Title
Understanding Variation of Intraoperative Disposable Supply Costs in Laparoscopic Cholecystectomy

Author
Childers, Christopher

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Understanding Variation of Intraoperative Disposable Supply Costs in Laparoscopic Cholecystectomy

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Health Policy and Management

by

Christopher Peter Childers

2019

ABSTRACT OF THE DISSERTATION

Understanding Variation of Intraoperative Disposable Supply Costs

in Laparoscopic Cholecystectomy

by

Christopher Peter Childers

Doctor of Philosophy in Health Policy and Management

University of California, Los Angeles, 2019

Professor Susan L. Ettner, Chair

The rising cost of healthcare in the United States is arguably one of the largest challenges of the next generation. In surgery, the operating room is the most resource dense component of a patient's care, estimated to cost $37 per minute. While much of this cost is not easily modifiable – such as wages and indirect costs - the supplies used by surgeons are tangible and potentially mutable. Preliminary research suggests significant variation in supply use and associated costs between surgeons, but little is known about surgeons' knowledge of the cost of instruments, how this knowledge affects instrument preferences, and how use of different instruments impacts patient outcomes. In this dissertation, we first conducted a multi-institutional survey of 83 attending surgeons at 3 academic medical centers in Southern California. We then linked survey data from one medical center to a comprehensive medical record-based administrative dataset that included over 1800 laparoscopic cholecystectomies performed between 2013 and 2018. Our analysis found that there was significant variation in disposable supply costs between surgeons and facilities, even after adjustment for patient case mix. Together, the surgeon and facility explained 34% of the variation in supply costs. Our analysis further suggested that the cornerstone of intraoperative supply costs appears to be the surgeon's preference card , with every 1 dollar increase in preference card cost associated with a 78 cents increase in actual case cost. Upstream, surgeons who were able to accurately discriminate the cost of common general surgery items may choose cheaper preference cards, but interestingly, passive exposure to instrument costs, such as through cost report cards, does not appear to increase cost knowledge and therefore may have little downstream effect. Finally, surgeons with more expensive preference cards do not appear to improve their patients' short-term outcomes. Taken together, the results suggest that the most successful efforts to reduce intraoperative supply cost variation will involve an active approach to reducing preference card cost, perhaps through standardization, and that cost reducing efforts may not adversely affect patient outcomes.

The dissertation of Christopher Peter Childers is approved.

Ron D. Hays

Gerald Kominski

Melinda Maggard-Gibbons

Susan L. Ettner, Committee Chair

University of California, Los Angeles

2019

To Kim -

For your endless love, support, and encouragement;

For tolerating my relentless need to delay gratification;

and For growing our beautiful Kaitlyn in the time it took me to write Chapters 6 & 7.

Table of Contents

List of Figures

List of Tables

List of Appendices

List of Abbreviations & Acronyms

ASA = American Society of Anesthesiologists
ASC = Ambulatory surgery center
BMI = Body mass index
CCR = Cost-to-charge ratio
CPT = Current Procedural Terminology
CSMC = Cedars Sinai Medical Center
DRG = Diagnosis Related Group
ED = Emergency department
FDA = Food and Drug Administration
GDP = Gross Domestic Product
ICU = Intensive Care Unit
LOS = Length of stay
NSQIP = National Surgical Quality Improvement Project
OR = Operating room
RR = Ronald Reagan
SM = Santa Monica
TDABC = Time Driven Activity Based Costing
UCLA = University of California, Los Angeles
UCSD = University of California, San Diego
U.S. = United States

Acknowledgements

This dissertation was supported by the Agency for Healthcare Research & Quality (F32HS02507), the Association for Academic Surgeons / Association for Academic Surgeons Foundation Clinical Outcomes Research Fellowship, and the University of California, Los Angeles (UCLA) Graduated Division Dissertation Year Fellowship.

I would like to recognize the following co-authors and contributors for their assistance with this dissertation:

Drs. Mark Zhao and Bryan Clary from the Department of Surgery at UC San Diego and Drs. Joshua Tseng and Rodrigo Alban from the Department of Surgery at Cedars Sinai Medical Center helped refine the survey instrument, including identifying appropriate supply comparisons at their local institution, and served as surgeon champions to enhance the survey response rate. Dr. Mark Zhao performed two cognitive interviews with UC San Diego faculty that were invaluable to the survey instrument generation process.

Drs. Ira Hofer, Drew Cheng, and Eilon Gabel in the Department of Anesthesiology at UCLA generated the administrative data extract using the UCLA Perioperative Data Warehouse that was the cornerstone of this dissertation. Mr. Victor Banh and David Apgar were invaluable resources at UCLA who provided insight into the hospital supply chain and electronic medical record systems, helped resolve missing and aberrant instrument prices, and provided access to surgeon preference cards.

Brief Curriculum Vitae

EDUCATION & TRAINING

2014-Present	**University of California, Los Angeles** Los Angeles, CA General Surgery Residency
2010-2014	**Northwestern University, Feinberg School of Medicine** Chicago, IL Medical Doctorate (M.D.)
2005-2010	**University of California, San Diego** La Jolla, CA B.S. in Human Biology with Distinction B.S. in Management Science with Distinction

PEER REVIEWED PUBLICATIONS

1. **Childers CP,** Dworsky JQ, Kominski G, Maggard-Gibbons M. A comparison of payments to a for-profit dialysis firm from government and commercial insurers. *JAMA Internal Medicine*. Accepted.

2. **Childers CP,** Maggard-Gibbons M, Nuckols T. California Teaching Hospitals Achieved Slower Growth in Operating Room Costs from 2005 to 2014. *Academic Medicine*. Accepted.

3. **Childers CP,** Dworsky JQ, Massoumi RL, et al. The contemporary appendectomy for acute uncomplicated appendicitis in children. *Surgery*. Accepted.

4. **Childers CP**, Dworksy JQ, Russell MM, Maggard-Gibbons M. Estimating the Bias of Using Hospital-Wide Cost-to-charge Ratios in Surgery. *JAMA Surgery*. Accepted.

5. Stamp M, Gordon O, **Childers CP,** Childers KK. Painting a Portrait: Analysis of National Health Survey Data for Cancer Genetic Counseling. *Cancer Medicine*. Accepted.

6. **Childers CP**, Dworsky JQ, Maggard-Gibbons M, Russell MM. The contemporary appendectomy for acute uncomplicated appendicitis in adults. *Surgery.* 2018.

7. **Childers CP**, Hofer IS, Cheng DS, Maggard-Gibbons M. Evaluating Surgeons on Intraoperative Disposable Supply Costs: Details Matter. *J Gastrointest Surg.* 2018.

8. **Childers CP**, Maggard-Gibbons M. Estimation of the Acquisition and Operating Costs for Robotic Surgery. *JAMA.* 2018;320(8):835-836.

9. Siletz A, **Childers CP**, Faltermeier C, et al. Surgical Technical Evidence Review of Hip Fracture Surgery Conducted for the AHRQ Safety Program for Improving Surgical Care and Recovery. *Geriatr Orthop Surg Rehabil.* 2018;9:2151459318769215.

10. O'Neill SM, Miake-Lye I, **Childers CP**, et al. Effectiveness of interventions to improve the efficiency of emergency department operations: An evidence map. *American Journal of Emergency Medicine.* 2018;36(12):2314-2317.

11. **Childers CP**, Showen A, Nuckols T, Maggard-Gibbons M. Interventions to Reduce Intraoperative Costs: A Systematic Review. *Ann Surg.* 2018;268(1):48-57.

12. Childers KK, Maggard-Gibbons M, Macinko J, **Childers CP**. National Distribution of Cancer Genetic Testing in the United States: Evidence for a Gender Disparity in Hereditary Breast and Ovarian Cancer. *Jama Oncology.* 2018;4(6):876-879.

13. **Childers CP**, Maggard-Gibbons M. Understanding Costs of Care in the Operating Room. *Jama Surgery.* 2018;153(4).

14. **Childers CP**, Siletz AE, Singer ES, et al. Surgical Technical Evidence Review for Elective Total Joint Replacement Conducted for the AHRQ Safety Program for Improving Surgical Care and Recovery. *Geriatr Orthop Surg Rehabil.* 2018;9:2151458518754451.

15. **Childers CP**, Maggard-Gibbons M, Ulloa JG, et al. Perioperative management of antiplatelet therapy in patients undergoing non-cardiac surgery following coronary stent placement: a systematic review. *Systematic Reviews.* 2018;7.

16. **Childers CP**, Childers KK, Maggard-Gibbons M, Macinko J. Reply to S. Nakamura et al and S. Narod et al. *J Clin Oncol.* 2018;36(5):517-518.

17. **Childers CP**, Childers KK, Maggard-Gibbons M, Macinko J. National Estimates of Genetic Testing in Women With a History of Breast or Ovarian Cancer. *J Clin Oncol.* 2017;35(34):3800-3806.

18. **Childers CP**, Maggard-Gibbons M, Shekelle PG. Antiplatelet Therapy in Patients With Coronary Stents Undergoing Elective Noncardiac Surgery: Continue, Stop, or Something in Between? *JAMA.* 2017;318(2):120-121.

19. O'Neill SM, **Childers CP**, Maggard-Gibbons M. Improving the value of same-day pediatric orthopaedic surgery by using ambulatory surgery centers: direct cost savings are important, but there is more to the picture. *Annals of Joint*. 2017;2(6).

Chapter 1: Introduction, Questions, and Hypotheses

In 2015, healthcare expenditures in the United States (U.S.) topped $3.2 trillion, accounting for 18% of all spending.[1] Surgical care expenses account for almost one-third of total healthcare spending and are projected to account for 1/14th of the U.S. gross domestic product (GDP) by 2025.[2]

The operating room (OR) is a resource dense environment estimated to cost $37 per minute to fully operate.[3] For a surgical episode, these costs are only exceeded by the cost of room and board.[4] Estimates vary, but between 10% and 50% of OR costs are attributable to the supplies used by surgeons.[3,5]

Before an operation begins, such as the removal of a gallbladder (i.e., cholecystectomy), supplies (e.g., gauze, sutures, and electrocautery devices) requested by the surgeon are opened and organized. This initial list of items is recorded a priori on a surgeon's preference card – a list typically generated when a surgeon joins a practice or first performs the procedure. As the surgery proceeds, additional items are opened to complete the case.

Surgical residents train under the supervision of attending surgeons during their residency. Technique and skill sets are subsequently refined throughout their career. While textbooks and online instructional videos outline the basic steps of an operation, the nuances of technique and instrumentation are not prescribed and must ultimately be acquired hands on.

This model of training and developing an individual surgical technique may result in significant variation in the instruments used to perform an operation. As with variation in other sectors of healthcare, the concern is that this may lead to added cost without patient benefit. Despite the immense expense of the OR, cost information is not routinely provided to surgeons.

