The Pennsylvania State University

The Graduate School

College of Engineering

VALUE-BASED PAYMENT IN TOTAL JOINT ARTHROPLASTY: A HEALTHCARE MARKET SEGMENTATION METHODOLOGY TO IMPROVE VALUE

A Dissertation in

Industrial Engineering and Operations Research

by

Eric R. Swenson

© 2017 Eric R. Swenson

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2017

The dissertation of Eric R. Swenson was reviewed and approved* by the following:

Harriet B. Nembhard
Professor of Industrial Engineering
Dissertation Advisor
Chair of Committee

Ravi Ravindran
Professor of Industrial Engineering

Conrad Tucker
Assistant Professor of Industrial Engineering

Naleef Fareed
Assistant Professor of Health Policy and Administration

Andrea Yevchak Sillner
Assistant Professor of Nursing

Janis P. Terpenny
Professor of Industrial Engineering
Head of the Department of Industrial and Manufacturing Engineering

*Signatures are on file in the Graduate School

ABSTRACT

The U.S. spends more on healthcare per capita each year than any other nation in the world, yet consistently underperforms other major developed nations in terms of healthcare quality, timeliness, cost, and access to care. Recognizing a need to reform, the U.S. passed the 2010 Patient Protection and Affordable Care Act (ACA) authorizing a Center for Medicare and Medicaid Innovation whose mission includes the development of new payment models that transform the nation from volume- to value-based care. In total joint arthroplasty (TJA), the most common and costly service covered by Medicare, the transformation is manifested in the comprehensive care for joint replacement (CJR). The CJR is a bundled payment initiative that holds hospitals financially accountable for a patient's outcome over a 90 episode of care. Because hospital reimbursement is tied to the total cost of care delivered by multiple independent care providers, hospitals are incentivized to take an active role in coordinating care through the episode to ensure a quality health outcome. The CJR necessitates a new approach to how hospitals manage joint patients from the pre-operative through the post-operative phase of care.

The overall objective of this research is to investigate how a healthcare market segmentation methodology can be implemented by any hospital subject to the CJR to improve healthcare value. The methodology applies machine learning, regression, and process improvement methods, and the analysis behind the methodology uses latent data stored in hospital electronic health records (EHRs). This new methodology has five pillars and addresses four separate but related questions: 1) how can data mining and market segmentation be used to identify unique and distinguishable patient segments (clusters), 2) how do hospitals accurately classify prospective TJA patients such that their assigned segment provides useful information for clinicians or other health professionals, 3) how can clinicians identify cost drivers in TJA that are tied to a patient's cluster, their attributes, or other clinical factors, and 4) how do hospitals

identify patient outcome drivers. Using operations research to build a methodology around these four questions will help hospitals adapt to the current transformation in healthcare payment models.

The main body of this dissertation is divided into three sections: 1) clustering and classification of patients, 2) identifying cost and outcome drivers, and 3) assessing interventions that reduce costs and improve healthcare value. In Chapter 3, clustering and classification models applied to EHR data divide a healthcare population into smaller segments for which health interventions can be applied. In Chapter 4, patient segments along with attributes such as gender, procedure code, and comorbidity burden are shown to be predictors of 30 and 90 day readmission and increased supply costs. Additionally, Chapter 4 highlights the impact that a small group of patients with complex bony or ligamentous deformities or infection risk have on a hospital under a value-based payment model. If not managed, this small subset of patients is shown to consume a disproportionate share of the overall implant budget for a hospital. In Chapter 5, two interventions are studied. A rapid recovery protocol focused on individual patients is shown to significantly improve length of stay and discharge to home rate while not impacting readmission rates. Finally, applying a clustering model using attributes related to cost and outcome drivers helps identify a small but high cost segment of the TJA population. A new approach to managing these high cost patients is introduced and modeled using simulation with significant cost savings. This dissertation provides a roadmap for hospitals seeking to improve healthcare value under value-based payment models.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENT

To my wife, Julie, and three daughters Caroline, Claire, and Emma. Thank you for your support, patience, and sacrifices over the past three years. I would also like to thank Dr. Harriet Nembhard, my doctoral adviser, for guiding me in my journey toward this degree. I appreciate your patience, friendship, mentorship, professionalism, and introduction to your colleagues at the Penn State Hershey Medical Center and NSF-CHOT. Dr. Nembhard and the National Science Foundation Center for Health Organization Transformation (NSF-CHOT) were the two forces that inspired me to undertake this important research. Additionally, I would like to thank my doctoral committee for their insight, mentorship, and candid feedback.

In addition to the members of NSF-CHOT and the faculty in my department, I would not have been successful without the support and mentorship of my colleagues at Hershey Medical Center. In particular, I would like to thank Dr. Charles Davis, Dr. Kevin Black, Dr. Andrew Resnick, Maria Hamilton, Lisa Zimmerman, Tiffany Gibbons, and all the other clinical and administrative staff who selflessly gave of their time and expertise to aid in my research.

I would be remised if I did not thank my officemates Yifeng Yu, LinLin Ma, Liz Welker, and Lisa Korman. Finally, to my great friend, colleague, and comrade in arms, Nathan Bastian. Your academic leadership and support over the past three years helped keep me focused on obtaining this objective.

Chapter 1

Introduction

Healthcare spending in the U.S. in 2014 topped $3 trillion which equates to a per capita spending of over $9,500; yet, despite spending more than any other country in the world on healthcare, the U.S. consistently underperforms other industrialized nations. In a recent report that compares healthcare in 11 industrialized nations, the U.S. has occupied that last position in five out of the past 10 years (Davis, Stremikis, Squires, & Schoen, 2014). Increasing costs, poor access to quality healthcare, and untimely care continue to plague the U.S. healthcare system (Cortada & Lenihan, 2012).

The causes and impacts of U.S. healthcare woes are well documented, but solutions were scarce until the signing of the 2010 Affordable Care Act (ACA) which authorized the creation of the Center for Medicare and Medicaid Innovation. The Innovation Center is leading the way in developing new payment models to increase cooperation and care delivery across an episode of care. This approach encourages, through financial incentives, cooperation between care providers. Considering that the CMS alone paid an estimated $18 billion on preventable readmissions in 2004, the government's largest insurer is motivated to improve their return on healthcare investment (Jencks, Williams & Coleman, 2009).

The U.S. healthcare industry is fundamentally transforming the way it pays for healthcare services. The transformation is manifested in the shift from a fee-for-service or volume-based payment model to a pay-for-outcome or a value-based model (Cortada & Lenihan, 2012). The transformation movement effectively started with the ACA, but has been going on for over a decade. Increased demand for healthcare services, shortages of beds, poor perceived quality, and high cost of healthcare per capita all contributed to the transformation. Details on this transformation were first published in a series of books by the National Academy of Medicine (formerly the Institute of Medicine) that include titles such as *Crossing the Quality Chasm*, *To Err is Human*, and *Building a Better Delivery System*.

In the U.S. there are many payment models for healthcare including fee-for-service (FFS), episode-of-care, traditional capitation, and comprehensive care payment. In FFS, also volume-based reimbursement, insurers reimburse for healthcare services based on a pre-determined amount for each service rendered. The more services rendered, the more reimbursement. An episode of care is defined as "all services provided to a patient with a medical problem within a specific period of time across a continuum of care in an integrated system (episode of care, n.d.)". An episode-of-care payment model pays a health service provider a fixed payment for a major condition (such as stroke or heart attack), but allows the provider to determine the frequency and number of services. Capitation models fix the number of episodes covered and the reimbursement amount per episode regardless of severity. Comprehensive care payments are the common ground between capitation and episode-of-care where a provider is paid for an episode, but the reimbursement is adjusted based on severity. With the exception of FFS, the three other common payment methods combine services over an episode into a single payment (Miller, 2009).

The most common payment model is FFS, but the other methods have been used dating back as early as 1983 with Medicare's inpatient prospective payment system (Miller, 2009). Single payments for care rendered by more than one independent care provider, also known as bundled payments, are not common (Miller, 2009). Over the years, government and private insurers alike developed and trialed different versions of a bundled payment, but in most cases, payments were made to individuals or groups of individuals controlled under one administrative unit (Miller, 2009). There are a handful of examples over the past 30 years where an insurer has bundled one payment for more than one independent provider to cover care delivered over a defined episode of care; all the examples involved a surgical procedure (Miller, 2009). It was not until lawmakers and the Obama Administration drafted the ACA, that the concept of bundled payments would reappear. Out of the legislation came the Innovation Center and in 2013, the first of a series of new bundled payment initiatives.

A volume-based reimbursement system that pays by procedure performed — without serious consideration for the short or long term patient health as a result of the procedure — places the post procedure (or discharge) recovery burden on the patient and payer and not the hospital or physician. In general, under a FFS payment model, clinicians are not reimbursed for taking the extra time and resources to follow up with patients as payers (insurers) pay for care that is codified and tied to a diagnosis. Under a value-based system clinicians would be indirectly compensated for non-billable expenses related to patient care if they minimized or eliminated care expenses later in the episode. Also a FFS payment policy dis-incentivizes physicians from ensuring safe and smooth care transitions from provider to provider or hospital to post discharge care location by paying the full reimbursement for readmission costs incurred greater than 24 hours after a discharge (Burton, 2012). In 2011, inadequate care coordination and poor care transitions are blamed for an estimated $25 to $40 billion in unnecessary readmission costs and avoidable complications (Burton, 2012).

To address some of the shortcomings from the patient and insurer's perspective of FFS payments, the Innovation Center developed a series of new payment models that penalize readmissions, reward quality metrics, and incentivize health outcomes. The first initiative, Bundled Payment for Care Improvement (BPCI, 2016), was designed to align care providers and improve healthcare value similar to patient centered medical homes (PCMH) which incentivize patient centered, comprehensive, coordinated primary care that is focused on quality and safety (AHRQ, 2016).

Although initially limited to only a few diagnosis related groups (DRGs), new value based bundled payment programs expanded and, in April 2016, Medicare introduced the Comprehensive Care for Joint Replacement (CJR) which targets total knee and hip arthroplasty covered under DRGs 469/470, major joint replacement or reattachment of lower extremity with and without major complications or comorbidities (CJR, 2016). DRG 469/470 is the most common and costly (as a percentage of annual budget) procedure for Medicare (CMS, 2016). In 2014, over 400,000 Medicare beneficiaries had a total hip or knee replacement at a cost well over $7B (CMS, 2016). Furthermore, the demand for TJA is growing and is projected to increase four fold by 2030 with the number of TJA surgeries performed each year expected to exceed 4 million (Kurtz et al., 2009, Kurtz, Ong, Lau, Mowat & Halpern, 2007).

Unlike the initial bundles, the CJR is a mandatory program and as of April 2016 started to impact hospitals in 67 regions of the U.S. that perform TJA. Although the goal of this program is to encourage coordinated and synchronized care across the care chain to include hospitals, physicians, and post-acute care providers, the burden of care coordination and financial risk fall on the hospitals that perform the surgeries. Under CJR, hospitals receive one fixed reimbursement that must cover the cost of care throughout the episode. The bundled payment is based on an average episode cost per hospital by DRG (469 or 470). With the exception of hip fracture patients requiring an arthroplasty, individual patient risk factors are not factored into the payment. All independent providers and suppliers bill Medicare and are paid for services rendered based on the current FFS schedule. At the end of the fiscal year, hospitals, without sharing risk with other care providers, either receive an additional payment or they get a bill (CJR, 2016). Hospitals included in the CJR must consider care related expenses from admission to discharge plus 90 days, the episode duration where they have full financial responsibility but limited patient control. The CJR incentivizes hospitals to improve health outcomes and minimize costs associated with readmission, discharge, supplies, and inpatient stay.

1.1 Statement of the Problem

The status of healthcare delivery and the need for healthcare reform in the U.S. cannot be understated. The signing into federal law of the Patient Protection and Affordable Care Act, or commonly called Affordable Care Act (ACA), and subsequent shift in how Medicare reimburses healthcare providers is a critical step in reforming healthcare. Although needed, the shift to value-based care in joint arthroplasty creates a new set of challenges. For hospitals and surgeons in the 67 regions of the U.S. subject to the comprehensive care for joint replacement (CJR), the mandatory shift to a new care bundle creates a dilemma that carries significant and immediate financial risk. Figure 1-1 illustrates the hospital's dilemma which is a gap created when hospitals lose direct control over patient health outcomes after discharge but retain financial responsibility.

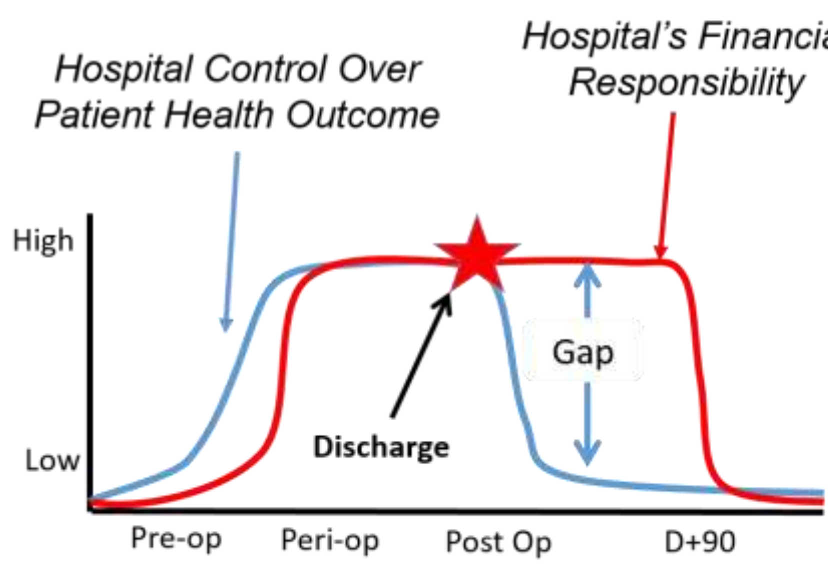


Figure 1-1. The Hospital Dilemma Created Under Comprehensive Care for Joint Replacement

Although the concepts of bundled and comprehensive care payments are not new, the manner in which the CJR holds hospitals accountable for care and associated care management costs are new. Under CJR, hospitals carry the burden to coordinate care and the financial risk, yet they only have direct control over the patient for a small window of time during the episode.

Two of the main cost drivers in any surgical procedure are post-acute care and readmission. Under CJR, hospitals cannot direct patient actions regarding potentially costly post-acute care options. In a recent study, the cost of post-acute care following joint arthroplasty ranged from $733 for a discharge home to over $16,000 for an inpatient rehabilitation stay (Ramos et al., 2014). Each of the many discharge options has its pros, cons, and costs. Hospital discharge teams can try to persuade or influence patients toward a certain location be it home or a hospital preferred post-acute care facility, but that decision remains with the patient (CJR, 2016). The 90-day episode of care also places hospitals at risk for costly readmissions. Although 90-day readmission rates for TJA average less than 10%, the cost of a readmission is significant (Ramkumar et al., 2015; Ramos et al., 2014; 2014; Peel et al., 2015). The lack of a national electronic health record (EHR) database makes care management across multiple settings challenging and costly. Hospitals subject to the CJR are incentivized to increase value by decreasing costs, increasing quality, and improving coordination of care throughout the episode.

1.2 Motivation and Challenges

Operations research methods have never been more needed in healthcare delivery than they are today. Broadly speaking, healthcare worker shortages, increasing demand for health services, changing payment models, the rise of chronic illness and comorbidities, and advances in personalized medicine are but a few of the major challenges the U.S. healthcare industry faces (Cortada, 2012). More narrowly, the rapid growth in demand for TJA, the changing landscape of payment models, and the added financial risk and coordination of care burden for hospitals necessitates a new approach to how hospitals perform total joint arthroplasty. The approach must be evidence based, data driven, and clearly focused on improving healthcare value across the spectrum of healthcare stakeholders including patients, providers, hospitals, and insurers. Porter and Teisberg (2006) define healthcare value as "the health outcomes achieved per dollar spent." Increased demand and new financial risk will both incentivize and challenge hospitals to change.

The biggest challenge for hospitals will be retaining the ability to make informed decisions in a complex environment of persistent change. The CJR is the newest mandatory program for hospitals in 67 regions of the U.S. and it encompasses the most common procedure for Medicare aged patients, but it is only covers one diagnosis. The biggest challenge in developing a new operations research-based framework is making it robust enough to adapt to new changes. Another challenge is to make it translatable to other diagnosis related groups.

1.3 Research Objective and Expected Contributions

The overall objective of this research is to investigate how a healthcare market segmentation methodology can be implemented by any hospital subject to the CJR to improve healthcare value. The methodology will apply machine learning, regression, and process improvement methods, and the analysis behind the methodology will use the latent data already captured and stored in hospital electronic databases. Modern hospitals record a vast amount of data regarding costs, patient visits, reimbursements, supply costs and other related data. Applying a data-driven operations research based methodology specifically to total joint arthroplasty presents a new approach to manage the challenges faced under the CJR.

This new healthcare market segmentation methodology has five pillars and will address four separate but related questions: 1) how can data mining and market segmentation be used to identify unique and distinguishable patient segments (clusters) in the total joint replacement patient population, 2) how do hospitals accurately classify prospective total joint replacement patients such that their assigned segment provides useful information for clinicians or other health professionals, 3) how can clinicians identify cost drivers in TJA that are tied to a patient's cluster, their attributes, or other clinical factors, and 4) how do hospitals identify patient outcome drivers. Building an operations research methodology around these four questions will help hospitals adapt to the current transformation in healthcare payment models.

The methodology is rooted in operations research methods and organized around a framework that considers patient composition and demand, classification of patients, and cost and outcome drivers. Figure 1-2 outlines the methodology.

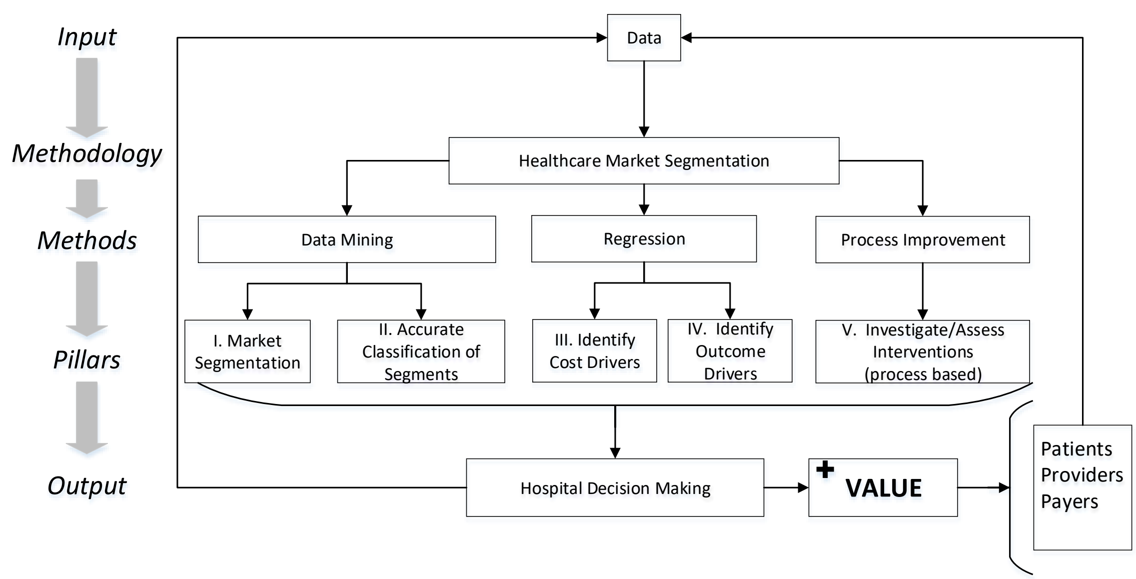


Figure 1-2. Five Pillar Healthcare Market Segmentation Methodology

The fifth pillar of this framework includes the investigation and assessment of interventions in the total joint arthroplasty process. It is beyond the scope of this dissertation to measure and prove the effectiveness of various known and undiscovered value improving interventions; however, applying this methodology to reduce the financial gap created by value-based payments will likely require hospitals to intervene with either process based or care based interventions. As part of this framework, non-clinical interventions related to process improvement or cost reduction will be included. This framework alone is not designed to solve the hospital's financial risk dilemma, rather it should enhance a hospital's ability to incorporate data analytics to improve healthcare value. Furthermore, though this methodology is framed around the payment models derived from the ACA, Medicare is not the only insurer developing and implementing bundled payments around joint replacement. Should the ACA be replaced, Medicare and private insurers will continue to seek value improvements that increase quality and decrease cost. This data driven methodology helps hospitals and clinicians achieve those goals regardless of payer.

This methodology does not assess relative value between different patients nor does it attempt to quantify value specifically from the patient's perspective. Individual patient value is a function of many hard to measure factors including potential lost wages, distance to home, travel time and access to the hospital, complications, discharge location, and readmission. For this research, reducing readmission, decreasing length of stay for healthier patients, and discharging patients home are considered quality metrics.

1.4 Outline

The remainder of this dissertation is organized around the five pillars shown in Figure 1-2 (above). Chapter 2 provides a systematic review of the current literature in healthcare market segmentation and data mining. The systematic review uncovered gaps in the published literature regarding the use of advanced data analytics in healthcare. Chapter 3 bridges the gaps by detailing the application of machine learning to healthcare data. A two-phase methodology is presented. The first phase aligns with the first pillar of the healthcare market segmentation methodology and investigates the application of unsupervised statistical learning to a total joint arthroplasty patient data set to cluster patients into unique and distinguishable market segments.

The second phase aligns with the second pillar of the framework and uses supervised statistical learning to accurately predict (classify) a patient's market segment from a given set of patient attributes.

Chapter 4 presents a data driven approach to uncovering the cost and outcome drivers in total joint arthroplasty. The first half of the chapter uses regression techniques to uncover the factors that impact supply cost and readmission. Supply cost in TJA comprise a significant portion of the episode of care costs and readmission is both a cost and outcome driver. High readmission rates impact patient outcomes and hospital quality scores which drive down patient satisfaction and reimbursements. Also, readmission costs are covered under the bundled payment effectively reducing hospital reimbursement. The second half of Chapter 4 is focused on high supply cost patient outliers and their impact on average supply cost per episode.

Chapter 5 is a compilation of the other four pillars that evaluates several interventions and a new physician led clinical and process improvement framework. The chapter represents the culmination of a 24-month study of the total joint arthroplasty process at the Penn State Hershey Medical Center (PSHMC), an academic teaching hospital in Pennsylvania. Market segmentation is revisited and applied to a new set of patient attributes to better understand post-acute care costs, which similar to supply costs and readmission, consume a large portion of the episode costs. After market segmentation and process improvement studies showed that patient length of stay has high variability between clusters, a rapid recovery protocol was implement to reduced length of stay and improve healthcare value. Market segmentation also showed that a specific homogeneous cluster was prone to high cost post-acute care. Interventions to mitigate post-acute care costs were extracted from recent articles and evaluated using simulation. The simulation is a tool that enables clinics to understand the financial impacts of decisions under certainty.

Chapter 2

Healthcare Market Segmentation and Data Mining: A Systematic Review

Consumers' behaviors and attitudes are critical information in a healthcare environment that is rapidly moving toward patient centered care. Patient centered care is premised upon individuals becoming more active participants in managing their health. A systematic review of the literature concerning healthcare market segmentation and data mining identified several gaps and opportunities in healthcare marketing research. Common themes included: 1) reliance on survey data, 2) clustering methods, 3) limited classification modelling after clustering, and 4) detailed analysis of clusters by demographic data. Opportunities exist to expand healthcare marketing research to leverage patient level data with advanced data mining methods.

2.1 Introduction

According the World Health Organization (WHO), "health promotion is the process of enabling people to increase control over, and to improve, their health. It moves beyond a focus on individual behavior towards a wide range of social and environmental intervention" (WHO, 2014). Further, the Centers for Disease Control and Prevention (CDC) state that healthcare marketing involves "creating, communicating and delivering health information and interventions using customer-centered and science-based strategies to protect and promote the health of diverse populations." Note that healthcare marketing draws from traditional marketing theories and principles and adds science-based strategies to prevention, health promotion and health protection (CDC, 2011). The purpose of market segmentation is to find specific well-defined, homogenous customer groups in a larger population, some of which are likely to respond positively to promotions or service offers (Woodside, Nielsen, Walters, & Muller, 1998).

Market segmentation offers insights into healthcare consumers' behaviors and attitudes, which is critical information in an environment where healthcare delivery is moving rapidly towards patient-centered care that is premised upon individuals becoming more active participants in managing their health. Awareness of patients' preferences and styles needs to be taken into consideration. Strategies to encourage and support consumer engagement in healthcare are important for health care organizations (e.g., providers, health plans, pharmaceutical companies, etc.). Increased access to health information can help patients make better and more informed decisions leading to better quality of care, health outcomes, and satisfaction with care. Providing individuals in a community with more useful information may change their behavior in a way that reduces health costs. Healthcare market segments may provide valuable clues as to how healthcare organizations may more specifically target and personalize products and services for healthcare consumers (Greenspun & Coughlin, 2012).

Many patients are motivated to increase control over and improve their health based on individual circumstances, to include experience with a new medical problem, loss of employer-sponsored coverage, or their inability to obtain effective medical treatment due to cost or denial of coverage. As these circumstances increase across the patient population and as healthcare costs force many to go without insurance, it is anticipated that consumer activist segments will increase (Greenspun & Coughlin, 2012). Individuals' self-care is positively correlated to education and cultural perspectives about what constitutes health and healthcare. Further, with the onset of the Affordable Care Act and changes to employer-sponsored insurance coverage, individuals may experience higher levels of price sensitivity, forcing them to become more actively involved in their medical treatment decisions (Greenspun & Coughlin, 2012).

As a means to improve health promotion for patients in a given community, effective healthcare marketing strategies should be developed and employed. Pires and Stanton (2008) discuss the application of marketing knowledge to healthcare services, arguing that social marketing has played a crucial role in acceptability and awareness regarding key health issues by campaigns (e.g., anti-smoking, anti-obesity, etc.). The authors proposed the importance of market segmentation in the healthcare services for better strategizing as per specific needs. As a result of improved information and communication technologies as well as health information technology (HIT), patients are now better empowered to improve their health.

Market segmentation is a critical step in healthcare marketing, which the CDC defines as a blending of social networking and health communication (CDC, 2015). Customer-based market segmentation provides the focus and precision required to enhance personalized healthcare by identifying the latent relationships between attributes found in individual health records, customer surveys, and or demographic data. These relationships help define patient clusters or segments which hospitals, health systems, insurers, and affiliated health agencies can use to refine healthcare marketing efforts. Understanding market segments can focus health communications, which are strategies to inform and influence health-based decision-making (CDC, 2015). Targeting health promotions to specific market segments increases efficiency, decreases health promotion costs, enhances patient-centered care and personalized healthcare goals, and is more likely to increase health consumer participation in managing their own health. Additionally, understanding the uniqueness of market clusters can identify underserved segments and may help link existing health promotions to yet unexplored segments.

Market segmentation studies hold the potential to be a critical component of the National Institutes of Health translational research initiatives. Although the definition of translational research varies, its essence involves the transfer of laboratory or benchtop research to larger audiences (Rubio et al., 2010). The NIH-funded Clinical and Translational Science Institutes provide and important platform for this improvement. Ideally, research investments at a local level spawn best practices that ultimately become standard operating procedures that are widely adopted across the healthcare continuum. Market segmentation allows translational researchers to efficiently locate desirable healthcare market segments to target with new laboratory research; this will allow new clinical research to proliferate more rapidly to patient segments most in need.

Tynan and Drayton (1987) discuss the importance of market segmentation techniques in overall marketing strategy. They emphasize that segmentation helps marketers improve precision of the prediction of consumer responses to a marketing stimuli. They suggested that the main market segmentation bases can be: geographic, demographic, psychological, psychographic or behavioral. They argue that market segmentation leads to closer association with the targeted set of consumers. In addition, strategic market segmentation plays a key role in discovery, innovation and development of medical products and services (MacLennan & Mackenzie, 2010). The authors argue that there are both driving and constraining forces acting for and against strategic market segmentation in any organization. These forces are mostly associated with limited resource availability and their optimum allocation along with the organizational culture.

There have been numerous healthcare marketing research studies done over the past few decades. Common clustering methods include hierarchical and non-hierarchical clustering, chi-squared automatic interaction detection, and classification and regression trees (CART). Additionally, market segmentation studies normally fall in to one two categories: a priori or data-driven (Wind, 1978). In healthcare, a majority of the papers also use either surveys or interviews to gather the data. In several papers, the concept of market segmentation is discussed without a formal model or the application of data analytics. In this chapter, I survey the data mining approaches to healthcare market segmentation. In addition to discussing the results and limitations, I provide recommendations for future opportunities in healthcare marketing research.

2.2 Methods

2.2.1 Systematic Search and Article Selection

In order to build the initial list of journal articles concerning healthcare market segmentation, a systematic literature search was performed using PubMed and PMC online database searches. Clustering and market segmentation are well-established and published methods; therefore, containing the search to medical related journals helped filter results. The search terms included clustering, market segmentation, health market segmentation, and healthcare market segmentation. Each was used individually. After filtering queries initially by publishing date and key word search, further filtering via abstracts and ultimately full-text reviews reduced the number of articles to 12. Figure 2-1 provides the article selection flowchart.

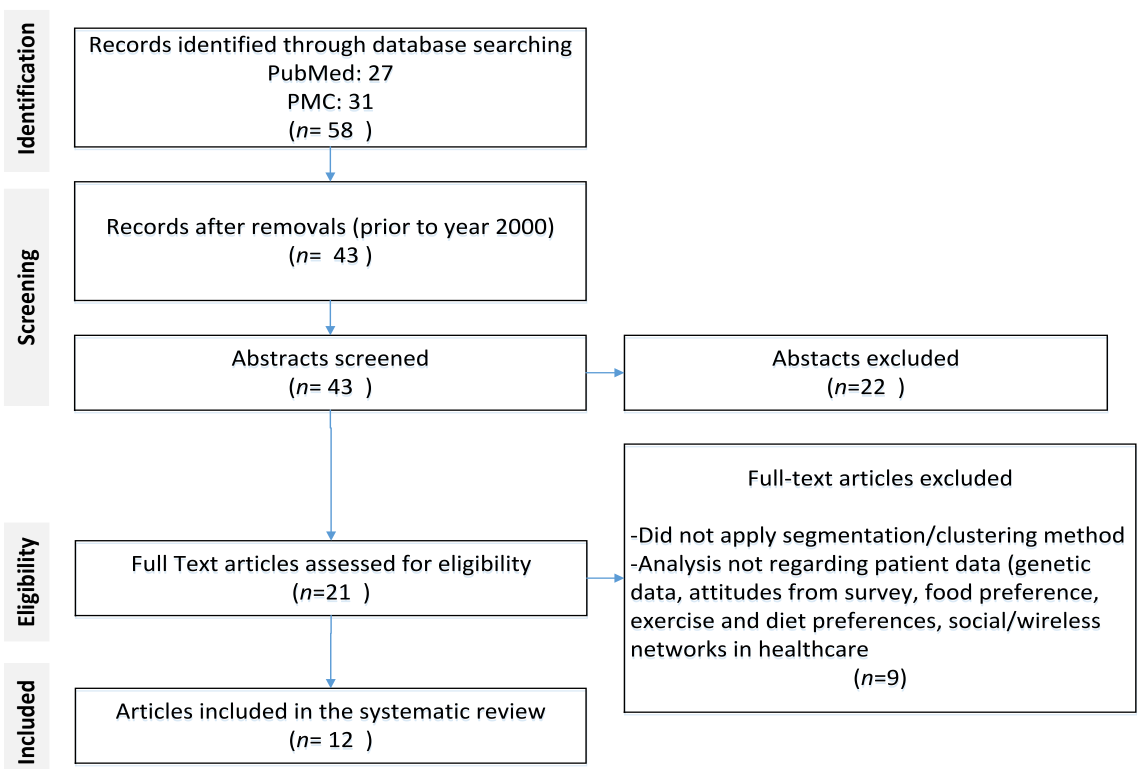


Figure 2-1. Article Selection Flowchart

Here are some descriptive statistics of the 12 selected studies. Country breakdown: United States (6), Sweden (1), Korea (2), Denmark (1), Taiwan (2). Primary data mining method: Latent cluster analysis (1), hierarchical clustering (6), *k*-means (4), other (1). Type of data: survey (5), patient data/secondary use data/combination (7). Type of study: prospective (4), retrospective (8).

2.2.2 Description of Data Mining Methods

2.2.2.1 Non-hierarchical Clustering

A priori clustering. In a priori clustering, specific variables such as demographic, state of being, and geographic, are predetermined as the basis for clustering decisions. After all data is collected, clusters are formed around these specific predetermined variables. As compared with other clustering techniques, a priori clusters are easier to interpret, measure, and act upon given the observations fit the cluster. When the segmentation variables are not predetermined, resulting clusters must be interpreted to understand why they formed and what types of observations fit the cluster.

***K*-means clustering.** *K*-means clustering is an unsupervised statistical learning technique that separates *n* multidimensional observations into *k* clusters based on the similarity between the observation and the centroid of the cluster. The technique requires an initial value of *k* from which *k* initial clusters are formed. Depending on the variant of the algorithm, each observation is either assigned a cluster number first or *k* observations are randomly selected as initial centroids. In either case, every observation is assigned to a cluster based on a similarity measure. The most common for continuous attributes is the squared Euclidean distance (Jain, Murty, & Flynn, 1999).

The *k*-means clustering algorithm is iterative and at each step calculates the centroid of each cluster, then compares each observation to the centroid based on a similarity measure. Observations are reassigned to clusters based on maximizing similarity between the observation and clusters centroid. The process repeats until a predetermined convergence criterion is achieved. Convergence criteria could be based on iterations, when no more reassignments occur, or when no significant change in squared error from one iteration to the next (Jain et al., 1999). *K*-means clustering is widely used due to ease of use and ability to handle large data sets. The *k*-means clustering algorithm is susceptible to initial starting conditions, which can prevent it from reaching a global minimum. It works best when multiple starting points are used.

2.2.2.2 Hierarchical Clustering

Hierarchical clustering covers both agglomerative and divisive clustering. In each case, the method starts with a set of *n*-multidimensional observations. The difference being that agglomerative hierarchical clustering starts with *n* clusters and terminates with one cluster and divisive clustering starts with one cluster and subdivides into *n* clusters. The methods are similar but approach the clustering from different sides, one being construction and one be division. Unlike *k*-means clustering, there is no predetermined value of *k*. The user must determine an appropriate value of *k*. The output from hierarchical clustering is displayed in a dendrogram which “represents the nested grouping of patterns and similarity levels at which groupings change.” (Jain et al., 1999).

In agglomerative clustering, clusters are traditionally joined based on a minimum distance measure or maximum similarity measure. The similarity between pairs of observations, one from each cluster, are compared and clusters are merged based on a maximum similarity criteria (normally minimum distance). Different algorithms use different methods to determine minimum distance; two common techniques are complete link which combines clusters that have the minimum of the maximum pairwise distance between any two points (from different clusters) and single link which combines two clusters if the distance between them is the minimum of the pairwise distances (Jain et al., 1999).

2.2.2.3 SPSS Two-Step Cluster Analysis

The cluster analysis approach is used in the SPSS software package. The clustering algorithm is a combination of several techniques. In the first step or pre-cluster phase, sequential clustering is applied to each observation (Theodoridis & Koutroumbas, 1999). Observations are passed down a decision tree and are either assigned to a cluster of similar observations or the observation forms a new cluster. This output of step one is a set of sub-clusters, *p*, where *p* is less than or equal to *n*, the number of observations. In step 2, agglomerative hierarchical clustering is applied to the *p* sub-clusters to form the desired number of *k* clusters. By design, the sub-clustering step places observations into at most 512 sub-clusters. This reduction in size make subsequent hierarchical clustering feasible. This technique can be applied to large data sets (SPSS White paper).

2.2.2.4 Latent Class Analysis

Latent class analysis (LCA) is a probability-based clustering technique that seeks to cluster observations based on unobserved variables. LCA uses a stochastic approach to find likely distributions with the data and the placement of observations within the distributions such that two or more observed variables are conditionally independent of each other based on the condition that they are in the same latent class (Kent et al., 2014). The cluster model is

$$P(y_n|\theta) = \sum_{j=1}^{S} \pi_j P_j(y_n|\theta_j) \qquad \text{2-1}$$

where *S* is the number of clusters, y_n is the *n*th observation of the observable (not latent) variable, and π_j is the prior probability of membership in cluster *j*. P_j is the probability of y_n given θ_j (cluster specific parameters) (Haughton et al., 2009). LCA takes a model based approach to clustering and has been used in market segmentation studies. It is fairly common in marketing, economics, and the social sciences and used as an alternative to the common distance based methods (hierarchical, *k*-means).

2.2.3 Description of Distance/Similarity Measures

Ward's method: Ward's method is also known as minimum variance criterion. This method is applied in hierarchical clustering algorithms where the objective is to minimize the total within cluster variance. The algorithm starts with *n* clusters representing the *n* observations. Then, *n*-1 clusters are formed out of *n* clusters by combining the pair of observations that results in the smallest increase in within cluster variance. Ward's method uses a squared Euclidean distance measure to determine minimum variance (Ward, 1963).

Gower's dissimilarity coefficient: A general similarity measure, S_{ij}, that Gower (1971) developed to determine similarity between two observations, i and j. This coefficient can be applied to ordinal, continuous, and dichotomous data. In determining Gower's coefficient, the similarity between two observations on the *k*th dimension are calculated for all *k* dimensions.

$$s_{ijk} = 1 - \frac{|x_{jk} - x_{ik}|}{R_k}$$

where R_k is the range of *k*. The overall similarity coefficient is

$$S_{ij} = \frac{\sum_{k=1}^{q} s_{ijk}}{\sum_{k=1}^{q} \delta_{ijk}} \text{ where } \delta_{ijk} = \begin{cases} 0 \; if \; there \; is \; a \; missing \; value \; in \; i \; or \; j \\ 1 \; \; otherwise \end{cases}$$

2.3 Results

A total of 12 studies were examined in significant detail based on the article selection diagram in Figure 2-1. Table 2-1 shows the summary of the 12 articles.

