four models. Compared with the base model, the correlations between predicted durations of the three models and the observed durations were higher and the ranges of the correlations are relatively small. The distributions of correlations between Model 1 and Model 2 are almost the same. Compared with Model 1 and Model 2, Model 3 has a higher mean correlation.

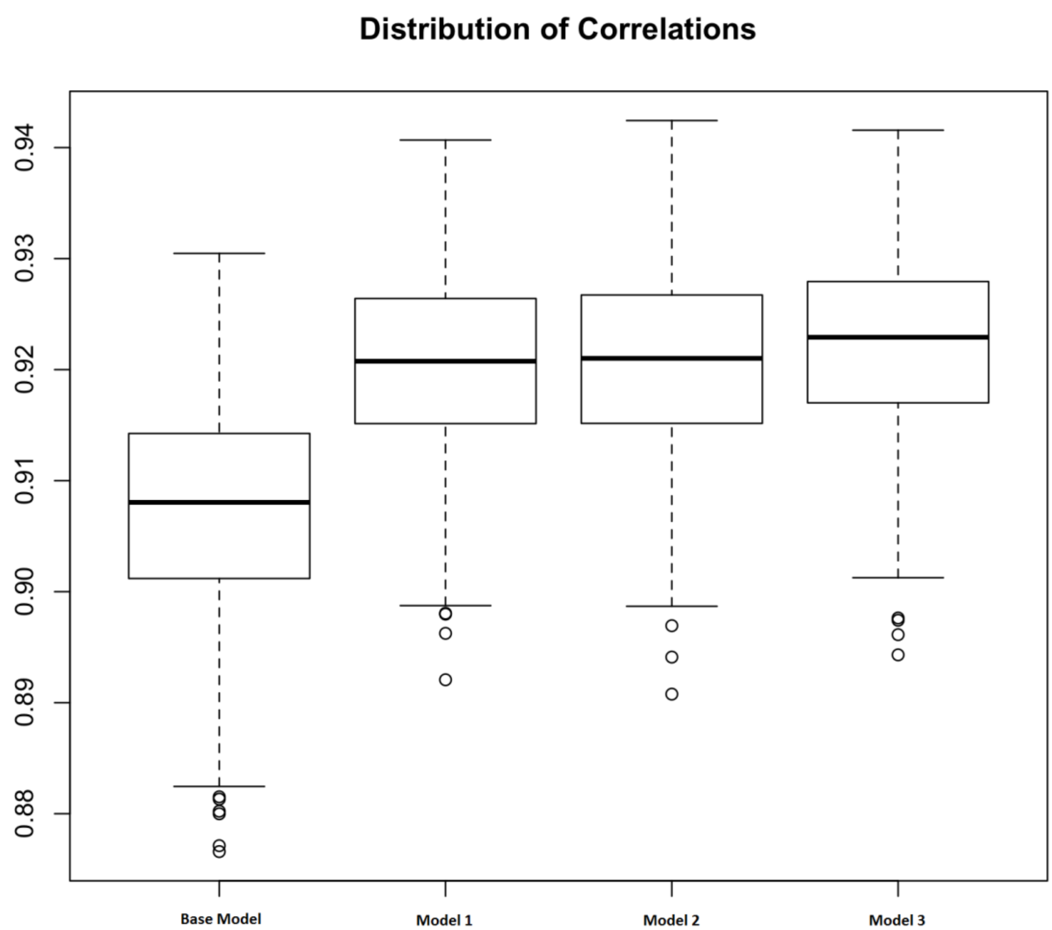


Figure 11. The distribution of correlations between fitted values and observed values.

The Wilcoxon signed rank paired tests confirm the observed results: the correlations between the predicted durations of the three models and the observed durations are significantly higher than the base model (p-value < 0.001). The correlation of Model 3 is significantly higher than Model 1 or Model 2 (p-value < 0.001). At alpha = 0.05, the difference between correlations in Model 1 and Model 2 is not evident (p-value = 0.24).