National efforts such as bundled payments are underway to reduces costs,[6] but surgeons and hospital administrators have little information to guide them in these efforts, especially in the OR. Research is needed to understand surgeons' knowledge of costs, how this knowledge impacts instrument preferences, variability in the cost of the actual instruments used, and how instrument variability impacts patient outcomes. Unfortunately, studies to date are limited, and the studies that do exist typically evaluate only one of these topics in isolation. This study seeks to further understand the complex interactions between these topics by answering the following four questions:

1. **What factors predict a surgeon's accuracy of knowledge of the relative cost of surgical instruments?**

   1a. Hypothesis: Surgeons with prior exposure to instrument pricing will have more accurate knowledge about the relative cost of surgical instruments.[1]

   Rationale: Instrument prices are not routinely disclosed to healthcare providers; previous exposure to prices should increase the likelihood of a surgeon knowing the relative cost of an instrument.

2. **How is accuracy of knowledge of the relative cost of surgical instruments associated with the cost of a surgeon's instrument preference card for laparoscopic cholecystectomy?**

[1] After adjustment for surgeon-level covariates (e.g., the relative importance surgeons place on cost versus effectiveness).

2a. Hypothesis: Surgeons with more accurate knowledge of the relative cost of surgical instruments will prefer a lower-cost combination of devices.

Rationale: When comparing two or more instruments that are designed to perform the same (or similar) task, surgeons with more accurate knowledge of the relative cost of those instruments will prefer the lower-cost instrument, all else equal.[2]

2b. Hypothesis: The association between a surgeon's accuracy of the relative cost of instruments and the cost of an instrument preference card will be stronger among surgeons who view cost minimization as more important.

Rationale: Cost knowledge is necessary but not sufficient for surgeons to choose less expensive instruments; to prefer a lower-cost instrument list, surgeons must both have cost knowledge and believe cost is an important consideration when choosing instruments.

3. For laparoscopic cholecystectomy, how much of intraoperative cost variation is explained at the patient, surgeon, and facility levels?

3a. Hypothesis: More of the intraoperative cost variation will be explained at the surgeon compared to the patient or facility level.

Rationale: With some exceptions,[3] surgeons are allowed to pick the combination of instruments they deem necessary for an operation. Surgeons generally have a

[2] Sensitivity analyses will be performed to assess the extent to which unmeasured perceived effectiveness may bias these associations.

[3] Facilities have a finite number of contracts and vendors from which they acquire supplies. Some supplies may only be available to certain surgeons. For example, a surgeon may request an expensive glue be available; the committee that makes this decision agrees to stock the product but only in limited quantities for that surgeon.

preferred technique for performing laparoscopic cholecystectomy that does not change significantly from one patient to the next.

4. How does a surgeon's a priori preference card for laparoscopic cholecystectomy influence intraoperative costs and patient outcomes?

4a. Hypothesis: There will be a significant positive association between the sum of instrument prices listed on the a priori preference card and the sum of instrument prices used in the OR.

i. *Rationale:* Unless an operation deviates substantially from average (e.g., an extremely complex case), surgeons will rely on the instruments outlined on their preference card.

4b. Hypothesis: Controlling for patient factors and facility, higher-cost preference cards will be associated with better intraoperative and postoperative outcomes.

i. *Rationale*: Surgeons may use more expensive instruments because they believe they will improve intraoperative outcomes (e.g., reduced operative time) or postoperative outcomes (e.g., decreased length of stay, fewer complications).

Chapter 2: Background and Definitions

2-1: The Cost of U.S. Healthcare and the Drivers of Increased Spending

Healthcare in the U.S. is expensive. In 2015, national health expenditures topped $3.2 trillion, accounting for 18% of all U.S. spending or $9,990 per person.[1] The growth of healthcare spending has been unbridled since the 1960's (Figure 1), outpacing the growth of the overall economy in 46 of the past 55 years (Figure 2).[1] Although healthcare is ubiquitously expensive across the globe, U.S. spending is disproportionately high – almost 50% greater than the next-highest developed country.[7] Despite this added cost, the U.S. has little to show for the investment with the lowest average life expectancy, highest infant mortality, and highest chronic disease burden among 13 high-income countries.[7]

Estimates of surgical expenditures are difficult to come by; however, a model developed by Dr. Francis Moore in the 1980's, and confirmed in a subsequent 2010 study, estimated surgical expenditures at 29% of total healthcare spending.[2,8] Projecting into the future, this model estimated that surgical spending alone would account for 1/14th of all U.S. spending by 2025.[2]

Numerous papers have explored the drivers of excessive spending in the U.S. A 2012 report from the Bipartisan Policy Center outlines a number of these drivers and describes the evidence associating each with increased spending.[9] Many are unavoidable - such as the aging population and growth in chronic disease – while others are structural and may require radical changes to reimbursement, medical malpractice, fragmentation of care, and the tax treatment of health insurance. Often cited as a dominant force behind increasing spending is the rapid advance in medical technology. Technologies that broaden the available options without increasing efficiency can be particularly problematic.[7,10,11] Lack of transparency related to the cost of these technologies prevents providers and patients from making cost-effective decisions.[9] The result is high unit prices, evidenced by the fact that the U.S. spends far more on prescription drugs, diagnostic imaging, and procedures than similarly situated countries.[9,12]

2-2: Surgical Technology and Innovation

Perhaps no field is more susceptible to the impact of medical technology and price opacity than surgery. Over 20 million invasive surgical procedures were performed in U.S. hospitals in 2014, at a cost of nearly $1 trillion.[2,13] New technologies are often rapidly adopted in surgery. An exemplar is laparoscopic cholecystectomy, the surgical removal of the gallbladder in a minimally invasive fashion. The gallbladder is a light bulb sized structure that sits on the underside of the liver. The primary purpose of the gallbladder is to store and concentrate bile, a fluid that, on the one hand, helps excrete toxins filtered by the liver, and on the other, acts as a powerful lipid emulsifier.[14] Approximately 10-15% of U.S. adults will develop stones in their gallbladder (i.e., gallstones) at some point in their life,[15] with 20% developing symptoms or complications potentially warranting surgical removal.[16] Prior to the 1980's, open cholecystectomy was the gold standard, involving a large incision in the upper right abdomen with significant risk of morbidity.[17] Patients stayed in the hospital for 5 days postoperatively and recovered over the span of 3-6 weeks.[18] A minimally invasive mechanism (i.e., laparoscopic) for performing this operation was introduced in 1989, substituting one large incision with several small incisions, a camera, and specially designed instruments. Laparoscopic cholecystectomy adoption was 99% within 3 years.[19] Subsequent studies demonstrated improved mortality, reduced length of stay, and shorter recovery times for patients undergoing laparoscopic compared to open cholecystectomy.[18,20] Interestingly, studies also found increased rates of cholecystectomy after the advent of the minimally invasive technique, perhaps because of a reduced threshold for proceeding with the operation.[20,21]

Since 1989, laparoscopic cholecystectomy has remained the gold standard, despite numerous attempts at innovation such as robotic,[22] single incision,[23] and natural orifice surgery.[24] Laparoscopic cholecystectomy is one of the most common operations in the U.S., with 950,000 performed in 2014.[13] At an estimated $22,000 per case (hospital cost, 2013 US dollars),[4] annual hospital expenses for this one operation are approximately $21 billion.

2-3: Surgical Instrumentation and the Medical Device Industry

Surgical residents learn how to perform operations from textbooks, simulation, online videos, training courses, and from attending surgeons. These resources describe the anatomy and the broad steps of the operation[14] – knowledge which is then used in simulators or on actual patients, under the close guidance of a faculty member. While the steps to perform an operation are typically well described, the tools used to complete each step are not prescribed and the number of combinations is vast. Take, for example, wound closure. At the end of the operation, after the gallbladder has been removed and the abdomen has been inspected, all of the laparoscopic instruments are removed and the small incisions need to be closed. Textbooks describe the three available techniques - staples, sutures, or adhesives – a seemingly short list. However, within each technique, there are numerous products/instruments that can be used. A review of one manufacturer's website provides insight. Ethicon, a subsidiary of Johnson & Johnson, is one of the largest medical device manufactures in the world with $25 billion in device sales in 2016.[25] Ethicon's wound closure website lists 6 staplers, 3 skin adhesives, hundreds of sutures (due to combinations of materials, suture lengths, needle shapes and sizes), along with dozens of novel devices such as knotless sutures, port closure devices, and suture-cannula combinations.[26] The breadth of medical devices becomes larger when considering the multiple steps in a given operation, the diversity of operations performed by surgeons, and the numerous medical device manufacturers.

This device saturation likely exists in part as a result of the U.S. Food and Drug Administration (FDA) approval process for medical devices. In brief, the FDA requires classification of all new medical devices into 3 classes, labeled I through III– proportional to their level of risk.[27] The process of approving class I and II devices (low and moderate-risk, respectively) was dictated in section 510(k) of the Federal Food, Drug, and Cosmetic Act ("510(k) clearance"). Most level I devices, and some level II devices, are exempt from all premarket evaluation and can go straight to registration; these devices include external staples, many of the instruments related to laparoscopy, and some sutures. Of the devices that are not exempt, 98% still receive expedited approval by proving substantial equivalence to a device already approved (i.e., predicate devices); these include mesh, internal stapling devices, and tissue adhesives.[28] There is significant criticism of this system,[29] including a 2011 Institute of Medicine report that highlighted that most items approved today use predicate devices that were never systematically assessed for safety and effectiveness and that the 510(k) process "lacks the legal basis to be a reliable premarket screen of the safety and effectiveness of moderate-risk deices and, furthermore, that it cannot be transformed into one."[30, pp. 2]

2-4: The Perfect Storm for Variation and Waste

Cost transparency is a ubiquitous problem throughout the U.S. healthcare system - patient costs are masked by third-party payers, insurance costs are confounded by contractual allowances, and hospital costs are kept under lock and key. For surgical instruments, each hospital contracts with numerous vendors and distributors and negotiates on prices. Prices change frequently as contracts are re-negotiated, new products (or upgraded products) come to market, and vendors are switched. Further, many vendors require non-disclosure agreements, such that prices cannot be disclosed between hospitals or even to physicians or staff within a hospital without fear of loss of contract or litigation.[31]

Instruments used in the OR therefore epitomize the Bipartisan Policy Center's concerns about advancing medical technology.[9] Metaphorically, surgeons are trained with a picture book without any technical or descriptive guidance of the tools required. This ambiguity is filled by medical device manufacturers providing a litany of instruments that are approved either via exemption or an abbreviated process that does not require proof of safety or efficacy. Surgeons have no information to compare these instruments on efficacy (as manufacturers are not required to conduct this type of research) and cost information is not routinely disclosed. The result is a system that may facilitate variation in surgeon instrument use resulting in excessive spending without tangible benefit.