Table 2-1. Summary of 12 Articles on Healthcare Market Segmentation

	Retrospective /Prospective	Setting	Sample Size	Country
1. Newcomer, Steiner & Bayliss (2011). Identifying subgroups of complex patients with cluster analysis.	Retrospective	HMO population; patients in the top 20% of care expenditures and with 2 or more chronic med conditions; data from CY2006/2007	15,480	USA
2. Kolodinsky & Reynolds (2009). Segmentation of overweight Americans and opportunities for social marketing.	Prospective	National level polling data; patient survey conducted by authors regarding food and lifestyle behaviors	581	USA
3. Berg et al. (2010). Using market research to characterize college students and identify targets for influencing health behaviors.	Prospective	Survey of college aged students from Minnesota; diverse sample	2,700	USA
4. Kent, Jensen, & Kongsted (2014). A comparison of three clustering methods for finding subgroups in MRI, SMS or clinical data.	Retrospective	Secondary use data; longitudinal studies; multiple data sets to include real data and randomly generated test data.	3x MRI data sets (412, 631, and 4162 patients); 1x self-reported lower back pain intensity data set (n=1121), clinical data set (n=543) based on patient responses to chiropractic care	Denmark
5. Axen et al. (2011). Clustering patients on the basis of their individual course of low back pain over a six month period.	Prospective (observational)	Outpatient chiropractic care; based on survey and clinical data.	176 patients with low back pain	Sweden
6. Liu & Chen (2009). Using data mining to segment healthcare markets from patients' preference perspectives.	Retrospective	US not for profit healthcare group; inpatients; telephone interviews/surveys	1,561	USA
7. Suragh et al. (2013). Psychographic segments of college females and males in relation to substance use behaviors.	Prospective online survey	6 college campuses in Southeast, USA.; diverse male and female student population; 230 question survey conducted in 2010.	3,469	USA
8. Moss, Kirby, & Donodeo (2009). Characterizing and reaching high-risk drinkers using audience segmentation.	Retrospective Survey data analysis	Combination of 2004 US survey data (BRFSS data plus Simmons Market Research Bureau data that consists of public records data, US Census data, etc.).	>30,000 people	USA
9. Kim, Oh, Cho, & Park (2013). Stratified sampling design based on data mining.	Retrospective	Single specialty clinics and hospitals that conduct either general surgery or ophthalmology; 2011 data; combination of hospital and insurance data merged into a single database for analysis	442 clinics/hospitals that did general surgery; 715 facilities with specialty of ophthalmology	Korea
10. Lee (2012). Data mining application in customer relationship management for hospital inpatients.	Retrospective	University hospital; data from Jan to Dec 2009.	14,072 discharge records	Korea
11. Wu, Lin & Liu (2014). Analyzing patients' values by applying cluster analysis and LRFM model in a pediatric dental clinic in Taiwan.	Retrospective	Pediatric Dental Clinic; data from July 2009 to June 2011	1,462 patients (under 18 years old)	Taiwan
12. Cheng, Chang, & Liu (2005). Enhancing care services quality of nursing homes using data mining.	Retrospective	Nursing Home; study period April to March 2003.	407 nursing home residents	Taiwan

	Outcome(s) measured	Factors Used	Results	Method
1	How patient's cluster around coexisting conditions, demographics	Obesity, mental health conditions, diabetes, cardiac disease, COPD, kidney disease, cancer, gastrointestinal bleeding, chronic pain stroke, skin ulcer, dementia, fall, abdominal surgery, orthopedic surgery, back surgery, hip fracture	10 clinically relevant clusters grouped around single or multiple anchoring conditions. Mental health and obesity prevalent in all clusters.	Agglomerative hierarchical clustering; Ward's Algorithm; used SAS V9.2 software
2	How patients clustered around food and lifestyle behaviors	Behavioral variables; personal and environmental factors. Ht, wt, computer use, smoker, gender, education, income, children, age, geographic region, residence location (urban/rural); knowledge of food pyramid; exercise	Five clusters (highest risk, at risk, right behavior/wrong results, getting best results, doing OK). 99% in highest risk were overweight	Two step cluster analysis with Schwartz's Bayesian Criteria. ANOVA and Chi squared used to determine whether cluster membership related to demographics. Used SPSS software.
3	Health related measures, confidence and motivation, market research; what are the influencers of behavior	Demographic, psychographic (attitudes and interests), health-related variables	Three clusters: stoic individualists, thrill seeking socialists, and responsible traditionalists.	Hierarchical cluster analysis using Ward's Method. Used Gower's general dissimilarity coefficient then clustered on distance matrix products. Used ANOVA and Chi-squared tests to compare variables across segments. Used SAS and SPSS software.
4	Consistency across methods	Number of subgroups detected, classification probability of individuals in a subgroup; reproducibility of results, ease of use of software	Number of subgroups detected varied by method; certainty of classifying individuals into subgroups varied; finding were reproducible; ease of use and interpretability varied; subjectively picked Latent Gold as best overall	Comparison of three common clustering methods using 9 data sets (five actual data sets and four artificial). Methods used are SPSS's TwoStep CA, Latent Gold Latent Cluster Analysis (LCA), and SNOB LCA.
5	Change in pain intensity over time	26 parameters reduced to four via spline (nonlinear) regression: slope and intercept of regression line in early course; difference in slope between two regression lines, intersection estimate	Four clusters with distinct clinical courses	Ward's method and hierarchical clustering; used SPSS, STATA, and Sleipner software.
6	How patients clustered; most important to least important attribute by cluster. Survey questions were demographic and statements that measured healthcare service preference.	24 Attributes reduced to five factors through cluster analysis: communication and empowerment, compassionate and respectful care, clinical reputation, care responsiveness, efficiency	Three clusters emerge: reputation driven, performance driven, and empowerment driven.	Hierarchical cluster analysis, Pearson correlation and average linkage to measure similarity. Used R software and a mix of hierarchical and nonhierarchical methods plus Enterprise Miner.
7	How students clustered according to psychographic characteristics and substance use behavior	15 psychographic measures (sensation seeking, personality traits (5), 9 measures adapted from tobacco industry	Thee psychographic distinct clusters: safe responsible, stoic individuals, thrill-seeking socializers.	Hierarchical clustering: Ward's method, Gower's dissimilarity coefficient (due to nominal and ordinal values). Used t-statistic to determine optimal number of clusters. Used SPSS software.
8	Clustering of population based on high risk drinking behaviors/attitudes	Self-reported drinking episodes, frequency, demographics.	66 audience segments with top ten analyzed in depth. Cyber-Millennial cluster has highest concentration of binge drinkers. Laid back towners, city producers, metro newbies are in descending order the clusters of highest risk.	Proprietary PRIZM^{TM} software. Audience segmentation that creates 66 clusters from nationwide database.
9	Classification of healthcare providers	Type of location, population density, number of specialists, number of beds, number of inpatients per specialist, lengthiness index, costliness index, case-mix index, rate of annual change in number of inpatients per specialists	Four clusters of ophthalmology facilities, three clusters of general surgery facilities.	K-means clustering and decision tree induction to segment and classify healthcare providers then to stratify them into five stratum. Used MATLAB software.
10	Customer loyalty, Customer relationship model	Recency, frequency, monetary: LOS, certainty of selectable treatment, surgery, number of accompanying treatments, kind of patient room, department from which discharged.	Customers were classified as either loyal or ordinary. Demographic characteristics were overlaid on two clusters. Decision tree showed most important factor is LOS.	K-means clustering, group comparison via t-test. Decision tree and logistic regression used to predict patients who were clustered as "loyal customers".
11	How pediatric dental patients cluster	LRFM (Length, Recency, Frequency ,Monetary model), gender, age	12 clusters based on LRFM.	K-means and self-organizing maps; used SPSS Modeler 14.2 software.
12	Patient clustering and CRM	K scale, LOS, Times of stay, discharge reason, # diseases, special passageways brought, # rehab outpatient visits, age	Four clusters of patients/residents. Used clusters to determine best care strategies based on expert opinion.	Demographic clustering using K-means; cluster size determined using MANOVA test of the discriminant analysis. Used SPSS and Intelligent Miner V6.1.

There have been numerous papers written on healthcare market segmentation over the past 40 years. The advent of powerful computers and statistical learning software have expanded opportunities for exploring market segments through the use of big data sets. The 12 papers reviewed include many of the market segmentation and cluster techniques that are used in the broader literature regarding marketing studies. *K*-means clustering and hierarchical clustering are the predominate methods in these studies. Other methods such as a priori clustering and chi-squared automatic interaction detection (CHAID) were cited in several of the articles published prior to 2000 (Carroll & Gagnon, 1983; Malhotra, 1989). The 12 papers included in this review started with a set a data and applied unsupervised learning techniques to find homogenous clusters or segments within the population.

2.3.1 Diversity of Studies

All 12 studies are published in peer-reviewed journals. They are authored by a mix of professionals to include medical doctors and PhD researchers from economics, healthcare science, industrial engineering, economics, and marketing. The studies range from analysis of clinical populations (Newcomer et al., 2012; Axen et al., 2011; Kim et al., 2013) to segmentation studies on survey data (Liu and Chen, 2009; Suragh et al., 2013; Moss et al., 2009; Kolodinski and Reynolds, 2009; & Berg et al., 2010). Of the papers that used survey data, two looked at college student substance abuse behaviors (Berg, 2010; Suragh, 2013), one looked at customer preference for healthcare service and clustered patients based on their preference and demographic attributes (Liu, 2009), and the last two used large survey data from a combination of the Behavioral Risk Factor Surveillance System (BRFSS), US Department of Agriculture funded nationwide polls, and a mix of public and U.S. Census data (Moss, 2009; Kolodinski and Reynolds, 2009).

Two of the studies based on patient data investigated RFM (or recency, frequency, and monetary models). Lee (2012) studied customer loyalty in a university hospital setting in Korea. He analyzed patient demographics and hospital visit data to understand which patient types were loyal or ordinary users. Wu et al. (2014) conducted a similar study in Korea where they looked at a tenth of the sample size as Lee (1462 vs 14,072), but studied LRFM which is RFM plus length. The goal of Wu et al. (2014) was to cluster the under 18 year old patient population in a dental clinic based on demographics, length of stay, frequency of visits, and proximity of recent visits.

2.3.2 Outcomes Measured

Two of the retrospective studies from Taiwan and Korea focused on customer loyalty and Customer Relations Management (CRM). Cheng et al. (2005) applied *k*-means clustering to demographic data regarding nursing homes. The goal was to cluster patients based on demographics, specialty care required, rehabilitation services, etc. and then develop care service strategies based on provider feedback. Lee (2012) conducted a similar study in Korea using a CRM. Lee (2012) also applied k-means clustering with *k* equal to two. The two clusters divided the population into loyal and ordinary patients. After clustering, Lee (2012) applied decision trees to stratify the loyal patients to determine which factors were most important in determining how a patient is classified.

Lee (2012) was not alone in his post-cluster stratification approach. Kim et al. (2013) used *k*-means clustering and decision tree induction to segment and classify healthcare providers. In this study of hospital providers, Kim et al. (2013) looked at location, population density, beds, patient to provider ratio and other costing data to segment both single specialty and hospitals that conduct either general surgery or ophthalmology services. After clustering both types of hospital services, they applied a stratification approach using decision trees to develop homogenous strata.

Determining homogenous strata allows for better sample approaches that aid in future policy studies (Kim et al., 2013).

Four of the papers that applied market segmentation to survey data measured health and behavior outcomes. Kolodinsky et al. (2009), Berg et al. (2010), Suragh et al. (2013), and Moss et al. (2009) all looked for influential behaviors with the end state of being able to identify distinct segments and then use specific techniques to target those segments in order to modify behaviors. Berg et al. (2010) and Suragh et al. (2013) conducted almost the same study in different regions in the US and arrived at the same number of clusters with strikingly similar names and cluster demographics. The prior study was in Minnesota and the latter was a larger study conducted at six universities in the Southeast. The congruency of results despite different timeframes, locations, statistical software programs, and sample sizes indicates the strength of cluster analysis to deliver repeatable findings given similar data sets. Although not specifically addressing college students, Moss et al. (2009) conducted a larger version of Suragh et al. (2013) and Berg et al. (2009) studies. Moss et al. (2009) used various large data sets from the CDC, publicly available data, and BRFSS to look at the attitudes and behaviors regarding high risk drinking. This study used a proprietary software called PRIZM that clusters large public data sets into 66 segments. The goal of this study is similar to the college surveys in that it tried to form homogenous subgroups, decompose each by the strength of their attributes, and then use that information to target at-risk segments with marketing strategies aimed at behavior modification.

Similarly, but on a much smaller scale, Kolodinsky et al. (2009) used national poll survey data to cluster based on behavioral, environmental, geographic, food knowledge, and education factors. Kolodinsky et al. (2009) was interested in obesity and the role of food and lifestyle behaviors on population health. A striking similarity in Kolodinsky et al. (2009), Berg et al. (2010), and Suragh et al. (2013) is how they use the same industry practices that created the problems they are studying to counter the problems. Both Suragh et al. (2013) and Berg et al.

(2009) borrow from the tobacco industry and Kolodinsky et al. (2009) borrow survey methods from the food industry.

Liu and Chen (2009) and Kent et al. (2014) conduct market segmentation using different approaches but each applies multiple clustering techniques to verify the results. The prior uses survey data while the latter is based on secondary use data from a variety of studies. Liu and Chen (2009) use a mix of hierarchical and non-hierarchical methods and ultimately settle on a hierarchical clustering method that reduces the attributes from 24 to 5 yielding 3 distinct clusters. Kent et al. (2014) apply and compare three different methods to five real data sets and 4 randomly generated data sets to test reproducibility, likeness of outputs, and ease of use.

The final two papers use patient data sets to cluster patient populations based on a specific condition or set of conditions. In Axen et al. (2011), a composite data set based on questionnaires and self-reported pain score data are analyzed. The self-reported data is via time series SMS text messages over a 26-week period. These patient pain progress scores are cleverly reduced to four parameters through the use of nonlinear spline regression. These four parameters (developed for all 176 patients) are segmented using hierarchical clustering. In Newcomer et al. (2011), hierarchical clustering is also used, however, in this study, the sample size is large (15,480 patients) and pulled from a health maintenance organization (HMO) database of patients with at least two chronic medical conditions that fall into the top 20% of care expenditures. The goal of the study is to further segment high risk and high cost patients to enable clinicians to target specific at risk populations with appropriate health interventions and care management plans.

2.3.3 Country of Origin, Time Frame and Statistical Software used in Studies

Half of the studies were conducted in the United States, four in Southeast Asia and two Scandinavian countries. The majority of the 12 papers were published after 2009 and apply current data analytic software including SAS, SPSS, STATA, and R. All studies use data collected after year 2000. Three studies use SAS, six studies use SPSS, and R, MATLAB, PRIZM, STATA, SNOB LCA, and Latent Gold LCA are used less frequently. See Table 2-1 for specifics.

2.3.4 Methods Used

The 12 papers in this review cover a breadth of subjects, methods, and outcomes. The common themes are market segmentation and understanding how patients, clinics, students, or adults align with others of like attributes. The goal of these studies is to provide insight and an angle to better understand a population. The number of clusters or segments varies across studies which is consistent with cluster analysis in general. In most cases, the user must define the number of clusters ahead of time or must identify a condition upon which the algorithm stops. Hierarchical clustering is used in five studies (Newcomer et al., 2011; Berg et al., 2010, Axen et al., 2011; Liu and Chen, 2009; & Suragh et al., 2009). In all but one of them, Ward's method is used as the distance/similarity measure. In Liu and Chen (2009) Pearson's correlation is the similarity measure. Four of the studies use *k*-means clustering (Kim et al., 2013; Lee, 2012, Wu et al., 2014, & Cheng et al., 2005).

2.4 Discussion

From the 12 articles investigated, I sought to learn how data mining techniques can be leveraged for conducting market segmentation with respect to patient preferences for healthcare attributes and exploring the patient segment demographic characteristics. The identification of gaps and opportunities provides the necessary direction for future healthcare marketing research. A detailed discussion of the surveyed articles follows.

Liu and Chen (2009) employed cluster analysis techniques to conduct healthcare market segmentation using complicated psychographic variables and to reveal the benefits of data mining to understand consumers' psychological needs for improving healthcare services. The authors used survey data for patients who received care from a non-profit healthcare group in 2006. Respondents were surveyed on 24 healthcare services attributes covering physiological care, psychological care, physical environment, and spiritual care. Factor reduction techniques reduced the number of factors to five and cluster analysis identified three segments. Factor reduction helped make the results more interpretable. Liu and Chen (2009) identified three healthcare market segments: reputation-driven, performance-driven, and empowerment-driven. Segments are subgroups with similar patient preferences in the whole healthcare market. Successfully identifying demographically well-defined consumer segments can assist hospital managers develop long-term business strategies and offer an optimal mix of products and services that meet customer needs and preferences (Ross et al., 1993; Woodside et al., 1998).

Kim et al. (2013) conducted a retrospective study using stratified sampling design based on k-means clustering and decision tree induction. Although their approach applied data mining techniques, they were focused on healthcare providers and not consumers. Their research was specific to general surgery and ophthalmology, where they identified three clusters of general surgery clinics and hospitals and four clusters of ophthalmology clinics and hospitals. The three general surgery clusters were divided based on whether they were private or public and the number of inpatients. The ophthalmology hospitals clustered similarly with the additional factor of whether there were multiple specialists in the hospital. The authors' motivation was to improve sampling efficiency by creating homogenous strata of clinics and providers based on several factors including size and ratio of patient to specialist. After clustering, decision trees were applied to the two sets of data to further stratify hospital and clinics. For each type of hospital/clinic, the decision trees resulted in five strata based on three variables: number of inpatients per specialist, population density, and lengthiness index. The results of this study are intended to help with future healthcare policy decision making. The authors did not compare their method against other well-known classification methods nor did they discuss the robustness of their method or stability of the clusters.

Lee (2012) applied data mining in a retrospective study to discover patient loyalty to a hospital and to model patient medical service usage. He studied customer relationship management marketing which is a process that segments customers to understand their behaviors with the goal of strengthening relationships with valuable customers. Patients were first classified into two groups: loyal and ordinary, based on recency, frequency, and monetary measures. Decision trees were then applied to each group (segment) to determine which factors/characteristics were most important in each segment. Logistic regression output was compared to the decision tree analysis and results were displayed on an ROC curve. This study is narrow on its approach to segmenting the market. It focuses on patient loyalty and uses frequency and monetary factors to determine segments. The author does not address why patients may use the same hospital frequently such as proximity to the next closest hospital, insurance considerations, or ability of patients to get to other facilities. Length of stay (LOS) is the leading factor in determining a patient's loyalty but LOS may be an unintended consequence of an unplanned hospitalization or a procedure gone wrong.

Chang et al. (2005) applied market segmentation, in particular *k*-means clustering, to a nursing home population in Taiwan to assist with customer relationship management. The goal of the study was to understand the characteristics of patient subgroups in a nursing home environment so that the staff can provide better, more customized, care to each patient. The authors use *k*-means clustering in combination with discriminant analysis to determine the appropriate number of clusters. Clustering was done with SPSS and Intelligent Miner V6.1. They showed that the population could be clustered into four unique subgroups. Each subgroup was then analyzed by a team of professionals to determine the best care service strategy. Given the wide range of patient care needs in a nursing care setting, understanding how patients segment according to their conditions and needs can help management tailor care to existing and future residents.

Newcomer et al. (2011) applied hierarchical clustering, namely Ward's algorithm, to a large HMO patient data base to identify clinically similar subgroups. The patient population included over 15,000 adult patients who had at least two comorbidities and ranked in the top 20% for cost expenditure per year. Using agglomerative hierarchical clustering, Newcomer et al. (2011) merged clusters based on Ward's distance. To assess the stability of their algorithm, they divided the data set in half, create a dissimilarity matric for each set using Jaccard's coefficient, then applied Ward's algorithm. Since the two data sets had similar cluster membership, the algorithm was applied on the entire data set. In 8 of the 10 resulting clusters with k=10 subjectively chosen, there was a clear dominate chronic condition that defined the segment. Newcomer et al. (2011) then analyzed each cluster by predominance of attributes and other comorbidities. The shortcomings in this study include the narrow focus on a single two-year data set and a lack of generalizability to other patient populations outside this HMO. Newcomer et al. (2011) did experiment with different clustering techniques but they do not show the results of the other methods nor how the outputs varied. The authors also do not discuss the relevance of their finding in mitigating chronic conditions or targeting at risk populations.

Kolodinsky et al. (2009) applied a social market segmentation approach in a behavioral study regarding peoples eating habits and the effect on body weight. The goal of this prospective study was to apply similar market segmentation techniques that the food industry uses to market products to understand people's behaviors and attitudes towards foods. Their survey questions were rooted in social learning theory and health belief model and interspersed with questions to understand socio-demographic attributes of the survey population. Kolodinsky et al. (2009) applied SPSS's Two-Step Cluster Analysis to the survey data initially excluding the demographic data. The 581 respondents clustered into five distinct segments primarily separated due to overweight risk. Segments were then analyzed using demographic data to better understand their composition. As in many of the healthcare market segmentation studies, the study ends with a list of clusters distinguished based on a factor or series of factors directed related to the goal of the study. What is missing is discussion on the relevance of the clusters and how machine learning can further help to classify new patients and match interventions to help with improved health outcomes.

Berg et al. (2010) and Suragh et al. (2013) each reported on the same topics with near identical results. Both considered college-aged students and segmented them based on survey questions specifically designed to assess health behaviors and substance abuse. They both used hierarchical clustering albeit from different software packages (SAS and SPSS respectively) and they used Gower's general dissimilarity coefficient and Ward's method. Gower's coefficient was applied to handle both nominal and ordinal values in the survey results. Each research team concluded that their respective student population, which was drawn from different regions within the US, segmented into the same three clusters: safe and responsible, stoics, and thrill seekers. Unfortunately, both studies conclude with three distinct segments. There is no discussion about the utility of each segment, what interventions could be used or have been used, and how statistical learning can further help classify new patients. Also, although Suragh et al. (2013) referenced the Berg et al. (2010) study, there were no parallels drawn or suggested.

Kent et al. (2013) is a comparative study of three different clustering methods on healthcare related data. In the study, the authors compare the clustering results of five real data sets and three artificial data sets across several criteria to include the number of segments or subgroups formed, the classification probability of observations into specific clusters, and the reproducibility of the clusters over 10 replications of each method on each data set. Kent et al. (2013) also compared methods for ease of use and interpretability of output. The methods tested in this paper included SPSS Two Step Cluster Analysis, Latent Class Gold, and SNOB latent class analysis. Although the results varied by methods and data set, the author's chose Latent Gold as the best method based on overall performance, sensitivity to determining the right amount of clusters, ease of use, and interpretability. All the methods provided highly reproducible results, but this could also be a function of starting seeds. The authors acknowledged that repeating the test with different starting seeds could negatively impact reproducibility.

Axen et al. (2011) provides another example of a prospective market segmentation study using a hybrid mix of survey and clinical data. This study is based in Sweden and focused on 176 patients with low back pain. The authors used a SMS messaging service to track pain scores of patients over 26 weeks. This time series data was reduced using non-linear spline regression to four measures that included the slope and intercept of the nonlinear regression line during the early part of the treatment course, the difference in slope between the early and late courses, and the intersection estimate. From this data, Axen et al. (2011) was able to cluster patients into four distinct segments. They used Ward's method, which is an agglomerative hierarchical clustering method. Given the small size of the data set, this technique is computationally efficient. Given the nebulous nature of non-specific lower back pain, providing a clustering tool to categorize and segment the treatment population based on the change of pain related factors over time is a unique approach and application of data mining. As in many of the healthcare-related segmentation studies, the details of how data analytics can be used in the treatment or monitoring of treatment and intervention planning is missing.

Similar to the Berg and Suragh papers, Moss et al. (2009) apply market segmentation in a study of high-risk drinking behaviors. They use a combination of data from the BRFSS and other private and publically available survey data. The authors use a proprietary software called PRIZM that segments the data into 66 subgroups. The article analyzes the top ten segments that are most likely to display highest risk behaviors. Each cluster is then dissected based on alcohol and tobacco use, digital communication use, sports and leisure activities, and media use to provide insight into how marketing strategies could be tailored to influence change in a subgroups behavior. Much of the details of the clustering technique are excluded from the paper.

Wu et al. (2014) conducted a market segmentation study of pediatric dental patients using SPSS's Modeler 14.2. The retrospective study applied *k*-means clustering and organizational maps to a sample of over 1,400 patients. The goal of the segmentation study was to understand how the patients clustered using attributes such as length of stay, recency of visits, frequency of visits, and monetary costs of visits. Demographic data such as age and gender were also included. The authors found 12 distinct clusters. The paper does not offer insight into how the clusters can or will be used to assist in better service or care delivery based on cluster assignment.

2.4.1 Gaps and Opportunities in Healthcare Market Segmentation

The predominance of healthcare market segmentation research over the past 26 years has focused on segmenting a healthcare population to identify segments for the purpose of behavior modification marketing and identifying subgroups within a larger but still specific group. There is a lack of studies based on patient-level EHR data. In the 12 papers that met the inclusion criteria for this review, five were based on survey data and a sixth used a combination of survey data and clinical data. Three papers used RFM data in conjunction with customer responsiveness models, one used specific hospital/clinic data on facility usage, one used service specific data from both chiropractic care and imaging services, and the final paper used patient level data. Although EHRs have been in existence for over a decade, only one study (Newcomer et al., 2011) took a large hospital data set and applied data mining techniques to cluster patients into meaningful segments. Understanding these segments will help health service providers, healthcare providers, and insurers target the right intervention and health services to "at risk" at "at benefit" subgroups.

Another gap in the healthcare market segmentation research is the lack of differentiation between market or audience segmentation and clustering. Many of the articles use clustering and segmentation interchangeably, whereas Liu et al. (2012) cite a few differences, namely, that clustering is a subset of segmentation that groups people or patients based on similarity (distance, likeness of needs, preferences, etc.). The clustering of people is a fundamental task of market segmentation and at one point in the late 1970s was synonymous with segmentation (Wind, 1978); however, market segmentation has evolved to include more than clustering or descriptive segmentation, and now includes predictive market segmentation (Liu et al., 2012). Furthermore, market segmentation research often involves multi-criteria optimization because the goal often includes the application of the descriptive clusters into economic criteria related to responsiveness, identifiability, profitability, and accessibility (Liu et al., 2012). With multiple objectives, there may be no single optimal solution.

In the majority of the 12 papers reviewed, the authors stopped at the clustering solution. They applied some form of cluster analysis to define homogeneous or near homogeneous subgroups, but they did not use those clusters to aid in predictive market segmentation. The gap in methods is the absence of supervised statistical learning applied after the unsupervised methods assigned a cluster to each patient or observation.

2.5 Conclusion

The importance of market segmentation studies applied to healthcare cannot be understated. In fact, Kennett et al. (2005) discuss the importance of healthcare market segmentation and assess how well hospital executives understand and use various marketing tools to include market segmentation. They conducted a survey of healthcare executives and mid to upper level healthcare managers to assess how hospital leaders rate the importance of and their current level of knowledge of marketing. They found that although market segmentation was considered to be very important for hospitals, it ranked in the top three tasks that that hospitals were least knowledgeable about (Kennett et al., 2005). Although Kennett et al. published their study over a decade ago, the lack of current healthcare market segmentation research indicates that there remains a lack of emphasis, knowledge, or use of market segmentation in healthcare.

The majority of healthcare market segmentation studies over the past twenty years focus on either survey data or specific data sets with the purpose of segmenting a specific population. Although these studies help define near homogenous clusters of patients, providers, or observations within the study, the studies end with defining the clusters. Market segmentation is more than just a study in defining a segment, it also includes predictive market segmentation in which the "decision maker seeks to optimize both within-segment homogeneity and segment level predictability" (Liu et al., 2012). Predictive segmentation is a key gap missing in most healthcare market segmentation papers.

Market segmentation is a well-known approach in marketing research and when applied to healthcare presents a great opportunity to identify subgroups of patients that share commonalities. In an era of skyrocketing healthcare costs and demand for services, understanding how patients cluster and respond to health promotions presents an opportunity to efficiently target segments of the market with health promotions tailored specifically to positively impact health outcomes. As healthcare costs increase, the trend for employers to shift more of the financial burden to individuals will continue and, as a result, will cause some consumers to seek personalized healthcare solutions to minimize their risks.

The widespread use of integrated EHR databases across the US presents an opportunity for healthcare providers to apply data mining methods to large healthcare data sets to enhance precision medicine. Hospitals, health systems and insurers already collect an enormous amount of patient data to include physical characteristics (age, weight, height), as well as past medical conditions, lab results, radiology reports and images, and a host of time-series data pertaining to each visit to a networked provider (those with access to the patient's EHR). Modern EHRs store all patient data in a centralized and searchable database. The EHR provides real-time access to providers in the clinical setting, but it also holds the potential to tell a much bigger story about a patient's past, current, and future health. This may include the types of treatments or health promotions patients may respond to, whether they value customer service, prefer messages via an interactive personal health record, or value routine care. In an era of unprecedented demand for hospital services and rising health care costs, the old adage that an "ounce of prevention is worth a pound of cure" is more relevant than ever. Healthcare market segmentation holds the potential to enhance personalized and precision medicine by allowing health providers to efficiently find and target at-risk or at-benefit market segments. At-benefit is defined as a segment of the population that can greatly benefit from preventative care or interventions to help sustain or strengthen current health.

This systematic review of healthcare market segmentation and data mining sets the conditions for the application of a two-phase healthcare market segmentation framework that is based on the analysis of patient EHR data. In the next chapter, I propose a two phase framework that consists of both unsupervised and supervised statistical learning using patient data. The phases represent pillars I and II from the healthcare market segmentation methodology shown in Figure 1-2. In the first phase (pillar I), patients are clustered into a set of distinct patient segments and in the second phase (pillar II), a series of classification models are run to predict patient cluster. The classification model that results from pillar II of the methodology is used to accurately predict a new patient's market segment based on eight patient attributes found their EHR. In subsequent chapters, patient market segments will be useful in understanding cost and outcome drivers as well as to help isolate segments of the population most at risk or at benefit for value improvement interventions.

This page is intentionally left blank.

Chapter 3

Data Analytics in Health Promotion: Healthcare Market Segmentation and Classification of Total Joint Replacement Surgery Patients

3.1 Introduction

To address the gaps found in the systematic review, a two-stage methodology is applied to a total joint arthroplasty data set. The first stage will apply unsupervised machine learning methods or clustering to patient level EHRs. The data is based on an approved Institutional Review Board (IRB) request from the Penn State Hershey Medical Center (see appendix A). After clustering each patient into the cluster for which they are most closely aligned or similar, supervised machine learning methods will predict market segment using patient attributes. The goal of this research objective is a classification model that accurately predicts the cluster assignment for out-of-sample patients, while offering insight into patient behaviors and attributes to help clinicians, healthcare marketers, and healthcare consumers move toward the goal of patient-centered and value-based healthcare.

3.1.1 Objectives for Pillars I and II of the Healthcare Market Segmentation Methodology

Numerous healthcare market segmentation studies were conducted over the past few decades, but few used EHR patient data and none applied supervised learning to help predict patients' healthcare market segment (i.e., cluster). In this research, I conduct healthcare market segmentation using EHR data on over 700 patients with total joint replacement surgery at PSHMC. The research objectives are: 1) to identify distinguishable patient clusters based on age, body mass index (BMI), pain scores, and other descriptive variables (pillar I); 2) to use those clusters to build predictive models that can accurately classify future patients into their healthcare market segment (pillar II); and 3) to discuss how patient clusters can provide useful information for helping project future patient outcomes. At a minimum, clustering patients into distinguishable segments will help clinicians, healthcare marketers, hospitals, and health insurers identify specific and targetable healthcare market segments. Efficiently targeting at-risk or at-benefit healthcare market segments will increase patient-centered care and improve value-based care initiatives.

The explosion of data in healthcare coupled with the need for real-time or near real-time decision support necessitates the use of advanced data analytics. This has set the conditions for new expert and intelligent systems healthcare-related applications. Although this is a small data set, it represents a large revenue-generating portion of a hospital's orthopedic service line. The demonstration of expert system tools applied to total joint arthroplasty could be adapted to other service lines such as oncology and cardiology. Healthcare data is vast, but with the proper tools from expert and intelligent systems it can provide great value to the healthcare industry.

3.2 Materials and Methods

3.2.1 Study Population

Unlike many of the previous health-related market segmentation studies, this study uses actual patient data from EHRs. The patients were treated at the PSHMC between 30 December 2013 and 30 September 2015, during which time they had total joint arthroplasty (joint replacement surgery) of the lower extremity (hip or knee). The de-identified clinical data are divided into two main categories: training and testing. The training data include 814 patients and over 75 attributes ranging from age, BMI, procedure code, length of stay, pain scores, ICD-9 codes (International Classification of Diseases and Related Health Problems, Version #9) and numerous other fields related to surgical side, approach, and times that various steps in the surgical process started and ended. There are 14 ICD-9 variables representing documented disease or health problems. In total, over 650 unique ICD-9 codes are referenced.

To account for common disease codes across the patient sample, each unique ICD-9 code is re-coded as a binary variable. Patients with that ICD-9 code are coded as a "1." The data include patients from diagnosis readiness groups (DRG) 461, 462, 469, and 470 (bilateral joint replacement or reattachment of the lower extremity without and with major complications and comorbidities (MCC) and major joint replacement or reattachment of the lower extremity without and with MCC) and both procedure codes 81.51 (hip) and 81.54 (knee). For this study, I excluded DRG 461/2 patients to focus on single surgical site patients. Removing DRG 461/2 reduced the patient count to 767 (loss of 47 records). Patient records that were missing key data elements for this analysis were also excluded, bringing the sample size down to 741 (26 more deletions). The final training data set includes 741 patients spanning 719 variables (columns). The testing data set includes 63 patients from the month of September 2015. After removing three DRG codes 461/2 and nine entries with missing data, the final testing data set contains 49 complete patient records. Table 3-1 below summarizes the data.

Table 3-1. Summary of Total Joint Replacement Surgery Patient Data

Original Data	Training Data	Testing Data
Dates	30/12/2013-30/08/2015	1/09/2015-30/09/2015
Sample Size	814	61
DRG 461	1	0
DRG462	46	3
DRG469	17	1
DRG470	750	57
Hip (81.51)	286	17
Knee (81.54)	528	44
Male	286	24
Female	528	37
Average BMI	33.21	35.17
Average LOS	65.21	56.12
Average Age	63.12	61.66
Cleaned Data	**Training Data**	**Testing Data**
Sample Size	741	49
Number removed due to DRG	47	3
Number removed due to incomplete data	26	9

3.2.2 Methodological Approach

This study investigates the use of a two-stage methodology that employs both unsupervised and supervised machine learning techniques applied to clinical patient data from the EHRs of total joint replacement surgery patients. The goal is to apply two different unsupervised learning (clustering) methods to a large subset of the patient data to identify unique clusters of patients. Once the clusters are identified, each patient's cluster label is appended to his/her record as a new variable. Supervised learning is then applied to the same training data set. The training data are further broken down into two subsets: a training set and a validation set. Four supervised learning (classification) methods are applied to the new training data (80% of data). The four methods include decision trees, random forests, bootstrap aggregation (bagging), and non-linear support vector machines. Each classification model is evaluated for accuracy in its ability to predict a patient cluster. The four models are compared based on their classification accuracy using the validation data.

3.3 Machine Learning Methods

3.3.1 Pillar I: Market Segmentation using Unsupervised Machine Learning

I applied both *k*-means clustering and hierarchical clustering to the patient data. *K*-means clustering can handle high-dimensional data and is generally computationally faster than hierarchical clustering with large data sets, but the user must specify a value of *k* up front. To determine the best value of *k*, I initially selected *k* such that the marginal change in the within-cluster sum of squares, as *k* increases, diminished. In hierarchical clustering, the data are clustered completely and the user must determine the best place to cut the resulting dendrogram to distinguish the number of clusters. The following subsections provide the details and theory on *k*-means and hierarchical clustering models.

3.3.1.1 K-means Clustering

K-means clustering is an unsupervised machine learning technique that separates *n* multidimensional observations into *k* distinct, non-overlapping clusters based on the similarity between the observation and the centroid of the cluster. The main idea behind *k*-means clustering is to find a good clustering where the within-cluster variation (i.e., dissimilarity) is as small as possible.

The algorithm starts by placing each of the *n* observations into exactly one of *k* pre-determined clusters. Let $C_1, C_2, \dots, C_k$ denote sets containing the indices of the observations in each cluster. The following two conditions must hold at every iteration of the algorithm: $C_1 \cup C_2 \cup \dots \cup C_K = \{1,2,\dots,n\}$ and $C_k \cap C_{k'} = \emptyset \quad \forall\, k \neq k'$ (James et al., 2013). That is, each observation *i* belongs to at least one of the *k* clusters, and the clusters are non-overlapping (i.e., no observations belongs to more than one cluster). After the initial placement of observations into one of the *k* clusters, a cluster centroid (i.e., the vector of the *p* feature means for the observations in the *k*th cluster) for each of the *k* clusters is computed, where the mean for feature *j* in cluster C_k is calculated in Equation 3-1:

$$\bar{x}_{kj} = \frac{1}{|C_k|} \sum_{i \in C_k} x_{ij} \;\; \forall\, k, j \tag{3-1}$$

where $|C_k|$ represents the number of observations in the *k*th cluster. In the next step of the algorithm each observation is assigned to the cluster whose centroid is the closest, which is equivalent to minimizing the total within-cluster variation; this objective function *Z* is denoted in Equation 3-2:

$$\min_{C_1, C_2, \dots, C_k} Z = \sum_{k=1}^{K} 2 \sum_{i \in C_k} \sum_{j=1}^{p} (x_{ij} - \bar{x}_{kj})^2 \tag{3-2}$$

The value of *Z* is guaranteed to decrease at each step.

At each iteration of *k*-means clustering, the algorithm first calculates the centroid of each cluster and then compares each observation to the centroid of each cluster as to minimize the total within-cluster variation. Note that observations are reassigned to clusters based on maximizing similarity between the observation and clusters centroid. The *k*-means clustering algorithm continues until some pre-determined convergence criterion is achieved. Convergence criteria could be based on iterations, when no more reassignments occur, or when no significant change in squared error from one iteration to the next (Jain et al., 1999).

K-means clustering is widely used due to its ease of use and ability to handle large data sets. The *k*-means clustering algorithm converges to a local optimum and is susceptible to initial starting conditions, which can prevent it from reaching a global minimum. It works best when multiple starting points are used.

3.3.1.2 Agglomerative (Hierarchical) Clustering

Agglomerative (hierarchical) clustering is another common approach to unsupervised machine learning. The method starts with a set of *n*-multidimensional observations which each comprise a cluster (of size one). Clusters are merged based on a similarity measure until there is one cluster. Unlike *k*-means clustering, there is no predetermined value of *k*. The user must determine an appropriate value of *k*. The output from hierarchical clustering is displayed in a dendrogram which "represents the nested grouping of patterns and similarity levels at which groupings change" (Jain et al., 1999).

3.3.1.3 Algorithm Input Matrix

As a prerequisite for clustering, a distance or similarity matrix is required. In the case of continuous variables, the predictor variables are standardized then clustered based on Euclidean distance in the case of the *k*-means algorithm or Ward's method in the case of hierarchical clustering. Ward's method is also known as the minimum variance criterion. This method is applied in hierarchical clustering algorithms where the objective is to minimize the total within-cluster variance. The algorithm starts with *n* clusters representing the *n* observations. Then, $n-1$ clusters are formed out of *n* clusters by combining the pair of observations that results in the smallest increase in within cluster variance. Ward's method uses a squared Euclidean distance measure to determine minimum variance (Ward, 1963).