The mean absolute errors for Model 1, 2, 3 are 22.55, 22.94, and 22.32 minutes. The difference is not prominent. However, the Wilcoxon signed rank paired tests show that the mean absolute error of Model 3 is significantly less than Model 1 (p-value < 0.001) and that mean absolute error of Model 1 is significantly less than Model 2 (p-value < 0.001). Compared with the scheduled duration, Model 3 can decrease the mean absolute errors by 8.47 minutes.

CHAPTER 5. CONCLUSION AND DISCUSSION

5.1 Research Questions

5.1.1 Is surgeon-scrub nurse collaboration frequency correlated with the operative times?

The permutation test is applied to answer this question. To address surgeon-scrub nurse collaboration frequency is a significant factor of operative time prediction, the base model using the log scheduled time, which didn't consider the physician-scrub nurse collaboration, and procedure types, which serve as dummy variables, is fit. To prevent the model from over-fitting, LASSO is used to identify the factors that contribute to the prediction. The mean residuals of the base model are used to compare the unexplained variation with the surgeon-scrub nurse collaboration frequency. The permutation test indicates that the surgeon-scrub nurse collaboration frequency is correlated with the unexplained variation of the base model and that the more frequent surgeon and scrub nurse collaborate, the less time the surgery will take.

5.1.2 Can staff information contribute to the accuracy of operative time predictions?

Four models are developed to answer this question. Based on the answers of the first question, surgeon-scrub nurse pairs are entered Model 1. In order to make a comparison between surgeon-scrub nurse pair effect and individual professionals' effects, Model 2 including individual surgeon ID and individual scrub nurse ID is fit. Concerning that there's a chance of a new surgeon-scrub nurse pair comes out for any new cases, Model 3 is developed to include both pairs and individual professionals.

Compared with the base model, which doesn't include any staff information, the three models can provide more accurate prediction of the operative times. The results show that Model 3, which includes comprehensive staff information, is more accurate than the other two models. In addition, the Wilcoxon signed rank paired test of mean absolute errors between Model 1 and Model 2 shows that treating surgeon and scrub nurse as a pair lead to less prediction errors than treating them separately.

5.2 Discussion

This study generated findings that are consistent with the existing literature - reinforcing the effect of team familiarity and decreasing procedure duration by including by choosing teams that frequently working together. Besides, we proved that using information gained one day before the surgery can improve the accuracy of operative time prediction. This result is corresponding with the finding of the systematic review on predicting operative times for pulmonary surgeries conducted by Dexter et al. (2008).

Different from the study conducted by Xu et al. (2013), whose study explored the effect of two team members in the same profession, this study explores the two different professionals' collaboration. The reason for us to analyze the surgeon-scrub nurse collaboration, which, to our knowledge, seldom has been studied by other researches in operative time prediction, is that their collaboration was regarded the most important by surgical team members (Undre, 2006). Besides, the observational study conducted by Gillespie et al. (2012) indicated that most communication failures were made between surgeons and scrub nurses. For example, scrub nurses sometimes failed to catch the major request by surgeons. This failure may be due to the limited times they have collaborated together. Another example provided by Gillespie et al. (2012) is that surgeons' requests were sometimes not straightforward so that the scrub nurses misunderstood the requests. If they collaborated several times, they may understand each other efficiently.

Another strength of our study is that we use Lasso regression to analyze the individual effects of surgeon-scrub nurse pairs, surgeons and scrub nurses on operative times. Lasso can select the important predictors in predicting operative times. So both our models' predictions and coefficients can be utilized by operating room managers in assigning staff to form an effective surgical team and in moving cases around the ORs with more accuracy. For example, this study finds that some pairs may prolong the surgery twice and some pairs may shorten the procedure duration by 0.38. Using this information, operating managers can not only have a more accurate prediction of procedure duration but also assign staff to decrease the input.

In addition, few studies used information gained one day before surgeries in predicting operative time. We used staff information that can be gained one day before the surgery in improving the prediction accuracy. Our results prove that using information gained one day before surgeries can improve the prediction accuracy of operative times.