2-5: Definitions of Select Financial and Supply Terminology

2-5A: Cost Perspective

When working with costs, the perspective of the analysis must be described. Whereas quality measures, such as mortality, often share meaning amongst patients, physicians, and hospitals – the financial burden differs depending on the perspective being analyzed. For this analysis, the perspective is that of the hospital (i.e., the healthcare perspective). In other words, we are interested in calculating the resources that must be expended by the hospital to deliver a service. For surgical instruments in particular, this is the purchase price; that is, the negotiated price between the hospital and the vendor/supplier.

Numerous other perspectives exist. For example, a patient perspective would include costs relevant to patients such as co-pays, deductibles, transportation expenses, and the opportunity cost of their time. A third-party payer (i.e., insurer) perspective would reflect the amount reimbursed for an episode of care or a service. This is typically a negotiated rate between the third-party payer and the facility. Finally, the broadest perspective is that of society which would include hospital, payer, patient and other associated costs, such as the cost of caregiver time. The cost-effectiveness literature, which typically compares two or more competing therapies, has traditionally advocated for a societal perspective. However, many analyses fall short of achieving a true societal perspective and recent updates from the Second Panel on Cost Effectiveness now encourage a healthcare perspective (in addition to the societal perspective), which is the one considered in this analysis.[32]

In summary, hereafter the word cost will refer to the hospital's cost of providing a service. We will use charge to refer to the bill a facility sends to a payer and reimbursement to refer to the amount a third-party pays to a facility. If additional perspectives are considered, we will be explicit about their use.

2-5B: Cost Accounting and Attribution

A brief introduction into hospital financial accounting is necessary. A small list of terms and their definitions are included in Appendix 1.

First, costs can be direct or indirect. Direct costs are those that can be attributed to a department or service. Another way to think of direct costs is to ask the question, "Would these costs exist if this department (or service, or patient) were not here?" Supplies used in the OR are a classic example – if a hospital did not perform surgery, there would be no need for OR supplies. Indirect costs refer to costs that cannot be directly tied to a department or service but rather are required as part of running the larger operation. An example is the hospital billing department. Hospital billing is required for a hospital to exist but is not specific to any one department or service line. The costs associated with running this department (e.g., salaries) have to be calculated and allocated to all of the other departments (e.g., based on the square footage of the department).

Second, costs can be considered fixed or variable. The simplest way to think of this terminology is to consider whether the costs change with patient volume. The lease payment for a hospital does not change with hospital volume, and is therefore "fixed." Disposable supplies, on the other hand, are variable. Every patient needing their gallbladder removed will need a new set of disposable supplies, and therefore the costs associated with these supplies will increase proportional to volume. Not all examples are as clean as these two. Take, for example, the cost of nursing care. On the one hand, nursing salaries are somewhat fixed. For example, a small hospital floor with 8 beds may only pay one nurse salary if there were 1, 3, or maybe even 5 patients on the floor. However, when there are 6 or 8 patients, a second nurse will be required. Costs that have this step-function property are sometimes called "semi-variable" reflecting the fact that they are, in some circumstances, variable, and in others, fixed.

Third, and finally, there are numerous accounting methodologies to assign costs. From a research perspective, often asked is how much does it cost to take care of a patient? Two popular ways to do this are cost-to-charge ratios (CCRs) and time-driven activity-based costing (TDABC). The former is more common due to the ease of calculation. Two frequently used administrative databases – the Healthcare Cost and Utilization Project and the Centers for Medicare and Medicaid Services datasets – include CCRs in their research files. It is calculated as follows. An encounter-level charge is generated for each patient and included on the billing extract. To convert this into a hospital cost, these charges are multiplied by a fraction, referred to as the CCR. This CCR, in turn, is generated as a ratio of the hospital's operating expenses (costs) and operating gross revenue (charges). The major assumption of this model is that costs are the same fraction of charges, regardless of actual resources consumed – an assumption that has been criticized.[33,34] For example, it would be easy to imagine two patients undergoing outpatient cholecystectomies. The first operation takes 45 minutes, the patient recovers for 3 hours requiring minimal care of only one nurse, and walks out of the recovery room with the help of her husband. The second operation takes 75 minutes, requires unusual and expensive surgical instrumentation due to inadvertent bleeding, stays in the recovery room for 3 hours but, due to uncontrolled pain, requires near constant attention from 1-2 nurses at a time. This patient then requests extensive hand-written documentation regarding disability coverage. The charges for these two patients would likely be comparable, yet the costs of delivering these services are clearly different. TDABC accounting would capture these differences. In TDABC, each step of a patient's care is carefully mapped and the unit cost of each step is calculated. For example, every minute of nursing time required to care for the patient would be assigned a cost (perhaps $1/minute); this would then be added together with the supplies and salary costs in the OR, medications given in recovery, etc., until a total cost was calculated. TDABC benefits from being more transparent and is likely more accurate and actionable, but it is also resource intensive.[35,36] For many hospitals, TDABC has not been widely adopted as simpler methods – such as CCR - are sufficient.

2-5C: Supplies

How do we reconcile these concepts as it relates to surgical supplies? One final distinction is required. Surgical supplies can be characterized as disposable and reusable. Disposable supplies, such as gauze, sutures, drains, and bandages, are one-time use and, once opened, can never be used for another patient. Reusable supplies can be re-sterilized (i.e., reprocessed) and used on more than one patient. This would include items such as scissors, instruments used to hold needles, dissecting devices, and retractors. Disposable supplies have direct variable costs – they are attributable to the OR and vary proportional to patient volume. Reusable supplies have direct costs but are fixed or semi-variable, depending on how often they need to be replaced as well as the costs associated with reprocessing and repair. A number of studies have attempted to report the marginal cost of a reusable instrument but estimates vary significantly – ranging from 86 cents to $57 per instrument and from $31-127 per tray of instruments.[37-43] In part, this may reflect comparisons of different instruments or instrument groups, but this heterogeneity also likely reflects the fact that no standardized accounting system exists to estimate these costs. TDABC estimates would require tracking each step for a single instrument – from the nurse handling and cleaning it in the OR, to the environmental staff transferring it to cleaning department, to the cost of the soap, the pressure cleaner, the probability of a device breaking and needing repair, the cost of each type of repair etc. Further confounding these estimates is that most devices are cleaned in bulk (as a tray) and removing one instrument would have a negligible effect on overall costs. As a result, we have chosen to focus this study on disposable supply costs. They have direct, variable costs, and by working with the purchase price paid for each instrument, we are effectively using TDABC accounting.

Chapter 3: Literature Review and Contribution of Dissertation to the Literature

This chapter is divided into two sections. In the first, Chapter 3-1, we will describe the existing literature as it relates to each of the four questions. The literature review for each question is supported with a detailed evidence table (Tables 1 & 2 and Appendices 2 & 3). In the second half, Chapter 3-2, we will describe how the dissertation analyses will contribute to the knowledge in this field.

3-1: Literature Review

3-1A: Question 1

What factors predict surgeons' knowledge of the relative cost of surgical instruments?

Four studies[44-47] have surveyed surgeons to assess knowledge of OR instrument costs (Table 1). All utilized cross-sectional online or mailed surveys and targeted resident or attending physicians within a single discipline, such as orthopedics or otolaryngology. Only two studies were conducted in the U.S., neither of which focused on general surgery.

The structure of the survey was relatively consistent across studies with each study including two components – the first elicited surgeon demographics and their opinions about cost/value, and the second assessed the surgeon's knowledge of instrument costs. The mechanism for assessing knowledge of instrument costs typically involved presentation of an item, asking the respondent to estimate the cost, and then considering an accurate response to be one that fell within a certain range. This range was typically +/- 25% of the actual cost (studies usually justified this choice by referencing other studies); for example, if an instrument's price was $20, a correct response would have been between $15 and $25.

Only one of the four studies included over 60 respondents. Covariates were heterogeneous. All included a variable for distinguishing resident (i.e., in training) from attending (i.e., independent) physicians. Other covariates included years of experience, practice setting (public versus private), subspecialty, frequency with which the surveyed surgeon used the item in question, previous participation on a value analysis committee, and self-rating of cost knowledge.

Studies reported low rates of accuracy, ranging from 19% to 30%. For covariates, two studies reported lower accuracy rates for trainees than attending physicians,[45,46] while two others found no difference based on level.[44,47] Studies did not report effect size and while statistically different, the practical significance of the magnitude is unclear – in one, the accuracy rate for residents was 17% while the accuracy rate for attendings was 21%. This begs the question – does this cost knowledge difference translate to downstream utilization and costs? No other covariate showed consistent findings across studies.

A larger body of literature about physician knowledge of costs exists outside of surgery. A 2008 meta-analysis[48] pooled 14 articles that surveyed physicians about their knowledge of cost for diagnostic/therapeutic tests/interventions. Their pooled estimate found overall accuracy to be 33% when using a range of correct answers +/- 20-25%, increasing to 50% when the range of correct answers expanded to +/- 50-100%. They evaluated several study-level covariates, but found no difference in accuracy across country, year of publication, or specialty.

Summary:

Very few studies have been published reporting on surgeons' knowledge of supply costs. Of the studies that exist, only two have been conducted in this country, neither of which assessed general surgery. Studies in this area have been limited to generating a point estimate for physician accuracy and understanding covariates that may help explain variation.

<u>3-1B: Question 2</u>

How does knowledge of the relative cost of surgical instruments impact the cost of a surgeons' instrument preference card for laparoscopic cholecystectomy?

No studies have directly looked at the association between surgeons' cost knowledge and their subsequent selection of items listed on a preference card.

There are nine studies (Appendix 2) that provide indirect evidence that cost information may affect surgeon preference cards and the actual instruments used in the OR.[49-56] All were conducted at a single site or within a single healthcare system (e.g., Northshore Healthcare in Illinois). Three studies, all reporting from Northshore, targeted general surgery, five studies focused on other surgical disciplines (e.g., pediatrics, orthopedics), and one assessed all surgical departments within a single teaching hospital. Studies adopted similar designs – almost all utilized a pre-post single group design with an intervention that provided feedback to surgeons about cost, either in the form of a cost receipt ("cost report card") or by meeting with surgeons to agree on a standardized set of instruments used for an operation.

All studies reported cost savings, with a range from $48 to $539 per operation. Three studies also surveyed surgeons or staff following the intervention. One showed high rates (~90%) of interest and use of the report card, yet only 56% stated that it changed the way they performed an operation.[49] A larger study found similar rates of utilization (86%), but with higher agreement (79%) that the data influenced supply use.[56] Interestingly, the former study demonstrated cost savings of $269 per operation while the latter showed cost savings of $91 per operation. None of these studies analyzed the relationship between the surgeon's opinion about the cost intervention and subsequent utilization.