When the variables are categorical or logical (binary) or a mix of types, using a standard distance measure such as Euclidean or Manhattan distance is not appropriate. As is the case in this patient data set, there is a mix of binary, categorical, and continuous variables which requires a method that first defines the similarity between patients. I apply Gower's coefficient to create a dissimilarity matrix which is then used in both *k*-means and hierarchical clustering. Gower's dissimilarity coefficient is a general similarity measure, S_{ij}, that Gower (1971) developed to determine similarity between two observations, i and j. This coefficient can be applied to ordinal, continuous, and dichotomous data. In determining Gower's coefficient, the similarity between two observations on the *k*th dimension are calculated for all *k* dimensions:

$$s_{ijk} = 1 - \frac{|x_{jk} - x_{ik}|}{R_k}, \tag{3-3}$$

where R_k is the range of *k*. The overall similarity coefficient is:

$$S_{ij} = \frac{\sum_{k=1}^{q} s_{ijk}}{\sum_{k=1}^{q} \delta_{ijk}} \ \forall\, i, j \ \text{ where } \delta_{ijk} = \begin{cases} 0 \ if\ there\ is\ a\ missing\ value\ in\ i\ or\ j \\ 1 \ \ otherwise. \end{cases} \tag{3-4}$$

3.3.2 Pillar II: Classification using Supervised Machine Learning

Classification modeling is a component of supervised machine learning that seeks to predict a response for an observation where the response falls into a particular class. At this point in the study, the first phase of the two-stage methodology is complete, where each patient in the training data set is now assigned to the cluster such that the total within-cluster variation is minimized. In order to assign out-of-sample patients to a particular segment, one potential approach could be to assign them to the closest cluster center. While this procedure would be relatively simple to implement, it lacks inherent flexibility and interpretability compared to traditional supervised machine learning models. Therefore, to predict patient segments in the second phase of the methodology, four classification models are tested, including decision trees, random forests, bagging, and non-linear support vector machines. The goal of the classification modeling is to identify the best method given the data that will accurately predict a patient's cluster using the validation data. Recall that the original training data are randomly split into two subsets: training (80%) and validation (20%). The training data are used to train each classification model and the validation data are used to evaluate each model against the others. The area under the (ROC) curve (AUC) is used to evaluate the four models. Note that an AUC of 0.5 indicates that a classifier performs no better than chance, where as an AUC of 1.0 perfectly classifies each observation.

3.3.2.1 Decision (classification) Trees

Decision trees, random forests, and bagging are all tree-based classification models, whereas support vector machines (SVM) seek to find a non-linear boundary between classes. In the tree-based methods, the predictor space is segmented into smaller regions through a series of rules that define a split at each level of the tree. Recursive binary splitting is used to grow the tree with one of three common criteria: classification error rate, Gini index, or cross-entropy, used for binary splits (James et al., 2013). The following three equations show each criterion:

Classification error rate: $E = 1 - \max_{k} \hat{p}_{mk}$ (3-5)

where $\hat{p}_{mk}$ is the proportion of observations from the *k*th class in the *m*th group

Gini index:

$$G = \sum_{k=1}^{K} \hat{p}_{mk}(1 - \hat{p}_{mk}) \tag{3-6}$$

Cross-entropy:

$$D = -\sum_{k=1}^{K} \hat{p}_{mk}\log(\hat{p}_{mk}) \tag{3-7}$$

Cross-entropy and Gini index are often referred to as node purity measures since values of $\hat{p}_{mk}$ near one or zero (observations predominately from a single class or pure) result in an index value near zero. This approach creates splits based on the one that maximizes the reduction in impurity (Cran, 2015) or classification error rate.

3.3.2.2 Bootstrap Aggregation (Bagging) and Random Forests

Although decision trees are easy to interpret, they commonly suffer from high variance. This can be offset using bootstrap aggregation and random forests. Both bagging and random forests are ensemble classifiers that bootstrap the training data set to create *Z* multiple training data sets. The average of the *Z* decision trees produce a classification model with lower variance than a single tree. Random forests proceed one step beyond bagging by regulating the number of predictors available for consideration at each split. Random forests overcomes the high correlation that can occur in bagging when a dominate predictor causes all subsequent bootstrapped trees to look similar (James et al., 2013).

In bagging (classification trees), *M* different training sets are generated by bootstrapping the original training data. Bootstrapping is random sampling with replacement. *M* decision tree prediction models, $\hat{f}^i(x)$, are developed, one for each bootstrapped data set i $(i \in 1,2, \ldots, \mathrm{M})$. For a given test set observation, the predicted class is determined based off the class most commonly returned by the *M* prediction models. In essence, bagging is an ensemble classifier that determines a classification based majority rules.

Random forests are a modification to the bagging method described above that reduces the correlation between bagged trees. A drawback of bagging occurs when one or two predictors dominate the rest resulting in very similar looking trees. Random forests deviate from bagging in that it restricts the number of available predictors that a given tree can use. For instance, if there are *p* dependent variables, bagging will build a tree considering *p* variables. In random forests, the user selects the number of variables (less than *p*) that are randomly selected for each bagged tree. Dominate variables will not be able to influence every bagged tree.

3.3.2.3 Support Vector Machines

Although tree-based methods can provide excellent classification results, they are not the only classification models. SVMs are a good classification approach especially when observations can be divided by a separating hyperplane. When this condition does not hold and the boundaries are non-linear, non-linear SVMs using kernels can accurately classify data.

SVM classification for multi-class problems starts with the understanding of how it applies to a two-class problem. In linear form, SVM seeks to find a maximum marginal classifier with decision rule, $D(x)$ such that if $D(x) > 0$, x is in class 1, and if $D(x) < 0$, x is in class 0. For an unknown sample, u, the maximal marginal classifier can be written in parametric form using a slope and intercept, as in Equation 3-8:

$$D(u) = \beta_0 + \sum_{j=1}^{P} B_j u_j = \beta_0 + \sum_{i=1}^{n} y_i \alpha_i x_i u; \qquad \alpha_i \geq 0 \tag{3-8}$$

This is performed in such a way that includes every point in the sample (Kuhn & Johnson, 2013).

In the case where the classes are not separable, the set of positive $\alpha's$ define points on the boundary of the margin. These points define the support vectors and the prediction equation (Kuhn & Johnson, 2013). If the boundary between classes is not linear, then a kernel function, *K*, is substituted as shown in Equation 3-9:

$$D(u) = \beta_0 + \sum_{i=1}^{n} y_i \alpha_i K(x_i u); \quad \alpha_i \geq 0 \tag{3-9}$$

A kernel is a function that describes the similarity between observations (James et al., 2013). In non-linear SVMs, there are several kernel functions to include polynomial, radial, hyperbolic tangent, and others (Kuhn & Johnson, 2013). By adjusting parameters within each function, the SVMs can be tuned to prevent over-fitting. In multiclass problems where *C*, the number of classes, is greater than two, there are several methods to extend the binary classifier such as one-verse-all and one-verses-one. In the prior, *C* SVMs are fit with each one considering two classes: a class *c* and a class consisting of all other classes. Test observations are assigned a classification based on the largest fitted value from the *C* SVMs evaluated at each observation. In one-verse-one, $\binom{K}{2}$ SVMs are created with each one comparing a pair of classes. Each test observation is evaluated by each SVM and the test observation is assigned a classification based on the most frequently occurring class (James et al., 2013).

3.4 Results and Discussion

3.4.1 Descriptive Statistics of the Patient Population

Before discussing the results obtained from applying unsupervised and supervised machine learning methods to the patient data, I first describe the EHR data elements to understand better the study population. Table 3-2 provides descriptive statistics of several patient characteristics, Hip/Knee Disability and Osteoarthritis Outcome Scores (HOOS/KOOS), and the most prevalent medical and disease diagnoses found in the patient data.

Table 3-2. Description of the Study Population

Clinical Data Overview		
Dates: 30/12/2013-30/08/2015		
Prevalence (or absence) of clinical conditions	**#**	**%**
Obesity (BMI>30 by ICD9 code)	544	66.8%
Hypertension	482	59.2%
Never smoked	457	56.1%
Osteoarthritis lower leg	450	55.3%
Esophageal Reflux	369	45.3%
Anemia acute (blood loss)	222	27.3%
Osteoarthritis pelvis	214	26.3%
High Cholesterol	148	18.2%
Obstructive Sleep Apnea	144	17.7%
Hypothyroid	123	15.1%
Depression disorder	118	14.5%
Diabetes	117	14.4%
Prior TJA	105	12.9%
Long term aspirin use	105	12.9%
Anemia	93	11.4%
Current Smoker	92	11.3%
Asthma	89	10.9%
Anxiety State	80	9.8%
Coronary Atherosclerosis of native coronary artery	59	7.2%
Retention of Urine	58	7.1%
Hip and Knee Scores (HOOS/KOOS) from patient survey	**Mean**	**StdDev**
Pain	29.99	24.46
Symptom	32.86	26.00
Function Daily Living	30.67	27.09
Function Sports, Recreational Activities	15.47	22.20
Quality of Life	16.17	18.03
Patient Characteristics	**Mean**	**StdDev**
Age	63.122	11.40
BMI	33.21	7.06

3.4.2 Results of Pilar I: Market Segmentation using Unsupervised Machine Learning

Given the number of available predictors, I first experimented with dimension reduction techniques in order to find the right combination of predictors that could accurately represent the key relationships between patients while still being interpretable. Of the 724 available variables, the majority of them are binary (over 650 are binary) and, of the ICD-9 disease codes, only 20 of them occurred with frequency greater than 55 (6%). To reduce the size of the initial patient data set, I combined several of the weight-related variables into an obesity variable. A patient is obese if either his/her BMI is greater than 30 or he/she is diagnosed as obese or morbidly obese (ICD-9 codes 278.0 and 278.01). Since there were five patients without a BMI score, creating a new obesity variable allowed us to keep those patients in the data set. I also combined four of the asthma related ICD-9 codes into one binary variable for asthma.

The EHR data had redundant fields for smoking and tobacco use, so I combined three into two: patient never smoked and patient is a current smoker. The two separate ICD-9 codes for prior knee and hip replacement were combined into one binary variable: prior joint replacement. Finally, I excluded all remaining ICD-9 codes that occurred with a frequency less than 55 times and removed the patients with missing HOOS/KOOS scores. The final data set included 29 variables and 746 observations (741 if you exclude patients without a BMI score).

I ran both *k*-means and hierarchical clustering using Ward's method on two data sets. All methods were performed in R 3.2.0 using the package "stats" version 3.3.0 (R-Core Team, 2016). Both data sets contained 746 patients. The first had all 29 variables discussed above and the second was a subset that only contained the continuous variables (age, BMI, LOS, and five HOOS/KOOS scores). Since the larger data set contained mixed variables, I converted the matrix of 29 variables into a dissimilarity matrix using Gower's coefficient. The best value of k that maximized average silhouette distance was $k = 2$. The resulting two clusters were different but not significantly different in factors such as age, BMI, or HOOS/KOOS scores. I then applied the hierarchical clustering algorithm using Ward's distance. I cut the dendrogram at a height that gave six clusters. The clusters are fairly balanced and have some separation between key demographic and HOOS/KOOS scores. Figure 3-1 shows the dendrogram with $k = 6$.

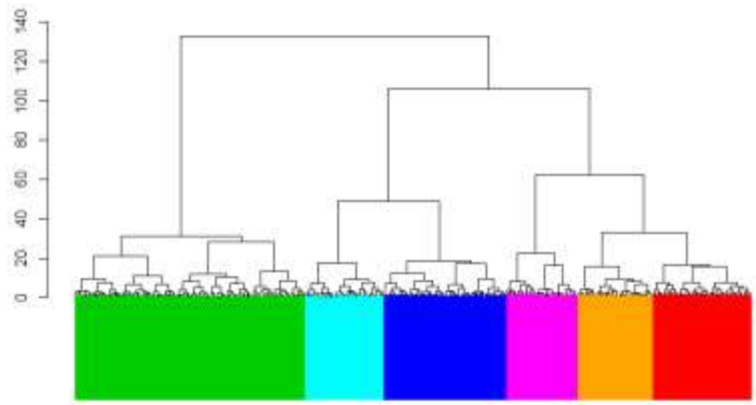


Figure 3-1. Dendrogram of Patient Data using Ward's Method with *k* = 6

With the exception of cluster 2 (green), the clusters are evenly balanced. Table 3-3 shows the summary of each cluster. LOS is measured in hours and represents the mean time a patient in a specific cluster spends in the hospital from admission to discharge. Pain, symptoms, function daily life, function sports recreation, and quality of life represent the six hip and knee osteoarthritis outcome score category averages for patients in each cluster. Scores can range from 0-100 with 0 being the most severe. The mean age per cluster ranges between 61 and 65. Every cluster has a high percentage of obese patients and clusters 1 and 2 are likely total knee patients given their high percentage with osteoarthritis of the lower leg, whereas clusters 3, 4 and 5 are likely comprised of hip patients as osteoarthritis of the pelvis is a prevalent medical condition. Also, clusters 2 and 5 contain the patients with the most painful conditions just prior to surgery as their HOOS/KOOS scores are very low (low score indicates significant deviation from normal lifestyle).

Table 3-3. Summary Data from Hierarchical Clustering using Ward's Method

Cluster	Freq.	Age	BMI	LOS	Pain	Symptoms	Function Daily Living	Function Sports Rec.	Quality of Life	Most Prevalent Medical Condition
1	109	65.62	32.61	63.36	46.03	48.98	47.14	27.28	28.23	Hypertension (50%) Reflux (51%) Osteoarthritis: lower leg (86%) Obese (57%)
2	254	63.54	34.64	63.59	36.45	39.75	37.48	19.15	19.65	Never smoked (72%) Not smoked in >1 year (98%) Obese (79%) Osteoarthritis: lower leg (88%) Reflux (44%) Hypertension (57%)
3	135	61.79	30.90	69.73	8.84	12.01	7.85	2.99	2.87	Obese (53%) Osteoarthritis: pelvis (63%) Hypertension (46%)
4	87	63.32	33.87	62.46	38.67	41.71	40.30	17.18	19.33	Not smoked in >1 year (98%) Never smoked (68%) Obese (76%) Osteoarthritis: pelvis (57%) Hypertension (66%)
5	79	63.19	30.92	64.24	35.18	38.61	36.30	19.78	21.52	Obese (48%) Reflux (47%) Osteoarthritis: pelvis (89%) Hypertension (66%)
6	82	61.51	34.69	64.83	8.44	8.72	8.91	2.36	2.52	Obese (76%) Reflux (62%) Hypertension (67%) Sleep Apnea (41%)

Although the results from the hierarchical clustering were distinguishable, there was not significant distance between age, BMI, and pain scores. As an alternative method, I also conducted clustering on a reduced variable matrix where I clustered on just 8 variables (age, BMI, length of stay, and the five HOOS/KOOS measures). I applied *k*-means clustering to the reduced data set and evaluated the within-group sum of squares over a range of *k* values. Figure 3-2 shows the results.

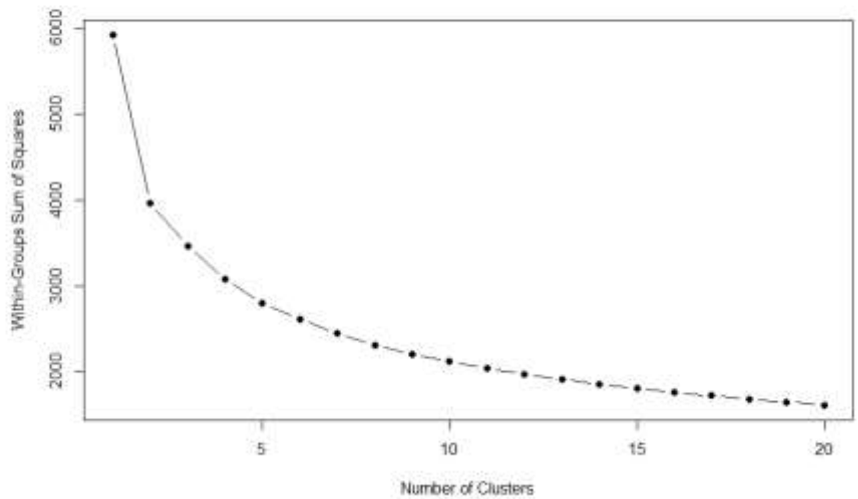


Figure 3-2. Within-Group Sum of Square Errors for *k*-means Clustering using Reduced Data Set

Based on Figure 3-2, I selected $k = 6$ despite the significant decrease in the marginal reduction in within-group sum of squares seen at $k = 2$. The results of the clustering are shown in Table 3-4. Again, I present six clusters. This time, I did not use any binary variables and were, therefore, able to use Euclidean distance versus Gower's dissimilarity measure. By using eight scaled continuous variables, six distinct clusters are presented below.

Table 3-4. Results of *K*-means Clustering on Reduced Data Set

Cluster	Freq.	Age	BMI	LOS	Pain	Symptoms	Function Daily Life	Function Sports & Rec	Quality of Life
1	84	60.14	32.58	58.10	59.08	56.45	63.83	63.68	40.25
2	124	73.06	29.25	65.35	6.66	9.75	5.43	1.61	1.92
3	130	53.75	33.78	55.61	3.50	5.29	4.31	0.98	1.11
4	43	67.67	33.74	136.86	14.11	18.79	11.32	6.57	8.58
5	200	68.96	31.12	61.14	48.38	54.07	48.22	18.83	29.94
6	160	56.63	38.52	59.82	35.57	38.10	37.51	10.87	11.60

Cluster 4 has the least number of patients at 43, but it also has the longest average LOS. This is a significant cluster for clinicians to understand as these patients' average age is within the coverage of Medicare, and they require the most resources of all the patients who have a joint replacement. Under value-based care policies such as the new comprehensive joint replacement program (CJR) by the Center for Medicare and Medicaid Services, hospitals will receive a fixed reimbursement for each patient regardless of the number of days they spend in the hospital. Cluster 4 patients also have relatively low HOOS/KOOS scores, indicating that their quality of life is severely impacted by their joint pain. Cluster 3 and cluster 6 patients are the youngest of the patients but differ in their BMI. Cluster 6 patients have an average BMI nearing morbid obesity. Cluster 3 patients are the youngest, obese on average, and have the lowest HOOS/KOOS scores. These are likely middle-aged men and women who suffer from joint pain but are also still relatively young and value their mobility. Their low length of stay indicates that they are motivated to return to work and their families. With the average age below 65, Cluster 3 patients may also be conscience of their stay in the hospital, as they likely have private insurance and high copays. Cluster 2 and Cluster 5 patients contain the oldest patients with average ages of 73 and 69 years old, respectively. These two clusters differ in that the older patients (Cluster 2) are in more pain and discomfort.

Table 3-5 shows the results of hierarchical clustering using Ward's method on the reduced data set. The accompanying dendrogram is in Figure 3-3. The results are similar to the output from the *k*-means clustering output in Table 3-4. Hierarchical clustering also found a small size cluster (6) with relatively low HOOS/KOOS scores and a high length of stay. Cluster 6 in Table 3-5 differs from cluster 4 in Table 3-4 by age, whereas the average age in the latter table is slightly lower. Cluster 1 and Cluster 2 in Table 3-5 correspond to Cluster 6 and Cluster 2 in Table 3-4, respectively. With a small data set and few predictor variables, the results from *k*-means and hierarchical clustering methods are similar; however, the *k*-means results have slightly better distribution amongst the clusters. Additionally, the computer processing time for both hierarchical clustering and *k*-means was less than 1.2 seconds when run on a Hewlett-Packard laptop with an Intel Core i7 processor. With a small patient data set, computational time is not a significant factor.

Table 3-5. Results of Hierarchical Clustering using Ward's Method

Cluster	**Freq.**	**Age**	**BMI**	**LOS**	**Pain**	**Symptoms**	**Function Daily Life**	**Function Sports & Rec**	**Quality of Life**
1	185	59.48	36.75	61.25	33.83	35.33	34.85	8.89	10.38
2	76	77.03	27.57	68.29	9.21	17.86	8.26	1.68	3.62
3	231	64.37	31.86	57.79	49.48	53.35	50.38	26.06	31.48
4	163	58.63	33.22	58.44	1.86	3.36	2.08	0.41	1.11
5	41	65.34	33.53	57.34	64.38	63.40	69.80	75.46	46.36
6	45	64.38	34.48	136.6	19.73	22.00	18.07	11.39	9.59

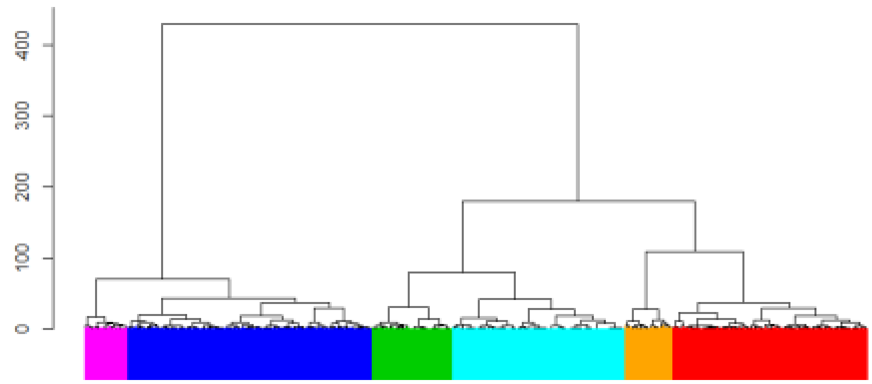


Figure 3-3. Dendrogram of Hierarchical Clustering on Small Data Set

3.4.3 Results of Pillar II: Accurate Classification of Segments using Supervised Machine Learning

There are many supervised machine learning models readily available and each has its own strengths. I chose four, three of which are tree-based classification models and common in the literature. Since there is no one best model for all data sets, I tested several and selected the one with the best AUC. As mentioned previously, AUC is the area under the receiver operator curve which is a plot that traces the performance of a classifier as the threshold value is changed. In other words, ROC is a plot of the true positive rate (y-axis) versus the false positive rate (x-axis) at different thresholds. An ideal predictor, or one that can perfectly separate the classes of a response variable, has an AUC of one, whereas a classifer that is no better than random guessing would have a AUC around 0.5. The ideal classifier hugs the upper left corner of the unit square. AUCs greater than 0.8 are generally considered good classifiers. I also divided the data into a training set and testing set. As in regression modeling, creating training, validation, and test data is an important step to prevent the problem over-fitting.

After performing unsupervised machine learning to determine homogeneous subgroups of patients within the TJA patient population, I proceeded with the two-stage methodology by appending the generated cluster label to each patient observation to then apply supervised machine leraning. As mentioned previously, I built four classification models namely, decision trees, random forests, bagging, and support vector machines for predicting which segment a new patient observation would fall in. Figure 3-4 shows a plot of the one-versus-all ROC curves derived from the four classification models that were trained on 80% (randomly assigned) of the original training data, then validated on the remaining 20% of the training data. All models performed very well, with the decision trees having the lowest multi-class AUC at 83.2% and random forests having the highest at 92.8%. Note that I improved the decision tree fit with a pruned version via cross-validation using a miss-classification scoring method. Table 3-6 provides the multi-class AUCs derived from the four classification models. Note that these same type of classification models can then be modified to predict other patient health outcomes, such as patient readmission risk, length of stay, and medical supply costs. Here, the patients' healthcare market segment identifier becomes a new predictor in the prediction model.

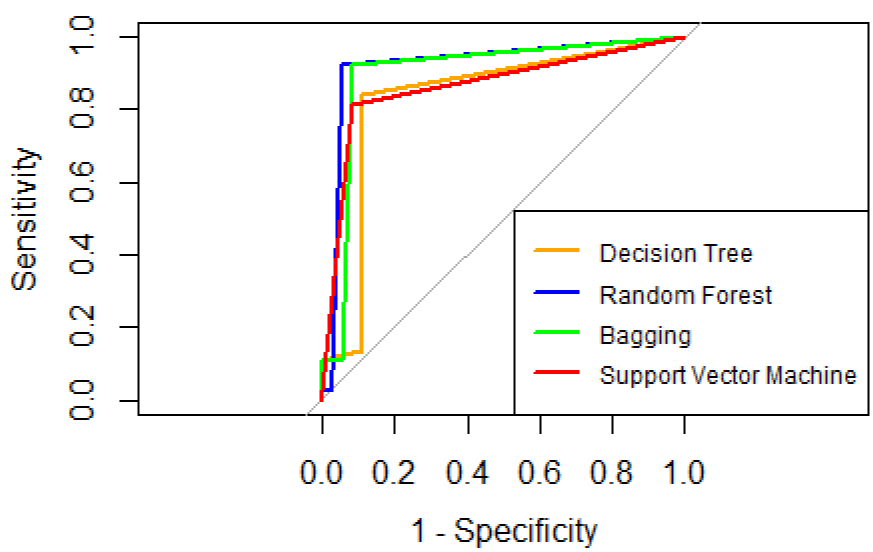


Figure 3-4. One-versus-all ROC Curves for Four Classification Models

Table 3-6. Multi-class AUC by Classification Model

Model	Decision Tree	Random Forest	Bagging	Support Vector Machine
AUC	0.8322	0.9278	0.9006	0.8673

Although applying supervised learning to predict a patient's healthcare market segment after performing *k*-means clustering may seem redundant, supervised learning methods, namely decision trees, provide a useful advantage in healthcare. Decision trees, such as the one shown in Figure 3-5, are user-friendly and help improve the interpretabilty and flexibility of the predictive model. Since decision trees are based on the actual predictor variables, clinicians can easily determine a patient's cluster by simply following the tree branches until the classification is evident. This does not require calculating distances between cluster centers and new patients. Additionally, since some patients have complex medical conditions and require care from multiple clinicians in a single visit, decision support tools must be easy-to-use and universally understood.

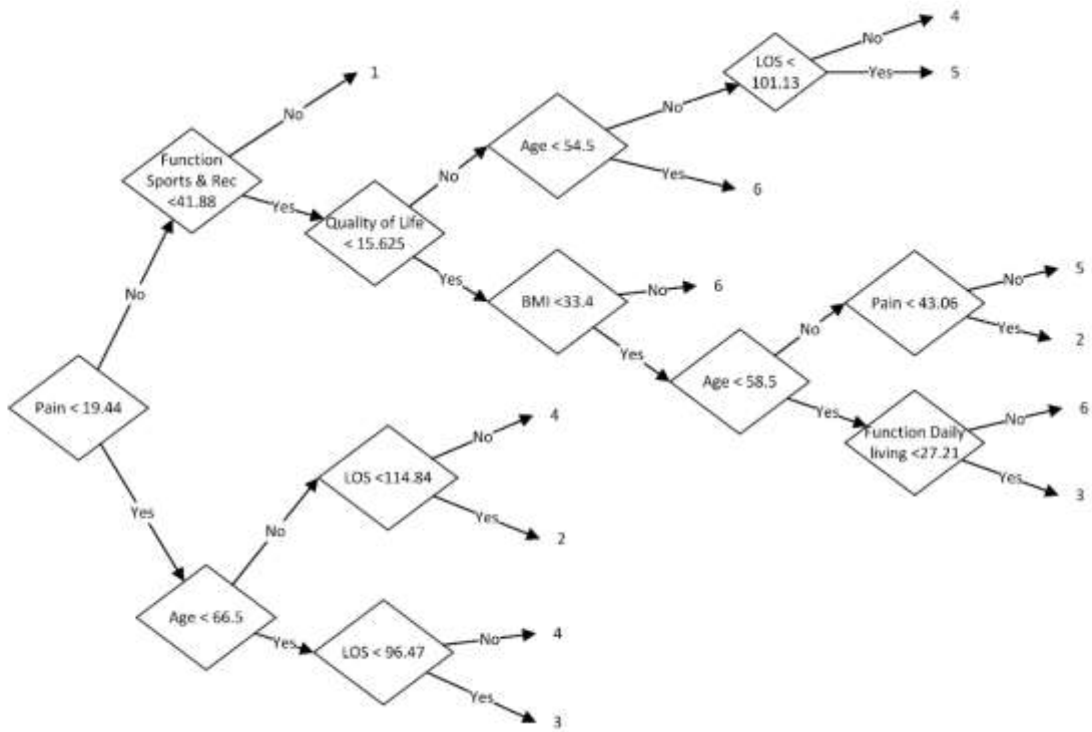


Figure 3-5. Decision Tree for Total Joint Replacement Surgery Patient

The decision tree in Figure 3-5 above provides a graphical representation of the classification model. For example, if a new patient is recovering from a total knee replacement surgery and his/her pre-surgery KOOS pain score is greater than 19.44, HOOS sports and recreation activity score is less than 41.88, HOOS quality of life score is less than 15.63, and BMI is greater than 33.4, then the patient is assigned to cluster 6.

During the clustering phase, all observations are assigned to the nearest cluster center. Since I am using the reduced data set, the closeness measure is based on Euclidean distance. In the classification phase, every patient's cluster assignment number is appended to their record and used as the response variable. Up to this point, the same data is used for both classification modeling and clustering.

To test the effectiveness of this two-stage methodology, I acquired a new data set (from September 2015), assigned every patient to the closest cluster center, and then predicted the cluster using the random forests model (since it had the largest AUC value compared to the other classification models). Table 3-7 below shows the results. The predicted cluster from the random forests model are on the left side, and the assigned clusters based on distance to nearest cluster center are across the top. Nine patients were misclassified out of 49, resulting in an accuracy rate of 81.6%.

Table 3-7. Predicted versus Assigned Segment for Testing Data

Predicted\Assigned	**1**	**2**	**3**	**4**	**5**	**6**
1	**11**	2	2	0	0	0
2	1	**14**	1	0	1	0
3	0	0	**3**	0	0	0
4	0	1	0	**1**	0	1
5	0	0	0	0	**2**	0
6	0	0	0	0	0	**9**

In this research, I demonstrated that EHR data is suitable for clustering total joint replacement patients into distinct clusters that describe the likeness of their attributes. This information alone is applicable to the healthcare industry, as it identifies patients' segments that may require post-operative interventions to improve discharge times. In Table 3-4, I show that patients in cluster 4 have an average BMI around 33, are on average Medicare patients (over 65), and have low HOOS/KOOS scores. Although this cluster represents only 6% of the total population, they spend over twice as long in the hospital. In an era of hospital overcrowding and ever-increasing healthcare costs, identifying cluster 4 patients during pre-op enables the hospital to use lower-cost interventions to speed discharge.

This study also determined that there are two large segments of the population (Cluster 3 and 6) that contain younger patients who are likely still working full-time, have private insurance, and want to get back to normal activities. The clusters differ in the degree of obesity and pain/discomfort levels, but they present an opportunity for healthcare organizations and health insurers as they look to increase preventative care and focus on wellness and value-based initiatives. The next step in analyzing these clusters is to research patient insurance types, frequency of routine care appointments, and other touchpoints that healthcare providers have with these patients.

After clustering each patient, I showed that classification models can accurately predict a patient's segment with the original data and with new data. Applying the best classification model (random forests) to a new 30-day sample of patient data accurately predicted a patient's cluster better than 80% of the time. This implies that a new patient's healthcare market segment may accurately predicted without performing clustering. A limitation, however, may stem from the small test size.

Future healthcare market segmentation research should include larger datasets that span multiple service lines within the hospital. Additionally, merging cluster information with patient attributes holds great potential for predicting various outcomes such as patient length of stay, discharge location, and expected costs of treatment. In an era of reduced reimbursements and increased demand for services, predicting patient outcomes offers insights to clinical decision makers that can inform pre- and perioperative healthcare decisions. With larger data sets comes the potential for longer computational time. As shown in Table 3-8, the methods used in this study were easily managed on a standard laptop computer. With big patient data, however, processing times will increase and could quickly become a discriminating factor in determining which methods to use.

Table 3-8. Processing Time for all Machine Learning Methods Applied to Reduced Data Set

Method	**Computational Time (sec)**
K-means Clustering	1.09
Hierarchical Clustering	0.03
Decision Trees	0.01
Bagging	0.44
Random Forests	0.42
Support Vector Machines	0.11

3.5 Conclusions

The advent and widespread use of the EHR not only enables clinicians to access patient data in real-time, but also provides the data warehouse to support healthcare market segmentation efforts. Through secondary use of EHR data, healthcare organizations can efficiently target specific segments of their populations without requiring additional patient information. Using these two complementary data-driven analytical approaches in total joint arthroplasty empowers patients, providers, and healthcare marketers with valuable healthcare market segment information. Clinicians can use this information to adopt specific protocols during the pre-, peri-, and post-operative phase of TJA. Health insurers can use this information to isolate high costs and high-risk patients and then incentivize them and their primary care providers to take appropriate, patient-customized interventions to reduce health risks. Patients who are interested in their health outcomes can learn from their segment what actions to take now to delay or prevent common ailments.

This study is small in scope and focuses on one specific surgical procedure. Healthcare market segmentation holds great potential if it can be applied to large patient populations without the need to conduct time-consuming surveys, and if it can provide detailed demographic, geographic, and clinical information regarding each segment. Additionally, exploiting the classification models for new patients and helping predict health outcomes for existing patients is a powerful tool for clinical decision support.

The two-stage data mining approach presented in this chapter comprise the first two pillars of the healthcare market segmentation methodology presented in Figure 1-2. This approach demonstrates the efficacy of using data analytics to increase knowledge about a specific patient population, which in turn can inform decision making. Furthermore, the results of this chapter show that the two separate, but related data mining methods, namely unsupervised and supervised machine learning, show congruence in their ability to predict a patient's segment. Applying the predictive classification model is computationally efficient and, when using a decision tree, provides an intuitive visual aid for clinicians. A healthcare extension to this research entails incorporating the predictive modeling directly into the EHR to provide clinicians with real-time access to a patient's cluster. Ideally, hospitals will integrate the real-time decision support with clinical, procedural, and administrative interventions to achieve the best health value for each patient. I believe that computationally efficient predictive models could easily run in the background of the EHR and update when new information or a new patient is added. Although larger patient databases would increase computation time to generate the initial clusters and predictive models, the application time of the predictive models to each patient should remain constant.

Chapter 4

Identifying Cost and Outcome Drivers

4.1 Introduction

Reducing the financial gap created by CJR requires hospitals to take a new approach to understanding the complexities surrounding the transformation from volume- to value-base care. The healthcare market segmentation methodology introduced in Chapter 1 is a key component that can help hospitals transform. The gaps identified in the literature specifically regarding data mining approaches applied to healthcare were advanced in Chapter 3. Chapter 3 presented two of the fundamental data mining pillars of the framework: market segmentation using unsupervised statistical learning and accurate classification using supervised statistical learning. This chapter advances the third and fourth pillars of the methodology which incorporate the identification of cost and outcome drivers. The Venn diagram shown in Figure 4-1 highlights the two pillars and several of the traditional cost and outcome drivers. For hospitals at risk under bundled payments, discharge location and readmission are both outcome drivers and cost drivers. They are related in that discharge location can impact the cost of post-acute care and readmission rate and these two drivers significantly impact the financial risk that hospitals have under the CJR.

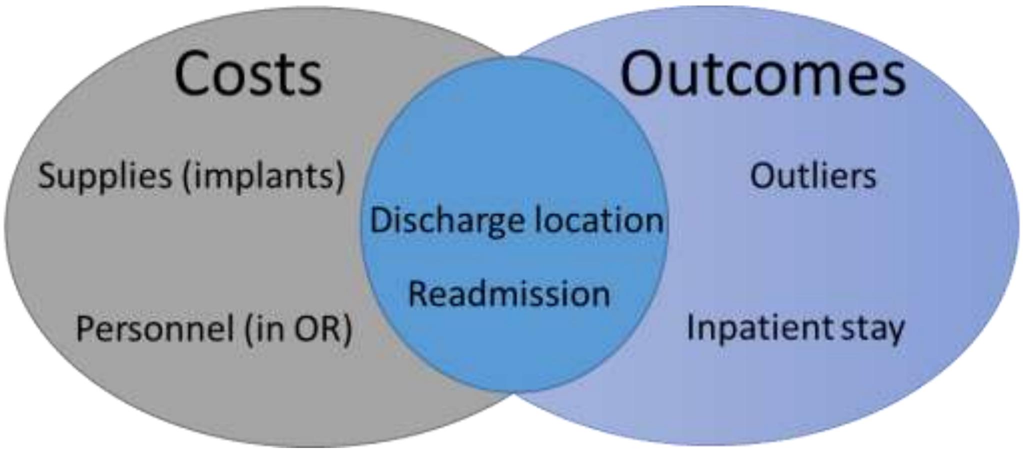


Figure 4-1. Venn Diagram of Cost and Outcome Drivers

The third pillar, identifying cost drivers, builds on the healthcare market segmentation methodology by using a regression-based approach to model cost drivers in total joint arthroplasty. There are several traditional cost drivers in total joint arthroplasty, as noted in the literature, to include supply cost, operating room cost, and inpatient stay. Under the CJR bundle, post discharge care management costs are added to the list of expenses covered under the single payment. These costs include those associated with preventable readmissions and post-acute care providers such as skilled nursing facilities and home healthcare. Depending on where patients go after discharge and how well they are able to recovery, post discharge costs can be significant (Ramos et al., 2014).

The regression models in pillar III of the methodology will help identify and explain which attributes (patient, process, and outcome) in the data contribute to higher costs. The models will assess the impact of these data on peri-operative supply costs and readmission costs. This data driven approach complements the first two data mining pillars of the methodology and help hospitals focus their value, thereby improving interventions.

The fourth pillar of the healthcare market segmentation methodology concerns the identification of outcome drivers in total joint arthroplasty. Similar to pillar III, this research aim takes a data analytics approach to understand which factors drive patient and cost outcomes. Patient outcomes can include specific readmission diagnoses, discharge locations, and lengths of stay. Outlier analysis applied to patients that have high supply costs can identify and isolate specific factors that lead to higher resource utilization.

Two case studies based off pillars III and IV are applied in this chapter to help hospitals subject to the CJR better understand cost and outcome drivers. The first investigates readmission and supply cost based on patient attributes and is grounded in linear and logistic regression modeling. The second study looks at outliers, with a focus on total knee arthroplasty cost outliers. Improving health value for the patient and hospital requires improving the outcome per cost of surgery. Despite the complexity of TJA, focusing on high supply cost outliers can help hospitals focus interventions on a small cluster of patients that disproportionately consume the narrow margins hospitals operate on under the bundle.