There are three limitations in the study. First, given that the data used in this study was secondary data, an inherent limitation is that it is very difficult to evaluate the accuracy of secondary data and to interpret some findings in the study. The data we used was time stamped. We don't know when the exact time people collected the records. Secondly, because we only analyzed less than two year's records in one health system in Middlewest, the results of our study may only apply to the similar health systems. In addition, the prediction of our models is based on the individual effect of each surgeon-scrub nurse pair or each professional's id. When a new surgeon or a new scrub nurse joins in the system, our models may lose its unique power in predicting the operative times.

Based on the findings in this study, there are three directions of future work that are of interests. First, the qualitative studies on surgeon-scrub nurse collaboration during a procedure can be conducted to further explore the factors that influence their collaboration. These factors can be used in team training and staff assignment. Secondly, an electronic decision support white board based on our Model 3 can be developed to assist operating room managers in assigning staff to form an effective surgical team and in moving cases around the ORs with more accuracy. By updating information on new surgeons or staffs, the prediction accuracy of Model 3 can further be improved. Thirdly, since this study proves that information gained one day before surgery is useful in predicting operative times. More kinds of information that can be gained one day before the surgery should be recorded and applied in improving the accuracy of prediction models for operative times.

Reference

1. Bühlmann, P. L., van de Geer, S. A., & Van de Geer, S. (2011). Statistics for high-dimensional data. Springer.

2. Cassera, M. A., Zheng, B., Martinec, D. V., Dunst, C. M., & Swanström, L. L. (2009). Surgical time independently affected by surgical team size. *The American Journal of Surgery*, *198*(2), 216-222.

3. Eijkemans, M. J., van Houdenhoven, M., Nguyen, T., Boersma, E., Steyerberg, E. W., & Kazemier, G. (2010). Predicting the unpredictable: A new prediction model for operating room times using individual characteristics and the surgeon's estimate. *Anesthesiology*, *112*(1), 41-49.

4. ElBardissi, A. W., Duclos, A., Rawn, J. D., Orgill, D. P., & Carty, M. J. (2012). Cumulative team experience matters more than individual surgeon experience in cardiac surgery. *The Journal of thoracic and cardiovascular surgery*.

5. Dexter, F., Willemsen-Dunlap, A., & Lee, J. D. (2007). Operating room managerial decision-making on the day of surgery with and without computer recommendations and status displays. *Anesthesia & Analgesia*, *105*(2), 419-429.

6. Dexter, F., Dexter, E. U., Masursky, D., & Nussmeier, N. A. (2008). Systematic review of general thoracic surgery articles to identify predictors of operating room case durations. *Anesthesia & Analgesia*, *106*(4), 1232-1241.

7. Dexter, F., Epstein, R. H., Bayman, E. O., & Ledolter, J. (2013). Estimating surgical case durations and making comparisons among facilities: identifying facilities with lower anesthesia professional fees. *Anesthesia & Analgesia*, *116*(5), 1103-1115.

8. Gillespie, B. M., Chaboyer, W., & Fairweather, N. (2012). Factors that influence the expected length of operation: results of a prospective study. *BMJ Quality & Safety*, *21*(1), 3-12.

9. Healey, A. N., Undre, S., & Vincent, C. A. (2006). Defining the technical skills of teamwork in surgery. *Quality and Safety in Health Care*, *15*(4), 231-234.

10. Healthcare Financial Management Association. (2005). Achieving operating room efficiency through process integration. Healthc Financ Manage, 57(3), 1-7.

11. Hebiri, M., & Lederer, J. (2012). How Correlations Influence Lasso Prediction.

12. Leach, L. S., Myrtle, R. C., Weaver, F. A., & Dasu, S. (2009). Assessing the performance of surgical teams. *Health care management review*, *34*(1), 29-41.