Summary:

No studies have directly associated surgeon knowledge of cost to the cost of instruments listed on preference cards or used in the OR. There is indirect evidence, from intervention studies, that providing cost information may reduce supply use, but the magnitude of effect is variable. Post-intervention surveys suggest surgeons have interest in cost information but it is unclear how the intervention affects this interest or what the effect of interest is on instrument utilization.

3-1C: Question 3

For laparoscopic cholecystectomy, how much of intraoperative cost variation is explained at the patient, surgeon, and facility levels?

Ten studies have evaluated intraoperative disposable supply variation (Table 2).[57-66] Of these, only two have looked at general surgery operations,[57,59] neither of which used multivariate analysis.

Nine studies were conducted in the United States and one was conducted in Taiwan. Of the nine U.S. studies, 6 were at single sites, with the remaining 3 conducted at multiple facilities within a single health system. The number of patients in a study ranged from 108 to 5,623 and the number of surgeons ranged from 3 to 72. Most studies (n=6) focused on one procedure such as laparoscopic cholecystectomy or pediatric adenotonsillectomy, while the remainder focused on a department or service line.

Operationalization of the dependent variable (intraoperative supply cost) was heterogeneous and the source and calculation of estimates were often not well described. As mentioned in Chapter 2, ideally, we would know the cost perspective (i.e., cost, reimbursement, or charge), the instruments included in estimates (i.e., disposables vs. reusables) and how costs were assigned (i.e., accounting methodology – CCR, TDABC). The largest general surgery study defined intraoperative supply cost as the "estimated acquisition cost for instruments on the surgical field,"[59, pp.114] and further "the estimated cost was used instead of the exact cost to account for inter-facility differences and the impact of purchasing agreements including tiered pricing, incentives, and rebates."[59, pp.114] It is unclear what instruments were included in these estimates or how these costs were adjusted between facilities. The other general surgery study was similarly vague, stating "surgeon preference sheets, hospital invoices, and operative reports were used to calculate mean variable supply costs for each surgeon based on distinct technique; each surgeon performed the critical steps to the case with varying equipment."[57, pp.2680] It is unclear if these estimates represent actual instrument usage or those forecasted on the preference card, and the authors do not describe the source of cost information.

Many of these analyses included broader measures of costs, such as total OR or total encounter costs. If studies are unable to detail the necessary measures to address disposable supply costs, it seems dubious that estimating costs as a higher level will be any more valid. What is the cost of a day in the Intensive Care Unit (ICU)? Or the price of a lab test? The importance of perspective and accounting mechanism are evidenced by the fact that estimates for a minute of OR time have varied in the literature from $12 to over $100 per minute.[67-71]

Finally, studies used a variety of covariates and analytic strategies. Five studies utilized only bivariate or graphical analyses. Of the studies that did conduct multivariate analysis, only one described a multilevel strategy. In this study of neurosurgical procedures at one large U.S. teaching hospital, the authors regressed surgical supply costs on patient factors (age, gender, payer, body mass index [BMI], admission source, comorbidity index), and case factors (procedure length, procedure date), with a random effect for each surgeon and surgeon case volume. They did not include procedure as a variable in their model. The number of cases within each procedure varied from a low of 6 to a high of 1,062. At the patient level, they found older females with more comorbidities and with Medicare insurance had higher supply costs. Longer procedures and surgeons with higher volumes also had higher supply costs. By excluding procedure as a variable from their model, there is concern of omitted variable bias as those undergoing more complex procedures may be both older with higher comorbidities and that these procedures may require more expensive supplies. The significance of Medicare in their model provides evidence of this bias as there should be no foreseeable reason that this *reimbursement* model would impact supply cost.

Summary:

Ten studies have addressed variation in intraoperative supply cost, however, only two have been conducted in general surgery. Both of these studies have challenging methodologic limitations with respect to the definition and acquisition of supply cost data and both were limited to bivariate analysis. Beyond general surgery, only one study has utilized a multilevel model to examine supply cost variation, but interpretation of the results is limited due to procedure heterogeneity and inadequate control of potential confounders.

3-1D: Question 4

How does a surgeon's a priori preference card for laparoscopic cholecystectomy influence intraoperative costs and patient outcomes?

No studies have looked at the relationship between the cost of instruments on a preference card and instrument costs or patient outcomes.

Five of the studies included in the literature review for question 3 included patient outcomes as a component of their analysis (Appendix 3). The primary concern about this analysis is the possible confounding from case complexity and the possibility of reverse causality/endogeneity. It is likely that more complex cases (i.e., those with high risk of complications) will both require more expensive instruments and have worse patient outcomes. Analyses must therefore control for these factors, but to our knowledge no one variable has ever been proven to capture case complexity in its entirety, leaving the possibility of omitted variable bias even after multivariate adjustment. Of these five studies, only one analyzed a general surgery operation.[4] This seven-site study reviewed administrative data and compared estimated intraoperative supply costs with 30-day patient outcomes using National Surgical Quality Improvement Program (NSQIP) data. They plotted surgeon-level costs and surgeon-level complication rates and found no correlation. However, they only had outcome data for a subset (proportion and sample size not specified) of patients and data were only analyzed at the surgeon level, effectively reducing their sample size to ~ 32 (derived from Figure 3) severely limiting power. Of the remaining studies outside of general surgery, only two conducted multivariate analyses, with one showing more costly surgeries were associated with worse outcomes[63] and the other finding no association.[66]

Summary:

No studies have directly associated the cost of instruments on a preference card with actual intraoperative costs or patient outcomes. Studies that have analyzed the association between intraoperative supply costs and patient outcomes often use rudimentary analyses which do not adjust for confounders, and those that have employed multivariate analysis are still susceptible to bias due to endogeneity between their predictor and outcome.

3-2: Contribution of Dissertation to the Literature

The literature addressing intraoperative disposable supply cost variation is limited. For each question, this study will improve upon prior research in meaningful ways (as detailed below). Perhaps more importantly, this dissertation will integrate these components into a unified picture that can be used by providers and hospitals as they embark towards value-based care.

3-2A: Question 1

What factors predict surgeons' knowledge of the relative cost of surgical instruments?

This study will expand assessment of surgeon cost knowledge to the field of general surgery and will utilize a different operationalization of the dependent variable (relative costs, instead of absolute). This operationalization conceptually mirrors the way we think about costs in other domains of life (see below).

3-2B: Question 2

How is accuracy of knowledge of the relative cost of surgical instruments associated with the cost of a surgeons' instrument preference card for laparoscopic cholecystectomy?

The association between surgeons' knowledge of relative costs and the cost of their preference cards has not previously been characterized. This study will also analyze the interaction between cost knowledge and perceived importance of cost, a potentially important moderator that may explain the modest results from previous studies trying to reduce supply costs.

3-2C: Question 3
For laparoscopic cholecystectomy, how much of intraoperative cost variation is explained at the patient, surgeon, and facility levels?

This study will expand upon previous research by utilizing multi-level models to differentiate the effects of the patient, surgeon, and facility on intraoperative costs. By analyzing only one operation, we will eliminate the confounding issue of multiple procedures, although at the expense of generalizability. We will also use cost data that is adjusted over time and corrected for missing or aberrant prices.

3-2D: Question 4
How does a surgeon's a priori preference card for laparoscopic cholecystectomy influence intraoperative costs and patient outcomes?

As mentioned above, associations between intraoperative costs and patient outcomes are likely confounded by case complexity – a variable that is unlikely to be adequately captured regardless of the granularity of an administrative database - leading to bias in the tested associations. By utilizing the exogenous preference card cost (i.e., a reduced form model, described later), we can avoid this confounding and generate an unbiased estimator for the effect of preference card cost on patient outcomes. From a policy perspective, modifying the preference card is more tenable than trying to modify instrument use while a case unfolds.

3-2E: Overall Impact
Finally, the integration of these questions will provide important insights that could not be answered through studies limited to one question. For example, the implicit assumption of previous studies that have shown low accuracy rates for surgeon cost knowledge is that this lack of knowledge has downstream effects on utilization. Our study will not only assess this relationship but will also explore the potential effect size. If there is no association, or the magnitude of difference required to impart a downstream effect is large, then administrators may

find marginal or no benefit in merely providing cost information to surgeons without taking additional steps.

Chapter 4: Conceptual Model

The conceptual model was generated de novo and is presented in Figure 3. The model flows from left to right, sequentially addressing the four questions of this proposal (the outcome of each question is indicated with Q1-Q4). The relationship being tested for each hypothesis is indicated in the Figure with 1a, 1b etc.

The following description of the model is guided by the four dependent variables, which are highlighted in light gray: (1) The accuracy of a surgeon's knowledge of the relative cost of instruments, (2) The cost of instruments on the a priori preference card, (3) The cost of instruments used in the OR, and (4) Patient outcomes. For each dependent variable, we will first provide a description of the dependent variable, followed by a description of the factors that influence the dependent variable.

4-1: Dependent Variable #1
The accuracy of a surgeon's knowledge of the relative cost of instruments

Dependent Variable Description:

The first dependent variable*, the accuracy of a surgeon's knowledge of the relative cost of instruments* (hereafter, *cost knowledge*), is located in the upper left corner of Figure 3. The ability for surgeons to choose instruments based on price requires surgeons to know the relative price between two or more competing items. Previous research assessing surgeon (and physician) knowledge of costs has focused on accuracy of actual costs instead of relative costs. Relative comparisons mirror decision making in the OR more appropriately for several reasons. First, while there are numerous items that could be used for a given step in a given operation, there are often only a handful of instruments that are commonly used. Second, the instruments available are constrained by contracts established with the purchasing department, to items in stock, and to items the nurse can find. This decision-making process is somewhat analogous to shopping in a grocery store. If a consumer wants to purchase mustard, they will go to the condiment aisle and compare prices of the various options. The decision of which mustard they choose will reflect a balance of their preferred mustard with the price of that mustard. Whether the mustard costs $6 or $60 is less relevant than whether their preferred mustard is significantly more expensive than the "next best thing." Decisions in the OR mirror this relative choice but are actually simpler because the option of going to a different grocery store (i.e., going to a different facility with different instruments) is not available, so surgeons must choose between the available items.

<u>Factors Influencing Dependent Variable #1:</u>

We identify one factor that directly influences the first outcome as well as 4 antecedent or moderating variables. Moving down and to the left, *cost knowledge* is directly influenced by *familiarity with instrument prices* (hereafter, *familiarity*). This familiarity is directly influenced by *prior exposure to instrument* prices (hereafter, *exposure)*. When surgeons are exposed to an instrument's price, they may retain that knowledge. Conversely, no exposure to price means no familiarity and therefore their ability to discriminate the relative price of instruments should be no better than chance or a semi-informed guess (i.e., they correctly discern that larger items are more expensive than smaller items).