In a case study of total knee arthroplasty patients, individual or small groups of high supply cost patients might otherwise be overlooked when in fact their unique cases significantly drive negative outcomes. Some hospitals have focused on optimizing implant pricing for primary total knee arthroplasty (TKA) but receive smaller discounts for high-priced implants used in revision surgery and for some complex primary TKAs. I hypothesize that the use of revision implants in a small group of DRG 469/470 TKA patients significantly increases total costs and average implant cost for the entire group. Revision implant use in a patient who is not scheduled for a revision surgery is a surgeon's decision and that decision may not be evident prior to surgery. Better understanding of the root causes for patients requiring high cost implants, especially at the data level, presents an opportunity to achieve better preoperative understanding of patient requirements and costs. This can help surgeons plan for and deliver better health outcomes.

4.2 Understanding Patient Readmission Risk and Supply Cost

Cost drivers and outcome drivers in healthcare are closely related and are best analyzed together. The aim of this study is to help surgeons and hospitals understand the important patient-level factors that influence bundled payment costs in the context of readmission risk and operating room supply costs for TJA patients. That information can inform surgeon decisions about what interventions to take to mitigate the risk of readmission post-discharge. Additionally, preoperative knowledge of the patient factors that increase supply costs can help physicians and hospitals manage expenses and develop cost saving contracts with vendors. Success in the CJR or similar bundled payment programs requires that all stakeholders including surgeons understand the "actual direct costs of providing their service" (Schutzer, 2016).

Surgical supply costs consisting primarily of implant costs are significant cost drivers in TJA. Depending on the method used to calculate the cost of TJA, either traditional accounting or time-driven activity based costing (TDABC), implants are one of the top three expenses (Akhavan et al., 2015; Bosco et al, 2014). The other two being hospital room and operating room. A fourth cost driver for CJR patients is readmission, which has been shown to account for 60% or more of the cost of the primary surgery (Peel et al., 2015). Prior to the CJR, hospitals were reimbursed for most if not all readmissions following TJA (Jencks et al., 2009). Doctors and hospitals did not focus on post-operative care, as rework was reimbursed under a fee-for-service payment model (Jencks et al., 2009). Since the CJR is a relatively new initiative, there are few studies that have looked at the readmission costs out to 90 days post discharge. In a recent study of primary TJA patients, the 30-day readmission rate was 5.5% and accounted for 11.2% of the post discharge payments (Bozic et al., 2013).

In a bundled payment model, hospital revenue is fixed-per-case and, therefore, reducing expenses without sacrificing quality improves health value. One expense that hospitals can measure and influence is the fixed costs associated with purchasing supplies in the operating room. The majority of the supplies are expendable or implant-related and are purchased at point-of-use. A second expense (and one that is harder to control and carries significant financial impacts) is readmission risk. Understanding patient risk factors that lead to high supply cost and high readmission risk can help surgeons in two ways: 1) to balance their surgical case load, and 2) to predict higher cost patients for which they should invest in interventions that will reduce readmission risk.

4.2.1 Background

There is a large body of literature on predictive modeling in healthcare especially in terms of readmission risk. A recent systematic review of predictive modeling in readmission found many of the published risk prediction models have mixed results and are not generalizable across departments, specialties, and patient populations. This may indicate that the best model is one tailored to a specific patient population (Kansagara, 2011). Furthermore, despite advances in computer-aided statistical analysis programs, risk prediction models that used patient-level factors were better at predicting mortality than readmission (Kansagara, 2011). A significant gap in readmission models could be attributed to the lack of hospital and health system factors such as number of care coordinators, number of follow up visits, effectiveness of medication reconciliation, and bed availability (Kansagara, 2011). Another complicating factor to understanding readmission data is the lack of information on the actual number of readmissions (Fry et al., 2015). Hospitals traditionally cannot get information on patient readmission visits to other hospitals; in the case of CMS patients, that information may not come until it is too late to intervene.

Since the passing of the ACA in 2010, the CMS started to publish 30-day readmission rates for select diagnosis related groups (DRGs) which in turn encourages hospitals to focus on 30-day readmission rates. In a recent meta-analysis studying 30-day orthopedic readmissions, the authors found that studies that began enrollment after 2009 had lower readmission rates than ones that started before 2009 (Bernatz et al., 2015). The study attributes some of the improvement in overall readmission rates to the Hawthorne effect, where merely drawing attention to readmissions has improved outcomes.

At the individual study level, certain risk factors have been identified with higher readmission risk. Age, hospital length of stay, discharge location, BMI, and American Society of Anesthesiologists (ASA) classification were positively associated with increased risk of 30-day readmission following orthopedic-related surgery (Bernatz et al., 2015). Other studies found that comorbidity burden, age, and prior medical service use are useful predictors of readmission (Kansagara, 2011). In a total knee arthroplasty study, the authors found that a history of transient ischemic attack/cerebrovascular accident, female sex, and general anesthesia were significant predictors of readmission (Belmont et al., 2015). Obesity has also been tied to higher readmission risk following total knee arthroplasty revision surgery (Hanly et al., 2016). In total hip arthroplasty, an increase risk of readmission has been attributed to procedure type, length of stay greater than 5 days, cardiac valvular disease, substance abuse, and diabetes with end organ complications (Schairer et al., 2014).

Although less studied than readmission, there have been several studies that investigate supply costs in the TJA operating room. These studies focus on the rising cost of implants and the reduction in margins for both primary and revision joint surgery as opposed to the patient factors that impact supply costs. The Lahey Clinic instituted cost containment methods to reduce TKA expenses (Healy et al., 2011). Bosco et al., (2014) successfully managed variability in implant costs by implementing price points that were not negotiable; thus forcing implant vendors to a fixed cost (Bosco et al., 2014). Gioe et al. (2011) studied premium versus standard implants and found no significant difference in revision rates at 7-8 years. The foremost mentioned studies did not assess how patient factors influenced surgical supply costs or contributed to the variation in supply costs across similar procedures. Two other inpatient cost studies found that certain comorbidities, such as diabetes and obesity, increase inpatient hospital costs, but neither investigated which components of cost were impacted (Pugely et al., 2014; Kremers et al., 2014). Understanding the patient factors that increase the cost of supplies could help surgeons and hospitals better manage the direct surgical costs of TJA.

4.2.2 Methods

4.2.2.1 Study Setting and Design

This case study is based on a TJA patient population located at the PSHMC. This retroactive study looks at patients who underwent TJA between December 2013 and September 2015. The data for this study were approved under PSU IRB STUDY00054. The initial data set included all DRG 469 and 470 patients who underwent total joint replacement surgery of the knee or hip joint. Patients who had bilateral procedures were excluded. For the majority of the study period there were three primary TJA surgeons. All patients who presented for surgery between December 30, 2013 and September 30, 2015 were initially included. There were 248 patient records excluded due to missing or incomplete data fields. The final set included 596 patients.

This study focuses on understanding the patient-level factors that are significant in determining surgical supply costs and readmission risk, which are two critical components to understanding and managing financial risk under value-based bundled payment models. At the time of this study, the hospital was not under the CJR. Additionally, 48% of the patients were eligible for Medicare (aged 65 or older).

I employed regression modeling techniques to understand the impact of various predictors, most of which are binary or categorical variables, on supply costs and readmission. To model supply cost, a continuous variable, I used linear regression, and to model readmission, a binary response variable, I used logistic regression. Each patient has a supply cost which consists of implant and non-implant costs and includes the cost of every accountable medical consumable used in the operating room. These include implant components, drapes, knife blades, etc. The readmission risk model is in two parts: 30-day readmissions and 90-day readmissions.

The CJR does recognize a limited number of readmission exemptions for which it does not hold the hospital financially accountable. The readmission exemption codes are listed by DRG of primary discharge diagnosis for readmission visit. Although the data did not contain the DRG for readmission, each readmission reason was compared to the exemption list. All but 14 of the 90 day readmissions would have been subject to the CJR; those not covered included cases where a patient readmitted for an arthroplasty on a different joint than the index admission.

4.2.2.2 Data Collection / Variable Selection

The patient data for this study came from several sources to include patient EHRs, billing records, and Surginet – the hospital data collection tool used in the operating room. The patient data set contains over 30 attributes to include age, body mass index, length of stay, hip and knee osteoarthritis outcome score (HOOS/KOOS), smoker status, ethnicity, discharge location, and ICD-9 diagnosis codes. Dummy variables were used to represent the multiple levels of many of the attributes such as ASA classification, smoking status, and discharge destination.

After data pre-processing, I calculated two primary comorbidity-based indices for use in modeling readmission risk and supply cost. A comorbidity index is a numeric score derived from the combination of documented patient comorbidities (Bjorgul et al., 2010). I calculated the Charlson Index and the count of Elixhauser's ICD-9 mappings to comorbidities. The Charlson Index was originally developed in 1987 to help classify comorbidity conditions that lead to mortality (Charlson et al., 1987). The index uses 19 comorbidity conditions that were selected based off their association with a one year mortality risk. Each comorbidity is weighted from one to six and the index, which ranges from 0 to 33, is the sum of the weighted comorbidities. The Charlson Index has been used to predict patient outcomes including mortality, readmission rates, and complications. For example, a Canadian study found that patients with a Charlson index of 2 or more increased both readmission and in-hospital complication risk from TKA by 2.1 percent (Kreder et al., 2003). A Danish study also found that high Charlson scores (>2) were strong indicators of post discharge hip replacement failure (Johnson et al., 2006). Although the Charlson index demonstrated success in some studies, it does have limitations. In an Australian study, the Charlson index did not perform well as a predictor of health-related quality of life outcomes following joint replacement surgery (Harse et al., 2005).

The second comorbidity index is the count of the Elixhauser's ICD-9 mapping to 30 comorbidities. Similar to the Charlson Index, Elixhauser's first maps administrative patient ICD-9 codes to 30 predetermined comorbidities, then calculates the weighted index score. Instead of using the weighted index, I summed the number of positive comorbidities and used the total number of comorbidities as a patient severity index that ranges from 0 to 30. I retained the 30 binary variables from the Elixhauser's comorbidity conversion and included them as binary variables (presence of condition =1, absence = 0). R software under the ICD-9 package was used to translate the ICD-9 codes into the Charlson score and to calculate the comorbidity count. Paralysis, weight loss, and blood loss did not occur in any cases and were therefore dropped from models where individual comorbidities were included. Table 4-1 summarizes the main attributes with either a count or mean depending on variable type.

Table 4-1. Summary of Attributes Considered for each Model

Age	Mean 63	Median 64	Min 22	Max 89	
Gender	Male 255	Female 367			
ASA	1 13	2 231	3 371	4 7	
Ethnicity	Hispanic 16	Non-Hispanic 606			
Smoking status	Non-smoker 339	Smoker 73	Former smoker 210		
Procedure	Hip 218	Knee 404			
LOS (average hours)	63.5				
Discharge location	Home 477	HHC 27	SNF 32	Rehab 76	Other 10
Diagnosis related group (DRG)	469 13	470 609			
Readmission count	30 days 20	90 days 48*			
HOOS/KOOS (average)	Pain 39.4	Symptoms 43.3	Function daily life 40.5	Function sports leisure 20.3	Quality of life 21.1
Supply cost	Mean 4889				
ICD-9 count	Mean 1.73	Min 0	Max 6		
Comorbidities	Total				
CHF	*10*				
Arrhythmia	*75*				
Valvular	*18*				
PHTN	*11*				
PVD	*12*				
HTN	*371*				
Paralysis	*0*				
NeuroOther	*5*				

Pulmonary	*39*
DM	*4*
DMcx	*29*
Hypothyroid	*96*
Renal	*22*
Liver	*17*
PUD	*3*
HIV	*2*
Lymphoma	*11*
Mets	*1*
Tumor	*64*
Rheumatic	*4*
Coagulopathy	*6*
Obesity	*50*
WeightLoss	*0*
FluidsLytes	*30*
BloodLoss	*0*
Anemia	*89*
Alcohol	*6*
Drugs	*5*
Psychoses	*10*
Depression	*86*

**14 cases readmitted for subsequent TJA not related to index admission arthroplasty*

A correlation matrix between the aforementioned variables did show some weak correlation ($\rho < 0.40$) between several of the regresssors namely ICD-9 count, Charlson score, and five of the comorbidities. To mitigate the correlation, models will either contain individual comorbidities or the comorbidity count. Also, ASA 2 and 3 are the two predominate ASA categories and therefore are highly negatively correlated ($\rho < -0.93$). ASA 3 is also weakly correlated ($\rho < 0.35$) with comorbidity count, which indicates that patients with a more severe ASA (poorer fitness for surgery) score also have more chronic conditions.

Instead of using the raw patient data for hip and knee osteoarthritis outcome score (HOOS/KOOS), patient age, body mass index (BMI), and length of stay (LOS), as individual predictors, I assigned each patient to their health market segment. From Chapters 2 and 3, market segmentation techniques applied to patient EHR data have shown that patients cluster into a small number of unique segments that share likeness between a subset of patient predictors such as age, LOS, BMI, etc. I applied the predictive cluster model from Chapter 3 to the patient data set and assigned each patient to one of six distinct clusters (i.e., health market segments). Table 4-2 summarizes the patient clusters. The values in the table are the means for each variable in each cluster.

Table 4-2. Summary of Patient Clusters

Cluster	Age	BMI	LOS (hr)	Pain	Symptoms	Function (daily life)	Function (sports & leisure)	Quality of life
1	68.82	32.36	59.25	52.53	57.56	52.69	19.85	33.16
2	68.08	29.24	66.15	18.13	30.82	16.52	5.56	7.72
3	62.91	33.19	58.89	62.60	61.24	66.18	68.84	43.10
4	51.08	30.19	55.23	37.11	34.70	44.19	24.32	16.57
5	59.17	41.25	59.33	33.47	34.91	33.30	7.89	10.43
6	68.00	33.73	146.42	25.32	32.14	20.14	10.34	13.70

4.2.2.3 Binary Logistic Regression

Binary logistic regression is based on the fundamentals of the multiple regression model (Nelder & Wedderburn, 1972; Rush, 2001; Sheather, 2009). In multiple regression analysis, I assume that the response variable is a linear combination of a set of predictor variables such that given a response Y and predictors $X_1, X_2, X_3, \dots X_k$:

$$E(Y|\boldsymbol{X}) = \beta_0 + \beta_1 X_1 + \beta_1 X_2 + \cdots + \beta_1 X_n + \varepsilon = \beta_0 + \overline{\beta X} + \varepsilon$$

In this model the β_o is the intercept, β_i's are the regression coefficients, and the predictors variables are represented as $X_1, X_2, \dots X_k$. The expected value of Y, $E(Y|X_1, X_2, \dots, X_k)$, also known as the conditional mean, is $Y - \varepsilon = Y'$. The foremost mentioned model works well under certain assumptions and when the response variable is continuous. When the response, Y, is binary, this model could result in values other than zero or one.

Binary logistic regression is a variant of multiple regression that accounts for the binary response variable by using a logistic function to model the relationship between the probability of success ($Y = 1$) and the predictors (***X***). Assuming $\theta(\bar{X}) = P(Y = 1|X_1, X_2, \dots, X_n)$, the logistic function takes the form $\theta(\bar{X}) = \frac{1}{1+e^{-(\beta_0 + \beta\bar{X})}}$, which can be written $\beta_0 + \beta\bar{X} = \ln\left(\frac{\theta(\bar{X})}{1-\theta(\bar{X})}\right)$. The logistic function creates (0,1) bounds on the conditional mean. The odds of a success are $\frac{\theta(\bar{X})}{1-\theta(\bar{X})}$ and the natural log of the odds or logit is $\ln\left(\frac{\theta(\bar{X})}{1-\theta(\bar{X})}\right)$. Once I solve for the regression coefficients, the logit is a linear function of x. Exponentiation of the log odds provides the increase in the odds for a one unit increase in the predictor variable. Next, I fit the binary logistic regression model to the data and use maximum likelihood estimation (MLE) to estimate the regression coefficients.

MLE is an effective method to estimate regression parameters. Since the distribution is not normal, as it is in linear regression, and a closed form solution does not exist, an approximation technique such as Newton's method is used. For each of the n training data points $(\bar{x}_i, y_i)$, the contribution to the likelihood function, L, is $\theta_i(\overline{x_i})$ if $y_i = 1$ and $1 - \theta_i(\bar{x}_i)$, if $y_i = 0$. Assuming independent observations, the equation for the full likelihood function is:

$$L(B) = \prod_{i=1}^{n} P(Y_i = y_i|x_i) = \prod_{i=1}^{n} \theta(x_i)^{y_i}[1 - \theta(x_i)^{1-y_i}]$$

This equation is more easily solved by taking the natural logarithm of both sides. To solve for the parameters, I differentiate with respect to the coefficients, $\beta_0, \beta_1, \beta_2, \dots, \beta_k$, set the resulting k+1 equations to zero, and solve using an iterative method. The regression coefficients were computed in R using RStudio.

The binary logistic regression output includes the odds ratio (OR), the standard error of the coefficient estimate, the p-value, and the 95% confidence interval around the OR. The marginal effect is also included in the output table. Additionally, AIC and McFadden's adjusted R-squared are used to evaluate the models.

4.2.2.4 Multiple Linear Regression

Linear regression is a form of multiple regression where the responsive variable Y is continuous and can be written as a linear combination of the predictors variables $\boldsymbol{X}$. In multiple linear regression, *n* observations are fit to the equation:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_1 X_{2i} + \cdots + \beta_1 X_{ki} + \varepsilon = \beta_0 + \overline{\beta X_\iota} + \varepsilon_i$$

where ε_i is the error or difference between the actual value of Y and the predicted value, such that $E(\varepsilon_i|X) = 0$. In multiple linear regression the response variable and the predictors are linear in the regression parameters.

The next step in linear regression involves estimating the coefficients, $\beta_0, \beta_1, \beta_2, \dots, \beta_k$, using ordinary least squares. To solve for the regression coefficients, minimize the residual sum of squares (RSS) by taking the derivative of RSS with respect to each coefficient.

$$RSS = \sum_{i=1}^{n} \hat{e}_i{}^2 = \sum_{i=1}^{n} (y_i - \hat{y}_\iota)^2$$

This creates *k*+1 equations in *k*+1 unknowns which is solvable by calculus or in most cases a statistical computer package. Linear regression calculations are performed in R using RStudio.

Linear regression is a common and powerful tool to model the relationship between a response variable and a set of predictor variables, but in order to apply linear regression, a few assumptions must hold. The four main assumptions are that Y is related to x such that $E(Y|X = x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$, and the error terms are independent, have constant variance, and are normally distributed with a mean of zero and variance σ^2.

The linear regression output from RStudio includes the coefficient estimate, which can be positive or negative, the standard error, t-value, and the p-value associated with the estimate. The output also includes the adjusted R^2 to measure the proportion of variation in the response variable explained by the predictor variables.

4.2.3 Results

After preprocessing the data, the study included 622 patients with 42 variables. By design, not all 42 distinct (excludes dummy variables) variables were used in the same regression model to avoid correlation, especially between the 30 attributes representing the presence or absence of a comorbidity and the sum of comorbidities attribute. To further ensure consistency between both the supply cost linear regression model and the readmission risk logistic regression model, additional data modifications were required to meet the linear regression model assumptions. The initial linear regression model for predicting supply cost suffered from heteroscedasticity. After completing an outlier analysis, patients whose supply costs or ICD9 count (variable capturing number of comorbidities or comorbidity burden) were greater than five standard deviations from the mean were removed. This reduced the size of the data set, but ensured the model met the required assumptions. The Breusch Pagan test supported the assumption of homoscedasticity, variance inflation factors were within tolerance indicating no significant multicollinearity, the Durbin Watson Test was used to show non-autocorrelated errors, and plots of the residuals indicated normally distributed standardized residuals. The adjusted R^2 for the model without the outliers was 0.5176, indicating that 51% of the variation in the model is accounted for by the predictor variables.

Two linear regression models (Model 1 and 2) were used to model supply cost. Model 1 considered a subset of the predictor variables that included the sum of the comorbidities and Model 2 considered each individual comorbidity as calculated using the Elixhauser's ICD-9 mapping to 30 comorbidities. Consistent to both models, supply cost (response) is regressed on gender, ASA score, ethnicity, smoking status, orthopedic surgeon, procedure code, DRG, discharge destination, Charlson index, and patient cluster.

In Model 1, there were numerous significant ($\alpha <= 0.05$) variables of which some increased supply cost and others decreased it. Table 4-3 shows the regression output. Female gender ($92), non-smoker ($101), surgeons MD1 and MD2 ($901 and $395), hip replacement ($392), and patients in market segment or cluster 6 ($235) all put upward pressure on the supply cost while patients coded as DRG469 (-$449) and those in market segment 3 (-$143) decreased supply cost. Furthermore, a one point increase in Charlson index score, which is normally associated with having an increased number of comorbidities (comorbidity burden), was associated with an increase in supply costs ($45). This is contrary to the decrease in supply costs seen when moving from DRG470 (patient without major complications and comorbidities-MCC) to DRG469 (without MCC). Although not significant at $\alpha = 0.05$, patients with ASA of 4 had a downward impact (-$471, p=0.055) on supply costs. Recall, these coefficients are solely based on supply costs and not total hospitalization or bundled payment expenses accrued during post discharge care.

Table 4-3. Linear Regression Output for Model 1.

Coefficients	**Estimate**	**Standard Error**	**t value**	**Pr(>|t|)**
(Intercept)	4250.22	142.7	29.783	**< 2e-16**
Female	92.29	39.54	2.334	**0.01991**
ASA2	-212.78	133.47	-1.594	0.11144
ASA3	-204.51	134.69	-1.518	0.12947
ASA4	-470.96	244.91	-1.923	0.05498
Latino	121.01	120.1	1.008	0.31407
Non-smoker	100.56	40.94	2.456	**0.01434**
Current smoker	69.31	65.58	1.057	0.29099
MD1	901.27	40.84	22.069	**< 2e-16**
MD2	395.01	54.66	7.227	**1.58E-12**
THA (hip)	392.38	39.71	9.88	**< 2e-16**
DRG469	-448.63	164.7	-2.724	**0.00665**
Discharge HHC	50.18	94.59	0.531	0.59596
Discharge SNF	69.18	89.99	0.769	0.44234
Discharge Rehab	101.85	61.48	1.657	0.09811
Discharge Other	-180.84	151.55	-1.193	0.23325
#Comorbidities	-22.56	18.51	-1.219	0.22322
Charlson score	45.17	22.57	2.002	**0.04577**
Cluster 2	-91.08	56.01	-1.626	0.10445
Cluster 3	-142.8	64.47	-2.215	**0.02715**
Cluster 4	11.07	58.47	0.189	0.8499
Cluster 5	-10.13	53.88	-0.188	0.85096
Cluster 6	235.21	114.62	2.052	**0.0406**

In Model 2, female gender, non-smoker, surgeon MD1 and MD2, THA and patients in cluster 6 remained significant and increased supply costs in a similar magnitude as in Model 1. ASA4 patients (-$595), those discharging to inpatient rehabilitation hospitals (-$135), and those in segment 3 (-$163) were significant and decreased supply costs. Additionally, there were six individual comorbidity conditions that had a significant impact on supply cost. CHF (+$360), renal (+$257), and drug use ($782) significantly raised supply costs whereas arrhythmia (-$125), HIV (-$1045), and fluidslytes (-$190) decreased supply cost. The adjusted R^2 for Model 2 improved to 0.5382. Table 4-4 shows the regression output for Model 2.

Table 4-4. Regression Output for Model 2

Coefficients	**Estimate**	**Standard Error**	**t value**	**Pr(>|t|)**
(Intercept)	4158.721	141.8098	29.326	**< 2e-16**
Female	117.3564	40.5694	2.893	**0.003971**
ASA2	-153.791	132.5281	-1.16	0.246375
ASA3	-124.735	134.6469	-0.926	0.354653
ASA4	-595.047	255.7155	-2.327	**0.020329**
Latino	62.0693	119.505	0.519	0.603701
Non-smoker	144.742	41.2748	3.507	**0.000491**
Current smoker	15.9301	67.0141	0.238	0.812193
MD1	904.0303	41.4324	21.819	**< 2e-16**
MD2	427.658	55.3713	7.723	**5.42E-14**
THA (hip)	402.8949	39.8771	10.103	**< 2e-16**
DRG469	-290.308	184.299	-1.575	0.115787
Discharge HHC	34.328	95.6204	0.359	0.719731
Discharge SNF	63.4352	89.6821	0.707	0.47966
Discharge Rehab	135.4856	62.4496	2.17	**0.030472**
Discharge Other	-115.936	149.6541	-0.775	0.438855
CHF	359.8052	172.7928	2.082	**0.037779**
Arrhythmia	-125.196	59.8713	-2.091	**0.036981**
Valvular	-0.2156	111.9764	-0.002	0.998464
PHTN	-192.741	146.4764	-1.316	0.188775
PVD	-48.0579	136.0797	-0.353	0.724104
HTN	-45.1379	39.8851	-1.132	0.258256
NeuroOther	23.9178	205.8308	0.116	0.907536
Pulmonary	116.6937	88.5686	1.318	0.188205
DM	-277.396	270.7479	-1.025	0.306025
DMcx	56.9401	95.7854	0.594	0.552453
Hypothyroid	23.7441	52.0502	0.456	0.648444
Renal	256.9237	124.0146	2.072	**0.038759**
Liver	198.2371	143.6675	1.38	0.168201
PUD	-400.806	259.9973	-1.542	0.123753
HIV	-1045.34	397.1882	-2.632	**0.008731**
Lymphoma	-10.9103	141.3731	-0.077	0.938513
Mets	-658.525	487.1052	-1.352	0.17696
Tumor	-97.213	61.2185	-1.588	0.11287
Rheumatic	275.9431	226.4781	1.218	0.223593
Coagulopathy	135.0668	220.205	0.613	0.539887
Obesity	52.159	69.4811	0.751	0.453159
FluidsLytes	-190.382	95.4967	-1.994	**0.046691**
Anemia	-41.3172	53.0057	-0.779	0.436029

Alcohol	41.4608	201.2687	0.206	0.836869
Drugs	782.0746	229.0146	3.415	**0.000685**
Psychoses	219.4143	162.9683	1.346	0.178742
Depression	-10.4286	55.1879	-0.189	0.85019
Charlson score	8.7871	35.3314	0.249	0.803681
Cluster 2	-90.9712	56.7029	-1.604	0.109213
Cluster 3	-162.919	64.3051	-2.534	**0.011569**
Cluster 4	-8.9098	58.0107	-0.154	0.877991
Cluster 5	-44.8049	55.447	-0.808	0.419402
Cluster 6	283.8611	118.716	2.391	**0.017135**

To predict readmission (a binary event), I applied a binary logistic regression model to the same inputs used in the first two supply cost models. Although the input variables are the same, the beta coefficients returned are in terms of log odds. The odds ratio is derived by exponentiating the coefficient for each variable. The output is reported as the change in the odds ratio when a significant attribute increases by one holding all others at their mean. For dummy variables, the odds ratio reported is the amount the odds increase or decrease when the variable changes (i.e., male to female) whereas with continuous variables the odds ratio reflects the change in odds ratio per one unit increase or decrease. The odds ratio can range from 0 to infinity and an odds ratio of one indicates neither an increase nor decrease in odds or risk of a readmission. Odds ratios are reported with a 95% confidence interval. Since an odds ratio of 1 indicates there is neither an increase nor decrease in probability of readmission, significant attributes are only considered if the 95% CI for the odds ratio does not include 1. Models 3 and 4 capture 30-day readmission and Models 5 and 6 capture 90-day readmission under both input variable options.

In Model 3, female gender, THA, discharge other (an inpatient facility not specifically for rehabilitation), the number of comorbidities, and market segment 5 are significant. Table 5 shows the regression output for Model 3 to include the OR, standard error, p-value, 95% confidence interval and marginal effect. The OR is the exponentiated version of the model estimate (not shown). All significant factors have positive odds ratios greater than one. The odds of a 30-day readmission are on average increased by 3.44 for females, 5.0 for hip replacement patients, and 3.13 for patients discharged to an inpatient facility other than rehabilitation. Additionally, 30-day readmission odds increase by an average of 2.2 for each for each comorbidity and by an average of 4.52 for patients in market segment 5. Table 5 contains the complete results of Model 3 including the 95% confidence intervals and marginal effects. The AIC and McFadden's adjusted R-squared value for Model 3 are 161.83 and 0.312, respectively. The results in Table 4-5 are based on using maximum likelihood estimators (MLE) to estimate the model parameters. Since the dependent variables, 30 and 90 day readmissions, were infrequent, a condition that can result in small-sample bias (Williams, 2016), I also evaluated the logistic models using Firth's (1993) penalized MLE (PMLE) method. The results using both methods are similar with the exception that PMLE did reduce the magnitude of the ORs and standard errors for non-significant factors. The output tables using PMLE are available in Appendix B.

Table 4-5. Logistic Regression Output for Model 3

Coefficients	**OR**	**Standard Error**	**Pr(>|t|)**	**2.50%**	**97.50%**	**Marginal Effect**	**Standard Error**	**z.ratio**
(Intercept)	3.21E-11	4622.301	0.996	NA	8.77E+112			
Female	3.44E+00	0.602	**0.040**	1.09E+00	1.20E+01	0.0313	0.037	0.8451
ASA2	1.22E+07	4622.301	0.997	0.00E+00	NA	0.413	0.0654	6.3105
ASA3	7.40E+06	4622.301	0.997	0.00E+00	NA	0.4002	0.0658	6.0805
ASA4	2.27E-01	8233.466	1.000	2.55E-87	5.98E+85	-0.0375	NA	NA
Latino	7.00E+00	1.310	0.137	2.82E-01	7.09E+01	0.0493	0.2607	0.1889
Non-smoker	1.50E+00	0.564	0.471	5.06E-01	4.75E+00	0.0103	0.0283	0.3645
Current smoker	3.12E-01	1.249	0.351	1.31E-02	2.55E+00	-0.0295	0.2091	-0.141
MD1	2.62E+00	0.635	0.130	7.89E-01	9.95E+00	0.0243	0.0602	0.4042
MD2	4.69E-01	0.914	0.408	6.64E-02	2.61E+00	-0.0191	0.1847	-0.1036
THA (hip)	5.00E+00	0.579	**0.005**	1.68E+00	1.69E+01	0.0407	0.028	1.4559
DRG469	1.85E-08	4847.285	0.997	NA	3.71E+215	-0.4506	0.1157	-3.8956
Discharge HHC	2.65E+00	0.946	0.303	3.23E-01	1.53E+01	0.0246	0.1511	0.163
Discharge SNF	7.23E-08	2633.472	0.995	NA	7.91E+58	-0.4161	0.0399	-10.4246
Discharge Rehab	2.37E+00	0.760	0.256	4.92E-01	1.03E+01	0.0219	0.0954	0.229
Discharge Other	3.13E+01	1.117	**0.002**	2.98E+00	2.74E+02	0.0872	0.2121	0.411
#Comorbidities	2.20E+00	0.260	**0.002**	1.33E+00	3.75E+00	0.0199	0.0124	1.6132
Charlson score	1.22E+00	0.259	0.436	7.16E-01	2.00E+00	0.0051	0.0207	0.2463
Cluster 2	1.50E+00	0.819	0.622	2.87E-01	7.75E+00	0.0102	0.1334	0.0766
Cluster 3	1.05E-07	1946.071	0.993	NA	6.25E+42	-0.4066	0.0816	-4.9819
Cluster 4	2.50E+00	0.884	0.300	4.07E-01	1.46E+01	0.0232	0.1469	0.1578
Cluster 5	4.52E+00	0.741	**0.042**	1.11E+00	2.15E+01	0.0381	0.0868	0.4393
Cluster 6	4.73E-07	3035.104	0.996	NA	5.32E+74	-0.3685	0.1002	-3.6793

Another way to look at the logit results is through a marginal effects analysis. The interpretation of the marginal effects varies depending on the type of independent variable. For binary variables, the marginal effect represents the change in the probability of the outcome (patient readmits) when a variable is increased by one (from 0 to 1). For continuous independent variables, the interpretation is more of an instantaneous rate of change in readmission status given a small change in the independent variable. The marginal effects for a given predictor are calculated based on holding all other predictors at their mean.

The results of Model 4 are similar to Model 3 with the exception that the market segments are not significant. Market segment 2's 95% CI on the OR does not contain 1, but the p-value is 0.056. Additionally, discharge with home healthcare and to a non-rehabilitation setting (discharge other) are significant and increase the odds of a readmission within 30 days. There were four comorbidities that were significant including arrhythmia, HTN, other neurological conditions, and lymphoma. The AIC and McFadden's adjusted R-squared value for Model 4 are 182.23 and 0.4997, respectively. Table 4-6 shows the regression output for Model 4.

Table 4-6. Logistic Regression Output for Model 4

Coefficients	**OR**	**Standard Error**	**Pr(>|t|)**	**2.50%**	**97.50%**	**Marginal Effect**	**Standard Error**	**z.ratio**
(Intercept)	1.03E-12	7093.7	0.99689	NA	2.79E+155			
Female	1.06E+01	0.9342	**0.01159**	1.99E+00	8.53E+01	0.0476	2.98E+12	0
ASA2	6.36E+06	7093.7	0.99824	1.25E-190	NA	0.3159	2.69E+12	0
ASA3	1.57E+06	7093.7	0.9984	9.23E-165	NA	0.2877	1.90E+12	0
ASA4	4.55E-03	11475.5	0.99963	3.01E-141	2.17E+129	-0.1088	NA	NA
Latino	2.16E+01	1.6249	0.05844	5.69E-01	5.03E+02	0.062	3.01E+12	0
Non-smoker	2.28E+00	0.7666	0.28304	5.38E-01	1.15E+01	0.0166	1.20E+12	0
Current smoker	1.49E-02	2.2409	0.06056	1.02E-04	5.50E-01	-0.0848	2.01E+12	0
MD1	2.03E+00	0.7756	0.36182	4.60E-01	1.03E+01	0.0143	1.36E+12	0
MD2	3.23E-01	1.2184	0.35377	2.35E-02	3.15E+00	-0.0228	1.19E+12	0
THA (hip)	1.05E+01	0.7993	**0.00326**	2.49E+00	6.16E+01	0.0474	1.85E+12	0
DRG469	8.16E-07	7555.9	0.99852	NA	2.66E+189	-0.2827	8.10E+12	0
Discharge HHC	2.35E+01	1.4372	**0.02795**	1.38E+00	4.53E+02	0.0637	4.52E+12	0

Discharge SNF	5.74E-08	3613.6	0.99632	NA	2.99E+79	-0.3363	2.83E+12	0
Discharge Rehab	1.19E+00	1.0967	0.87331	1.17E-01	9.53E+00	0.0035	1.60E+12	0
Discharge Other	4.81E+01	1.6453	**0.01855**	1.72E+00	1.25E+03	0.0781	4.84E+12	0
CHF	2.93E-10	6305.1	0.99722	NA	9.37E+169	-0.4427	4.74E+12	0
Arrhythmia	1.13E+01	1.0498	**0.02095**	1.32E+00	9.68E+01	0.0489	3.40E+12	0
Valvular	1.52E+00	1.5822	0.79227	4.12E-02	2.77E+01	0.0084	1.68E+12	0
PHTN	9.12E-02	2.2796	0.29344	3.10E-04	4.43E+00	-0.0483	4.70E+12	0
PVD	7.13E-09	7160.7	0.99791	NA	4.91E+163	-0.3783	6.13E+12	0
HTN	2.36E+01	1.1988	**0.00837**	3.15E+00	3.95E+02	0.0638	1.79E+12	0
NeuroOther	1.83E+02	2.0783	**0.01223**	2.76E+00	1.50E+04	0.105	NA	NA
Pulmonary	8.94E+00	1.2797	0.08695	6.95E-01	1.19E+02	0.0442	1.14E+12	0
DM	6.73E-06	11514.3	0.99917	NA	Inf	-0.2402	NA	NA
DMcx	4.01E+00	1.2577	0.26986	3.13E-01	4.79E+01	0.028	2.42E+12	0
Hypothyroid	5.25E+00	0.883	0.06031	8.73E-01	3.08E+01	0.0335	2.58E+12	0
Renal	4.23E-08	5063.5	0.99732	0.00E+00	1.43E+95	-0.3424	1.23E+13	0
Liver	2.09E+01	1.82	0.09512	5.34E-01	8.29E+02	0.0613	4.47E+12	0
PUD	1.48E-06	15974.9	0.99933	NA	Inf	-0.2708	NA	NA
HIV	4.92E+04	20721.8	0.99958	7.33E-237	2.19E+242	0.2179	NA	NA
Lymphoma	2.36E+01	1.3502	**0.0192**	1.40E+00	3.61E+02	0.0638	5.52E+12	0
Mets	2.08E-09	29232.44	0.99945	NA	Inf	-0.4032	NA	NA
Tumor	1.76E+00	0.9481	0.55057	2.38E-01	1.10E+01	0.0114	1.13E+12	0
Rheumatic	5.31E-08	10906.1	0.99877	NA	Inf	-0.3378	NA	NA
Coagulopathy	8.81E-08	10218.59	0.99873	NA	1.04E+252	-0.3276	NA	NA
Obesity	5.91E+00	1.238	0.15132	4.74E-01	6.88E+01	0.0358	2.24E+12	0
FluidsLytes	8.03E+00	1.2541	0.09664	6.00E-01	9.60E+01	0.042	2.69E+12	0
Anemia	1.33E+00	1.0679	0.78808	1.39E-01	9.86E+00	0.0058	8.25E+11	0
Alcohol	4.39E-09	10799.08	0.99858	NA	1.25E+291	-0.3881	3.58E+12	0
Drugs	8.95E-07	12040.09	0.99908	NA	Inf	-0.2809	NA	NA
Psychoses	1.27E-07	7643.975	0.99834	NA	8.45E+199	-0.3202	3.54E+12	0
Depression	1.07E-01	1.6604	0.17822	1.85E-03	1.62E+00	-0.0451	3.61E+12	0
Charlson score	2.07E+00	0.5221	0.16255	7.54E-01	6.12E+00	0.0147	4.55E+11	0
Cluster 2	8.26E+00	1.1071	0.05645	1.01E+00	8.74E+01	0.0426	2.63E+12	0
Cluster 3	9.68E-09	2438.313	0.99396	NA	3.35E+51	-0.3722	4.64E+12	0
Cluster 4	4.72E+00	1.0824	0.1514	5.47E-01	4.30E+01	0.0313	1.98E+12	0
Cluster 5	6.63E+00	1.0315	0.06675	9.68E-01	6.03E+01	0.0381	2.72E+12	0
Cluster 6	1.70E-07	4442.46	0.9972	NA	5.09E+119	-0.3143	2.77E+12	0

In Models 5 and 6, 90-day readmission is the response variable of interest. The 90-day readmission risk model will also capture 30-day readmissions. The fourteen patients who readmitted within 90 days for non-index admission reasons were not treated as readmitted patients. This condition would occur for patients who schedule a second arthroplasty on an opposite joint within their 90 day recovery period. Model 5 is identical to Model 3 with the exception of the response variable. In Model 5, there were four significant regressors: Latino, THA, discharge other, and number of comorbidities. The confidence interval for the odds ratio for discharge other does include one. The AIC and McFadden's adjusted R-squared for Model 7 are 254.14 and 0.146, respectively. The regression output for Model 5 is in Table 4-7.