13. Leach, L. S., Myrtle, R. C., & Weaver, F. A. (2011). Surgical teams: role perspectives and role dynamics in the operating room. *Health Services Management Research*, *24*(2), 81-90.

14. Macario, A. (2010). What does one minute of operating room time cost?. *Journal of clinical anesthesia*, *22*(4), 233-236.

15. May, J. H., Strum, D. P., & Vargas, L. G. (2000). Fitting the Lognormal Distribution to Surgical Procedure Times*. *Decision Sciences*, *31*(1), 129-148.

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

17. Spangler, W. E., Strum, D. P., Vargas, L. G., & May, J. H. (2004). Estimating procedure times for surgeries by determining location parameters for the lognormal model. *Health care management science*, *7*(2), 97-104.

18. Stepaniak, P. S., Heij, C., & De Vries, G. (2010). Modeling and prediction of surgical procedure times. *Statistica Neerlandica*, *64*(1), 1-18.

19. Strum, D. P., Sampson, A. R., May, J. H., & Vargas, L. G. (2000A). Surgeon and type of anesthesia predict variability in surgical procedure times. *Anesthesiology*, *92*(5), 1454-1466.

20. Strum, D. P., May, J. H., & Vargas, L. G. (2000B). Modeling the uncertainty of surgical procedure times: comparison of log-normal and normal models. *Anesthesiology*, *92*(4), 1160-1167.

21. 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), 232-240.

22. Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological), 267-288.

23. Undre, S., Sevdalis, N., Healey, A. N., Darzi, S. A., & Vincent, C. A. (2006). Teamwork in the operating theatre: cohesion or confusion?. *Journal of evaluation in clinical practice*, *12*(2), 182-189.

24. Viapiano, J., & Ward, D. S. (2000). Operating room utilization: the need for data. *International anesthesiology clinics*, *38*(4), 127-140.

25. 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), 1235-1245.

26. Xu, R., Carty, M. J., Orgill, D. P., Lipsitz, S. R., & Duclos, A. (2013). The Teaming Curve: A Longitudinal Study of the Influence of Surgical Team Familiarity on Operative Time. *Annals of surgery*.

27. Zheng, B., Panton, O. N., & Al-Tayeb, T. A. (2012). Operative length independently

affected by surgical team size: data from 2 Canadian hospitals. *Canadian Journal of Surgery*, *55*(6), 371.

Appendix

Pair ID: 1~771

Surgeon ID: 772~845

Scrub Nurse ID: 846~876

The Coefficients (>1.2 or <0.8) of surgeon-scrub nurse pairs in Model 1.

pairID	coefficient	exp(coef)
1	0.50405444	1.66
3	0.18253906	1.2
15	0.35353254	1.42
20	0.2512437	1.29
21	0.28312648	1.33
22	0.23215535	1.26
24	0.21942739	1.25
26	0.37751656	1.46
29	0.23266329	1.26
54	0.28371747	1.33
58	0.514018	1.67
64	0.27933853	1.32
75	0.57777161	1.78
76	0.33261774	1.39
79	0.19247857	1.21
82	0.2131372	1.24
94	0.22572207	1.25
99	0.36834553	1.45
103	0.25891952	1.3
112	0.20718976	1.23
122	0.22291559	1.25
126	0.22277906	1.25
148	0.18052024	1.2
152	0.26348303	1.3
161	0.23407768	1.26
163	0.4436012	1.56
170	0.25946399	1.3
172	0.23384045	1.26
175	0.23481542	1.26
176	0.22594967	1.25
191	0.37583501	1.46
195	0.19401365	1.21
198	0.33631787	1.4
199	0.47142502	1.6
200	0.27086935	1.31
203	0.31081121	1.36
206	0.18470706	1.2
213	0.69955458	2.01
215	0.75076359	2.12
217	-0.2503124	0.78
240	0.32900576	1.39
242	0.27231407	1.31
246	0.38472278	1.47
248	0.30703494	1.36
249	0.20320355	1.23
250	0.22706617	1.25
251	0.42267565	1.53
254	0.18204436	1.2
255	0.3893992	1.48
256	0.28166966	1.33
257	0.19152555	1.21
260	0.30271796	1.35
261	0.53104875	1.7
263	0.24260983	1.27
264	0.27665802	1.32
265	0.24480982	1.28
266	0.28292059	1.33
268	0.18971955	1.21
270	0.20789565	1.23
271	0.250339	1.28
272	0.23358308	1.26
273	0.22201348	1.25
277	0.21557783	1.24
281	0.32783308	1.39
288	0.18493513	1.2
298	0.22754123	1.26
308	-0.248113	0.78
310	0.24804585	1.28
312	0.28564117	1.33
326	0.24672978	1.28
327	0.20665853	1.23
334	-0.3131889	0.73
335	-0.2543487	0.78
341	0.22867001	1.26
344	0.21195648	1.24
345	0.42023841	1.52
346	0.20331521	1.23
348	0.33350803	1.4
353	0.23992633	1.27