However, this relationship is complex. Surgeons exposed to prices may differ in their ability to retain this knowledge. The effect of *exposure* on *cost knowledge* will likely be heightened or attenuated by the *relative importance of cost versus effectiveness to the surgeon* (hereafter, *importance of cost* or *cost importance*). This moderation effect is represented in Figure 3 by an arrow going from the *importance of cost* to the arrow between *exposure* and *cost* *familiarity*. If a surgeon believes cost is an important component of instrument choice, they may retain cost knowledge more readily than those who do not value cost as important.

A further complexity is the relationship between *exposure* and the *importance of cost* to the surgeon. This relationship is marked with a two-way arrow indicating uncertainty in the direction. On the one hand, surgeons who place increased emphasis on cost may, in turn, seek price information (and increase their *exposure*), while on the other hand, greater *exposure* may, over time, cause surgeons to increase the relative weight they place on cost.

Two final variables to consider when evaluating a surgeon's cost knowledge are the *recency of their surgical training* and *institutional policies*, as illustrated in the bottom left corner of Figure 3. The importance of controlling healthcare costs is a contemporary issue, resulting reflexively from the increased burden of healthcare expenses over the past 20 to 30 years. It therefore seems possible that surgeons trained more recently would be more attuned to healthcare costs than those trained more remotely. For surgeons, in particular, the rapid introduction of new technologies may also contribute to increased cost awareness. Laparoscopic techniques were mainstreamed in the early 1990's with robotic surgery only receiving FDA approval in 2001. With this national emphasis on cost and the rapid introduction of surgical technologies, surgeons trained contemporaneously may have more *exposure* to prices and may place more weight on the *importance of cost*. The converse may also be true, with surgeons trained more remotely having more opportunity for increased exposure and therefore increased knowledge. Nevertheless, the impact of *recency of surgical training* on *cost knowledge* is likely mediated both through *exposure* and the *importance of cost.* Finally, some institutions have started to introduce policies or initiatives to purposefully increase the *exposure* of surgeons to prices. An example would be a cost report card given to a surgeon after the operation listing the items they used and their associated prices. These policies may directly increase *exposure*, with downstream effects on *familiarity* and *cost knowledge*.

4-2: Dependent Variable #2
The cost of instruments on the surgeon's a priori preference card

Dependent Variable Description:

The second dependent variable, the *cost of instruments on the a priori preference card*, is located in the middle of the conceptual model. The preference card is a list of instruments surgeons believe they will need for a typical operation. The cost of this preference card is the sum of the purchase price for each item on this list. This list is specific to the operation, such that the instruments needed to take out an appendix will differ from the instruments needed to take out a gallbladder. This analysis only includes laparoscopic cholecystectomy so we have not included a variable for operation.

Factors Influencing Dependent Variable #2:

We identify three factors that directly influence the items on the preference card (and therefore the *cost of instruments on the a priori preference card*) – (1) *instrument available at the* facility, (2) *surgeon's perceived cost of instruments*, and (3) *surgeon's perceived effectiveness of instruments*. Four additional factors will be discussed in this section. The first of these additional factors, *actual cost of instruments* will be described in this section. The other three – *cost knowledge, cost importance* and *recency of surgical training* have been introduced in dependent variable #1 but their effects on this outcome will be expanded upon.

The first factor with direct influence is located beneath the outcome and is whether or not the desired instrument is *available at the facility*. First is a note on nomenclature. In this dissertation we will use the terms health system or institution to refer to distinct organizations such as the University of California, Los Angeles (UCLA) or the University of California, San

Diego (UCSD). We will use facility or site to refer to the different settings within a health system, either because the system has multiple locations or because they have more than one sets of ORs within a given location. For example, at UCLA there are two locations – Ronald Reagan (RR) and Santa Monica (SM) – each of which has a main OR and a standalone ambulatory surgery center (ASC). Most hospital systems will have a main OR to care for both inpatients and outpatients, while some will have an additional ASC to care for lower-acuity patients undergoing outpatient surgery. As it relates to the conceptual model, hospital systems will have different contracts. In general surgery, the two largest medical device manufacturers are Covidien and Ethicon. The products supplied by these two manufacturers overlap significantly, so hospitals will generally contract with one or the other (supplies proprietary to the non-contracted manufacturer are purchased on an ad hoc basis). Facilities may also have different instruments available to them. For example, the supplies stocked at each facility may differ – with more complex and intricate devices only available at the larger hospital (at UCLA this would be RR over SM). Conversely, ASCs may stock fewer supplies than the main OR. Further, a common practice is to bundle basic supplies together. Certain supplies are a part of almost every operation – such as gowns, drapes, basins, and light handles. Instead of the nurse having to individually find these at the beginning of an operation, they are bundled into a pack in order to expedite the process. It is possible that these packs may differ in their contents from one facility to the next within a healthcare system.

The second and third factors that will directly influence the *cost of instruments on the a priori preference card* are the counterbalancing factors of the *surgeon's perceived cost of instruments* (hereafter*, perceived cost*) and the *surgeon's perceived effectiveness of the* *instruments* (hereafter*, perceived effectiveness*). These two factors are located up-and-to-the-left and down-and-to-the-left of the dependent variable, respectively.

Perceived cost should mirror the *actual cost of the instruments* (indicated in the Figure directly above *perceived cost*), but this relies on surgeons having accurate *cost knowledge* (discussed in dependent variable #1). As such, *accuracy of cost knowledge* should influence the *cost of instruments on the a priori preference card* through *perceived cost*. Without accurate knowledge, surgeons may *perceive cost* as higher or lower if they were provided inaccurate information, or if they otherwise make assumptions about the cost of an instrument. It is conceivable that surgeons may assume larger or seemingly more intricate devices would be more expensive, even if that is not the case.

Counterbalanced with *perceived cost* is *perceived effectiveness*. Conceptually, this captures the idea that among a set of devices, surgeons will view each as having a certain intrinsic value and could rank order which they would prefer to use. Reflecting back on our grocery store metaphor, when choosing between various mustards, a consumer would be able to rank order their preference. There are likely infinite underlying reasons behind this decision; for mustard, it may be based on factors such as taste, smell, the aesthetic of the bottle it is stored in etc. An entire dissertation could be written exploring the factors surgeons may consider when assessing instrument effectiveness, but presumably this assessment includes factors such as instrument familiarity, tactile features, the ability to reduce OR time, reliability/consistency, and the ability to reduce or prevent complications (i.e., expectations about outcomes). The concept we are aiming to capture here is not each of these individual factors, but the broader concept that a surgeon can rank order the possible instruments based on their own internal effectiveness calculations. We identify one surgeon-level factor that directly influences this *perceived* *effectiveness* - the *recency of surgical training* (located in the bottom left hand corner with an arrow pointing to *perceived effectiveness)*. Given the rapid expansion of medical devices available on the market, surgeons trained more contemporaneously may have been exposed to a wider array of instruments during training. There has also been an increased emphasis on quality metrics and efficiency which may have altered the way newly trained surgeons think about instrument effectiveness.

Finally, c*ost importance* may strengthen or attenuate the effect of *perceived cost* and/or *perceived effectiveness* on the *cost of instruments on the a priori preference card*. This is indicated by two arrows extending from *cost importance* to the arrows between *perceived cost* and dependent variable #2 and *perceived effectiveness* and dependent variable #2. The following example will help illustrate the underlying concept (formally described as the quadrants of a cost-effectiveness plane).[72] Imagine two surgeons picking between two potential instruments. For simplicity, we assume the two surgeons perceive the cost and effectiveness of the two instruments similarly. Item 1 has a perceived effectiveness of X and item 2 has a perceived effectiveness of Y. Item 1 has a perceived cost of $50 and item 2, $75. If X > Y then both surgeons would pick item 1 as it is both more effective and less costly. However, if Y > X then the two surgeons may pick different items based on each surgeon's cost-effectiveness threshold.

4-3: Dependent Variable #3
The cost of instruments actually used in the operating room

Dependent Variable Description:

The third dependent variable, located just to the right of dependent variable #2, is the sum of the unit prices of the instruments actually used in the OR, hereafter referred to as *actual costs*.

Factors Influencing Dependent Variable #3:

It seems likely that the list of instruments on the a priori preference card reflects what the surgeon will actually use during an average operation. Therefore, the *cost of instruments on the a priori preference card* should drive the *actual costs* of the instruments used (indicated with an arrow between dependent variable #2 and dependent variable #3). However, there are several additional factors that may influence *actual costs* depending on the situation. We describe four of these factors – (1) *instrument available at facility* (introduced in dependent variable #2), (2) *resident involvement,* (3) *patient comorbidities and case complexity, and* (4) *surgeon's technical skill.*

Similar to dependent variable #2, the *instruments available at a facility* will constrain the available options for a surgeon. One could imagine a scenario where excessive bleeding is encountered during an operation. A surgeon calls for an instrument but that instrument is not available (e.g., out of stock, not stocked at that facility) and therefore the surgeon would have to pick another instrument.

The second factor that may directly influence *actual costs* is whether a resident physician is involved (*resident involvement).* The presence of a resident physician may alter the instruments opened before the operation begins, during the operation, or both. Attending surgeons are usually aware of whether or not a resident will assist with the case before the operation begins. While the attending physician dictates the approach, technique, and instruments used for an operation, the resident physician is often responsible for the physical performance. It is conceivable that if an attending surgeon has designed their preference card with the assumption that a resident will not be present, they may alter the items opened before the case when a resident is present (by contacting the nurse in the OR who is responsible for opening the supplies). Further, during an operation, if an unusual situation is encountered, such as dense scar tissue or more bleeding than average, the decision of which additional instruments to open may be influenced by *resident involvement*. An attending physician may exercise more caution and choose to open different instruments out of an abundance of caution, compared to if they were operating alone.

The third factor that will directly influence *actual costs* is the *patient's medical comorbidities and the complexity of the case*. Examples of patient comorbidities include obesity, diabetes, and steroid use. All of these conditions may predispose a patient to wound complications; a surgeon may anticipate this additional risk and alter their wound closure technique accordingly. Case complexity captures the idea that cases vary from simple to extremely difficult, over and beyond the patient's comorbidities. Factors such as scar tissue, anatomy, and predisposition to bleeding are included and may or may not be influenced by the patient's comorbidities.

Finally, a *surgeon's technical skill* (hereafter, *skill*) may influence their choice of instruments during an operation. One could imagine an intraoperative complication – such as intraoperative bleeding - that, to an adept surgeon, could be managed with existing or simple instruments - yet a surgeon less skilled may believe they need a novel, expensive, hemostatic agent.