Table 4-7. Logistic Regression Output for Model 5

Coefficients	**OR**	**Standard Error**	**Pr(>|t|)**	**2.50%**	**97.50%**	**Marginal Effect**	**Standard Error**	**z.ratio**
(Intercept)	1.90E-09	1806.29	0.9911	NA	1.80E+44			
Female	1.89E+00	0.42636	0.1364	8.19E-01	4.42E+00	0.0287	0.0265	1.0848
ASA2	3.18E+06	1806.29	0.9934	9.91E-34	1.41E+232	0.6769	0.051	13.2848
ASA3	3.90E+06	1806.29	0.9933	3.20E-45	NA	0.6862	0.0459	14.9462
ASA4	4.56E-01	3161.432	0.9998	1.33E-30	2.11E+29	-0.0355	NA	NA
Latino	5.78E+00	0.88957	**0.0487**	7.61E-01	2.90E+01	0.0793	0.3193	0.2483
Non-smoker	9.41E-01	0.41778	0.885	4.15E-01	2.17E+00	-0.0027	0.0275	-0.0993
Current smoker	4.29E-01	0.81396	0.298	6.23E-02	1.77E+00	-0.0383	0.2437	-0.1571
MD1	1.19E+00	0.44656	0.6938	4.98E-01	2.92E+00	0.0079	0.032	0.2484
MD2	8.53E-01	0.58261	0.7855	2.55E-01	2.59E+00	-0.0072	0.0374	-0.1919
THA (hip)	2.64E+00	0.40955	**0.0176**	1.19E+00	6.03E+00	0.0439	0.0251	1.7484
DRG469	1.63E-07	1804.654	0.9931	NA	1.04E+82	-0.7067	0.1041	-6.7886
Discharge HHC	1.76E+00	0.8394	0.5022	2.47E-01	7.74E+00	0.0255	0.2445	0.1041
Discharge SNF	1.74E+00	0.87562	0.528	2.31E-01	8.25E+00	0.025	0.2673	0.0935
Discharge Rehab	1.97E+00	0.55924	0.2271	6.18E-01	5.71E+00	0.0305	0.0927	0.3295
Discharge Other	7.72E+00	0.94556	**0.0307**	9.37E-01	4.40E+01	0.0924	0.3209	0.2879
#Comorbidities	1.53E+00	0.18597	**0.0219**	1.06E+00	2.21E+00	0.0193	0.0099	1.9523
Charlson score	1.02E+00	0.206	0.9239	6.59E-01	1.49E+00	0.0009	0.0141	0.0631
Cluster 2	1.62E+00	0.56525	0.3935	5.32E-01	5.04E+00	0.0218	0.0349	0.6242
Cluster 3	4.08E-01	1.09668	0.4137	2.12E-02	2.47E+00	-0.0405	0.3653	-0.1109
Cluster 4	2.07E+00	0.62797	0.2462	5.86E-01	7.18E+00	0.0329	0.0334	0.9857
Cluster 5	1.69E+00	0.56819	0.3581	5.48E-01	5.28E+00	0.0236	0.0326	0.7235
Cluster 6	2.39E-07	1231.674	0.9901	4.96E-184	1.43E+16	-0.6893	0.0566	-12.1756

Similar to Model 4, Model 6 looks at 90-day readmission risk using individual comorbidities as regressors. Female gender, Latino, THA, discharge other, and the comorbidity pulmonary were significant. The AIC and McFadden pseudo R-squared value are 279.16 and 0.2568, respectively. Table 4-8 shows the regression output.

Table 4-8. Logistic Regression Output for Model 6

Coefficients	**OR**	**Standard Error**	**Pr(>|t|)**	**2.50%**	**97.50%**	**Marginal Effect**	**Standard Error**	**z.ratio**
(Intercept)	1.63E-10	4.73E+03	0.9962	NA	4.86E+137			
Female	2.78E+00	5.22E-01	**0.0499**	1.02E+00	8.05E+00	0.0423	0.0525	0.8058
ASA2	2.11E+07	4.73E+03	0.9972	1.71E-77	Inf	0.6983	0.0707	9.8722
ASA3	1.80E+07	4.73E+03	0.9972	1.50E-88	Inf	0.6917	0.0796	8.6934
ASA4	1.17E-01	7.23E+03	0.9998	1.05E-60	1.32E+59	-0.0889	NA	NA
Latino	9.09E+00	9.84E-01	**0.0248**	1.04E+00	5.72E+01	0.0914	0.3182	0.2872
Non-smoker	9.38E-01	4.86E-01	0.8947	3.64E-01	2.48E+00	-0.0027	0.0532	-0.05
Current smoker	1.52E-01	1.01E+00	0.0629	1.55E-02	8.89E-01	-0.078	0.3217	-0.2426
MD1	9.49E-01	4.95E-01	0.9156	3.59E-01	2.55E+00	-0.0022	0.0471	-0.0461
MD2	5.77E-01	6.75E-01	0.415	1.42E-01	2.06E+00	-0.0228	0.0679	-0.3354
THA (hip)	3.04E+00	4.58E-01	**0.0152**	1.26E+00	7.70E+00	0.046	0.0599	0.7685
DRG469	2.22E-07	4.95E+03	0.9975	0.00E+00	7.63E+92	-0.6343	0.2547	-2.4904
Discharge HHC	1.38E+00	9.64E-01	0.7371	1.56E-01	7.80E+00	0.0134	0.3182	0.0421
Discharge SNF	3.61E+00	9.02E-01	0.155	4.67E-01	1.87E+01	0.0531	0.2859	0.1858
Discharge Rehab	1.72E+00	6.43E-01	0.4008	4.52E-01	5.84E+00	0.0224	0.1718	0.1303
Discharge Other	8.90E+00	1.10E+00	**0.0468**	7.95E-01	6.75E+01	0.0905	0.332	0.2726
CHF	2.32E-09	4.43E+03	0.9964	0.00E+00	2.26E+81	-0.823	0.1991	-4.1346
Arrhythmia	3.04E+00	6.43E-01	0.0844	7.82E-01	1.02E+01	0.046	0.1372	0.3349
Valvular	2.25E+00	9.55E-01	0.3955	2.61E-01	1.26E+01	0.0336	0.2911	0.1154
PHTN	6.59E-01	1.48E+00	0.7783	2.03E-02	9.55E+00	-0.0173	0.4607	-0.0375
PVD	1.03E-08	4.74E+03	0.9969	0.00E+00	2.06E+82	-0.7612	0.1003	-7.5915
HTN	1.78E+00	4.89E-01	0.2365	7.10E-01	4.94E+00	0.024	0.0552	0.4348
NeuroOther	1.52E+01	1.50E+00	0.0699	4.86E-01	2.75E+02	0.1126	NA	NA
Pulmonary	8.83E+00	8.34E-01	**0.009**	1.69E+00	4.63E+01	0.0902	0.1681	0.5363
DM	1.38E-07	8.61E+03	0.9985	0.00E+00	9.27E+173	-0.654	NA	NA

DMcx	3.84E+00	7.96E-01	0.0909	7.39E-01	1.77E+01	0.0557	0.1742	0.32
Hypothyroid	9.27E-01	6.32E-01	0.9049	2.35E-01	2.96E+00	-0.0031	0.18	-0.0174
Renal	1.74E-08	3.42E+03	0.9958	0.00E+00	2.50E+60	-0.7398	0.2977	-2.4853
Liver	5.80E+00	1.25E+00	0.1589	4.32E-01	6.58E+01	0.0728	0.3117	0.2335
PUD	2.02E+01	1.54E+00	**0.0514**	6.12E-01	3.88E+02	0.1244	NA	NA
HIV	1.96E+06	1.24E+04	0.9991	0.00E+00	4.15E+241	0.5998	NA	NA
Lymphoma	5.56E+00	1.01E+00	0.0885	6.21E-01	3.79E+01	0.0711	0.3438	0.2067
Mets	5.94E-08	1.77E+04	0.9993	NA	Inf	-0.6889	NA	NA
Tumor	2.14E+00	5.93E-01	0.2004	6.17E-01	6.54E+00	0.0314	0.0733	0.4287
Rheumatic	3.65E-08	8.36E+03	0.9984	NA	Inf	-0.709	NA	NA
Coagulopathy	5.29E-08	7.00E+03	0.9981	NA	Inf	-0.6937	NA	NA
Obesity	2.54E+00	8.40E-01	0.2668	4.20E-01	1.22E+01	0.0386	0.1714	0.2253
FluidsLytes	1.57E+00	9.66E-01	0.6398	1.82E-01	9.16E+00	0.0187	0.3159	0.0593
Anemia	1.99E+00	5.77E-01	0.2316	5.96E-01	5.94E+00	0.0286	0.0461	0.6204
Alcohol	4.47E-09	6.70E+03	0.9977	NA	1.02E+275	-0.7959	0.2012	-3.9557
Drugs	6.52E-08	7.84E+03	0.9983	NA	Inf	-0.685	NA	NA
Psychoses	3.87E-08	5.55E+03	0.9975	0.00E+00	2.74E+96	-0.7065	0.2654	-2.662
Depression	1.38E+00	6.37E-01	0.6093	3.55E-01	4.51E+00	0.0135	0.1236	0.109
Charlson score	9.32E-01	3.42E-01	0.8362	4.47E-01	1.77E+00	-0.0029	0.088	-0.0333
Cluster 2	3.19E+00	6.51E-01	0.0747	9.06E-01	1.20E+01	0.0481	0.0571	0.8415
Cluster 3	4.96E-01	1.15E+00	0.5426	2.43E-02	3.46E+00	-0.029	0.4076	-0.0711
Cluster 4	2.79E+00	6.86E-01	0.1348	7.15E-01	1.10E+01	0.0425	0.0613	0.6934
Cluster 5	1.91E+00	6.64E-01	0.3291	5.14E-01	7.18E+00	0.0268	0.0433	0.6192
Cluster 6	5.17E-08	2.96E+03	0.9955	0.00E+00	1.91E+56	-0.6946	0.1225	-5.671

4.2.4 Discussion

As one of the main costs in TJA, understanding the patient factors that impact supply costs can inform surgeon decision making in terms of how to negotiate new contracts, how to balance patient surgical load to be financially neutral, and how to price bundled payments with non-CMS insurers. The results of model 1 show that there are numerous factors that affect supply cost including female gender, surgeon, hip surgery, Charlson comorbidity score, and patients in cluster (market segment) three. Surgeon selection is the biggest factor with one surgeon in particular increasing the cost as much as $901. This supports the literature that physician preference regarding implant use creates variability in surgical costs, and it highlights the disconnect between who pays for the implant and who selects the implant (Christo et. al., 2000; Bosco et al., 2014). Procedure code is also significant in that patients having THA cost more than those having TKA. It was assumed that non-implant supply costs (drapes, knifes, etc.) were consistent across physicians and based on their surgical preferences rather than patient attributes.

Other significant factors leading to higher supply cost are non-smokers, patients who go to rehabilitation hospitals, and patients with a higher comorbidity burden. It is counter intuitive to think that non-smokers increase supply costs especially in healthcare where smoking is associated with many negative health outcomes. In both model 1 and 2, non-smokers had a small, but upward influence on supply costs which could be attributed to surgeon selection of premium implants for younger patients. Although only based on one large joint registry study, Gioe et al. (2010) found no difference in revision rate for primary TKA between standard and premium implants. He did find that the average age for recipients of premium implants was younger than that of standard implants and that premium implants were more expensive (Gioe et al., 2010). These younger healthier patients would likely include non-smokers. Patients in cluster 6 (the patients with the longest average LOS) and patients discharged to a rehabilitation hospital have an increasing impact on supply costs which indicates that they may have required additional supplies or implant components resulting from a more complex procedure.

The most surprising result was that patients coded as having major complications and comorbidities had a downward impact on supply cost, yet patients with a greater comorbidity burden increased supply cost. This could be attributed to the low number of DRG469 patients in the sample or could follow from the same argument where patients with major comorbidities and complications (DRG469) have such severe medical conditions that they would not benefit from a more expensive specialty implant.

The results of Model 2 were similar to Model 1 except in Model 2, the comorbidity count was removed. Of the 30 comorbidity conditions evaluated, 6 of them were significant. CHF, renal, and drug use all increased the supply costs. Drug use alone had the largest impact with a beta coefficient of $781. CHF and renal were lower averaging $360 and $257. FluidLytes and Arrhythma put minor downward pressure on supply costs.

As expected, the common regressors between Models 1 and 2 are consistent in magnitude and direction. In both models, the biggest impacts on supply cost are attributed to the type of procedure and the surgeon. In this study, the three surgeons each use a different vendor for primary TJA. Vendor selection drives implant costs across similar patient types. In both models, there is clearly a cost separation between surgeons (MD 1 vs 3: $904 and MD 2 vs 3: $428). To contain the high cost of implants, hospitals can switch to a sole source supply contract to reduce cost variation and potentially gain a volume discount. Alternatively, hospitals can renegotiate contract prices with each vendor to mitigate variation or determine a cost cap and have vendors compete on price (Bosco et al., 2014).

Contrasting the models shows that comorbidity burden or the sum of a patient's comorbidities may not be an effective way to assess supply cost. This result supports previous research that found that patient factors have little impact on supply costs (Gioe et al., 2011, Bosco et al., 2014). In Model 1, comorbidity burden was not significant yet in Model 2, six of the comorbidities were. This indicates that certain comorbidities impact costs more than others. This is reinforced in Model 1 where the Charlson index was significant. Since the Charlson index is a weighted score, it puts more weight on specific comorbidities and only looks at a subset of the 30 used in Model 2. This subset of comorbidities likely has more impact on costs. The challenge with using comorbidities as regressors is that several of them are rare events occurring in less than one percent of patients.

In addition to supply costs, readmission is a significant concern for hospitals performing TJA under the CJR. In the two 30-day readmission models, there were several statistically significant variables. Similar to the supply cost models, gender and procedure code are significant factors to readmission. In Model 3, the factors that increased the odds ratio of 30-day readmission were female gender (OR: 3.44, 95% CI: 1.09, 12.0), hip surgery (OR: 5.0, 95% CI: 1.68, 16.9), discharge-to-other (OR: 31.3, 95% CI: 2.98, 274), number of comorbidities (OR: 2.2, 95% CI: 1.33, 3.75) and patients in cluster 5 (OR: 4.52, 95% CI: 1.11, 21.5). The distinguishing characteristic of a patient from market segment 5 (cluster 5) is their extremely high BMI (greater than 40). This is likely the contributing factor to increased 30-day readmission risk.

Model 4 results were similar to Model 3 in that gender, hip surgery, and discharge to an inpatient setting (non-rehab) were significant and increased the risk of readmission. Also, discharge to home with home healthcare was also significant (OR: 23.5). Cluster two and five were just over the significant factor threshold (0.05). Patients in clusters two and five are at increased odds of readmission (OR: 8.26 and OR: 6.63 respectively). There were four comorbidities that also increased the odds ratio for 30-day readmission. Arrhythmia (OR: 11.3), HTN (OR: 23.6), neuro-other (OR: 183), and lymphoma (OR: 23.6) all increased a patient's odds of 30-day readmission. As seen in both models, comorbidities did impact readmission odds.

Models 5 and 6 measured 90-day readmission. In Model 5, Latino (OR: 5.78), hip patients (OR: 2.64), discharge other (OR: 7.72) and number of comorbidities (OR: 1.53) were significant and increased the odds of readmission. The results of Model 6 were similar to Model 5 with the addition of two individual comorbidities. Pulmonary (OR: 8.83) was significant and peptic ulcer disease (PUD) nearly significant (p-value: 0.051; OR: 20.2). Compared to Model 4, the number of significant individual comorbidities decreased as the length of time increased.

In summary, gender and procedure significantly impact both supply cost and readmission. Female gender and hip replacement increased supply cost and readmission risk while comorbidity burden mainly impacts readmission risks. In measuring short term readmission, discharge location and a few specific comorbidities increase the odds of readmission. There are fewer comorbidities that impact the odds of readmission at 90 days. Patients in market segments two and five increase the odds of 30-day readmission which is likely attributed to their long length of stay or extremely high BMI. Pulmonary and peptic ulcer disease are the primary comorbidities that significantly increase the odds of readmission at 90 days.

The six models evaluated above underscore the importance that individual patient factors and comorbidities can have in understanding cost drivers in TJA. In evaluating the factors that impact supply cost and readmission, the role of individual comorbidities in a small study is limited by the infrequency with which they occur in the sample. They can however provide insight to surgeons who are debating interventions and polices as they relate to modifiable patient conditions. For instance, patients with pulmonary conditions might benefit from a specialty consult or care intervention prior to scheduling surgery especially knowing that the patient is at a high risk for readmission. In addition to the negative impacts of readmissions on hospital revenue under bundled payments, decreasing readmission risk will improve the patient’s outcome.

4.2.5 Limitations

This study has several limitations, some of which are noted by other authors. First, this study focused on a small sample size from one academic hospital in one region of the US. As shown by others, translating these results to other hospitals may not be effective which underscores the need for hospitals to understand similar research using their specific patient population. A second limitation of this work is that readmission data only captures readmissions to the same hospital. Although EHRs have improved the speed with which clinicians can access a patient’s current and past medical history, EHR data is not universally accessible and, therefore, the only mechanism to confirm readmission data is via insurers or asking the patient. At this time, neither option is optimal. A final limitation is that the data sample included patients with various types of insurance. It was not possible to isolate traditional Medicare patients who are subject to the CJR.

4.3 High Supply Cost Outlier

This section focuses on the clinical factors that lead to high supply costs and interventions that could reduce the economic burden of outliers. Implant costs, a large subset of supply costs, are one of the main cost-drivers for TJA accounting for 13 to 87% of the total in-hospital costs (Robinson, J., Pozen, A., Tseng, S., & Bozic, K., 2012). To reduce costs and improve margins for TJA procedures, many hospitals have focused on optimizing implant pricing for primary total knee arthroplasty (TKA). Hospitals have focused less on optimizing pricing for implants primarily used for revision procedures, as these procedures are less common and receive higher reimbursement. With the advent of ICD-10 some patients who require more expensive revision implants (re-implantations for infection) and previously coded to higher–reimbursed DRG 467/468, now track to lower-reimbursed DRG 469/470. More expensive revision implants are also sometimes required for complex primary TKAs which track to DRG 469/470 as well. I hypothesized 1) that the use of revision implants in a small group of DRG 469/470 TKA patients would significantly increase total costs and average implant cost for the entire group and 2) that the placement of re-implantations for infection into DRG 469/470 would result in large financial losses for hospitals on these cases.

4.3.1 Methods

I reviewed both total implant costs and total implant plus cement costs for each patient (838 cases) receiving a TKA that tracked to diagnosis related group (DRG) 469/470 over a 31 month period starting at the end of December 2013. The total group was separated into high-cost and low-cost cohorts based on total implant costs. The high-cost cohort consisted of patients with implant costs 1.5 times above the average contracted primary implant price, and the low-cost cohort included those patients with implant costs less than 1.5 times the average contract price. Total implant costs and average per case implant costs were compared for the entire cohort, the high-cost cohort and the low-cost cohort.

The average implant price was based on the average cost across six basic primary total knee replacement packages (three vendors). Each package consists of four components: femoral component, tibial component, articular surface, and patella. There is an $842 difference between the most expensive and the least expensive. Since vendor contract pricing has changed since December 2013, the component prices used were adjusted to the price as of February 2016. Total implant costs included the four components listed above plus any sleeves, augments, stems, and restrictors that were utilized. Implants utilized for the high cost patients in this study are listed in Table 4-9. One patient required a distal femoral placement hinged prosthesis. All supply costs represent the cost that the hospital paid and not the reimbursed amount. Cases that did not include a minimum of a tibial component, femoral component, and articular surface were not considered.

Table 4-9. Supply Component Categorization and Frequency of Use

	Complex primary TKA (n=37)	2nd stage re-implantation (n=8)
more constrained femoral component	27 (73%)	8 (100%)
femoral augment	3 (8%)	4 (50%)
femoral stem	16 (43%)	6 (75%)
more constrained polyethylene liner	29 (78%)	7 (88%)
antibiotic cement	25 (68%)	7 (88%)
revision tibial tray	14 (38%)	8 (100%)
tibial augment	4 (11%)	5 (63%)
tibial stem	21 (57%)	8 (100%)
tibial sleeve	1 (3%)	3 (38%)
distal femoral replacement	1 (3%)	0 (0%)

I determined the economic impact of tracking patients undergoing re-implantation for infection to DRG 469/470 instead of DRG 467/468 by comparing the total implantable costs (implants plus cement) for these patients to the costs for the entire cohort and to the average Medicare re-imbursement for DRG 468 and 470.

This study was approved by the Pennsylvania State University institutional review board (IRB#: STUDY000002554).

4.3.2 Results

4.3.2.1 Cohort:

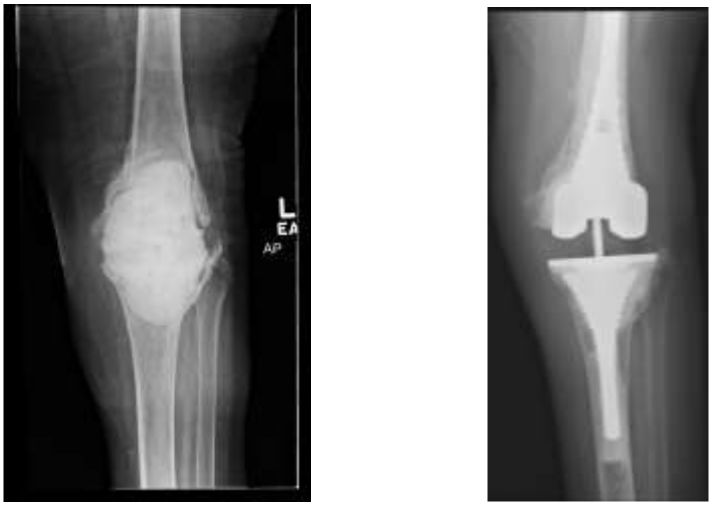
From 30 December 2013 until 25 July 2016, 847 patients received a total knee replacement coded to DRG 469/470 at our institution. Of the 847 patient supply records, 8 were discarded due to incomplete supply records and one was recoded from DRG470 to DRG468 leaving 838 cases in the study cohort. 793 of the 838 had implant costs (minus cement) less than 1.5 times the average implant cost while 45 (5.4 %) had implant costs greater than 1.5 times the average. Two of the high cost patients were mapped to DRG 469. Eight of the high-cost patients underwent re-implantation as part of a two-stage treatment for infection, 30 patients had a complex primary TKA requiring implants typically reserved for revision cases due to bony deformities and ligamentous laxity and 7 cases utilized a primary implant not on our primary implant contracts (Table 4-10, Figure 4-2, and Figure 4-3).

Table 4-10. Breakdown of High Cost Cases by Condition and Frequency

Condition	**Number of patients**
Complex bony or ligamentous deformities	30
Primary implants not purchases on contract	7
2nd Stage Spacer Re-implantation	8

A B

Figure 4-2A-B. 69 Yo Female with a 26 Degree Valgus Deformity and Severe MCL Dysfunction.



A B

Figure 4-3A. 86 YO Female with a Myobacterial Prosthetic Infection and Severe Tibial and Femoral Bone Loss Treated with a 2-Stage Approach.

4.3.2.2 Implant costs:

The total implant cost for all patients was just over $3.4 million with the low cost group accounting for $2.87 million (84.1 %). The 45 high cost patients (5.4 %) accounted for $543,074 in implant costs (15.9 %) and averaged $12,068 per case. This is over $8,450 more than the average cost of the low cost patients. The added costs for high cost patients increased the average implant cost for all patients by $454 or 12.6%. If the high-cost patients had received implants at the average cost, the savings would have been $380,399 (Table 4-11).

Clearly, the high cost group puts upward pressure on average cost of implants, but within the low cost patient group, there is still high implant cost variability. Figure 4-4 is a histogram showing the frequency of patients by total implant cost. Total implant costs are normalized based on the average contracted implant cost. As expected, the majority of surgeries fall around a normalized cost of 1, which represents the average contract price.

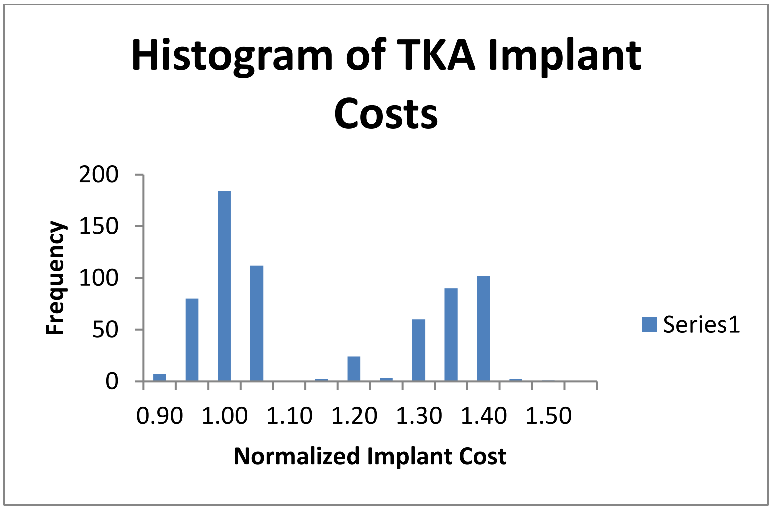


Figure 4-4. Histogram of Low Cost Patient Implant Costs

Figure 4-4 also highlights the bimodal nature of the distribution showing a second peak at 1.4 times the standard contract price. Despite having vendor contracts for primary TKA, deviations from the contract are putting upward pressure on the average cost.

Hospital reimbursement for TKA depends on the patient's insurance type. 73.3% of the high cost patients were covered by Medicare or Medicare Advantage and 26.7% had private insurance. Seven out of 12 with private insurance had a pass-through for high cost implants. Of the 45 high cost outliers, 32 of 33 patients covered by Medicare or Medicare Advantage mapped to DRG 470 and one to DRG 469. The hospitals average FY16 CMS inpatient perspective payment system (IPPS) hospital payment was $15,898. The implant costs for the average high cost DRG 470 Medicare cases accounted for 74.3% of the total hospital reimbursement. That percentage decreased to 31% for the DRG 469 patient as the average IPPS reimbursement is higher for patients with documented major complications and comorbidities.

Table 4-11. Summary of Implant Costs.

TKA Cohort	All TKA	High Cost Cases	Low Cost Cases
Case Count	838	45	793
Total Implant Cost	$3,409,764	$543,074	$2,866,691
Average	$4,069	$12,068	$3,615
		Impact of high cost cases on average	$454
		Total implant cost if no outliers	$3,029,366
		Savings generated if all average cost	$380,399
		% of high cost patients	5.4%
		% of total costs consumed by high cost cases	15.9%
	MEDICARE Patient Results by DRG	469	470
	# high cost Medicare cases Dec13-Jul16 by DRG	1	32
	Average implant + cement cost	$10,023	$11,808
	Average FY16 total hospital payment (Medicare IPPS)*	$32,409	$15,898
	Implant + cement cost as % of CMS reimbursement	30.93%	74.3%

*The average total payments to all providers for the DRG including the MS-DRG amount, teaching, disproportionate share, capital, and outlier payments for all cases. Also included in average total payments are co-payment and deductible amounts that the patient is responsible for and any additional payments by third parties for coordination of benefits.

4.3.2.3 ICD-10:

With implementation of ICD-10 in October 2015, patients undergoing second-stage TKA re-implantation for infection tracked to DRG 469/70 rather than DRG 467/468. At our hospital all spacer re-implantation patients mapped to DRG470. From October 2015 through July 2016, this group constituted 8 out of 288 TKA patients with 4 having Medicare or Medicare Advantage insurance. The average implantable cost (implants plus cement) for these patients was $20,923. For FY2016, the average hospital CMS re-imbursement for DRG 468 was $23,890 and the average CMS reimbursement for DRG 470 was $15,898. Only one of the spacer patients had a pass-through for high cost implants. Thus, considering only the four patients on Medicare, the hospital lost an average of $5,025 per spacer patient on cement and implant cost. That loss does not take into account rest of the hospital cost normally covered under the IPPS. Had those four patients mapped to DRG468, the hospital would have covered implant and cement costs leaving an average of $2,967 for other DRG related expenses. The average implantable cost for these 8 cases (2.78% of cases) represented 13.47% of the total implant cost across all FY16 DRG 470 patients and when compared to the average total Medicare hospital reimbursement, the implant plus cement costs per spacer case was nearly 132% of the average total reimbursement. For comparison, had the hospital been reimbursed under DRG468, the average spacer re-implantation case implant supply costs would have comprised just 87.6% of the total reimbursement. The FY16 average cost for all non-spacer patients was $3,704 per case comprising just over 23% of the total hospital reimbursement (Table 4-12).

Table 4-12. Summary of Patients who had 2^{nd} Stage Re-implantation

DRG*	468	470
# cases Oct 1, 2015-July 31, 2016		288
# cases with spacer re-implantation (SR)		8
% of spacer pat		2.78%
Implantable supply costs of SR cases		$167,388
Percent of implant costs utilized by spacer patients		13.47%
# of spacer patients with high cost pass through		1
Avg Medicare FY16 total hospital reimbursement **	$23,890	$15,898
Average implant + cement cost of spacer patient		$20,923
SR implant + cement cost as % of reimbursement***	87.6%	131.6%
Average non SR implant + cement cost		$3,704
Non SR implant + cement cost as % of reimbursement***		23.3%

*There were zero DRG469 spacer re-implantation patients
**The average total payments to all providers for the DRG including the MS-DRG amount, teaching, disproportionate share, capital, and outlier payments for all cases. Also included in average total payments are co-payment and deductible amounts that the patient is responsible for and any additional payments by third parties for coordination of benefits.
***comparing average cost in FY16 to average Medicare reimbursement. Assumes all patients reimbursed at Medicare rates.

4.3.3 Discussion

As the healthcare market has evolved, there has been increased emphasis on improving value and decreasing costs. Several studies have demonstrated that implant costs significantly vary by surgeon and can account for as much as 50 % or more of the total in-hospital costs for total joint arthroplasty (Bozic, et al., 2005; Healy & Iorio, 2007, Robinson, Pozen, Tseng & Bozic, 2012; Bosco et al., 2014). As a result, cost-reduction efforts have often focused on reducing implant costs. Many approaches have been described including surgeon cost awareness, vendor contracts, preferred vendor discount agreements, and price caps or ceiling (Healy, Rana & Iorio, 2011; Healy & Iorio, 2007; Belatti, et al., 2014; Healy et al., 1997; Healy et al., 1998; Healy et al., 2000). In general, most success has been achieved when hospitals and physicians are well-aligned (Bosco, et al., 2014, Healy & Iorio, 2007). Most implant cost-reduction efforts have focused on the implants used for primary total knee arthroplasty as these are higher volume, and the reimbursement for revision total knee arthroplasty is typically higher which offsets the increased costs of revision implants. Occasionally, more expensive revision implants are required for primary total knee arthroplasty (TKA) to ensure a high-quality outcome particularly in cases involving post-traumatic deformity, ligament insufficiency, and substantial bone loss. Since October 2015 when ICD-10 was implemented this group also included second stage re-implantations for infection. In these circumstances, implant costs may be 2-7 times greater than for typical primary TKA and may account for most or all of the hospital DRG payment unless there is a pass-through to insurance carriers for more expensive implants.

Although high cost implants drive up the average, cost variation within the low cost implant cohort also impacts overall implant costs. This is highlighted in the bimodal distribution shown in Figure 4-12 and could be attributed to surgeon preference and upcharges caused by patients who need abnormally large components, but may also be attributed to a lack of provider knowledge of the cost of implants or the difference in costs between comparable implants. Gioe et al. (2011) studied the differences in costs and outcomes between specialty and standard implants and found that specialty implants cost about $1,000 more, yet did not provide better survival 7-8 years after the procedure.

In the operating room, the surgeon should be focused on the patient and on achieving a quality health outcome first. However, cost and cost awareness are important and “surgeons cannot watch from the sidelines during the intensifying debate over national healthcare reform” (Gioe et al., 2011). In a recent survey of orthopedic surgeons and residents, Okike et al. (2014) found that a majority of orthopedic surgeons consider cost as an important factor (less than 1% thought cost was not important). They also found that 50% of the survey respondents assessed their implant device cost awareness as average and 36% as below average or poor. Okike et al. (2014) also reported that surgeons overestimated the cost of lower cost devices and underestimated higher cost devices which could lead surgeons to discount cost saving implants when choosing between comparable devices. Lack of surgeon awareness about the cost of the components they use, along with their colleagues, could be the reason that the surgeon was a significant factor in the linear regression model at the beginning of this chapter. Raising cost awareness for surgeons can inform their decision making and potentially help hospitals reduce expenses without sacrificing quality.

In this study I examined the financial impact of primary TKA cases that coded to DRG 470 but required implants typically used in revision TKA. I found that 45 of 838 primary TKAs (5.4 %) had implant costs that were greater than 1.5 times the average primary implant costs. Implants used for these cases varied but included stems, augments, intramedullary sleeves and more constrained components. The average cost for these cases was $12,068 which was more than 3 times greater than the average cost for low-cost cases. Overall the high-cost cases increased the total implant spend over this period by $380,399, accounted for 15.9 % of total implant costs and added an average of $454 (12.6 %) to the implant costs of each case. The implants for the high-cost cases accounted for 76 % of our average 2016 CMS DRG 470 payment ($15,898) while the implants for the average low-cost case accounted for 23% of the payment. The impact of the second-stage re-implantation for infection cases was even more dramatic. Between October 1, 2015 and July 31, 2016, I performed 8 re-implantations for infection which accounted for 2.8 % of TKA cases coded to DRG 470 during that time. Those 8 cases accounted for 13.5 % of total implant costs during that time. Three of the patients were covered by Medicare and one by a Medicare Advantage PPO. The average implant cost for these cases was $20,923 (including cement) which constituted 132 % of the actual hospital reimbursement. This issue should be improved somewhat as second-stage re-implantation for infection codes to DRG468 as of October 1, 2016.

Total knee arthroplasty provides tremendous functional improvements for patients but costs need to be managed in order to improve value. Most institutions have focused on managing primary implant costs however, I have found that revision implant costs, especially when used for primary TKA, can have a significant impact on overall implant spending and deserve more focus. Surgeons and hospitals should:

1) Carefully consider when these higher-costs implants are necessary using evidence based study results to inform surgeon decision making concerning implants when possible (AAOS, 2014; Healy & Iorio, 2007).

2) Improve surgeon education regarding the cost of implants (Okike et al., 2014) and incentivize surgeons who choose reduced cost implants when medically appropriate (Wilson et al., Healy and Iorio, 2007).

3) Re-negotiate revision implant prices.

4) Create high-cost implant pass-through to insurance carriers.

5) Advocate for risk-stratification based on orthopedic complexity, which could allow for appropriate reimbursement to cover the additional implant costs of more complicated cases (Bozic, et al., 2005).

4.4 Conclusion

Under value-based payment models, the role of data analytics and economic modeling will increase. Hospitals responsible for patient costs post-discharge must take into account patient-level factors and demographics in order to understand the potential costs associated with various types of patients. The value of understanding cost and outcome drivers, namely supply costs and readmission risk, is not to limit or shield an organization's risk and should not be used to restrict access to high cost patients. Instead, knowing the factors that can drive-up costs can inform a hospital's decision making regarding which interventions to apply to a patient to mitigate their risk of readmission or propensity for a high cost implant. The CJR is not designed to penalize hospitals, rather it is designed to incentivize them to extend care beyond their borders and into a patient's home where cost-saving interventions will reduce costly and preventable readmissions.

Cost and outcome drivers can negatively impact healthcare value. Hospitals and surgeons should not ignore the factors that contribute to these drivers and should actively work to identify and mitigate them in their patients.

Chapter 5

Investigating and Assessing Interventions

5.1 Introduction

The final pillar of the healthcare market segmentation methodology investigates and assesses interventions in TJA that reduce process variability, improve clinical care, and manage cost and outcome drivers. This pillar will not seek to prove the effectiveness of one intervention over another, but rather highlight interventions derived from the insight gained in the other four pillars. The interventions will include process and clinical improvement strategies based on the application of industrial engineering tools such as lean, six sigma, simulation, and time driven activity based costing. This framework alone is not designed to solve the hospital's financial risk dilemma, rather it should enhance a hospitals ability to incorporate data analytics to improve decision making and ultimately lead to improved healthcare value.

Physicians can improve value and outcomes by reducing unnecessary costs and through improvements in clinical efficiency. By streamlining the process and identifying and removing waste, resource time is more efficiently used. It is reported that, on average, 35% of clinicians' time is spent on non-value-added activities, which include delays, rework, errors, and redundancies (Murphy, 2003). These non-value-added activities cause clinical inefficiencies, and, consequently, have a negative impact on the provision of safe and quality care. These clinical inefficiencies are seen as major barriers to the provision of care. According to Sherman (2006), in healthcare settings there are two types of inefficiencies: (1) waits, delays, and defects, and (2) flow related issues.

This chapter investigates and assesses interventions in TJA that could help reduce cost drivers and improve outcomes. It starts by introducing a value improvement framework called DM-AIM or *define-measure-analyze-intervene-measure*, designed for physician leaders to integrate into their medical practice to drive continuous quality improvement. After introducing the framework, two interventions are proposed and evaluated using data and expert opinion from the Penn State Hershey Medical Center total joint arthroplasty division. The first involves the application of a new rapid recovery care pathway designed to incorporate the latest best practices in hip and knee perioperative surgical care. The second is a post-acute care assessment intervention that draws on market segmentation techniques developed for pillar II and cost and outcome drivers developed in pillars III and IV. The market segments are used to isolate cost and outcome drivers. A decision support simulation is presented as a tool to help clinicians evaluate improvement that clinical interventions would have on healthcare value. Both interventions are interrelated and stem from an exhaustive value improvement study documenting costs, quality, and patient flow through the TJA process from a patient's first visit to see a surgical specialist regarding sever osteoarthritis of the lower extremity through surgery and ultimately full recovery.

5.2 Value Improvement

Improving value in TJA under alternative payment models comes down to improving outcomes at a decreased overall cost. Improving outcomes and reducing cost are not always complementary as described by many healthcare publications over the past two decades. The U.S. spends more on healthcare than any other nation but lags other developed nations in many healthcare quality metrics. The shift to value based case is the government's attempt to reverse the trend and force hospitals to provide better long term value.

Now that surgeons and hospitals are at financial risk for outcomes beyond discharge, they are looking inward to their processes. This includes how those processes intersect with their clinical practice to impact patient health and wellness across the spectrum of pre-, peri-, and post-operative care. In terms of improving value, the area in which clinicians have the most control is during the pre- and perioperative phase when the patient is more or less under the direct supervision of the hospital and physician.