355	0.2745247	1.32
358	0.48333515	1.62
361	-0.3435394	0.71
362	0.1946025	1.21
365	0.2304151	1.26
366	0.19011445	1.21
367	0.32753515	1.39
368	0.19969009	1.22
370	0.20089708	1.22
371	0.18071567	1.2
389	0.26394401	1.3
391	0.3005483	1.35
393	0.50244916	1.65
394	0.20959223	1.23
395	0.2943485	1.34
396	0.20042603	1.22
397	-0.3353318	0.72
398	0.21836599	1.24
402	0.27011122	1.31
403	0.28487943	1.33
404	0.31260749	1.37
405	0.20432819	1.23
406	0.27994284	1.32
407	0.45568608	1.58
409	-0.3624179	0.7
411	0.29045591	1.34
412	0.21489499	1.24
415	0.21693409	1.24
416	0.38928178	1.48
417	0.22817744	1.26
419	0.38830492	1.47
420	0.40901263	1.51
421	0.24733239	1.28
422	0.24964637	1.28
423	0.25791906	1.29
424	0.21724228	1.24
438	-0.3138665	0.73
459	0.2042564	1.23
462	0.19505104	1.22
468	0.20566629	1.23

469	0.32259804	1.38
470	0.48865996	1.63
472	0.25299638	1.29
473	0.2446937	1.28
476	0.22254451	1.25
477	0.29552228	1.34
478	0.39149309	1.48
480	0.19705254	1.22
482	0.31703245	1.37
484	0.1911567	1.21
488	0.60898371	1.84
489	0.45592066	1.58
491	0.2819735	1.33
493	0.33739886	1.4
494	0.27787115	1.32
495	-0.3169127	0.73
502	0.18437993	1.2
503	0.91561604	2.5
507	0.27982472	1.32
511	0.2194035	1.25
514	0.33278768	1.39
517	0.33978345	1.4
518	-0.2849758	0.75
520	0.37074293	1.45
521	0.25804192	1.29
523	-0.2719556	0.76
531	0.30511431	1.36
537	0.18803205	1.21
552	-0.2244564	0.8
573	-0.2641321	0.77
576	-0.3681745	0.69
588	0.27539775	1.32
589	0.24762971	1.28
590	0.27548759	1.32
592	-0.9767908	0.38
599	0.17970666	1.2
606	0.330908	1.39
608	0.46604691	1.59
609	0.18428964	1.2
636	0.31203645	1.37

649	0.19232987	1.21
652	-0.2862911	0.75
679	0.26680469	1.31
684	-0.3870413	0.68
685	0.4796792	1.62
686	0.25877287	1.3
688	0.25124218	1.29
689	0.44199494	1.56
690	0.2887329	1.33
693	0.22602058	1.25
694	0.25405999	1.29
707	-0.4641368	0.63
722	0.29451924	1.34
731	0.33222078	1.39
736	0.18280447	1.2
750	0.27807714	1.32
757	0.25204027	1.29
763	0.19010561	1.21
766	0.1911635	1.21
767	0.23641182	1.27
768	0.49172579	1.64
771	0.19512546	1.22