4-4: Dependent Variable #4
Patient outcomes

<u>Dependent Variable Description:</u>

Patient outcomes are located to the far right of the conceptual model. While there are virtually infinite potential patient outcomes to consider, perhaps most relevant when considering the instruments used in the OR are intraoperative outcomes such as procedure time and postoperative clinical and utilization measures such as complications and length of stay.

Factors Influencing Dependent Variable #4:

The instruments actually used in the OR, and therefore the *actual costs,* should directly influence patient outcomes (indicated with an arrow between dependent variable #3 and dependent variable #4). It seems logical that surgeons would choose more expensive instruments (over less expensive instruments) if they have improved efficacy and therefore improve patient outcomes. However, several variables beyond the instruments used (*actual costs)* may influence *patient outcomes*. Broadly, there are patient-level, surgeon-level, and facility-level factors worth considering.

At the patient level, three groups of factors directly influence *patient outcomes*. First, are *medical comorbidities* and *complexity* (located above and between dependent variable #3 and dependent variable #4), described in detail for dependent variable #3. To this, we add two new patient-level concepts. The first is located directly above dependent variable #4, *patient motivation,* and captures the idea that some patients will be more proactive in their recovery than others. This *motivation*, in turn, directly influences *patient behaviors* such as early ambulation, use of their breathing machine, and cooperating with their healthcare providers, which in turn, improves patient outcomes. The final patient-level factor, located in the top right corner, is consideration of *patient financial resources*. While recovery will differ from one patient to the next, recovery from surgery often requires great expense both directly, in the form of paying for prescriptions and supplies, and indirectly, by taking time off of work. Patients who cannot afford their prescriptions or who push themselves to go back to work too quickly to generate income may suffer from adverse events, such as wound disruptions, unnecessary pain, or postoperative infections.

Surgeons can also influence *patient* outcomes, which we describe as *skill* (previously introduced for dependent variable #3). As introduced in dependent variable #3, this captures a surgeon's technical adeptness which should correlate with intraoperative outcomes, such as the length of the operation, or the amount of bleeding. In challenging cases, surgeons with good judgment will be able to make the decision that will most benefit their patient. However, when it comes to *patient* outcomes, *skill* extends beyond the OR. As patients are managed postoperatively, a number of important decisions must be made as it relates to timing of discharge, if and when to start antibiotics, and how to manage evolving complications, all of which may influence *patient outcomes*.

At the facility level, the *quality of postoperative care* will also directly influence patient outcomes. Evaluations of quality by third party organizations and algorithms (e.g., Leapfrog, Hospital Compare) have suggested there are variations in hospital quality. This variation may reflect myriad of factors, such as the quality of nursing and allied care (e.g., licensed vocational nurses, therapy, pharmacy), nurse-to-bed ratios, financial resources, the layout of the hospital, the presence of rapid response teams, etc.

In summary, there are many factors to consider when assessing *patient outcomes*. The instruments used (*actual costs*) in the OR are undoubtedly one part of this, but understanding the effect of *actual costs* on *patient* outcomes without considering other patient, surgeon, and facility factors will lead to inaccurate estimates.

Chapter 5: Methods

This analysis includes both primary and secondary data. Question 1 utilizes primary data collected from a three-institution web-based surgeon survey. Question 2 merges survey data from the UCLA respondents with UCLA medical record ("administrative") data. Finally, questions 3 and 4 rely exclusively on a secondary analysis of UCLA administrative data.

The order of this Methods section is as follows: (1) Human subjects research approval, (2) Data sources & study cohort, (3) Measurement model, (4) Variable definitions, and (5) Regression specification, statistical considerations, and missing data.

5-1: Human Subjects Research Approval

Creation and analysis of the patient-level clinical database was approved by the UCLA IRB with a waiver of informed consent (IRB #16-001327). The surgeon survey, including cognitive interviewing, was approved by the IRB at each site (UCLA, UCSD, and Cedars Sinai Medical Center [CSMC]) including a waiver of signed informed consent (IRB#18-000477).

5-2: Data Sources & Study Cohort

Below we detail the data source used for this analysis, including inclusion/exclusion criteria, and sample sizes. A summary is included in Table 3 and sample size flows for each question are included in Figure 4. Because this analysis includes multiple data sources, missing data, as well as variation in the sample from one hypothesis to the next, we have limited the discussion of sample to our primary questions/hypotheses. Detailed flow diagrams and Tables that discuss sample for each sensitivity analysis are included in Appendix 4 for the interested reader.

5-2A: Surgeon Survey

A web-based survey was administered to attending surgeons at UCLA, UCSD, and CSMC. An attending and resident surgeon "champions" within the Division of General Surgery were identified at UCSD and CSMC (see acknowledgements) based on prior collaborative work related to supply costs. These champions were asked to identify a list of surgeons that perform laparoscopic cholecystectomy or who use laparoscopy frequently in their practice. While this ultimately resulted in both general surgeons and subspecialists (e.g., colorectal, transplant, surgical oncology), all of the included surgeons had familiarity with laparoscopic cholecystectomy as it is required during general surgery training. Even if a surgeon has not performed a laparoscopic cholecystectomy in 20 years, the instruments used in these related disciplines are consistent across operations. For example, surgeons who perform colectomies (removing part of the colon) are familiar with laparoscopic instruments including electrothermal, clipping, and stapling devices. At UCLA we further expanded our pool of respondents by surveying surgeons who had performed at least five laparoscopic cholecystectomies in the study time period, regardless of the location of their current appointment. Three additional surgeons were identified this way using a preliminary administrative data set.

Using the conceptual model as a guide, we first identified the 6 areas we wanted the survey to address - (1) Recency of surgical training, (2) Prior exposure to instrument prices, (3) Familiarity with instrument prices (4) Relative importance of cost versus effectiveness to the surgeon, (5) Accuracy of a surgeon's knowledge of the relative cost of instruments, and (6) Surgeon's perceived effectiveness of instruments. We then developed one or more questions to address each area. Because there were no existing items to draw upon, questions were developed de novo. Questions were refined through six cognitive interviews (four at UCLA and two at UCSD). The four cognitive interviews at UCLA were performed by the PhD candidate (CC) and two additional interviews were performed by a Mark Zhao, a resident physician at UCSD. The faculty members participating in the cognitive interviews were not eligible for the main survey study (e.g., surgeons that operate at the county hospital or the Veterans Affairs). The cognitive interviewing protocol is included in Appendix 5. Procedures for cognitive interviewing were adapted from Dillman et al,[73] including the use of a "think-aloud" technique whereby respondents were asked to describe their thought process as they completed the survey. Probing questions provided insight into issues related to wording, order, and visual design. The respondents completed the survey on their own computer allowing identification of issues with platforms and web browsers. The interviews were recorded and notes were taken in order to revise the instrument after each interview.

The survey was distributed via Qualtrics – a web-based survey platform - to 100 potential respondents on May 29th, 2018. Data collection was cut off on June 28th, 2018 with the last response recorded the same day. Eighty-eight surgeons started the survey and answered at least one question, with 83 answering at least half of eligible questions. These 83 are hereafter referred to as "responders"[74] and determine the overall response rate of 83%. This is summarized, by institution, in Table 4. At UCLA, 29 surgeons completed the survey and were potentially eligible for merging with the UCLA administrative data.

<u>5-2B: UCLA administrative data - Surgeon preference cards for laparoscopic cholecystectomy</u>

Surgeon preference cards for laparoscopic cholecystectomy are stored electronically and were provided by OR staff for this analysis (see acknowledgements). There are two types of preference cards: (1) Surgeon and Site preference cards (hereafter, "surgeon/site" preference cards), and (2) Service and Site preference cards (hereafter, "generic" preference cards).

For common operations, such as laparoscopic cholecystectomy, surgeon/site specific preference cards are available. For example, if surgeon A performs laparoscopic cholecystectomy at the RR-OR and RR-ASC, he/she will have a preference card for laparoscopic cholecystectomy at each site. In the event a surgeon does not have a site-specific preference card, a generic preference card was used that is site specific but *not* specific to that surgeon. This may occur when a surgeon performs an operation at a site where they seldom work (e.g., a surgeon that typically operates at RR that performs an operation at SM). Conceptually, for surgeons who do not adopt a site-specific preference card, we considered the generic card their preferred item list (2 surgeons, 2 preference cards). It seems likely these cards are either close enough to their preferred list, or, if not, they are used infrequently. We describe analytic steps later in this chapter to account for differential utilization of preference cards across sites.

At UCLA, the process of generating a site/specific preference card is as follows. When a new surgeon joins the practice and schedules an operation, the surgical nurse will ask the surgeon which instruments they would like to use. When this occurs, the surgeon can either generate a preference card de novo, or, more likely, will copy an existing preference card from one of their colleagues. Following the initial and subsequent operations, the surgeon can ask the OR nurse to update their preference card with alternative instrumentation. This is an iterative process throughout the surgeon's career. At UCLA, surgeons do not have the ability to change their preference card - this can only be done by a designated nurse at each facility - and therefore there exists the possibility that a preference card may not reflect the desired changes of the surgeon immediately.

An example of a surgeon/site preference card is found in Appendix 6. The focus of this analysis is on the data on the first page, under the header **Supplies**. This is the list of the disposable items that the surgeon would like opened or available for the operation. For each item, there is a location (this allows the nurse to find the item), the name of the item, the EMPAC ID and manufacturing number, and columns titled "open" and "PRN" (*pro re nada*, when necessary). The open column indicates that those items should be physically opened before the case begins. For example, in the sample preference card (Appendix 6), the second row has the number 2 in the open column indicating that 2 monocryl sutures (sized 4-0) should be opened. Opened items cannot be used on other patients regardless of whether or not they are used. PRN items should be readily available but *not* opened. In theory, the surgeon will tell the nurse in the OR if/when that item should be opened. For the preference card analysis we only considered the cost of items that are opened. Interesting questions could be addressed by analyzing the PRN items, but this is beyond the scope of this study.

There are a number of additional components to a surgeon's preference card (Appendix 6). Beyond supplies, the card includes extensive details about dressings, cleaning and positioning the patient, technical notes, and a revision history. Finally, the preference card also includes items such as drugs (typically local anesthetic agents), instruments, and equipment. These latter two fall under the category of "reusable" items (discussed in the Definitions – Supplies section of Chapter 2) and will not be considered in this analysis.

Preference cards served as both a dependent variable (Question 2) and as a primary predictor (Question 3 and 4). Samples varied across questions based on data available (see Table 3). For question 2, we limited our analysis to UCLA surgeons who completed the survey and had laparoscopic cholecystectomy preference cards (26 surgeons and 38 preference cards). For questions 3 and 4 we limited the analysis to surgeons with 4 or more "relevant" (defined below) operations during the study time period with preference card data (23 surgeons and 36 preference cards).