Evidence based care pathways and protocols are well documented strategies to improve value by decreasing costs (supply, personnel, length of stay) and improving overall quality through fewer complications and readmissions (Van Citters et al., 2014; Walter et al., 2007). Reducing LOS has been another focus in healthcare as it decreases inpatient bed utilization, reduces backlog in overcrowded hospitals, and creates opportunity to improve throughput by increasing surgeon utilization in the operating room. Under value based care, the role that physicians take in discharge planning is of increasing importance as post-acute care is a significant driver of the overall episode. Discharge planning for patients without adequate home care providers or with complex medical conditions can result in sending patients to more expensive post-acute care settings. Post-acute care settings are costly and not without significant readmission risk (Lavernia et al., 2013; Nichols, et al., 2016). Bozic et al. (2014) attribute over one third of episode of care costs to post discharge care and Ramos et al. (2014) found no significant correlation between discharge costs (location) and improved outcomes.

I posit that

1) a rapid discharge protocol that consists of best practices in TJA patient care can help decrease cost and improve quality as measured by length of stay and readmission rate, and

2) discharge planning that takes into account patient market segment can inform clinician decision and policy making with regards to care management both in the perioperative and post-

acute care setting. Better decision making will result in greater healthcare value for the patient and hospital alike.

The challenge for hospitals and physicians who perform total joint arthroplasty is how to reduce costs and create an efficient care umbrella that extends beyond the episode. Hospitals and physicians cannot directly control patients post discharge, but they can influence their post-surgery recovery, preparedness to return home at discharge, and the cost of supplies and resources consumed during their hospitalization.

5.2.1 Process and Value Improvement Literature Review

Lean Six Sigma (LSS) is a commonly used business strategy to improve process efficiency and quality, and reduce cost. LSS is the combination of two complementary strategies, Lean and Six Sigma. *Lean* is attributed to Taiichi Ohno and his work in developing the Toyota Production System (Ohno, 1988). The goal of *Lean* is to maximize value for the customer while minimizing resource usage which is achieved by the removal of *muda*, or waste, in a process (Ohno, 1988). *Lean* healthcare is the extension of the principles of the Toyota Manufacturing System to healthcare organizations (Graban, 2011). *Lean* healthcare seeks to remove eight forms of healthcare waste: transportation, overproduction, motion, defects, waiting, inventory, over processing, and lost human potential, creativity and opportunities (Graban and Swartz, 2012). Six Sigma is a powerful business technique initially developed to align business needs with quality and process improvement efforts. This technique was developed by Motorola in the mid-1980s, and was successfully expanded and applied in the mid-1990s by General Electric (Pande, Neuman, & Cavanagh, 2000. The set of statistical and non-statistical tools composing the Six Sigma toolkit has been used in various industries including healthcare. Some of the main advantages of Lean Six Sigma in healthcare settings are a better utilization or resources, elimination of wastes, improved working conditions, increased clinician satisfaction, and reduction of costs (Taner et al., 2007).

In order to support the understanding of the factors that affect efficiency and ultimately drive outcomes and cost, I present a continual healthcare value improvement methodology that is a hybrid of the six sigma DMAIC framework. Before defining DM-AIM, I define healthcare value, TDABC, and traditional DMAIC. Healthcare value "is measured in terms of patient outcomes achieved per dollar expended" (Kaplan & Porter, 2011). The costs of TJA are calculated using time driven activity based costing or TDABC. The cost per patient can be broken down into direct costs such as supplies, and resource (personnel) costs as measured by the cost per resource minute multiplied by the total minutes spent with a patient. I will identify and investigate key improvement areas in the total joint arthroplasty process by uncovering root causes of inefficiencies, which is necessary to provide safe, cost effective, and quality care. After developing a holistic understanding of the problem, tailored interventions are generated and evaluated to improve value.

DMAIC or Define-Measure-Analyze-Improve-Control is the core framework for process improvement used in Lean Six Sigma. Since the early 2000s, the DMAIC framework has been used for a variety of healthcare implementations. Improta, et al. (2015) used it manage costs and reduce average length of stay in total hip replacement by 44%. Warner et al. (2013) identifies the causes of delayed vascular surgery. After the first year of the implementation of the recommended solutions, they were able to improve on-time start of surgeries from 39% to 86%. Sievers et al. (2014) used the DMAIC framework to enhance diabetes management. They identified two main issues that needed improvement; timing of bedtime blood glucose measurement, and snack administration and documentation. Education interventions were proposed to overcome those issues resulting in considerable improvements.

Other applications of the DMAIC framework in healthcare settings include: decreasing patient identification band errors (Walley et al., 2013), reducing liver transplant length of stay (Toledo et al., 2013), improving timely bleed reporting (d'Young et al., 2014), improving inpatient pharmacoterapeutic process (Font Noguera et al., 2013), increasing reach rates for depression outreach (Beard, 2008), improving satisfaction of hospital based outpatients (McDonald & Kirk, 2012), improving efficiency of the resident rounding process (Chand, 2011), measuring quality care in a PICU (Sussmane et al., 2012), improving efficiency in a cystic fibrosis clinic (Smith et al., 2011), improving patient flow (Aakre et al., 2010), improving patient satisfaction (DuPree et al., 2009), reducing turnover rate of physicians (Taner & Sezen, 2009), reducing patient waiting time (Kim et al., 2009), and minimizing medical errors (Kumar & Steinebach, 2008).

TDABC in healthcare: The roots of ABC date back to George Staubus's work in the early 1970s and more formally in 1987 by Robert Kaplan and William Burns. Kaplan extended his earlier work and in 2004 developed TDABC which is a modification of ABC that requires only two parameters: the cost per unit time of each resource that impacts a patient and the total time that a resource encounters the patient in a given episode of care (Kaplan & Anderson, 2004). As in traditional ABC, direct costs are added to the TDABC costs to get the total cost per episode of care.

There have been numerous studies published in the last few years that apply TDABC in a healthcare setting. Akhavan et al. (2015) found that traditional hospital cost accounting methods overestimate the cost of TJA surgery, but acknowledge that there are inherent differences in how two methods account for costs. Further, they suggest that TDABC is a valuable tool in uncovering variability in resource utilization (cost) across physicians in the same practice. This level of cost granularity helps drive process improvement events. The Hoag Orthopedic Institute and Rothman Institute in Philadelphia applied TDABC to measure costs across a patient's care cycle. Understanding actual resource costs provides insight into bundled care contract negotiation (Kaplan, 2015a). Scottsdale Healthcare in Arizona used TDABC to determine direct costs in a study that measured value from three alternate prostate cancer treatments (Kaplan, 2015b). McLaughlin et al. (2014), applied TDABC in two pilot studies at UCLA Health showing that TDABC is an effective tool to engage "healthcare providers in process assessment and costing activities."

Prior TDABC healthcare studies consider TDABC as the process improvement method, whereas I employ TDABC in a larger value improvement framework. Furthermore, TDABC is not used as a hospital wide cost accounting method nor do I attempt to capture all indirect costs. Instead, TDABC is a tool that physicians can use to measure costs as they relate to resources within their control. TDABC requires two parameters and is scalable, therefore physicians can implement it without requiring extensive research into the financial aspects of reimbursements and hospital finance. Physicians can, but do not need to consider indirect costs such as hospital administration, utilities, and real property; rather they can just account for the time that a resource spends with a patient and the cost per unit time of that resource. Process improvements that decrease resource utilization reduce cost. Direct costs for supplies and implants are added to resource costs to determine total cost per patient.

5.3 Method and Framework

This chapter presents a hybrid healthcare value improvement intervention framework that is based on TDABC and the LSS DMAIC framework. Similar to the DMAIC framework, I define and measure the baseline process, then apply AIM, analyze-intervene-measure, to improve healthcare value in TJA and reduce the financial gap created by CJR. DM-AIM is a continuous improvement methodology that is applied to improve healthcare value in total joint arthroplasty. This approach is unique in that it is physician directed, based on a continuous improvement methodology, and focuses on costs that physicians can influence. Clinical and process related inefficiencies impact quality measures and the numerator of the value equation while cost efficiencies impact the denominator. TDABC measures the cost component of the value equation. By applying a simple cost accounting approach that only considers direct costs and resource costs directly attributable to a patient's care, a physician led quality improvement team can uncover variability, quality issues, and inefficient use of resources. The quality improvement is rooted in Lean Six Sigma principles of reducing waste and increasing clinical efficiency. This methodology can be tailored to different surgical settings, applies an intuitive resource costing method based on tangible resource utilization, and provides the clinical team with the data and visualizations required to make clinical and process changes to improve value.

DM-AIM Framework: The first and most important step is to define the problem and key sub processes that a patient completes along their joint replacement journey. Next, the team applies TDABC to calculate a surrogate patient cost at each step in the process. At this point, we start an iterative A-I-M or analyze-intervene-measure approach. The team recommends interventions based on observation, experience and documented best practices. The team implements the intervention(s), measures outcomes and compares results against prior data. As Figure 5-1 shows, the A-I-M phase is iterative and never enters a sustainment phase as the goal is to make continuous improvement. Once all objectives are met, new objectives are established and the team continues to improving healthcare value. DM-AIM is different from other LSS methods in that it does not require a certified black belt or LSS trained process engineer. The physician leader owns the process and is responsible for driving both clinical and process improvements.

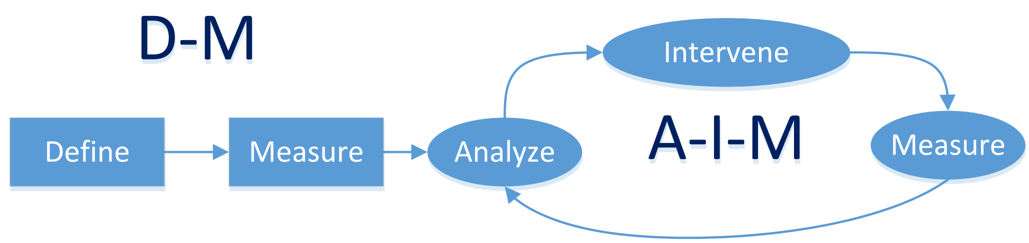


Figure 5-1. DM-AIM Continuous Improvement Framework

Key Performance Indicators: In TDABC, improving process and clinical efficiency translates into cost savings as fewer resource hours are spent on each patient. Increased throughput with the same resource allocations can also increase the number of patients per day which decreases the cost per patient. As part of the define phase, the team conducted a strategic goals crosswalk (Table 5-1) to determine what outcomes to measure and improve to achieve the project goals. Identifying, defining and measuring quality outcomes comprises the numerator of the value equation while TDABC measures the denominator. Recommended KPIs include patient length of stay, 30 and 90 day readmission rate, and implant costs.

5.4 Application of DM-AIM

This study is centered on the Hip and Knee Joint Arthroplasty Division at the Penn State Hershey Medical Center in Pennsylvania. There are three main orthopedic physicians that perform TJA and together perform over 1200 surgeries per year. Approximately 50% of the patients are covered under Medicare or Medicare Advantage and the hospital is part of the CJR bundle. Understanding the root causes of clinical and cost inefficiency were set as key priorities by the TJA stakeholders. The goal is to improve health value.

5.4.1 Phase One (Define & Measure)

The initial phase of the value improvement framework seeks to understand the scope of the problem under analysis. The team conducted a strategic goals crosswalk to develop actionable and measurable tasks that are nested in strategic and operational goals and aims set forth by the Institute of Medicine and hospital leadership. We formed a project team sponsored by the chief quality officer and chief of orthopedic surgery and lead by the chief of the total joint arthroplasty division. Other team members included an anesthesiologist, a project manager of the hospital's Quality Programs/Quality Academy, a physician assistant and finance manager from orthopedics, and a data analyst from Data Warehouse and Decision Support division.

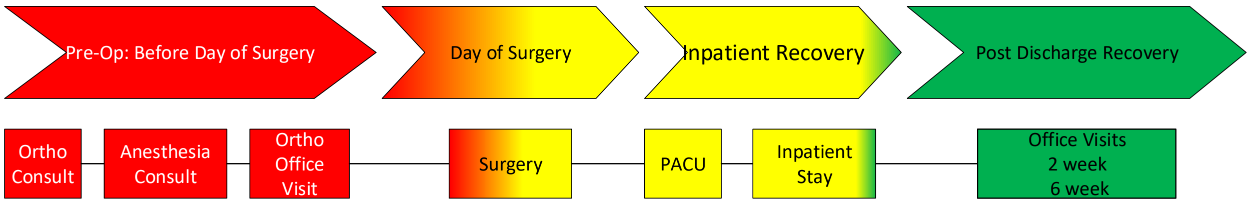
Table 5-1. Strategic Goals Crosswalk

Strategic Goal Cross Walk

Strategic	Institute of Medicine's "Six Aims for Improvement • Safe • Effective • Patient Centered • Timely • Efficient • Equitable
Operational	Hospital Leadership Imperatives for Patient Centered Healthcare • Patient-family centered care • Imperatives: Quality, Safety, Value • Eliminate Waste: reduce all cause 30 day readmission rate • Create extraordinary patient experience -Compassionate care experience -Improved access to care • Achieve top 10 percentile ranking in UHC
Tactical	Project Aims and Goals: • Improve patient health outcomes -Reduce readmissions -Improve patient readiness for surgery and recovery • ID/explain/reduce variability in care delivery process • Improve efficiency: -Reduce redundancy and waste in process -ID and mitigate cost drivers • Incorporate best practices: -Surgical, administrative, care coordination • Measure the cost of care for TJA from decision to operate to discharge +90 days

In January 2014, the team started meeting monthly to map the process and measure resource utilization in the current state. The team drafted a charter, high level process map, and timeline for the project along with the project scope. The high level process map is shown in Figure 5-2. Critical to success of this project is a clearly defined scope and current state. The scope of the project mirrored that of the CJR, but considered all patients admitted under DRG 469 and 470. Time studies and interviews with subject matter experts were essential for capturing the current state and the processes where electronic time stamps are not used such as pre-surgery appointments (orthopedic and anesthesia). The retrospective portion of this study included EHR data for patients admitted under DRG 469/470 starting in December 2013. Each month, new data was added until October 2016.

Figure 5-2. High Level Process Map of TJA



At the inception of the study, a small team conducted time studies of each step in the process. Using a simple stopwatch and notepad, the steps and resources utilized during each sub process were recorded and logged. The result was a detailed process map that captured the time and resource by type that interacted with the patients at each process. An example of the detailed process map is shown in Figure 5-3. De-identified timestamp data from the day of surgery augmented the time study results and corrected for sample bias. Electronic time stamp data was only available for patient transitions on the day of surgery and to document significant transitions such as admission and discharge.

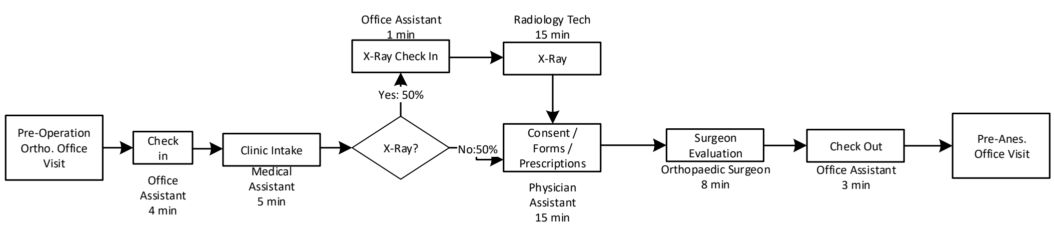


Figure 5-3. Detailed Process Map of the Pre-operation Orthopedic Visit Sub-process.

The next step was to calculate the cost per unit time of each resource (human) in the TJA process. Unit resource cost took into account non-clinical time, vacation time, and benefits. Since salaries vary by contract, experience and seniority, average Central Pennsylvania salaries by position type were used. Applying the tenets of TDABC, patient costs were calculated by sub process and by procedure code (knee or hip). Total cost calculations are surrogates for the actual cost since indirect costs and many resource costs outside of the control of the surgeon were not estimated. Supply costs per patient were extracted from electronic supply records. Based on the process map and time studies, total hip and total knee replacement consume different quantities of resources and are treated individually.

The process map allowed each team member to share a common operating picture of the TJA patient journey. Overlaid with the TDABC results, the process map told the story of cost, resource utilization, and opportunities for improvement.

5.4.2 Phase Two (AIM)

Phase two started with a detailed analysis of the process map, costs, and variability in performance measures discovered in phase one. Based on an assessment from phase one, phase two focused on two AIMs: decreasing LOS through a rapid recovery protocol and improving discharge planning to manage post-acute care costs and outcomes.

5.4.2.1 AIM1: Decreasing LOS through a Rapid Recovery Protocol

Analyze: In hospital post-surgery care includes a short window in the post-acute care unit (PACU) followed by a longer stay in the in-patient ward. A large component of the bundle cost under TJA is post-surgical care. Whether it is inpatient care or a short term stay at a post-acute care facility such as a skilled nursing facility or rehab hospital or both, post-surgical care costs are included in the bundle. To address the first stage of post-surgical care, I evaluated patient LOS. Of concern was the impact of PACU time on LOS. I hypothesized that a long PACU stay was positively correlated with a long LOS as patients who are slow to start recovery and discharge protocols in the inpatient ward spend more time in the hospital. Despite long PACU times, there was no correlation between PACU times and LOS (ρ=0.014).

Prior to this study, patients were averaging 67 hours per stay but the discharge time was highly variable. From Chapter 3, the market segmentation research showed that some patient segments had short stays while others were staying twice as long on average. In the TJA literature, there are a few surgeons, mostly private practice, who perform same day surgery or outpatient TJA with great outcomes (Parcells et al., 2016; Bertin, 2005; Aynardi et al., 2014). Based on discussions with clinicians, patients who track to market segment four, short length of stay, lowest average age, and low BMI, would make good candidates for a rapid discharge; but all patient could benefit from a patient centered recovery protocol. As an intervention to decrease LOS, the chief TJA surgeon applied a rapid recovery protocol to all patients. Not all patients recover at the same rate, but the protocol seeks to optimize recovery. The goal of the rapid recovery protocol is to reduce LOS to an average of one day and improve patient recovery to allow for a discharge home.

Intervene: The rapid recovery protocol consisted of six items:

1) Comprehensive pre-operative education class. This class included presentations from physical therapy, occupational therapy, discharge care coordinator, and orthopedic nurse navigator. Family care givers were highly encouraged to attend and learn how to manage care at home. Pre-operative education has been shown to be positively associated with discharge to home (Mahomed et al., 2000) and in a study by Heine et al. (2004), discharge readiness was positively associated with patient concern and feeling safe.

2) Team based approach. Patients were placed on a pathway with prescribed orders processes but managed individually. As patients transitioned from pre-op to peri-op to post-operative care, they were handed off by team members to team members.

3) Setting expectations. All patients were told to prepare for a one-night stay with discharge early on post-operative day one. Setting expectations was not about rushing patients,

rather preparing them and building their confidence. All joint team members reiterated the one-day discharge message. This follows from research by Heine et al. (2004) who analyzed discharge preparedness from the patient's perspective. Her team found that a patient's number one concern regarding discharge was feeling safe and that among others, confidence and expectations were important components of feeling safe (Heine et al., 2004).

4) Improved pain control. The surgery team worked with anesthesia to adjust pain medicine to enhance pain management while reducing nausea, vomiting, and muscle fatigue. This allowed patients to regain strength faster on the surgical day.

5) Reduced blood loss during surgery. Patients were screened for anemia prior to surgery and were administered tranexamic acid during surgery.

6) Early physical therapy following surgery. Post-operative patients must meet specific function and walking criteria prior to discharge. This protocol is linked to physical therapy such that patients receive PT as soon as the patient is deemed physically ready.

Measure: The rapid recovery protocol was implemented starting in August 2015. To assess the value improvement of the protocol, two samples, each containing approximately 200 consecutive primary total joint arthroplasty patients, are sampled from the data. The pre-intervention sample contains 199 patients who underwent surgery between August 2014 and April 2015 and the post-intervention sample contained 207 patients undergoing TJA from January 2016-May 2016. All patients are from the same two surgeons. Descriptive statistics from both samples are shown below in Table 5-2.

Table 5-2. Descriptive Statistics for Rapid Recovery Patient Samples

	LOS pre	LOS post	Age pre	Age post	BMI pre	BMI post	Supply cost pre	Supply cost post	ASA pre	ASA post
Mean	66.3	42.91	63.61	62.46	33.67	33.82	$4,849	$4,976	2.6	2.6
Median	57	34	66	62	33	34	$4,886	$4,094	2	3
Max	293	129	25	18	67	13	$16,511	$23,185	4	4
Min	32	12	87	87	20	71	$2,805	$3,066	1	2

	Pre RRP	Post RRP
male	82	97
female	117	110
Surgeon1	124	107
Surgeon2	75	100
30day readmit	4	7

Discharge	Pre RRP	Post RRP
Home	135	178
HHC	8	11
SNF	15	5
Rehab	38	10
Other	2	3

Using a Pearson's Chi Squared test, the proportion of patients discharged home is significantly greater in the post intervention period. Using a significance value of 0.05, the chi-squared value is 17.915 with a p-value of 2.31e-05. The 95% confidence interval for the difference does not contain zero (-0.2667, -0.0963). Similarly, there was a statistically significant decrease in LOS. Applying the Mann-Whitney U test to length of stay, the mean LOS between the pre-intervention sample and post intervention sample are statistically different (p=2.2e-16) indicating that the intervention reduced length of stay. Given two similar samples of primary total joint arthroplasty patients (DRG470) without major comorbidities or complications, applying a rapid discharge protocol significantly reduced average length of stay from 66 hour to 43 hours which effectively saves an average of one day as an inpatient. Also, the rapid recovery protocol had a significant impact on discharge. Prior to the RRP, a majority of patients were discharged home (~68%) but after the RRP, that majority climbed to 86% which is a significant improvement. Although these findings are significant, before concluding success, quality of care must be analyzed. In the absence of patient reported outcomes, readmission rate will be used to assess clinical quality.

Thirty-day readmission rate did increase under the rapid discharge protocol from 4 patients out of 199 (~2%) to 7 patients out of 207 (~3.4%). Given that readmission is a rare event and the sample size is small, determining statistical significance is challenging. In analyzing the readmissions both pre-intervention and post, in the pre-intervention sample, one was from discharge home, one from SNF, and two from rehab. Four of the post intervention patients readmitted after being discharged home, one from home with home healthcare, and two from SNF. In these two samples, with more patients discharged home after the RRP, there were more readmissions from home.

Although the actual percentage of readmissions from home increased in the post intervention sample, the overall impact of the rapid discharge protocol represents a significant value improvement for the hospital. Using readmission and post-acute care costs based on actual Medicare costs from patients undergoing TJA in FY2014, the hospital would have saved $5,473 per patient (Table 5-3). This analysis assumes that all patients are covered under a Medicare bundle or a commercial equivalent. The cost estimates are from actual hospital reimbursements make in FY2014 and do not reflect costs in the actual month and year that the patient was treated. Additionally, this does not include any potential readmission penalties levied by insurers for a lower quality score as a result of increased readmissions.

Table 5-3. Summary of Pre- and Post-Intervention Data

Pre intervention (199 consecutive TJAs under DRG470, Oct14-Mar15)				
	#patients	Avg LOS (HR)	LOS(day)	total cost
LOS cost	199	66.3	2.7625	1,264,396
Average $/day*	$2,300			
discharge	# patients	Average cost	Total Cost	
Home	135	0	0	
HHC	8	$3,281	$26,248	
SNF	15	$8623	$129,345	
Rehab	41	$19,925	$816,925	
30 day Readmission	4	$6,875	$27,499	
90 day Readmission	5	$6,875	$34,373	
		total cost	$2,298,786	
		cost/pat	$11,552	
Post intervention (207 consecutive TJAs under DRG470, Jan16-May2016)				
	#patients	Avg LOS (HR)	LOS (day)	total cost
LOS cost	207	42.91	1.7879	$851,227
Average $/day	$2,300			
discharge	# patients	Average cost	Total Cost	
Home	178	0	0	
HHC	11	$3,281	$36,091	
SNF	5	$8,623	$43,115	
Rehab	13	$19,925	$259,025	
30 day Readmission	7	$6,875	$48,123	
90 day Readmission	3	$6,875	$20,624	
		total cost	$1,258,205	
		cost/pat	$6,078	
	Post intervention savings		**$5,473**	
* http://www.beckershospitalreview.com/finance/average-cost-per-inpatient-day-across-50-states-2016.html				

Additionally, using propensity score matching, all 199 of the samples in the control group matched to post intervention samples. Using paired t-tests, there was not a statistically significant difference between the covariates age, BMI, and ASA score. Post intervention LOS was statistically greater than in the pre-intervention group.

5.4.2.2 AIM 2: Post Discharge Planning and its Impact on Value

Introduction: The readmission rate following the rapid recovery protocol (RRP) remained unchanged indicating that some patients are still reemitting following discharge. Additionally, despite a significant decrease in PAC discharges in the post RRP intervention sample, nine percent of the cases went to an inpatient facility. This intervention will apply market segmentation to cluster patients based on basic demographic data and discharge location. The goal is to identify a subset(s) of patients with a high utilization of skilled nursing facility and inpatient rehabilitation hospital services. It is also to explore the impact on hospital costs for specific patients given a longer index admission length of stay and reduction in the reliance on expensive PAC options in favor of lower cost options. The RRP focused on patient recovery, which led to shorter LOS for healthier patients, but LOS costs are only one of many costs included in the bundle. Decreasing LOS at the expensive of longer inpatient LOS and readmission may increase the cost burden for some patients. Also, this intervention will explore how the Medicare rules on payments to skilled nursing facilities and inpatient rehabilitation hospitals may increase PAC costs without improvements in quality.

Background: Discharge is the critical transition point for all patients and providers at the conclusion of a hospital stay. At discharge, the physician and in most cases the hospital lose their ability to directly influence patient health outcomes. Prior to value-based alternate payment models, post discharge medical expenses were not financially tied to the index hospitalization or the reimbursement. Now that payment for several DRGs including TJA are tied to an episode of care that includes discharge plus 90 days, the discharge related decisions that physicians and hospitals make following surgery can directly impact their reimbursement. Additionally, studies have shown that post-acute care, the care setting that select patients receive to ease their transition home, can contribute more than one third of the cost of the episode of care (Bozic et al., 2014).

The choice of post-acute care (PAC) setting is ultimately up to the patient but influenced by insurance coverage and physician recommendation. In the case of Medicare, prior to CJR, all discharges to PAC required a minimum of a three day LOS following surgery. Under CJR, Medicare has relaxed that constraint in an effort to encourage active involvement from the hospitals who now bear the financial risk.

The goal of discharge planning is to create the post discharge care conditions that lead to a cost effective and successful recovery. Home based discharge options are the least expensive and shown to average as low as $733 while inpatient care in a rehabilitation hospital can average well more than $16,000 (Ramos et al., 2014). For patients requiring in-patient services, the trend since 2005 has been to skilled nursing facilities over inpatient rehabilitation facility (IRF). This is in part to a change in Medicare’s IRF 75% rule (Haghverdian, Wright & Schwarzkofp, 2016). Under the CJR, hospitals would benefit if all patients discharged home without incident, but the reality is that some patients require additional care post discharge. The dilemma for hospitals is determining who needs PAC, what type of care setting is best, and how to best prepare patients for discharge so that their recovery time is minimized.

The individuality of patients, each with their unique set of past and present medical conditions, coupled with the burden of managing the time-space-distance challenges of a disparate surgical population makes discharge planning complex. As demonstrated in Chapter 3, market segmentation is a powerful tool that can divide a large population into subgroups or clusters of patients such that there is great similarity within each cluster and great dissimilarity between cluster centers. An advantage of using market segmentation in intervention planning is that it focuses the intervention on the segment most in need. Recall that patients in cluster one and three had the shortest average length of stay and lowest age whereas clusters two and four, especially four, were the oldest and had the longest length of stay. Also, as shown in Chapter 4, outliers can negatively impact a hospital’s narrow margin. High cost outliers in discharge planning are the patients who have a long hospital length of stay followed by a long inpatient rehabilitation stay. Identifying the patient attributes that define a cluster of patients with high utilization of expensive post-acute care services and high quality outcomes would help clinicians explore interventions that manage discharge location, time, outcome, and cost.

As shown in intervention one, the rapid recovery protocol (RRP), applied to all patients regardless of age or condition, significantly improved overall average LOS. The RRP did not eliminate LOS outliers as some patients require a longer LOS due to social, homecare, medical, and or surgical factors. The high LOS patients after implementing the RRP are the patients in cluster four who had an average LOS that was nearly twice as large as the other five clusters. The challenge for hospitals is determining who needs a longer LOS so that those patients are not rushed out only to return with complications or spend extended stays in post-acute care.

Early discharge under straight fee-for-service mitigated hospital costs, but now under value-based payment models, hospitals are at risk for post-acute care. This creates a risk-based decision for clinicians. Early discharge reduces immediate hospital costs but can increase bundle expenses with longer PAC stays, higher readmission rates, and poor patient outcomes (Mauerhan, Loneran, Mokris, & Kiebzab, 2003; Weingarten et al., 1998). Compounding the problem, higher readmission rates signal poorer quality and can lower the CMS readmission reduction factor, which reduces reimbursement across all primary joint replacements. Tied to the decision of when to discharge is where to discharge. As mentioned previously, readmission rates vary with discharge location.

There are several common discharge locations each with their own average readmission rate, quality, and cost. The first and least expensive is a discharge home. In order of average cost, the other common options include home with home healthcare, a skilled nursing facility, and a rehabilitation hospital. A fifth option, a subject of this intervention is to remain as an inpatient, then be discharged home. The decision of where a patient is discharged is often dependent on the type of insurance, the patient's willingness to be discharged home, caregiver accommodations, and other medical factors that impact a patient's ability to recover (Haghverdian, 2016; Heine, 2004; Mahomed, et al., 2000). The recent national trend (prior to CJR) was to discharge earlier from hospitals (Barsoum et al., 2010). Some authors have argued that early discharge reduces hospital costs at the expense of increased costs for PAC (Mauerhan 2003; Forrest, 1999; Weingarten, 1998). With shorter stays following TJA, hospital discharge coordinators have less time to plan for discharge, but no less incentive. PAC is expensive and under the CJR, a discharge home carries the least short term costs; but not all patients are good candidates for a discharge home or an immediate discharge to home. Furthermore, studies have shown little to no statistical difference in outcomes between home health, skilled nursing, and rehab facilities (Ramos et al., 2014; Mahomed et al., 2000). Readmission rates from SNF and Rehab have been shown to exceed those from home and home healthcare options (Nichols & Vose, 2016; Lavernia, Villa & Iacobelli, 2013). Higher readmission rates, especially when the readmission is from PAC, cost a hospital in three ways: the cost of PAC, the cost of readmission, and the reimbursement penalty for having a high readmission rate (low quality score).

Given the variety of conditions that cause a patient to readmit, the cost of a readmission, and the cost of PAC, this intervention seeks to assist hospitals with making smart discharge recommendations that consider health outcomes and cost. As shown in pillars III and IV of this methodology, discharge destination and readmission are related. Combining the outputs of the four pillars into the final pillar will enable clinicians to use the power of market segmentation and data mining of the medical record to identify and isolate patient clusters that are driving up costs and poor outcomes, and lowering healthcare value.

Analysis: In this intervention, market segmentation techniques are applied to a patient data set from the Penn State Hershey Medical Center total joint arthroplasty population. The 610 consecutive DRG469/470 patients all presented for surgery between 29 September 2015 and 27

September 2016. All patients were treated by one of three surgeons. Table 5-4 shows the summary data for the patient population.

Table 5-4. Descriptive Statistics for Intervention 2 Patient Sample

	Mean	***Median***	***Min***	***Max***
Age	62.9	63	18	91
BMI	33.3	32.5	16.1	71.4
ASA	2.6	3	1	4
LOS	44.4	34.3	12.2	273.4
Implant cost	$4,862.0	$4,428.3	$2,497.2	$23,380.0
	Male	***Female***		
Gender	254	356		
	Home	***Home with Home Care***	***SNF***	***IP Rehab***
Discharge	526	26	13	46
	Non smoker	***Prior smoker***	***Recently quit***	***Smoker***
Smoking	351	163	16	80
	Hip	***Knee***		
Joint	357	253		
	469	***470***		
DRG	13	597		
	#	***%***		
Outpatient therapy	357	58.5%		
30 day readmit	16	2.6%		
90 day readmit	22	3.6%		

The market segmentation analysis uses 11 attributes from EHR data. The attributes are age, BMI, ASA score, LOS, discharge destination, smoking status, gender, joint location, DRG, outpatient therapy, and current/prior fracture. Market segmentation was conducted using R software and the PAM package. PAM or partitioning around mediods is a robust version of the k-means algorithm used in Chapter 3. Unlike *k*-means, PAM will accept a dissimilarity matrix which is essential when using mixed variables. Binary and categorical variables were assigned a numerical value (i.e., 0,1 or 1,2,3,4) before using Gower's distance to create a dissimilarity matrix. Table 5-5 shows the categorical attribute conversion.

Table 5-5. Conversion of Categorical Variables

Gender		DRG		Smoking status	
Male	0	469	0	Never smoked	1
Female	1	470	1	Former smoker	2
Joint		Outpat Therapy; current/prior hip fracture		Recent quitter	3
Hip	0	Yes	1	Active smoker	4
Knee	1	No	0		

I experimented with a variety of different *k* values until settling on *k* = 7 clusters. Setting *k* =2 maximized silhouette distance, a measure of how well the clustering algorithm performed, but did not leave useful clusters with regard to discharge locations. With seven clusters, each cluster was unique and distinguishable. The clustering results are shown in Table 5-6 below. The attribute values for each cluster represent the mean value.

Table 5-6. Results of Clustering using PAM

Cluster	#	Age	Gender	BMI	ASA	Smoker	Joint	DRG	LOS (hr)	Discharge	Outpatient Therapy	Fracture
1	36	49.4	0.75	31.8	2.2	3.7	0	1.00	38.67	1.11	0	1.00
2	95	65.7	1	32.4	2.5	1.3	0	0.98	43.41	1.04	0	0.95
3	119	64.6	1	37.3	3.0	1.6	1	0.99	44.84	1.39	1	1.00
4	91	59.3	0	32.4	2.6	1.6	0	0.97	48.20	1.20	0	0.96
5	89	61.8	1	32.7	2.0	1.6	1	1.00	36.65	1.10	1	1.00
6	149	62.9	0	33.2	2.6	1.8	1	0.97	39.63	1.18	1	1.00
7	31	77.4	0.84	29.0	2.8	1.4	0	0.90	80.32	3.61	0	0.68

The clustering results show that cluster one is the youngest cluster composed primarily of females who smoke or recently quit. They are also all non-fracture hip patients who discharge early and often to a home setting. Clusters three, five, and six are similar with older patients and slightly different discharge, smoking patterns, gender, and joint type. Cluster seven is the most unique cluster. It is the smallest at just 31 mostly female patients who all underwent a hip replacement. 32% of patients were in surgery due to a hip fracture (non-elective surgery) or had a prior hip fracture (listed as an ICD-10 code). Additionally, 0 discharged to home under self-care, 3 discharged to home with home healthcare, 6 to a skilled nursing facility, and 22 (71%) to an inpatient rehab center. Surprisingly, none of the patients in cluster 7 readmitted to the same hospital that performed the surgery within 90 days. The average LOS for these patients was significantly higher resulting in an average of one more night stay in the hospital. Given the wide patient coverage area of the hospital, some patients could have readmitted to a different hospital and therefore their treatment record is not accessible for this analysis. For Medicare patients covered under the CJR, all episode costs regardless of location of service would be accounted for in the bundled price.

Given the number of hip fractures in these patients' medical history, their gender, average LOS, and their average age, it is not surprising that many of the patients in cluster 7 discharged to an inpatient post-acute care setting. Multiple studies have found some or all of the following conditions contributed to increased odds of discharge to extended care facilities (EFC): long length of stay, older age at time of surgery, increased comorbidity burden, Medicare insurance, non-elective surgery, and female gender (Bozic et al., 2006; Barsoum et al., 2010; Gholson et al., 2016, Keswani et al., 2016; Sharareh et al., 2014). Many of the aforementioned risk factors are not modifiable or easily modifiable especially in the case of the non-elective hip fracture joint replacements. The patients in cluster 7 likely require additional care post-surgery, but from the quality and cost perspective, an inpatient rehabilitation hospital setting may not be the best option.

Prior to the CJR, Medicare TJA patients discharged to an ECF were managed by the EFC and Medicare. If the ECF was part of the hospital, billing and financial risk were not attributed to the index admission. Medicare patients discharged to SNFs were covered for up to 20 days and those to rehabilitation or long term treatment hospitals for even longer (MEDPAC, 2016; CMS, 2016a). Now that hospitals are at financial risk under CJR, type and duration of PAC can drive up costs. Reimbursement tied to length of stay vs outcome is an area worth investigating and raises questions regarding the timeliness of discharges from PAC and whether they are tied to outcomes and readiness to discharge home or reimbursement. In a recent study, Haghverdian et al. (2016) reported that in a small study, Medicare patients had significantly longer SNF lengths of stay than a comparable cohort of managed care patients, yet had similar outcomes at discharge. They further reported that the Medicare patients achieved similar physical therapy outcomes at LOS day 14 that the managed care group achieved at LOS day 12, yet the Medicare patients stayed on average 24 vs 12 days. For Medicare patients, there was no correlation with longer length of stay and increased distance ambulated (Haghverdian et al., 2016).

In addition to long length of stay contributing to higher cost of PAC, post-acute care and in particular, skilled nursing facilities, has recently come under scrutiny. In two Department of Health and Human Services Inspector General Reports, SNFs as a whole have failed to meet Medicare requirements for quality of care, care planning, medication management, and discharge planning (Levinson, 2013 & 2014). Levinson (2014) found that poor quality of care resulted in 22% of patients suffering adverse events and 11% undue harm. He further found that about half of the patients harmed readmitted to the hospital. These reports demand that hospitals take an active role in discharge planning to ensure that their patients are placed in the location that will best achieve a quality outcome while minimizing cost.

Cluster 7 patients represent an opportunity for the hospital to investigate improvements in post-acute care savings. These 31 patients and future patients who represent the same cluster characteristics cost a disproportion amount of money and may not be getting the best care for the healthcare dollar. Given the average cost of an inpatient rehab hospital stay is nearly double the cost of a SNF and over four times the cost of HHC, these patients represent a high cost segment of the total joint population. The question for clinicians is why these patients are all requiring inpatient facilities and is there an intervention that the hospital can take to improve their odds of discharge to home both with and without homebased care. If they require resident care, how can the hospital or surgical team prepare them for a rapid discharge from ECF. There are numerous interventions posed in the literature that could help improve discharge to home rates: better patient education and confidence in their ability to self-care (Heine et al., 2004; Oldmeadow et al., 2003); setting discharge expectations with patient and family early, early follow ups of ECF discharges, and internal auditing of ECF facilities (Gholson et al., 2016); and pre-operative physical therapy for the non-hip fracture patients (Snow et al., 2014). Additionally, hospitals could extend the length of stay if the reasons for discharge to PAC are due to a lack of patient confidence in their ability to self-care or the inability to prepare for discharge that occurs with hip fracture patients who enter surgery via the emergency department.