The Coefficients (>1.2 or <0.8) of surgeon or scrub nurse ID in Model 2.

indivID	Coef	exp(coef)
775	0.369027	1.45
776	0.218367	1.24
777	0.290414	1.34
789	-0.27496	0.76
790	0.188922	1.21
793	-0.2331	0.79
801	-0.28278	0.75
803	0.476951	1.61
805	0.365761	1.44
810	0.537113	1.71
811	0.669476	1.95
812	0.57672	1.78
813	0.462851	1.59
816	0.563454	1.76
820	-0.37271	0.69
821	0.194083	1.21
825	0.222734	1.25
830	-0.31841	0.73
831	0.38926	1.48
832	0.181866	1.2
834	-0.34868	0.71
835	-0.49272	0.61
837	-0.23099	0.79
845	0.196666	1.22
850	0.190103	1.21
858	0.254984	1.29
871	0.219397	1.25

The Coefficients (>1.2 or <0.8) of pair ID & surgeon or scrub nurse ID in Model 3.

ID	coef	exp(coef)
1	0.40437	1.5
15	0.374067	1.45
20	0.227055	1.25
26	0.365312	1.44
29	0.192559	1.21
43	-0.23081	0.79
54	0.196114	1.22
58	0.462154	1.59
75	0.508112	1.66
76	0.282549	1.33
94	0.257631	1.29
99	0.377505	1.46
103	0.284369	1.33
163	0.345789	1.41
170	0.269492	1.31
175	0.182873	1.2
198	0.238002	1.27
199	0.367968	1.44
203	0.249328	1.28
213	0.652018	1.92
215	0.544427	1.72

217	-0.29125	0.75
240	0.189653	1.21
241	-0.24314	0.78
242	0.179676	1.2
246	0.274959	1.32
251	0.282298	1.33
255	0.251751	1.29
260	0.205646	1.23
277	0.195569	1.22
281	0.283204	1.33
308	-0.30184	0.74
310	0.211125	1.24
326	0.200849	1.22
335	-0.3076	0.74
345	0.392832	1.48
348	0.235605	1.27
355	0.231496	1.26
358	0.464434	1.59
365	0.179032	1.2
367	0.302864	1.35
389	0.211368	1.24
393	0.394748	1.48
397	-0.50166	0.61
407	0.331583	1.39
409	-0.41132	0.66
411	0.294376	1.34
416	0.187995	1.21
419	0.283557	1.33
420	0.206239	1.23
428	-0.21802	0.8
438	-0.35512	0.7
462	0.183501	1.2
469	0.276649	1.32
470	0.448058	1.57
472	0.185545	1.2
477	0.24044	1.27
478	0.293393	1.34
482	0.236452	1.27
489	0.378458	1.46

491	0.212451	1.24
493	0.257132	1.29
495	-0.46658	0.63
496	-0.29953	0.74
514	0.263913	1.3
517	0.281573	1.33
520	0.254008	1.29
521	0.380713	1.46
529	0.204787	1.23
531	0.201527	1.22
552	-0.23786	0.79
572	0.246854	1.28
576	-0.23734	0.79
592	-0.96576	0.38
606	0.194267	1.21
608	0.374644	1.45
636	0.283542	1.33
641	-0.23121	0.79
652	-0.26504	0.77
684	-0.37648	0.69
685	0.329413	1.39
689	0.254054	1.29
694	0.211036	1.23
707	-0.30633	0.74
722	0.284274	1.33
731	0.308615	1.36
759	-0.27899	0.76
768	0.234075	1.26
795	0.195151	1.22
803	0.454302	1.58
805	0.278007	1.32
810	0.515843	1.68
811	0.405931	1.5
812	0.290913	1.34
820	-0.28358	0.75
831	0.186248	1.2
834	-0.2732	0.76
845	0.269679	1.31