5-2C: UCLA administrative data – the Perioperative Data Warehouse

In April 2013, UCLA-health introduced the Epic Systems Corporation electronic medical record. Along with the ability to store electronic health information, the introduction of this data system also enabled retrospective queries of patient-level medical records for clinical research projects. Recently, the Anesthesia Department at UCLA developed the Perioperative Data Warehouse.[75] The details of the underlying bioinformatics structure are beyond our scope, but the algorithms translate a wealth of raw data into clinically-meaningful outputs. The algorithms are generated and validated by clinicians. This system allows researchers to extract patient-level information including details about the patient (e.g., age, sex, race), the case (e.g., type of procedure, length of procedure, surgeon), and postoperative events (e.g., length of stay, return to emergency department [ED]).

We queried the perioperative data warehouse for all laparoscopic cholecystectomies performed on adult (aged 18 years or older) patients at any UCLA facility between April 1st, 2013 (i.e., database inception) and March 31st, 2018. This included four facilities: RR-OR, RR-ASC, SM-OR, and SM-ASC. The timeframe was chosen to maximize sample size at the patient, surgeon, and preference-card level.

The inclusion and exclusion criteria for identifying relevant cases are found in Table 5. This analysis is restricted to adult patients for two reasons. First, while pediatric patients do undergo cholecystectomy, the overwhelming majority of cholecystectomies are performed in adults. Second, the age cutoff for pediatric varies. A 16-year-old at one institution may be managed by a pediatric surgeon, but they may be managed by a general surgeon elsewhere. Pediatric surgeons often have different techniques than general surgeons, introducing an additional confounder. Since gallbladder disease in adult patients is managed by general surgeons, this restriction limits this potential confounding without significant impact on sample size. We similarly excluded other non-general surgeons, such as liver transplant surgeons, that are not comparable due to different techniques and patient populations.

Cases are initially identified in the data warehouse using preoperative booking slips. At UCLA, a sheet of paper (booking slip) that includes the name and Current Procedural Terminology (CPT) code of the requested procedure is sent to OR staff to request an operation. The data warehouse queries these text fields for the phrase "cholecystectomy." Because cholecystectomy may be a part of a larger operation, we excluded cases that were scheduled to have an additional major operation (e.g,, colorectal resection, herniorrhaphy). We also excluded robotic cholecystectomies as the costs associated with the robotic platform are not well itemized in hospital systems and likely under-estimate the true costs of these devices.[76]

Identifying cases by booking slip is sensitive but not specific to the actual procedure performed. There are reasons to believe that the preoperative booking slip may differ from the operation performed. For example, the pathology identified during the operation may require additional or substitute procedures. Further, the booking slip is often not written by the attending surgeon, but instead, by a staff member in the clinic or by a resident or medical student. Since these individuals are not supervising the operation (and may not be involved at all), they may omit important components of the case.

To validate the actual procedure performed, we further restricted our analysis based on primary CPT and Diagnosis Related Group (DRG) codes. Both are assigned after the billing cycle has finished. The algorithm assigning primary CPT code has been refined by the Anesthesia team at UCLA and was the primary method for determining the procedure performed. In the event multiple CPT codes are assigned to a case, the data warehouse algorithm extracts the most complex CPT code, assigning it as the primary CPT. When equally complex codes are identified, the algorithm extracts the most commonly performed procedure at UCLA. We limited our analysis to cases with a primary CPT code of laparoscopic or open cholecystectomy, with or without intraoperative cholangiography or common duct exploration (see Table 6). Operations with cholangiography, common duct explorations, and open cholecystectomies were included as these potentially represent intraoperative complications. Routine cholangiography is not performed at UCLA. For the few operations with missing primary CPT codes, we looked at the primary DRG assigned to the encounter and limited our analysis to cases with a cholecystectomy-related DRG (see Table 6).

Finally, we restricted cases to surgeons who performed at least four operations during the study time period. Our goal of assessing surgeon-level costs is to generate a reliable and valid way of comparing surgeons. Surgeons who do not perform the operation as a common part of their practice should not be included for two reasons – first, measures with few observations are unreliable (i.e., standard errors are large), and second, with the ultimate goal of decreasing costs in the future, targeting very low volume surgeons will be low yield.

5-2D: Ancillary database – National Surgical Quality Improvement Program (NSQIP)

NSQIP is a national clinical registry organized by the American College of Surgeons. The purpose of NSQIP is to collect clinical data for surgical procedures that can be used to provide feedback to programs about their relative performance. At each participating site, one or more clinical abstractors (typically registered nurses) are hired to manually review and enter data at their institution including a review of the medical record, billing data, and, if necessary, contacting the patient or family to fill in missing values. NSQIP can therefore capture events that occur outside the hospital, and, because the data is manually entered by a trained clinician, adverse events that are relevant to surgeons – such as surgical site infections and reoperations – are consistently coded.

One of the primary limitations of NSQIP, however, is the expense of collecting the data. As a result, only a fraction of the surgical cases performed at an institution are captured. To improve generalizability, NSQIP designed an eight-day cycle with random sampling to ensure representative diversity of patients undergoing a given operation. However, hospital leadership can determine the overall focus. At UCLA, the focus is primarily on inpatient operations. Because many cholecystectomies are performed outpatient, the result is a biased sample of operations.

For analyses involving NSQIP data, our focus was generating a cohort of patients that was representative of the inpatients undergoing cholecystectomy. Of the 1817 relevant cases identified in the perioperative data warehouse (with 23 surgeons and 36 preference cards), 992 were undergoing inpatient operations, of which 329 had NSQIP data (with 19 surgeons and 22 preference cards).

5-3: Measurement Model

The Measurement Model is shown in Figure 5. The description of the Measurement Model will follow the structure used to describe the Conceptual Model. For each dependent variable, we will first describe the measurement of the dependent variable, followed by the measurement of the factors hypothesized to influence it.

<u>5-3A: Dependent Variable #1</u>
Accuracy of surgeon's knowledge of the relative cost of instruments (Cost Knowledge)

Cost knowledge was measured during the survey with a sequence of questions that asked the responder to choose the more expensive of two instruments or instrument groups.

Prior exposure to instrument prices was assessed as part of the survey (*exposure measure*). Surgeons were asked about prior exposure to the purchase price of disposable surgical supplies.

Familiarity with instrument prices was assessed as part of the survey (*familiarity measure*). Surgeons were asked how familiar they felt with the purchase price of disposable surgical supplies.

The *relative importance of cost vs. effectiveness to the surgeon* was assessed as a part of the survey (*C/E measure*). The surgeon was asked to scale how important cost is relative to effectiveness when choosing surgical instruments. As a sensitivity analysis, surgeons were also asked two questions related to "willingness to pay."

The measure of *recency of surgical training* was based on when the surgeon finished their clinical training.

The measure of *institutional policies* was measured using fixed effects for the 3 institutions (CSMC, UCLA, UCSD)

5-3B: Dependent Variable #2
Cost of instruments on the a priori preference card

Cost of instruments on the a priori preference card was measured as the purchase price (to the hospital) of acquiring the supplies listed on the preference card.

Perceived effectiveness was measured as a part of the surgeon survey (*Relative effectiveness sensitivity analysis)*. As described in more detail in the *Regression Specification* section, there is no ideal way to capture perceived effectiveness. However, to understand the potential bias of omitting perceived effectiveness, surgeons were asked during the survey to categorize the perceived cost and effectiveness of two instrument comparisons.

Perceived cost is not measured; instead we utilized a reduced-form model to assess the relationship between *cost knowledge (accuracy composite)* and *cost of instruments on the a priori preference card* (details under *Regression Specification*).

Actual cost of possible instruments is unmeasured but should not bias the estimated association between *cost accuracy* and the *cost of instruments on the preference card* as long as we adequately captured *cost knowledge* as part of the survey.

The measurement of *relative importance of cost vs. effectiveness to the surgeon* was discussed under dependent variable #1.

For *instrument available at facility* we used indicators for the 4 facilities (RR-OR, RR-ASC, SM-OR, SM-ASC).

5-3C: Dependent Variable #3
Cost of instruments used in the operating room (actual costs)

How *actual costs* was calculated is discussed in detail under *Variable Definitions*. Briefly, this measure reflects the purchase price of all disposable instruments used or wasted during an operation. Wasted in this context refers to items that were opened but unused, such as an item opened inadvertently or an item that was dropped on the floor. We included these wasted items as they are costs to the hospital, regardless if they were used or not.

Measurement of *instruments available at a facility* was discussed under dependent variable #2.

Resident involvement was measured by the presence or absence of a resident in the OR according to the OR log.

Numerous measures were used to capture *patient medical comorbidities and the complexity of the case*. Patient preoperative risk was measured using both demographic factors (age, sex, race/ethnicity) and pre-existing medical conditions (e.g., obesity, medical comorbidities). The inclusion of patient race/ethnicity reflects the underlying pathophysiologic mechanism of gallbladder disease. For example, Hispanics have a higher incidence and different composition of gallstones than non-Hispanics.

Case complexity is a challenging concept but the essence is that for any given patient characteristic (demographics, comorbidities), one could imagine cases varying from simple to very complex. We utilized three proxies. The first is the American Society of Anesthesiologists (ASA) Physical Status Classification score (henceforth, ASA score), which captures the severity of illness for a patient as they enter the OR. This ranges from describing patients as very fit to patients who are moribund and are likely to die within the next 24 hours. This should capture the overall severity of the patient's condition. Second, we included the urgency of the operation. Prior evidence has shown increased risk of morbidity and mortality during emergency procedures as opposed to elective procedures. Finally, we included indication for operation. There are multiple reasons someone may want/need their gallbladder removed that may influence the complexity of the case. For example, patients presenting for outpatient cholecystectomy for symptoms of biliary colic may have mild or absent inflammation allowing better delineation of the anatomy and less risk of adverse events. Conversely, patients with acute cholecystitis may have dense scar tissue, friability of the gallbladder wall, or gangrene, which can grossly distort anatomy and place them at risk of major complications. All the aforementioned variables – demographics, comorbidities, ASA score, urgency, and indication have previously been shown to predict patient outcomes in laparoscopic cholecystectomy.[77]

While we have no direct measure of *surgeon's technical skill*, as a sensitivity check we included the surgeon's operative volume for laparoscopic cholecystectomy (during the study time period) as a potential proxy. Intuitively, surgeons who perform an operation more frequently should be more adept at that operation and may also make better postoperative management decisions. There is evidence in the surgical literature supporting the association of higher surgeon volume with better patient outcomes[78]. Ideally, we would have included a fixed effect in our regression models to adjust for unmeasured surgeon technical skill, however, our models utilize the surgeon/site specific preference card as a primary regressor, in which case a fixed effect would be perfectly collinear with our predictor.