Intervention:

For this healthcare intervention, a high level TJA decision support cost model with modifiable parameters is presented. The model will account for variability in patient conditions, discharge options, and outcomes. The objective is to measure changes in overall episode costs based on changing input parameter under uncertainty. This model is focused on market segment 7 patients as they represent the largest user of inpatient post-acute care services.

Given the high variability in healthcare systems, a Monte Carlo simulation is used to model patient flow through the TJA process. Figure 5-4 below illustrates the model with parameters.

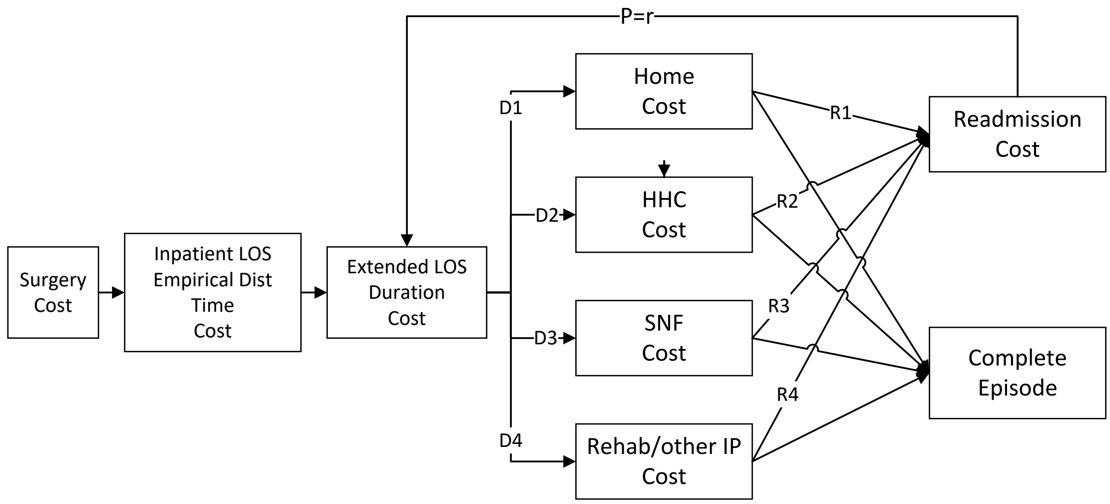


Figure 5-4. Model of Patient Flow through Episode of Care

Table 5-7 below illustrates the costs and transition probabilities for the model. In the base model, arrivals are bootstrapped from the actual data in cluster 7. At each iteration, a new patient is randomly selected with replacement from the cluster. The length of stay in hours and the supply cost for the bootstrapped patient is retained for the first two nodes. Table 5-8 shows the readmission frequencies and costs for all readmitted TJA patients in a recent study by Sebia et al. (2016).

Table 5-7. Cost and Transition Probabilities for Base Model

Node	Cost	Transition prob	New node	Notes
Surgery	Based on bootstrapped sample	100%	Inpatient LOS	Arrivals: bootstrap sample from cluster 7 (LOS and supply cost)
Inpatient LOS	LOS (hr) x Cost/hr	D1= 0	Home	Based on cluster 7 results; cost= $20/hr (TDABC)
		D2=.1	HHC	Based on cluster 7 results
		D3=.19	SNF	Based on cluster 7 results
		D4= .71	Rehab/IP	Based on cluster 7 results
Home	0	R1= 0.131	Readmission	Cost based on FY14 hospital average and readmission rates are based on a national database study by Nichols et al. (2016). Simulated costs were drawn randomly from actual PAC costs from patients in cluster 7.
HHC	$3,281	R2= 0.116	Readmission	
SNF	$9,465	R3= .22	Readmission	
Rehab/IP	$16,604	R4= .22	Readmission	
Readmission	$6,899	100%	Inpatient LOS	Readmission costs for simulation sampled from Table 5-7.

Table 5-8. Readmission Costs and Frequencies from TJA Outcomes Study by Sibia et al. (2016)

Diagnosis cost	Frequency	Cost
Altered mental status	4.55%	$1,142.00
Atrial fibrillation	4.55%	$11,663.00
Cholecystectomy	4.55%	$15,524.00
Hematoma	4.55%	$15,849.00
Hyperglycemia	4.55%	$2,706.00
Paresthesia	9.09%	$2,773.00
Postoperative ileus	22.73%	$3,707.00
Prosthetic joint infection	4.55%	$4,449.00
Pulmonary embolism	9.09%	$6,371.50
Sepsis	4.55%	$7,058.00
Syncope	4.55%	$7,215.00
Urinary tract infection	4.55%	$7,230.00
Wound infection	18.18%	$8,252.75

Model Assumptions:

1. The cost per readmission and the readmission rate per discharge location are based on an average taken from the published literature.
2. Costs for post-acute follow up care does not significantly impact the cost of the bundle (excluded from model).
3. The cost of outpatient post-acute follow up care (office visits, labs, x-rays) is the same regardless of discharge location.
4. The opportunity cost benefit to the hospital of a shorter length of stay is zero.
5. The cost per hour of inpatient stay is based on the average inpatient costs per stay across all cases divided by the average LOS across all cases. Costs based on TDABC estimates.
6. The cost of imaging, laboratory services, and specialty consults are excluded.
7. Readmission rates from PAC are based on a nationwide retroactive review by Nichols et al. (2016) of nearly 483,000 primary TJA patients. The actual readmission rates from the patient sample (cluster 7) were not used give the small sample size.

Monte Carlo simulation pseudo code:

1. Patient arrival: bootstrap a patient with their actual LOS and supply costs
2. Calculate LOS costs: LOS in hours * cost per hour
3. Generate a random uniform variable (0,1)
4. Assign discharge location (1-4) based on empirical distribution for discharge
5. Calculate the cost of PAC.
6. Generate a random uniform variable (0,1)
7. If RV < probability of readmission, calculate readmission cost.
8. Total cost of episode = supply cost + LOS cost + PAC cost + readmission cost
9. Repeat steps 1-7, n times, calculate average cost per patient.
10. Repeat steps 1-9, m times, calculate average cost per patient.

Base Model Results: Based on running 250 trials with each trial consisting of 100 patient samples, the average cost per patient is $22,799.43 with a 95% CI ($22707, $22891). The maximum average over the 250 trials was $25,050. This average underestimates the actual hospital cost of TJA based on the model assumptions. This model only takes into account readmissions, post-acute care, supply costs, and actual human resource costs. These are the costs the clinical team has the most control over. Breaking down the average patient cost by its components (Table 5-9), PAC and supply costs represent the biggest cost drivers across all patients. Readmission costs are not as significant based on an average readmission rate of 20.7% for all patients in the cluster.

Table 5-9. Episode Cost by Component

Episode	Cost	% of total
Total	$22,820	
Supply cost	$4,523	19.8%
First day cost	$1,395	6.1%
LOS cost	$1,607	7.0%
PAC cost	$13,949	61.1%
Readmission cost*	$1,346	5.9%
*overall average per patient		

As modeled and discussed in Chapter 4, supply costs have been steadily decreasing based on better contracting and physician awareness of costs. Therefore, assuming that patient outcomes remain constant or improve, several scenarios are modeled to assess the impact on healthcare value.

Scenario one: Longer inpatient recovery time reduces need for inpatient post-acute care

Hospitals have been discharging patients faster over the past decade or more as a result of inpatient bed shortages and the inpatient prospective hospital payment (Bozic et al., 2006). Some authors have raised concerns that early discharge places some patients at undue risk for readmission and longer PAC stays (Mauerhan et al., 2003; Forrest et al., 1999; Weingarten et al., 1998). During the 2nd quarter of fiscal year 2016, Hershey Medical Center had 12 Medicare primary TJA patient readmissions. One quarter of them (4) were within seven days of discharge. This could indicate that some patients are being discharged too quickly or are unprepared for discharge. In FY14, the average inpatient rehabilitation (IP rehab) patient stayed 12.3 days. In scenario one, inpatient LOS is increased by seven days and discharge to IP rehab is reduced by 75%. Those patients who would have gone to IP rehab will instead receive home healthcare. The assumption in scenario one is that a longer inpatient hospital stay stabilizes the patient and prepares them for a discharge home with home healthcare. This assumption follows from the literature and expert opinion from joint surgeons at Hershey Medical Center who feel that many older female patients with hip replacement (with or without hip fracture), either lack confidence or are unprepared for an immediate discharge home.

Reducing the percentage of patients who discharge to IP rehab by trading off for a longer inpatient LOS decreases the overall average episode cost for cluster 7 patients to $18,723 (95%CI: 18636,18811). This is significantly less than the base model ($22,800) but it does not include the lost opportunity cost of tying up inpatient beds for seven extra days. Based on intervention one, the rapid recovery protocol, the orthopedic joint service saved an average of one day per patient stay. In the 610 consecutive cases used in the market segmentation analysis, reducing the LOS by one day on average would save 610 inpatient bed days. This intervention would cost roughly one-third of the prior savings or 217 additional inpatient bed days for the 31 patients in cluster 7; however, it would potentially decrease the number of transitions that patients make after surgery and help the hospital manage financial and outcome risk for their most at-risk patients (elderly, fracture patients).

As discussed throughout this methodology, the challenge for hospitals under the CJR is to reduce the gap created at discharge when they retain financial risk but lose direct control over outcomes. Given patient demand, many hospitals cannot afford to extend the length of stay for all patients; however, cluster 7 patients are heavy users of post-acute care, contain the highest percentage of fracture patients, and have an average age over 77 years old making them Medicare eligible. In-hospital recovery guarantees patients access to specially trained orthopedic nurses, physicians, physical therapists, and other hospital services such as blood transfusions, that would otherwise be delayed from an outpatient or post-acute care setting. For instance, anemic Medicare patients whose blood hemoglobin levels decrease post discharge as a result of the surgery, would likely be readmitted or sent to the emergency department. In a 2009 prospective study, Carling et al. found that 10% of TJA patients received blood transfusions greater than 24 hours after surgery; for hip patients, odds of a transfusion are greater in older female patients who are not obese. Lavernia et at. (2013) found that anemia and blood disorders were the second most frequent reason for readmission from total joint arthroplasty. The patients in cluster 7 fit many of the risk factors for transfusion, thus reinforcing the benefits of a longer LOS for select patients.

Scenario two. Reducing reliance on IP rehab in favor of skilled nursing facilities (SNF) and home healthcare agencies (HHA).

The reliance on IP rehab care over SNFs and HHAs for post discharge patients generally increases a hospital's investment in a bundled payment. The analysis and results from 27 academic teaching hospitals enrolled in the bundled payments for care improvement initiative (BPCI) showed that hospitals that had previously favored inpatient rehabilitation over nursing care were able to shift some patients to lower cost care settings without increasing readmissions or mortality (Kivlahan et al., 2016). The shift saved one hospital an average of $3,000 per episode of care when relying more on HHC (Kivlahan et al., 2016).

In this scenario, I simulate an intervention that decreases the percent of discharges to IP rehab by 50% in favor of an equal increase in SNF and HHC. Recall that 71% of cluster 7 patients discharged to rehabilitation centers vs 19% for skilled nursing facilities. Under the assumption that the readmission rates by PAC location remain the same, the total cost per patient in segment 7 would be $19,142 (95% CI 19048, 19236). The scenario resulted in a statistically significant reduction (-$3,678) in the average total patient cost. The saving (>$58,000) represents a 16% reduction in episode costs for cluster 7 patients and is primarily due to the shift away from shorter duration but more expensive care. In FY14, Hershey Medical Center's average cost of post-acute care per day was $1215, $454 and $203 for Rehab, SNF and HHC respectively. Additionally, the average number of days per episode was similar at 12.3, 13.6, and 15.2. If the quality of skilled nursing facilities and home healthcare is similar to rehabilitation, then shifting away from the rehabilitation hospital will help improve healthcare value.

Another recent change in CMS policies regarding quality of post-acute care is the requirement for skilled nursing facilities to submit quality metrics and be graded on their outcomes (CMS, 2016b). Hospitals can use the market segmentation results to identify the patient clusters most likely to require IP rehab and then work pre-operatively with the patient to find high quality and reasonably prices skilled nursing facilities or home health agencies that are compatible with the patient. For non-elective hip fracture patients, the SNF rating system will allow discharge coordinators to influence patients towards high quality nursing care.

There are additional interventions that can be modeled using the clinical decision support simulation tool. In this analysis, patients in market segment 7 were the focal point as post-acute care costs are one of the top cost drivers in an episode of care following joint replacement and segment 7 has the highest proportion of discharges to inpatient facilities. The simulation is an efficient tool to measure changes in healthcare value based on changes in discharge disposition, readmission rates, and costs. It could also be used to model the success of interventions and justify future investments that will yield increased outcomes at a decreased cost.

5.5 Conclusion

Hospitals under alternate payment models such as the CJR are at financial risk for patient health outcome throughout an episode of care following the index admission. In total joint arthroplasty that risk extends 90 days of which most is spent outside the hospital. Hospitals can combat their risk through clinical and process related improvement in care delivery, management, and decision making that start in the preoperative phase. The DM-AIM framework is an effective tool for physician leaders to incorporate into their clinical practice to drive both clinical and process related changes that improve value. The rapid recovery protocol is one such improvement that reduced in-hospital LOS by approximately one day. Speeding patient recovery and getting many patients home faster without increasing readmission rates improves value. Additionally, applying market segmentation that includes discharge destination helped isolate a small segment of post-acute care high cost outliers. Simulating the impact of two interventions from the literature on this cluster resulted in decreased overall episode costs with fewer transitions. Improving healthcare value is not simply a function of reducing short term costs as the goal is to improve outcomes over the entire episode.

The impact of these interventions may be limited as they are based on a single hospital and rely on market segmentation to help identify the costliest segments. Clustering algorithms applied to a different data set might alter the results but the impact of identifying the highest cost segments cannot be understated. Furthermore, clustering is a tool, but the ultimate decision on who and how to intervene should be ethically based and made in partnership between clinicians and patients.

This page is intentionally left blank.

Chapter 6

Contributions and Future Work

This chapter summarizes the main contributions of this dissertation as well as areas of extension and future work. The three main contribution areas of this work include 1) market segmentation classification modeling with secondary use EHR data, 2) the identification of cost and outcome drivers in TJA, and 3) a new continuous improvement framework that places physician leaders in charge of driving clinical and process improvement with the help of data analytics and rudimentary financial accounting methods. These three areas of contribution spring from a healthcare market segmentation methodology comprised of five pillars. Applying the methodology can help hospitals improve healthcare value as they seek to identify specific subsets of their population for which they can target appropriate interventions that mitigate healthcare cost drivers and enhance outcomes. Improving healthcare value will reduce the gap created by bundled payments and the shift from a fee-based to a value-based healthcare payment model.

The U.S. healthcare woes are well documented, but until the Affordable Care Act, the government did not have a policy or program in place to incentivize change. Changing national level payment models along with the transition to widespread use of EHRs are forcing healthcare providers to adjust how they delivery care. The contributions to improving healthcare value demonstrated in this dissertation provide a framework for healthcare providers to adopt as they make the transition from fee-for-service to value-based care. Like many technological advances there are downside risks to health and public wellbeing when new tools are applied specifically to maximize profit. This methodology is not intended for hospitals, physicians, and payers to identify and exclude high cost or high risk patients. It is intended to aid in clinical and process related decision making so healthcare leaders can make ethical decision regarding the care options that enhance patient centered care and value. The contributions were shown using data from a total joint arthroplasty clinic in an orthopedic department of an academic teaching hospital, but could be adapted for use by any health system where EHRs and basic costing data are available.

6.1 Contribution 1

Applying advanced data analytics to include market segmentation and classification modeling to healthcare using the latent EHR data that hospitals already maintain fills a significant gap in the healthcare literature. Hospitals maintain vast stores of patient health data, but without advanced data mining and analytics tools, and a clear approach to using them, the data's potential is not fully realized. Previous market segmentation in healthcare relied on survey data whereas this contribution applied a two-stage clustering and classification method to EHR data to identify unique segments of the population that share commonalities amongst a set of known attributes. The classification models that follow the clustering allow hospitals to accurately assigned market segments to each new patient upon arrival or discharge. These segment classifications provide valuable information to the health provider and can help the provider intervene to improve healthcare value. Furthermore, applying data mining techniques to EHR data helps providers identify which attributes are most useful recognizing that not all attributes are modifiable. Those that are modifiable such as BMI or smoking status provide the healthcare community with an opportunity to market interventions to improve a patient's health, whereas those that are not modifiable provide the community with a basis for how to modify how care is delivered given the inherent conditions of the patients in those segments.

The market segmentation and classification portion of the healthcare market segmentation methodology was applied to a TJA patient population at PSHMC. The results showed that patients clustered into six unique and distinguishable clusters and that new patients could be accurately classified into one of the six segments. The approach demonstrates the efficacy of using data analytics to increase the knowledge about a single service line (orthopedic joint replacement) which in turn can inform surgeon decision making. Applied to other service lines in the hospital, this contribution could uncover similar patterns, further enhance provider decision making, and ultimately enhance patient health and health value. This could prove extremely helpful as patients often present with multiple conditions that span several service lines and individual providers. Furthermore, market segmentation classification modeling can be adapted to a variety of input factors. As seen in Chapter 5, incorporating a different set of attributes such as discharge destination, gender, smoking status, and ASA score helped isolate the set of patients who consume a disproportionate share of post-surgical costs. This two-stage approach provides clinicians, hospitals, and health systems with a flexible tool to identify and classify patients based on latent data.

6.2 Contribution 2

As CMS, the nation's largest insurer, moves from fee-for-service to value-based payment models that incentivize surgeons and hospitals alike to extend care beyond discharge, the role of data analytics and economic modeling will increase. Hospitals responsible for patient costs post-discharge must take into account patient-level factors and demographics in order to understand the potential costs associated with various types of patients. The value of understanding supply costs and readmission risk is not to limit or shield an organization's risk and should not be used to restrict access to high cost patients. Instead, knowing the factors that can drive-up costs (supply or readmission) can inform a hospital's decision making regarding which interventions to apply to a patient to mitigate their risk of readmission or propensity for a high cost implant. The CJR is not designed to penalize hospitals, rather it is designed to incentivize them to extend care beyond their walls and into a patient's home where cost saving interventions will reduce costly and preventable readmissions.

Changing healthcare payment models that include bundled payments with long episodes of care place hospitals and providers at financial risk for their patient's health. The increased financial risk motivates hospitals to improve patient healthcare value while retaining a larger portion of the payment. This dissertation demonstrates how a data analytics approach based on regression and descriptive statistics can identify major cost and outcome drivers. Cost drivers negatively impact the denominator of the value equation while outcome drivers impact the numerator. Although pillars III and IV of the methodology focus on TJA, with sufficient EHR data the methods may be deployed to other service lines.

In TJA, the major cost drivers are readmission, supply cost, inpatient stay, and discharge location. As shown throughout Chapter 4, these cost and outcome drivers are interrelated and impacted by gender, procedure code, comorbidities, and market segment. A patient's gender and procedure code significantly impacted both supply cost and readmission. A patient's comorbidity burden increased their readmission risk. In the short term (within 30 days of discharge), discharge location increased the odds of a readmission. Also, extending the work of the first two pillars, patients in specific market segments were at an increased risk of 30-day readmission.

Implant costs, a component of supply cost, are also cost drivers in TJA. Contracts and standardization of TJA protocols have mitigated much of the variation in implant costs but as shown in the outlier analysis of Chapter 4, there is a small subset of patients that are consuming a disproportionate share of the total implant cost budget. Prior to this study, the Penn State Hershey Medical Center, like other large hospitals, did not focus on the outliers. This type of deliberate analysis of outlier patients demonstrates the importance of identifying the types of patients that drive up implant costs for hospitals in bundled payments. By identifying these high cost outliers, hospitals can negotiate better contracts, improve surgeon education regarding the use of high cost implants, better negotiate with insurers to cover cost overruns, and better advocate with national healthcare insurers such as Medicare to provide orthopedic risk-stratification based reimbursement.

As hospitals seek to reduce the financial risk gap under bundled payment models, they must understand their patient population and how a patient's condition impacts the value equation. This contribution is not designed for hospitals to turn away high costs or high risk patients but rather to allow them to explore the patient factors that impact costs and outcomes and develop new protocols, contracts, or mitigating interventions that reduce risk and improve healthcare value for all.

6.3 Contribution 3

The final contribution of this dissertation is a continuous improvement framework that places physician leaders in charge of driving clinical and process improvement with the help of data analytics and rudimentary financial accounting methods. The continuous improvement framework is not novel but a derivative of the well-known DMAIC framework used in six-sigma. Although process improvement methodologies such as lean and six-sigma have been in healthcare for over a decade, they have not received as much attention as they have in other industries such as supply chain and manufacturing. It is widely accepted that healthcare has highly variable processes which can be blamed for high costs, delays, and poor outcomes. This research demonstrates that there is high variability in the TJA process but that the variability is manageable with proper leadership and a data driven modeling approach.

By applying the DM-AIM process and TDABC at PSHMC, the total joint arthroplasty division leadership was able to identify highly variable processes and then deliberately intervene to improve healthcare value. In Chapter 5, the DM-AIM framework was used to study both length of stay following TJA and a provider's discharge decision. Following the discussion in Chapter 4 on cost and outcome drivers, implementing a rapid recovery protocol improved patient flow through TJA, maintained high quality outcomes (low readmission rates), reduced length of stay, and decreased the discharge rate to higher cost post-acute care facilities. This single intervention, the rapid recovery protocol, had a significant impact on improving healthcare value.

The second intervention was the use of simulation modeling and market segmentation to identify and manage the highest cost patients. Through market segmentation, a small cluster of high cost patients is identified. The high cost patient segment is then modeled using several scenarios such as extended length of stay and altering discharge decisions. Published cost and outcome data from both the hospital and current literature are incorporated into the model. The combination of advanced data analytics and simulation provide clinicians with a decision support tool to aid in discharge planning. In one intervention, the trend for shorter length of stay is challenged for the high cost segment and is shown to cost more than a longer hospital stay followed by a shorter, less intense post-acute care stay. Although the focus was on discharge planning this tool could be used to study other cost drivers.

6.4 Future Work

The role of data analytics in healthcare is not fleeting and will only continue to grow in importance as major insurers and patients themselves demand higher quality care at a lower cost. From this dissertation, there are several extensions that would benefit healthcare providers, payers, and patients.

Based on pillars I and II of this methodology, there are several foreseeable extensions: 1) increasing the number and type of factors used in both clustering and classification, 2) increasing the size of the population and number of service lines used in the study, and 3) developing a market segmentation tool for health promotion as a method to encourage patients to improve the controllable factors that impact their health. As the number of sensors and data input opportunities increases, researchers will potentially have a much more robust data set from which to cluster patients. Adding additional well defined and unique clusters will better enable healthcare providers and marketers to efficiently reach subsections of the population that are at risk or at need. Furthermore, as data processing speeds increase and new algorithms are developed, hospitals, insurers, and health systems can expand these models to encompass larger populations and additional DRGs. Another future research line would be to investigate the feasibility of developing an online health application that tracks patients' health segment over time and links them to health promotions to enhance or maintain their health. This would allow the patient to be more involved in their own health and wellness at a time when health insurance costs are rising and individuals are looking to cut costs.

From pillars III and IV, some of the extensions to this research could include: 1) a multi-hospital or large health system study that looks at differences in outcomes across various hospitals and surgeons and benchmarks performance based on cost and outcome drivers, and 2) a decision support tool that assesses readmission risk based on the key factors that drive readmissions. Additionally, the contributions to identifying and addressing cost and outcome drivers were only evaluated in a one hospital setting. A future research line could expand this work to a larger health system where more generalizable results could be attained. After expanding this research to multiple hospitals, a subsequent future line of research could include an evaluation of why certain cost and outcome drivers impact healthcare value. Several of the factors included in this analysis are controllable and others could help surgeons develop guidelines on the treatment of osteoarthritis using TJA. For all but hip fracture patients, TJA is elective surgery and carries certain costs and risks to the patient and hospital. Developing a decision support tool for physicians to use to aid in their discussion with a patient on the merits of the surgery would be valuable.

Future work concerning the work surrounding pillar V could include: 1) an integrated decision support tool that predicts the total cost of a patient's episode based on their attributes and discharge decisions and 2) an evaluation of pre-habilitation such as physical therapy prior to surgery as a precursor to the rapid recovery protocol. It is well documented that physicians in general do not fully know the cost of surgery. Under FFS, readmission was normally a chargeable admission and in the OR there are no price tags on the items in the storage shelves. A future line of research could involve designing a decision support tool that takes inputs such as supply costs, patient attributes, discharge costs, and readmission costs and creates an episode of cost prediction tool. This would allow hospitals and surgeons to have an open and honest talk with patients regarding the cost of care. Raising cost awareness is critical to cost control. Addressing variation in physician practices within a hospital presents an opportunity for future research as hospitals look to standardize the delivery of care and establish practice guidelines.

A final future line of research is expanding the rapid recovery protocol to include pre-habilitation. There is scant published research on the benefits and cost of physical therapy before surgery, but given that the majority of TJA is elective, patients could benefit from enhanced therapy prior to surgery. If the intervention works, then hospitals at risk under the bundle might be willing to pay for it as a means to reduce the financial gap created after discharge.

The aforementioned future lines of research are tied to advances in the relationship between advanced data analytics and healthcare value. As the volume of healthcare data continues to grow, the reimbursements shrink, and the financial risk increases, hospitals will need to use data as a healthcare value multiplier.

REFERENCES

Aakre, K.T., Valley, T. B., & O'connor, M. K. (2010). Quality Initiatives: Improving Patient Flow for a Bone Densitometry Practice: Results from a Mayo Clinic Radiology Quality Initiative. *RadioGraphics*, 30(2), 309-315. doi:10.1148/rg.302095735 AHRQ (2016). Transforming the organization and delivery of primary care. Retrieved from https://pcmh.ahrq.gov/.

Akhavan, S., Ward, L., & Bozic, K. J. (2015). Time-driven activity-based costing more accurately reflects costs in arthroplasty surgery. *Clinical Orthopaedics and Related Research*,474(1), 8-15. doi:10.1007/s11999-015-4214-0

American Academy of Orthopaedic Surgeons. "98 percent of total knee replacement patients return to life, work following surgery." *ScienceDaily*. Available at: www.sciencedaily.com/releases/2013/03/130321082857.htm. Accessed October 30, 2016.

American Association of Orthopaedic Surgeons. 2014 Position Statement: Value driven use of orthopedic implants. Available at: http://www.aaos.org/uploadedFiles/PreProduction/About/Opinion_Statements/position/1104%20Value%20Driven%20Use%20of%20Orthopaedic%20Implants(1).pdf . Accessed September 15, 2016.

Axén, I., Bodin, L., Bergström, G., Halasz, L., Lange, F., Lövgren, P., & Jensen, I. (2011). Clustering patients on the basis of their individual course of low back pain over a six month period. *BMC Musculoskeletal Disorders*, 12:99.

Aynardi, M., Post, Z., Ong, A., Orozco, F., & Sukin, D.C. (2014). Outpatient surgery as a means of cost reduction in total hip arthroplasty: a case-control study. *HSS J*., 10(3), 252–255. doi:10.1007/s11420-014-9401-0

Barsoum, W. K., Murray, T. G., Klika, A. K., Green, K., Miniaci, S. L., Wells, B. J., & Kattan, M. W. (2010). Predicting patient discharge disposition after total joint arthroplasty in the United States. *The Journal of Arthroplasty*, 25(6), 885-892. doi:10.1016/j.arth.2009.06.022

Beard, G. (2008). Quality Toolbox: Improving Clinical Interventions Through Successful Outreach Using Six Sigma Quality Improvement. *Journal For Healthcare Quality*, 30(1), 38-43. doi:10.1111/j.1945-1474.2008.tb01126.x

Belatti, D. A., Pugely, A. J., Phisitkul, P., Amendola, A., & Callaghan, J. J. (2014). Total Joint Arthroplasty: Trends in Medicare Reimbursement and Implant Prices. *The Journal of Arthroplasty*, *29*(8), 1539-1544. doi:10.1016/j.arth.2014.03.015

Belmont, P. J., Goodman, G. P., Rodriguez, M., Bader, J. O., Waterman, B. R., & Schoenfeld, A. J. (2015). Predictors of hospital readmission following revision total knee arthroplasty. *Knee Surgery, Sports Traumatology, Arthroscopy*. doi:10.1007/s00167-015-3782-6

Berg, C., Ling, P., Guo, H., Windle, M., Thomas, J., Ahluwaia, J., & An, L. (2010). Using market research to characterize college students and identify targets for influencing health behaviors. *Social Marketing Quarterly*, 16(4), 41-69.

Bernatz, J. T., Tueting, J. L., & Anderson, P. A. (2015). Thirty-day readmission rates in orthopedics: a systematic review and meta-analysis. *PLOS ONE PLoS ONE*, 10(4). doi:10.1371/journal.pone.0123593

Bertin, K. C. (2005). Minimally invasive outpatient total hip arthroplasty: a financial analysis. *Clin Orthop Relat Res*, 435,154–163. doi:10.1097/01.blo.0000157173.22995.cf

Bjorgul, K., Novicoff, W. M., & Saleh, K. J. (2010). Evaluating comorbidities in total hip and knee arthroplasty: available instruments. *Journal of Orthopaedics and Traumatology*, 11(4), 203-209. doi:10.1007/s10195-010-0115-x

Bosco, J. A., Alvarado, C. M., Slover, J. D., Iorio, R., & Hutzler, L. H. (2014). Decreasing total joint implant costs and physician specific cost variation through negotiation. *The Journal of Arthroplasty*, 29(4), 678-680. doi:10.1016/j.arth.2013.09.016

Bozic, K.J., Katz, P., Cisternas, M., Ono, L., Ries, M.D., & Showstack, J. (2005). Hospital resource utilization for primary and revision total knee arthroplasty. *J Bone Joint Surg Am*., 87(3), 570-6.

Bozic, K.J., Wagie, A., Naessens, J.M., Berry, D., & Rubash, H.E. (2006). Predictors of discharge to an inpatient extended care facility after total hip or knee arthroplasty. *J Arthroplasty*, 21(6), 151-156.

Bozic, K. J., Ward, L., Vail, T. P., & Maze, M. (2014). Bundled payments in total joint arthroplasty: targeting opportunities for quality improvement and cost reduction. *Clinical Orthopaedics and Related Research*, 472(1), 188-193. doi:10.1007/s11999-013-3034-3

Bumpass, D. B., & Nunley, R. M. (2012). Assessing the value of a total joint replacement. *Current Reviews in Musculoskeletal Medicine Curr Rev Musculoskelet Med*, 5(4), 274-282. doi:10.1007/s12178-012-9139-6

Bundled Payments for Care Improvements (BPCI, 2016). CMS. https://innovation.cms.gov/initiatives/bundled-payments/.

Burton, R. (2012). Health Policy Brief: Care Transitions, *Health Affairs*.

Carroll N., & Gagnon J. (1983). Identifying consumer segments in health services markets. An application of conjoint and cluster analysis to the ambulatory care pharmacy market. *Journal of Health Care Marketing*, 3(3), 22-34.

Centers for Disease Control, Prevention. National Center for Health Statistics, Inpatient surgery. Available at: http://www.cdc.gov/nchs/fastats/inpatient-surgery.htm. Accessed July 28, 2016.

Center for Medicare and Medicaid Services (CMS), 2016a. https://www.medicare.gov/coverage/hospital-care-inpatient.html. [accessed 8 December 2016].

Center for Medicare and Medicaid Services (CMS), 2016b. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/NursingHomeQualityInits/Skilled-Nursing-Facility-Quality-Reporting-Program/SNF-Quality-Reporting-Program-Measures-and-Technical-Information.html [accessed 8 December 2016].

Centers for Disease Control, Prevention. National Center for Health Statistics, Inpatient surgery. Available from http://www.cdc.gov/nchs/fastats/inpatient-surgery.htm. Accessed 28 July 2016.

Centers for Disease Control and Prevention (2011). What is health marketing? Accessed 13 November 2014, available from http://www.cdc.gov/healthcommunication/toolstemplates/whatishm.html.

Centers for Disease Control and Prevention (2015): Gateway to health communication & social marketing practice 2015. Accessed 19 September 2015, available from http://www.cdc.gov/healthcommunication/healthbasics/whatishc.html.

Center for Medicare and Medicaid Services (CMS), 2016. https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/medicare-provider-charge-data/inpatient.html

Chand, D. V. (2011). Observational Study Using the Tools of Lean Six Sigma to Improve the Efficiency of the Resident Rounding Process. *Journal of Graduate Medical Education*, 3(2), 144-150. doi:10.4300/jgme-d-10-00116.1

Charlson, M. E., Pompei, P., Ales, K. L., & Mackenzie, C. (1987). A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *Journal of Chronic Diseases*, 40(5), 373-383. doi:10.1016/0021-9681(87)90171-8

Cheng, B., Chang, C., & Liu, I. (2005). Enhancing care services quality of nursing homes using data mining. *Total Quality Management & Business Excellence*, 16(5), 575-596.

Christ, A., Bargar, W., & Morris, E. (2000). Prosthesis cost containment in total joint replacement: a physician-driven free-market approach. *Orthopedics*. 2000: 23(5): 439-42.

Comprehensive Care for Joint Replacement (CJR, 2016). Retrieved April 30, 2016, from https://innovation.cms.gov/initiatives/cjr

Cran (2015). R website. Accessed 5 December 2015 at https://cran.r-project.org/web/packages/tree/tree.pdf

Davis, K., Stremikis, K., Schoen, C., & Squires, D. (2014). Mirror, Mirror on the Wall, 2014 Update: How the U.S. Health Care System Compares Internationally, The Commonwealth Fund.

Dupree, E., Martin, L., Anderson, R., Kathuria, N., Reich, D., Porter, C., & Chassin, M. R. (2009). Improving Patient Satisfaction with Pain Management Using Six Sigma Tools. *The Joint Commission Journal on Quality and Patient Safety*, 35(7). doi:10.1016/s1553-7250(09)35048-5

D'young, A. I., Young, L., Ockelford, P. A., Brasser, M., Slavin, K., Manson, L., & Preston, S. (2013). The use of a co-design model in improving timely bleed reporting by adults with haemophilia living in the Auckland region of New Zealand. *Haemophilia*, 20(3), 388-397. doi:10.1111/hae.12336

Episode of care. (n.d.) *Farlex Partner Medical Dictionary*. (2012). Retrieved July 2 2016 from http://medical-dictionary.thefreedictionary.com/episode+of+care

Firth, D. 1993. "Bias Reduction of Maximum Likelihood Estimates." *Biometrika* 80(1):27–38.

Forrest, G. P., Roque, J. M., & Dawodu, S. T. (1999). Decreasing length of stay after total joint arthroplasty: effects on referrals to rehabilitation units. *Arch Phys Med Rehabil*., 80,2.

Fox, J. & Weisberg, S. (2011). An {R} Companion to Applied Regression, Second Edition. Thousand Oaks CA: Sage. URL: http://socserv.socsci.mcmaster.ca/jfox/Books/Companion

Fry, D. E., Pine, M., Locke, D., & Pine, G. (2015). Prediction models of Medicare 90-day post discharge deaths, readmissions, and costs in bowel operations. *The American Journal of Surgery*, 209(3), 509-514. doi:10.1016/j.amjsurg.2014.12.005

Gholson, J. J., Pugely, A. J., Bedard, N. A., Duchman, K. R., Anthony, C. A., & Callaghan, J. J. (2016). Can We Predict Discharge Status After Total Joint Arthroplasty? A Calculator to Predict Home Discharge. *The Journal of Arthroplasty*,*31*(12), 2705-2709. doi:10.1016/j.arth.2016.08.010

Gioe, T. J., Sharma, A., Tatman, P., & Mehle, S. (2011). Do "premium" joint implants add value? Analysis of high cost joint implants in a community registry. *Clinical Orthopaedics and Related Research*, 469(1), 48-54. doi:10.1007/s11999-010-1436-z

Graban, M. (2011). *Lean hospitals: Improving quality, patient safety, and employee engagement*. Boca Raton, FL: CRC Press.

Graban, M., & Swartz, J. E. (2012). *Healthcare kaizen: Engaging front-line staff in sustainable continuous improvements*. Boca Raton: Taylor & Francis/CRC Press.

Greenspun, H. & Coughlin, S. (2012). The U.S. health care market: a strategic view on consumer segmentation. Deloitte Center for Health Solutions. Accessed 20 November 2015, available from http://www2.deloitte.com/content/dam/Deloitte/us/Documents/life-sciences-health-care/us-lhsc-mhealth-in-an-mworld-103014.pdf.

Griffin, P. M., Nembhard, H. B., DeFlitch, C., Bastian, N. D., Kang, H., & Muñoz, D. A. (2016). *Healthcare systems engineering*. Hoboken, NJ: John Wiley & Sons.

Haghverdian, B. A., Wright, D. J., & Schwarzkopf, R. (2016). Length of Stay in Skilled Nursing Facilities Following Total Joint Arthroplasty. *The Journal of Arthroplasty*. doi:10.1016/j.arth.2016.07.041.

Hanly, R. J., Marvi, S. K., Whitehouse, S. L., & Crawford, R. W. (2016). Morbid obesity in total hip arthroplasty: redefining outcomes for operative time, length of stay, and readmission. *The Journal of Arthroplasty*. doi:10.1016/j.arth.2016.02.023

Harse, J. D., & Holman, C. D. (2005). Charlson's index was a poor predictor of quality of life outcomes in a study of patients following joint replacement surgery. *Journal of Clinical Epidemiology,* 58(11), 1142-1149. doi:10.1016/j.jclinepi.2005.02.017

Haughton, D., Legrand, P., & Wooford, S. (2009). Review of three latent class cluster analysis packages: Latent GOLD, poLCA, and MCLUST. *The American Statistician*, 63(1), 81-91.

Healy, W. & Iorio, R. (2007). Implant selection and cost for total joint arthroplasty: conflict between surgeons and hospitals. *Clin Orthop Relat Res* 457: 57–63. doi: 10.1097/blo.0b013e31803372e0

Healy, W., Iorio, R., Ko, J., Appleby, D., & Lemos, D. W. (2002). Impact of cost reduction programs on short-term patient outcome and hospital cost of total knee arthroplasty. *J Bone Joint Surg Am*., 84-A(3):348–53.

Healy, W. L., Iorio, R., Lemos, M. J., Patch, D.A., Pfeifer, B.A., Smiley, P.M., & Wilk, R.M. (2000). Single price/case price purchasing in orthopaedic surgery: experience at the Lahey Clinic. *J Bone Joint Surg Am*., 82, 607–612.

Healy, W. L., Iorio, R., & Richards, J.A. (1997). Opportunities for control of hospital cost for total knee arthroplasty. *Clin Orthop Relat Res*., 345, 140–147. doi: 10.1097/00003086-199712000-00019.

Healy, W.L., Iorio, R., Richards, J.A., & Lucchesi, C. (1998). Opportunities for control of hospital costs for total joint arthroplasty after initial cost containment. *J Arthroplasty*, 13, 504–507. doi: 10.1016/S0883-5403(98)90048-1

Healy, W.L., Rana, A.J. & Iorio, R. (2010). Hospital economics of primary total knee arthroplasty at a teaching hospital. *Clinical Orthopaedics and Related Research*, 469(1), 87-94. doi:10.1007/s11999-010-1486-2

Heine, J., Koch, S., & Goldie, P. (2004). Patients' experiences of readiness for discharge following a total hip replacement. *Aust J Physiother*, 50(4): 227-33.

Hinarejos, P., Guirro, P., Puig-Verdie, L., Torres-Claramunt, R., Leal-Blanquet, J., Sanchez-Soler, J., & Monllau, J. C. (2015). Use of antibiotic-loaded cement in total knee arthroplasty. *World Journal of Orthopedics*, *6*(11), 877–885. http://doi.org/10.5312/wjo.v6.i11.877

Improta, G., Balato, G,. Romano, M., Carpentieri, F., Bifulco, P., Russo, M.A., & Cesarelli, M. (2015). Lean Six Sigma: A new approach to the management of patients undergoing prosthetic hip replacement surgery. *Journal of Evaluation in Clinical Practice*, 21(4), 662-672. doi:10.1111/jep.12361

Institute for Operations Research and the Management Sciences: INFORMS (2016): What is Operations Research? Accessed 30 June 2016 at https://www.informs.org/About-INFORMS/What-is-Operations-Research

Jain, A., Murty, M., & Flynn, P. (1999). Data clustering: a review. *ACM Computing Surveys*, 31(3), 264-323.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning: With applications in R*. New York, NY: Springer.

Jencks, S., Williams, M., & Coleman, E. (2009). Rehospitalizations among Patients in the Medicare Fee-for-Service Program. *Journal of Vascular Surgery*, *50*(1), 234. doi:10.1016/j.jvs.2009.05.045

Johnsen, S. P. (2006). Patient-related predictors of implant failure after primary total hip replacement in the initial, short- and long-terms: a nationwide Danish follow-up study including 36,984 patients. *Journal of Bone and Joint Surgery - British Volume*, 88-B(10), 1303-1308. doi:10.1302/0301-620x.88b10.17399

Kansagara, D., Englander, H., Salanitro, A., Kagen, D., Theobald, C., Freeman, M., & Kripalani, S. (2011). Risk prediction models for hospital readmission. *JAMA*, 306(15), 1688. doi:10.1001/jama.2011.1515

Kaplan, R. S., & Anderson, S. R. (2004, November). Time-Driven Activity-Based Costing. *Harvard Business Review*.

Kaplan, R.S., Haas, D.A., Reid, D., Warsh, J., & West, M.E. Getting bundled payments right in healthcare. *Harvard Business Review* (October 19, 2015a). (A collaboration of the editors of Harvard Business Review and the New England Journal of Medicine.)

Kaplan, R.S., Blackstone, R.P., Haas, D.A., Thaker, N.G., & Frank, S. Measuring and communicating health care value with chart. *Harvard Business Review* (October 26, 2015b).

Kaplan, R.S. & Porter, M.E. The big idea: how to solve the cost crisis in healthcare. *Harvard Business Review* (September 2011).

Kennett, P., Henson, S., Crow, S., & Hartman, S. (2005). Key tasks in healthcare marketing: assessing importance and current level of knowledge. *Journal of Health and Human Services Administration*, 24(4), 414-427.

Kent, P., Jensen, R., & Kongsted, A. (2014). A comparison of three clustering methods for finding subgroups in MRI, SMS or clinical data: SPSS TwoStep Cluster analysis, Latent Gold and SNOB. *BMC Medical Research Methodology*, 14(113).

Keswani, A., Tasi, M.C., Fields, A., Lovy, A., Moucha, C., & Bozic, K. (2016). Discharge destination after total joint arthroplasty: an analysis of postdischarge outcomes, placement of risk factors, and recent trends. *J Arthroplasty*, 31(6), 1155-1162.

Kim, Y., Oh, Y., Park, S., Cho, S., & Park, H. (2013). Stratified sampling design based on data mining. *Healthcare Informatics Research*, 19(3), 186-195.

Kim, Y.K., Song, K.E., & Lee, W.K. (2009). Reducing Patient Waiting Time for the Oupatient Phlebotomy Service Using Six Sigma. *Korean Journal of Laboratory Medicine*; 29(2): 171-7.

Kivlahan, C., Orlowski, J. M., Pearce, J., Walradt, J., Baker, M., & Kirch, D. G. (2016). Taking Risk. *Academic Medicine*, 91(7), 936-942. doi:10.1097/acm.0000000000001121 Kumar, S., & Steinebach, M. (2008). Eliminating US hospital medical errors. *International Journal of Health Care Quality Assurance*, 21(5), 444-471. doi:10.1108/09526860810890431

Kolodinsky, J. & Reynolds, T. (2009). Segmentation of overweight Americans and opportunities for social marketing. *International Journal of Behavioral Nutrition and Physical Activity*, 6:13, 1-13.

Kreder, H.K., Grosso, P., Williams, J.I., Jaglal, S., Axcell, T., Wal, E.K., & Stephen, D.J. (2003). Provider volume and other predictors of outcome after total knee arthroplasty: a population study in Ontario. *Can J Surg*, 46, 15–22.

Kremers, H. M., Visscher, S. L., Kremers, W. K., Naessens, J. M., & Lewallen, D. G. (2013). Obesity increases length of stay and direct medical costs in total hip arthroplasty. *Clinical Orthopaedics and Related Research*, 472(4), 1232-1239. doi:10.1007/s11999-013-3316-9

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*. New York: Springer.

Kurtz, S., Ong, K., Lau, E., Mowat, F., & Halpern, M. (2007). Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. *The Journal of Bone and Joint Surgery*, 89(4), 780-5. doi:10.2106/jbjs.f.00222

Kurtz, S. M., Lau, E., Ong, K., Zhao, K., Kelly, M., & Bozic, K. J. (2009). Future Young Patient Demand for Primary and Revision Joint Replacement: National Projections from 2010 to

2030. *Clinical Orthopaedics and Related Research*, 467(10), 2606-2612. doi:10.1007/s11999-009-0834-6

Lavernia, C., Lee, D. J., & Hernandez, V. H. (2006). The Increasing Financial Burden of Knee Revision Surgery in the United States. *Clinical Orthopaedics and Related Research, 446*, 221-226. doi:10.1097/01.blo.0000214424.67453.9a

Lavernia, C., Villa, J., Iacobelli, D. (2013). Readmission rates in the state of Florida: a reflection of quality. *Clin Orthop Relat Res*, 471(12), 3856.

Lee, E. (2012). Data mining application in customer relationship management for hospital inpatients. *Healthcare Informatics Research*, 18(3), 178-185.

Levinson, D.R. *Skilled Nursing Facilities Often Fail to Meet Care Planning and Discharge Planning Requirements (Executive Summary).* 2013. http://oig.hhs.gov/oei/reports/oei-02-09-00201.asp [accessed 8 December 2016]

Levinson, D.R. *Adverse Events in Skilled Nursing Facilities: National Incidence Among Medicare Beneficiaries (Executive Summary).* 2014. https://oig.hhs.gov/oei/reports/oei-06-11-00370.pdf [accessed 8 December 2016]

Liaw, A. & Wiener, M. (2002). Classification and regression by random forest. R News 2(3), 18--22.

Lie, S., Kiang, M., & Brusco, M. (2012). A unified framework for market segmentation and its applications. *Expert Systems with Applications*, 39, 10292-10302.

Liu S. & Chen, J. (2009). Using data mining to segment healthcare markets from patients' preference perspectives. *International Journal of Health Care Quality Assurance*, 22(2), 117-134.

Lynn, J., Straube, B., Bell, K., Jencks, S., & Kambic, R. (2007). Using population segmentation to provide better health care for all: The Bridges to Health Model. *The Milbank Quarterly*, 85(2), 185 – 208.

MacLennan, J. & Mackenzie, D. (2000). Strategic market segmentation: An opportunity to integrate medical and marketing activities. *International Journal Medical Marketing*, 1(1), 40 – 52.

Mahomed, N., Koo Seen Lin, M., Levesque, J., Lan, S., & Bogoch, E. (2000). Determinants and outcomes of inpatient versus home based rehabilitation following elective hip and knee replacement. *J Rheumatol*, 27(7), 1753-8.

Malhotra, N. (1989). Segmenting hospitals for improved management strategy. *Journal of Health Care Marketing*, 9(3), 45-52.

Mauerhan, D.R., Loneran, R.P., Mokris, J.G., & Kiebzab, G.M. (2003). Relationship between length of stay and dislocation rate after total hip arthroplasty. *J Arthroplasty*, 18(8), 963-7.

McDonald, A.P. & Kirk, R. (2013). Using lean Six Sigma to improve hospital based outpatient imaging satisfaction. *Radiology Management,* 35(1): 38-45.

McLaughlin, N., Burke, M.A., Setlur, N.P., Niedzwiecki, D.R., Kaplan, A.L., & Saigal, C. (2014) Time-driven activity based costing: a driver for provider engagement in costing activities and redesign initiatives. *Neurosurg Focus,* 37(5). doi: 10.3171/2014.8.Focus14381

MEDPAC: The Medicare Payment Advisory Commission. (March 2016). Report to Congress: Chapter 7 of Medicare Payment Policy: Skilled nursing facility services. Retrieved from http://www.medpac.gov/docs/default-source/reports/chapter-7-skilled-nursing-facility-services-march-2016-report-.pdf?sfvrsn=0 [accessed 8 December 2016]

Moss, H., Kirby, S., & Donodeo, F. (2009). Characterizing and reaching high-risk drinkers using audience segmentation. *Alcoholism Clinical and Experimental Research*, 33(8), 1336-1345.

Murphy, M. (2003). Eliminating wasteful work in hospitals improves margin, quality and culture. Murphy Leadership Institute Research Briefing.

Nelder, R.W. & Wedderburn, R.W. (1972). Generalized linear models. *Journal of the Royal Statistical Society*, 135(3), 370-384.

Newcomer, S., Steiner, J., & Bayliss, E. (2011). Identifying subgroups of complex patients with cluster analysis. *The American Journal of Managed Care*, 17(8), e324-332.

Nichols, C. I., & Vose, J. G. (2016). Clinical Outcomes and Costs Within 90 Days of Primary or Revision Total Joint Arthroplasty. *The Journal of Arthroplasty*, 31(7). doi:10.1016/j.arth.2016.01.022

Noguera, F., Fernandez Megia, M.J., Balasch Parisi, S., Edo Solsona, M.D., & Poveda Andres, J.L. (2013). Improving inpatient pharmacotherapeutic process by lean six sigma mythology. *Rev Calid Asist*, 28(6): 370-80.

Ōhno, T. (1988). *Toyota production system: Beyond large-scale production*. Cambridge, MA: Productivity Press.

Okike, K., O'toole, R. V., Pollak, A. N., Bishop, J. A., Mcandrew, C. M., Mehta, S., . . . Lebrun, C. T. (2014). Survey Finds Few Orthopedic Surgeons Know The Costs Of The Devices They Implant. *Health Affairs, 33*(1), 103-109. doi:10.1377/hlthaff.2013.0453

Oldmeadow, L.B., McBurney, H., Robertson, V.J. (2003). Predicting risk of extended inpatient rehabilitation after hip or knee arthroplasty. *J Arthroplasty*, 18, 775.

Pande, P. S., Neuman, R. P., & Cavanagh, R. R. (2000). *The Six Sigma way: How GE, Motorola, and other top companies are honing their performance*. New York: McGraw-Hill. Sharareh B, Le NB, Hoang MT, Schwarzkopf R. Factors determining discharge destination for patients undergoing total joint arthroplasty. *J Arthroplasty* 2014; 29(7): 1355-1358.

Parcells, B., Giacobbe, D., Macknet, D., Smith, A., Schottenfeld, M., Harwood, D., & Kayiaros, S. (2016). Total Joint Arthroplasty in a Stand-alone Ambulatory Surgical Center: Short-term Outcomes. *Orthopedics,* 39, 223-228. doi: 10.3928/01477447-20160419-06

Peel, T. N., Cheng, A. C., Liew, D., Buising, K. L., Lisik, J., Carroll, K. A., . . . Dowsey, M. M. (2015). Direct hospital cost determinants following hip and knee arthroplasty. *Arthritis Care & Research*, 67(6), 782-790. doi:10.1002/acr.22523

Penn State Hershey Medical Center (PSHMC, 2015) website, accessed on 10 December 2015 at http://www.pennstatehershey.org/web/guest/home/aboutus/history

Pires, G. & Stanton, J. (2008). Marketing issues in healthcare research. *International Journal on Behavioral and Healthcare Research,* 1(1), 38 – 60.

Porter, M. & Teisberg, E. (2006). *Redefining health care: Creating value-based competition on results*. Boston (Mass.): Harvard Business School Press.

Pugely, A. J., Martin, C. T., Gao, Y., Belatti, D. A., & Callaghan, J. J. (2014). Comorbidities in patients undergoing total knee arthroplasty: do they influence hospital costs and length of stay? *Clinical Orthopaedics and Related Research,* 472(12), 3943-3950. doi:10.1007/s11999-014-3918-x

Quan, H., Sundararajan, V., Halfon, P., Fong, A., Burnand, B., Luthi, J., . . . Ghali, W. A. (2005). Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Medical Care*, 43(11), 1130-1139. doi:10.1097/01.mlr.0000182534.19832.83

Ramkumar, P., Chu C., Harris, J., Athiviraham, A., Harrington, M., White, D., Berger, D., Naik, A. & Li L. (2015). Causes and rates of unplanned readmissions after elective primary total joint arthroplasty: a systematic review and meta-analysis. *American Journal of Orthopaedics*, 44(9), 397-405.

Ramos, N. L., Wang, E. L., Karia, R. J., Hutzler, L. H., Lajam, C. M., & Bosco, J. A. (2014). Correlation Between Physician Specific Discharge Costs, LOS, and 30-day Readmission Rates: An Analysis of 1,831 cases. *The Journal of Arthroplasty*, 29(9), 1717-1722. doi:10.1016/j.arth.2014.04.005

Robinson, J.C., Pozen, A., Tseng, S., & Bozic, K.J. (2012). Variability in costs associated with total hip and knee replacement implants. *J Bone Joint Surg Am*, 94 (18), 1693–8. Doi: 10.2106/JBJS.K.00355.

Ross, C., Steward, C., & Sinacore, J. (1993). The importance of patient preferences in the measurement of health care satisfaction. *Medical Care,* 31(12), 1138 – 1149.

Rubio, D., Schoenbaum, E., Lee, L., Schteingart, D., Marantz, P., Anderson, K., Platt, L., Baez, A., & Esposito, K. (2010). Defining translational research: implications for training. *Academic Medicine : Journal of the Association of American Medical Colleges*, 85(3), 470-475.

Rush, S. (2001). Logistic regression: the standard method of analysis in medical research. Technical Report Mathematics #S3, Trinity University.

R-Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

R-Core Team and contributors worldwide. (2016). The R Stats package. R package version 3.3.0. Retrieved January 28, 2016.

RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.

Santaguida, P. L., Hawker, G. A., Hudak, P. L., Glazier, R., Mahomed, N. N., Kreder†, H. J., Coyte, PC, & Wright, J. G. (2008). Patient characteristics affecting the prognosis of total hip and knee joint arthroplasty: a systematic review. *Canadian Journal of Surgery*, *51*(6), 428–436.

Schairer, W. W., Sing, D. C., Vail, T. P., & Bozic, K. J. (2013). Causes and frequency of unplanned hospital readmission after total hip arthroplasty. *Clinical Orthopaedics and Related Research*, 472(2), 464-470. doi:10.1007/s11999-013-3121-5

Schutzer, S. (2016). Practical tips for implementing bundled payments in your practice [American Association of Orthopedic Surgeons lecture notes]. Retrieved from http://www.aaos.org/.

Sheather, S. J. (2009*). A modern approach to regression with R*. New York: Springer.

Sherman, J. (2006). Achieving real results with Six Sigma. "Six Sigma to the rescue," declared the title of a June 2002 article in the technology section of Health Care Finance. Almost four years later, has Six Sigma helped healthcare organizations achieve the promised break-through improvement in their operations? *Healthcare Executive,* 21(1), 8-10.

Sievers, B. A., Negley, K. D., Carlson, M. L., Nelson, J. L., & Pearson, K. K. (2013). Enhancing Diabetes Management While Teaching Quality Improvement Methods. *The Journal of Continuing Education in Nursing*, 45(1), 14-19. doi:10.3928/00220124-20131223-02

Smith, C., Wood, S., & Beauvais, B. (2011). Thinking lean: implementing DMAIC methods to improve efficiency within a cystic fibrosis clinic. *Journal for Healthcare Quality*, 33(2), 37-46. doi:10.1111/j.1945-1474.2010.00130.x

Snow, R., Granata, J., Ruhil, A., Vogel, K., McShane, M., & Wasielewski, R. (2014). Associations between preoperative physical therapy and post-acute care utilization patterns and cost in total joint replacement. *J Bone Joint Surg Am*, 96(e165), 1-8.

SPSS. (2001). The SPSS TwoStep Cluster Component: A scalable component enabling more efficient customer segmentation. Technical Report. Accessed on 26 November 2015 from http://www.spss.ch/upload/1122644952_The%20SPSS%20TwoStep%20Cluster%20Component.pdf.

Suragh, T., Berg, C., & Nehl, E. (2013). Psychographic segments of college females and males in relation to substance use behaviors. *Social Marketing Quarterly*, 19(3), 172-187.

Sussmane, J. B., Torbati, D., & Gitlow, H. S. (2011). Measuring the quality of therapeutic apheresis care in the pediatric intensive care unit. *Journal of Clinical Apheresis*, 27(2), 43-50. doi:10.1002/jca.20318

Swenson, E., Bastian, N. & Nembhard, H. (2016a). Data analytics in health promotion: health market segmentation and classification of total joint replacement surgery patients. *Expert Systems with Applications*, 60: 118-129.

Swenson, E., Bastian, N. & Nembhard, H. (2016b). Healthcare market segmentation and data mining: a systematic review. *Health Marketing Quarterly*. Forthcoming.

Taner, M., Sezen, B., & Anthony, J. (2007). An overview of six sigma applications in healthcare industry. *International Journal of Health Care Quality Assurance*, 20(4): 329-40.

Taner, M. T., & Sezen, B. (2009). An application of Six Sigma methodology to turnover intentions in health care. *International Journal of Health Care Quality Assurance*, 22(3), 252-265. doi:10.1108/09526860910953520

Toledo, A., Carroll, T., Arnold, E., Tulu, Z., Caffey, T., Kearns, L., & Gerber, D. (2013). Reducing liver transplant length of stay: A Lean Six Sigma approach. *Progress in Transplantation*, 23(4), 350-364. doi:10.7182/pit2013226

Tynan, A. & Drayton, J. (1987). Market segmentation. *Journal of Marketing Management*, 2(3), 301-335.

Vaishya, R., Chauhan, M., & Vaish, A. (2013). Bone cement. *Journal of Clinical Orthopaedics and Trauma*, *4*(4), 157–163. http://doi.org/10.1016/j.jcot.2013.11.005

Van Citters, A.D., Fahlman, C., Goldmann, D.A., Lieberman, J.R., Koenig, K.M., DiGioia, A.M., O'Donnell, B., Federico, F.A., Bandowitz, R.A., Nelson, E.C., & Bozic, K.J. (2014). Developing a Pathway for High-value, Patient-centered Total Joint Arthroplasty. *Clinical Orthopaedics and Related Research*, 472(5), 1619-1635. doi:10.1007/s11999-013-3398-4

Walley, S. C., Berger, S., Harris, Y., Gallizzi, G., & Hayes, L. (2013). Decreasing Patient Identification Band Errors by Standardizing Processes. *Hospital Pediatrics*, 3(2), 108-117. doi:10.1542/hpeds.2012-0075

Walter, F.L., Bass, N., Bock, G., & Markel, D.C. (2007). Success of Clinical Pathways for Total Joint Arthroplasty in a Community Hospital. *Clinical Orthopaedics and Related Research*, 457(4):133-7. doi:10.1097/01.blo.0000246567.88585.0a

Ward, J. (1963), "Hierarchical grouping to optimize an objective function", *Journal of the American Statistical Association*, 58, 236–244.

Warner, C.J., Walsh, D.B., Horvath, A.J., Walsh, T.R., Herrick, D.P., Prentiss, S.J., & Powell, R.J. (2013). Lean principles optimize on-time vascular surgery operating room starts and decrease resident work hours. *Journal of Vascular Surgery*, 58(5), 1417-1422. doi:10.1016/j.jvs.2013.05.007

Wasey, J. (2015). icd9: Tools for working with ICD-9 codes, and finding comorbidities. R package version 1.3.http://CRAN.R-project.org/package=icd9

Weingarten, S., Riedinger, M.S., Sandhu, M., Bowers, C., Ellrodt, A.G., Nunn, C., Hobson, P., & Greegold, N. (1998). Can practice guidelines safely reduce hospital length of stay? Results from a multicenter interventional study. *American Journal of Medicine*, 105(1), 33–40

Wilson, N. A., Schneller, E. S., Montgomery, K., & Bozic, K. J. (2008). Hip And Knee Implants: Current Trends And Policy Considerations. *Health Affairs*, *27*(6), 1587-1598. doi:10.1377/hlthaff.27.6.1587

Wind, Y. (1978). Issues and advances in segmentation research, *Journal of Marketing Research*, 15, 317-338.

Woodside, A., Nielson, R., Walters, R., & Muller G. (1998). Preference segmentation of health care services: the old-fashioneds, value conscious, affluents, and professional want-it-alls. *Journal of Health Care Marketing*, 8(2), 14-24.

World Health Organization. (2014). Health promotion. Accessed 13 November 2014, available from http://www.who.int/topics/health_promotion/en/.

Wu, H., Lin, S., & Liu, C. (2014). Analyzing patients' values by applying cluster analysis and LRFM model in a pediatric dental clinic in Taiwan. *The Scientific World Journal*, 2014, 1-7.

Zeileis, A. & Hothorn, T. (2002). Diagnostic checking in regression relationships. R News 2(3), 7-10. URL http://CRAN.R-project.org/doc/Rnews/

This page is intentionally left blank.

Appendix A

IRB Document

PENNSTATE HERSHEY Milton S. Hershey Medical Center

PENNSTATE HERSHEY College of Medicine

APPROVAL OF SUBMISSION

Date: September 18, 2015

From: Heidi Watts, IRB Analyst

To: Eric Swenson

Type of Submission:	Initial Study
Title of Study:	Understanding Process Flow in total joint care
Principal Investigator:	Eric Swenson
Study ID:	STUDY00002554
Submission ID:	STUDY00002554
Funding:	Not Applicable
IND,IDE, or HDE:	Not Applicable
Documents Approved:	• HRP-585 (17.02), Category: Consent Form • Dr Davis memo (10 September 2015), Category: Other • HRP 598 (12 August 2015), Category: IRB Protocol • retrospective data fields.xls (18 August 2015), Category: Other • STUDY2554_Protocol_IRBedits_17 sep 2015 (18.01), Category: IRB Protocol • Interview sheet (12 August 2015), Category: Data Collection Instrument
Review Level:	Expedited

On 9/18/2015, the IRB approved the above-referenced Initial Study. This approval is effective through 9/17/2016 inclusive. You must submit a continuing review form with all required explanations for this study at least 45 days before the study's approval end date. You can submit a continuing review by navigating to the active study and clicking 'Create Modification / CR'.

If continuing review approval is not granted before 9/17/2016, approval of this study expires on that date.
To document consent, use the consent documents that were approved and stamped by the IRB. Go to the Documents tab to download them.

In conducting this study, you are required to follow the requirements listed in the Investigator Manual (HRP-103), which can be found by navigating to the IRB Library within CATS IRB (http://irb.psu.edu). These requirements include, but are not limited to:

- Documenting consent
- Requesting modification(s)

IDO0000003

Penn State Milton S. Hershey Medical Center • Penn State College of Medicine
Human Subjects Protection Office, Institutional Review Board, Mail Code A115, Academic Support Building, Room 1140
90 Hope Drive, P.O. Box 855, Hershey, PA 17033-0855 • Tel: 717-531-5687 • Fax: 717-531-3937 • www.hmc.psu.edu/irb

PENNSTATE HERSHEY
Milton S. Hershey Medical Center

PENNSTATE HERSHEY
College of Medicine

- Requesting continuing review
- Closing a study
- Reporting new information about a study
- Registering an applicable clinical trial
- Maintaining research records

This correspondence should be maintained with your records.

Penn State Milton S. Hershey Medical Center • Penn State College of Medicine
Human Subjects Protection Office, Institutional Review Board, Mail Code A115, Academic Support Building, Room 1140
90 Hope Drive, P.O. Box 855, Hershey, PA 17033-0855 • Tel: 717-531-5687 • Fax: 717-531-3937 • www.hmc.psu.edu/irb

Appendix B

Logistic Regression Output for Models 3-6 using Firth's Penalized MLE.

Table 5a: Logistic Regression Output for Model 3 using Firth's Penalized MLE.

Coefficients	**OR**	**Standard Error**	**Pr(>|t|)**	**2.50%**	**97.50%**
(Intercept)	0.002	1.769	**0.000**	0.000	0.047
Female	2.958	0.489	**0.041**	1.047	9.168
ASA2	0.540	1.537	0.721	0.041	79.660
ASA3	0.355	1.560	0.571	0.025	54.086
ASA4	0.343	2.310	0.659	0.001	98.433
Latino	7.573	1.060	0.107	0.591	54.837
Non-smoker	1.376	0.470	0.531	0.510	3.917
Current smoker	0.509	0.905	0.483	0.043	2.949
MD1	2.140	0.522	0.178	0.713	7.197
MD2	0.554	0.727	0.452	0.101	2.523
THA (hip)	3.855	0.467	**0.008**	1.426	11.509
DRG469	0.404	1.801	0.641	0.002	12.717
Discharge HHC	2.692	0.807	0.256	0.443	12.632
Discharge SNF	0.457	1.450	0.633	0.003	7.225
Discharge Rehab	2.196	0.637	0.268	0.523	8.276
Discharge Other	22.928	0.979	**0.005**	2.898	154.228
#Comorbidities	2.012	0.217	**0.003**	1.265	3.275
Charlson score	1.164	0.209	0.520	0.726	1.825
Cluster 2	1.331	0.687	0.701	0.294	5.879
Cluster 3	0.458	1.340	0.570	0.003	4.658
Cluster 4	2.177	0.738	0.337	0.425	10.720
Cluster 5	3.476	0.629	**0.056**	0.970	14.134
Cluster 6	2.347	1.372	0.639	0.017	38.502
AIC	-1.09				

Table 6a: Logistic Regression Output for Model 4 using Firth's Penalized MLE.

Coefficients	**OR**	**Standard Error**	**Pr(>|t|)**	**2.50%**	**97.50%**
(Intercept)	0.0019	1.6757	**7.28E-06**	0.0000	0.0426
Female	3.5052	0.5195	**3.59E-02**	1.0832	17.8691
ASA2	0.3657	1.4723	5.68E-01	0.0235	50.6902
ASA3	0.1851	1.5059	3.75E-01	0.0102	26.1906
ASA4	0.2885	2.1923	6.11E-01	0.0003	78.3604
Latino	7.1122	1.1410	1.47E-01	0.4459	68.1616
Non-smoker	1.3859	0.4741	5.42E-01	0.4870	4.2537
Current smoker	0.3107	0.9583	2.52E-01	0.0143	2.0887
MD1	1.4928	0.5141	5.02E-01	0.4644	5.3240
MD2	0.4290	0.7496	3.27E-01	0.0706	2.3246
THA (hip)	3.9274	0.4730	**9.84E-03**	1.3789	13.7540
DRG469	0.3817	1.8795	6.54E-01	0.0010	106.8986
Discharge HHC	7.5755	0.8592	**4.66E-02**	1.0338	53.4423
Discharge SNF	0.8545	1.1607	9.16E-01	0.0072	11.0522
Discharge Rehab	1.2028	0.6811	8.18E-01	0.2144	5.5115
Discharge Other	17.0027	1.0671	**2.03E-02**	1.6136	151.2806
CHF	0.2978	1.6942	4.98E-01	0.0007	7.3774
Arrhythmia	3.9823	0.6303	**9.22E-02**	0.7789	17.7525
Valvular	1.7166	1.0548	6.44E-01	0.1311	13.9721
PHTN	0.4857	1.2848	6.05E-01	0.0156	5.9754
PVD	0.9900	1.5788	9.95E-01	0.0043	13.5565
HTN	4.7197	0.6043	**1.62E-02**	1.2952	28.3608
NeuroOther	29.9192	1.3381	**2.86E-02**	1.5034	548.9108
Pulmonary	2.6689	0.8098	2.76E-01	0.4382	19.0402
DM	3.5230	2.2886	7.12E-01	0.0019	1602.3780
DMcx	3.0953	0.7847	2.18E-01	0.4887	17.5170
Hypothyroid	3.1665	0.5527	8.69E-02	0.8329	11.2393
Renal	1.2911	1.2798	8.66E-01	0.0098	16.4759
Liver	5.3583	1.2018	2.04E-01	0.3820	77.3504
PUD	25.7577	1.9898	1.84E-01	0.1091	1003.2950
HIV	2.0773	3.2435	8.40E-01	0.0006	6056.6150
Lymphoma	8.2401	1.0039	**5.41E-02**	0.9584	63.1755
Mets	2.7670	3.0873	7.15E-01	0.0057	648.1056
Tumor	1.7288	0.6413	4.59E-01	0.3690	6.6648
Rheumatic	3.0291	1.9318	5.81E-01	0.0161	85.6733
Coagulopathy	1.3076	1.9322	9.20E-01	0.0002	65.3360
Obesity	3.7107	0.7440	1.35E-01	0.6364	19.0024
FluidsLytes	3.6201	0.8049	1.63E-01	0.5459	23.8830
Anemia	1.2813	0.6248	7.28E-01	0.2880	4.8654

Alcohol	1.2850	1.8200	8.89E-01	0.0079	25.7417
Drugs	13.5489	1.9016	2.50E-01	0.0673	428.8524
Psychoses	5.0204	1.5170	4.20E-01	0.0241	78.2661
Depression	0.2874	0.8701	1.75E-01	0.0199	1.6205
Charlson score	1.4294	0.3392	3.56E-01	0.6485	3.0107
Cluster 2	2.2411	0.6871	3.15E-01	0.4647	16.3794
Cluster 3	0.3299	1.3444	4.66E-01	0.0011	4.4135
Cluster 4	2.5997	0.7329	2.56E-01	0.4857	13.9992
Cluster 5	3.2122	0.6541	1.15E-01	0.7492	15.8470
Cluster 6	2.5556	1.2498	5.93E-01	0.0011	42.0592
AIC	30.87				

Table 7a: Logistic Regression Output for Model 5 using Firth's Penalized MLE.

Coefficients	**OR**	**Standard Error**	**Pr(>|t|)**	**2.50%**	**97.50%**
(Intercept)	0.0109	1.5818	0.0001	0.0001	0.1402
Female	1.7611	0.3803	0.1619	0.7960	3.9537
ASA2	0.8405	1.4834	0.9111	0.0848	114.4039
ASA3	1.0079	1.4869	0.9959	0.1006	138.0492
ASA4	0.8803	2.2089	0.9534	0.0040	196.5511
Latino	5.7852	0.8044	**0.0527**	0.9761	25.1928
Non-smoker	0.9272	0.3727	0.8484	0.4275	2.0398
Current smoker	0.5400	0.6711	0.3653	0.1024	1.9236
MD1	1.1527	0.3985	0.7372	0.5032	2.6981
MD2	0.8575	0.5175	0.7776	0.2770	2.4366
THA (hip)	2.4188	0.3647	**0.0226**	1.1322	5.2969
DRG469	0.6448	1.7138	0.8041	0.0034	11.2972
Discharge HHC	2.0502	0.7292	0.3662	0.3724	7.7991
Discharge SNF	1.8182	0.7527	0.4742	0.3023	7.6590
Discharge Rehab	1.8647	0.5063	0.2515	0.6257	5.1258
Discharge Other	7.5011	0.8834	**0.0366**	1.1541	36.8040
#Comorbidities	1.4894	0.1678	**0.0267**	1.0476	2.1174
Charlson score	1.0182	0.1815	0.9264	0.6787	1.4651
Cluster 2	1.5004	0.5131	0.4490	0.5216	4.4051
Cluster 3	0.5760	0.8526	0.5233	0.0596	2.7582
Cluster 4	1.9363	0.5652	0.2693	0.5892	6.2531
Cluster 5	1.6113	0.5145	0.3730	0.5590	4.7210
Cluster 6	0.5008	1.3062	0.6223	0.0038	5.3192
AIC	12.49				

Table 8a: Logistic Regression Output for Model 6 using Firth's Penalized MLE.

Coefficients	**OR**	**Standard Error**	**Pr(>|t|)**	**2.50%**	**97.50%**
(Intercept)	0.0106	1.512508	**8.21E-05**	6.98E-05	0.1380575
Female	2.1281	0.408258	9.59E-02	8.76E-01	5.389863
ASA2	0.812	1.408285	8.94E-01	8.19E-02	110.253041
ASA3	0.7996	1.42412	8.87E-01	7.51E-02	110.071134
ASA4	0.2714	2.473581	5.94E-01	9.06E-04	79.1574464
Latino	6.5864	0.839541	**4.64E-02**	1.03E+00	32.4853912
Non-smoker	0.8665	0.385338	7.36E-01	3.77E-01	2.0243939
Current smoker	0.307	0.719431	1.05E-01	4.73E-02	1.2529072
MD1	0.9417	0.407807	8.93E-01	3.93E-01	2.2864013
MD2	0.6937	0.538954	5.27E-01	2.03E-01	2.1009708
THA (hip)	2.467	0.369811	**2.31E-02**	1.13E+00	5.5397405
DRG469	1.7691	1.687145	7.85E-01	4.80E-03	40.8693866
Discharge HHC	1.5337	0.793721	6.03E-01	2.59E-01	6.5155359
Discharge SNF	3.0234	0.705678	1.83E-01	5.48E-01	12.4433305
Discharge Rehab	1.6564	0.529158	3.82E-01	5.10E-01	4.8107807
Discharge Other	7.0251	0.961318	6.47E-02	8.74E-01	40.6571604
CHF	0.0925	1.991473	1.82E-01	4.60E-04	2.6960245
Arrhythmia	2.4201	0.516893	1.46E-01	7.16E-01	7.1381878
Valvular	2.2117	0.801653	3.63E-01	3.43E-01	9.7325195
PHTN	0.8317	1.147495	8.76E-01	5.90E-02	7.5719478
PVD	0.381	1.513858	4.95E-01	2.52E-03	3.9127411
HTN	1.4813	0.391955	3.48E-01	6.61E-01	3.5905987
NeuroOther	7.5059	1.185654	1.10E-01	5.81E-01	79.0387444
Pulmonary	4.9463	0.659295	**2.85E-02**	1.19E+00	21.4816066
DM	3.0208	1.820601	5.82E-01	1.25E-02	90.3656959
DMcx	3.1484	0.653934	1.06E-01	7.68E-01	11.4120727
Hypothyroid	1.045	0.499811	9.36E-01	3.21E-01	2.8618068
Renal	0.378	1.441589	5.08E-01	2.66E-03	4.6578912
Liver	4.7486	1.040852	1.55E-01	5.28E-01	38.1040563
PUD	13.429	1.38915	6.51E-02	8.27E-01	154.566525
HIV	0.8746	3.535367	9.71E-01	7.27E-04	3915.97793
Lymphoma	4.3066	0.885577	1.13E-01	6.75E-01	22.2901714
Mets	8.3018	2.882297	3.99E-01	2.73E-02	903.615114
Tumor	2.0468	0.484482	1.80E-01	7.00E-01	5.3647804
Rheumatic	2.1872	1.757215	6.59E-01	1.45E-02	33.603503

Coagulopathy	1.4259	1.631465	8.42E-01	4.78E-03	22.2258864
Obesity	2.1058	0.64365	3.09E-01	4.65E-01	7.9098903
FluidsLytes	1.4726	0.764977	6.33E-01	2.60E-01	6.4292723
Anemia	1.7201	0.468159	2.93E-01	6.04E-01	4.4364945
Alcohol	0.4412	1.817849	6.14E-01	2.46E-03	6.9053621
Drugs	3.0347	1.845807	5.68E-01	1.74E-02	58.2729103
Psychoses	1.3651	1.497832	8.47E-01	9.08E-03	15.6109599
Depression	1.207	0.506627	7.31E-01	3.81E-01	3.3222484
Charlson score	0.9631	0.282334	8.98E-01	5.14E-01	1.6649687
Cluster 2	2.4185	0.533789	1.21E-01	7.92E-01	7.7948163
Cluster 3	0.7225	0.807257	7.15E-01	7.56E-02	3.5809879
Cluster 4	2.2086	0.573278	1.98E-01	6.53E-01	7.4642172
Cluster 5	1.8182	0.548677	3.09E-01	5.69E-01	5.8585905
Cluster 6	0.4873	1.382023	6.65E-01	1.25E-03	7.3623905
AIC	43.88				

VITA

Eric R. Swenson

Eric R. Swenson was born and raised in Churchville, Pennsylvania where he attending elementary, middle, and high school in the Council Rock School district. He played high school football for four years before graduating and starting his military career at the United States Military Academy at West Point, New York. Eric graduated from West Point in 1998 and was commissioned a 2nd Lieutenant in the United States Army Corps of Engineers. He graduated with a Bachelor of Science in Systems Engineering. In 2006, he moved to State College, PA and started his graduate studies. In 2008, Eric graduated from the Pennsylvania State University with a Dual Master's Degree in Operations Research and Industrial Engineering. From 2008-2010, he served as an instructor in the Department of Mathematical Sciences at the United States Military Academy. From 2010-2011, Eric attended the Naval Command and Staff College and graduated in June 2011 with a Master of Arts in National Security and Strategic Studies. In 2014, Eric returned to Penn State to work under the tutelage of Dr. Harriet Nembhard in the area of healthcare delivery science.