5-3D: Dependent Variable #4
Patient Outcomes

This analysis evaluated a variety of *patient outcomes* that may conceivably be influenced by the instruments used in the OR. Our primary data set was auto populated using computer algorithms which limited our analysis of outcomes to those readily extracted without manual chart review. We considered one interoperative outcome (procedure length) as well as several postoperative outcomes:

- Procedure length; we presume that surgeons care about the efficiency of their operation and therefore would choose instruments that improve this efficiency, all else being equal.
- Length of stay (LOS); we included LOS for 3 reasons – first, it is an important measure of utilization which may juxtapose or justify added costs in the OR; second, there may be technical components of the operation that may expedite recovery (e.g., adequate local anesthesia), and third, short LOS often reflects a patient recovering well from an operation and may portend improved longer-term prognosis (i.e., reduce risk of complications or readmission).

- Post-procedure escalation of care – this variable captured whether the patient required an unexpected transfer to a higher level of care; for example, a patient that is scheduled for an outpatient operation in the ASC requires transfer to the main hospital or a patient scheduled for an inpatient operation requires transfer to the ICU following a floor admission. Unexpected escalations of care represent deviations from normal recovery and may signify an imminent complication. Higher levels of care also require additional hospital costs which may negate savings from lower intraoperative supply costs.
- 30-day and 90-day return to the ED– while the utility of these measures as quality metrics is often debated, from a surgical perspective they are of value for three reasons. First, cholecystectomy usually reflects an acute episode which should not require further inpatient care after discharge; second, return to the ED may signify an evolving complication; third, return to the ED represents increased utilization which must be juxtaposed with OR expenses. A limitation of using a single health system data set is that we did not have access to ED visits to hospitals outside of UCLA. However, especially for elective surgical patients (unlike other chronic conditions, such as congestive heart failure exacerbations) it seems reasonable to assume that many patients would choose to come back to the operating surgeon/hospital, if given the chance.

Additional outcomes evaluated for the fraction of inpatient cases with NSQIP data included 30-day rates of surgical site infection, reoperation, and readmission, including whether the readmission was believed to be related to the initial operation.

Patient motivation and *patient behaviors* are unmeasured.

Surgeon's technical skill was discussed under dependent variable #3.

Patient comorbidities and case complexity were described under dependent variable #3

Quality of postoperative care; there is good evidence that quality varies between facilities. As described in the conceptual model chapter, this may reflect a variety of underlying factors such as quality of nursing care, infrastructure, resources etc. We do not have direct measures of these individual constructs but instead included fixed effects for facilities in our models to account for this.

Patient financial resources were measured based on the patient's insurance. We do not have direct measures of the patient's means (i.e., ability to pay out-of-pocket expenses, disposable income etc.). Whether or not the patient is insured, and the broad type of insurance (i.e., Medicare), does provide a measure of the patient's out-of-pocket burden and may also reflect underlying socioeconomic status. It is also likely that financial means vary by demographic characteristics (e.g., age, sex, race/ethnicity) and therefore these variables may also reflect financial resources.

5-4: Variable Definitions

5-4A: Dependent Variables
(Indicated in gray boxes in the Measurement Model)

A copy of the complete survey is included in Appendix 7. Truncated descriptions of each measure are discussed here.

Accuracy of surgeon cost knowledge – Surgeons were shown 10 instrument (or instrument group) comparisons. The surgeon was asked to identify the more expensive of the two using a categorized scale (Instrument A >>>B, Instrument A>>B…. Instrument A<<<B). The primary

operationalization of each comparison was a dichotomous variable indicating whether or not the surgeon correctly identified the more expensive item. The composite score was operationalized as the proportion of correct answers out of the total questions answered; i.e., if surgeon A answered 7 of 10 questions correctly, their score was 70%. Item-rest correlations and Cronbach's alpha were calculated for the scale to assess internal consistency.

Instrument comparisons were generated by this writer and then iteratively changed through discussions with the dissertation committee as well as investigators at the other sites. Items were further refined through cognitive interviews (described earlier). Instruments were picked that are commonly used in general surgery, including laparoscopic cholecystectomy, to be familiar to a broad range of surgeons. The goal was to generate a range of comparisons that differed in difficulty based on intuitiveness (i.e., one would be expected to correctly guess that large complex instruments are more expensive than small simple instruments) and the magnitude of cost differences (i.e., one would expect comparisons of items that differ by only a few percent to be more challenging that items that differ by orders of magnitude). Generating instrument comparisons that were relevant to each site introduced several challenges. While some items were common to all three institutions, several of the comparisons had to be tailored to the site because of different suppliers or utilization patterns. For example, one institution used Covidien as the primary supplier of their Endoclip devices, while the other two used Ethicon. Further, even with a common supplier, institutions had different utilization patterns. For example, for two institutions the primary hemostatic agent was traditional Surgicel, while the third used a more specialized hemostatic agent called Fibrillar. When suppliers or utilization patterns differed between the three institutions, institution-specific comparisons were generated. The final comparisons were confirmed by investigators at each institution as being congruent with the items available and often used.

Once the final comparisons were decided upon, we obtained current cost information for each item at each institution in order to assign a correct answer. These costs were obtained by study investigators at each institution. These costs cannot be disclosed due to sensitivity of the information. While the absolute costs did vary from one institution to the next, the direction and overall magnitude of the relationships were preserved from one site to the next. For example, if one institution paid $50 and $100 for instruments A and B, the next institution might pay $55 and $110. A list of the final comparisons is included in Appendix 7.

The cost of instruments on the a priori preference card was operationalized as follows. For each item listed on the preference card we had access to the EMPAC number, the name of the item, and the desired number opened. For each item, we multiplied the number opened by the unit price (details of how this price is determined are described below) and then the sum of these rows was the total cost of the disposable items on the a priori preference card.

Operationalizing the cost of instruments used in the operating room (*actual costs)* requires a detailed discussion, and part of this discussion has been published elsewhere.[79]

As mentioned earlier, the data warehouse allows the extraction of a list of items used in the OR along with the *total price* of those items. The process that allows this to happen is as follows.

First, a list of the items used in the OR is generated. At UCLA, the a priori preference card is loaded into the electronic health record at the time of surgery. This preference card serves as a template for intraoperative item usage. During the operation, as additional items are opened, the circulating nurse is charged with updating the list, either through barcode scanning or manual entry. The final report generated by the circulating nurse contains a list of item IDs/names, the number used, and the number wasted.

The second step is the assignment of costs. A third-party purchasing system is the source of truth with respect to instrument prices. This system is continuously updated by the purchasing department as new items are added, old items are re-ordered, and prices are re-negotiated. Real-time price data from this system flows into the electronic health record without transformation. In other words, when a retrospective report is pulled, an item's price is based on the purchasing department's price at the time of the operation. Finally, the unit prices of the items are summed together to generate a total supply cost.

While we could pull this data in its native format (i.e., total cost per case), there are a number of reasons this may produce inaccurate estimates. First, the price of individual items changes frequently due to renegotiated contracts. As a result, a single item may have different costs at different time points. Previous studies have attempted to control for this by adjusting the costs to a national index, such as the consumer price index, but a recent study showed that OR costs do not grow in-line with these indices and therefore using these indices may produce inaccurate estimates.[6] Second, during our preliminary analysis, we found that a number of items had missing or seemingly aberrant prices. This introduces a significant amount of noise in the analysis that needs to be dealt with. The concern is that surgeons may be inappropriately assigned a high cost because of when they performed the operation or because of an accounting error.

To account for changing prices and the missing/aberrant prices, we extracted detailed price data for each operation including not just the total supply cost but also the list of items used and the price of individual items. From this list, we generated a "master price list." This list contained all items used to perform laparoscopic cholecystectomy during the study time period. In adjacent columns we extracted all the prices that have been assigned to that item. For items with minimal variation over time (ratio of highest to lowest price ≤2.0), we assigned the latest price to the item. For items with large variations (ratio >2.0), we contacted our purchasing department for clarification (See Acknowledgements). Through discussions with the purchasing department, we learned that large variations often reflected the assignment of a box versus an individual unit price (i.e., the price was for a box of staplers instead of a single unit). These prices were manually corrected. Finally, a few items had missing prices, typically because the circulating nurse manually typed in the item without an item ID, preventing the flow of cost information from the purchasing department into the electronic health record. These were also manually corrected. The result of this effort was the assignment of a single price for each item. The mechanism for doing this is less important (i.e., the cutoff for our minimal variation and higher variation is arbitrary) than the assignment of a consistent price. This will ensure surgeons are being compared consistently and will eliminate noise from our analysis.

<u>Patient outcomes</u> were operationalized as follows:

- Operation length – we analyzed two variations– (1) room time and (2) procedure time; the former is measured as the minutes between when the patient enters and leaves the OR with the latter measured as the minutes between when the first incision is made and the procedure is finished.

- LOS – we analyzed post-procedure LOS, defined as the number of hours between leaving the OR and discharge.
- Post-procedure escalation of care – we first separated patients based on the admission class they were assigned to; for example, "outpatient," "surgery outpatient," and "overnight recovery" patients were grouped together, as were patients scheduled for "emergency," "inpatient," and "same day admit." Broadly, this separated patients into elective outpatient operations and inpatient operations. For the elective outpatient operations, we then generated a dichotomous escalation of care value if they stayed in the hospital for >24 hours. For inpatient operations, we generated a dichotomous escalation of care value if they transferred at any point from the floor to the ICU during the admission.
- 30-day and 90-day return to the ED– dichotomous variable; we planned to exclude inpatient deaths from the denominator, but there were none.
- For the subset of observations with NSQIP data we also analyzed:
    - Surgical site infection – dichotomous variable if the patient had any type of surgical site infection within 30 days; this included a composite of superficial, deep, and organ space surgical site infections.
    - Reoperation – dichotomous variable if the patient had an unplanned additional operation within 30 days of the index operation.
    - Readmission – dichotomous variable if the patient was readmitted for any reason within 30 days of the index operation.

- Related readmission – dichotomous variable if the patient was readmitted within 30 days of the index operation and the clinical abstractor believed the reason was related to the index operation.

5-4B: Survey Independent Variables
(Indicated with square boxes in the Measurement Model, See Appendix 7 for more details)

Prior Exposure was assessed with the following question:

> Surgeons vary in their exposure to instrument prices. They can be exposed to prices by participating on supply purchasing committees, seeing cost report cards, or asking about the price of instruments.
> How much exposure have you had to the prices of instruments you use in surgery?
> Answer: 4 category scale ranging from *none at all* to *a great deal.*

Familiarity was assessed with the following question:

> How familiar are you with the purchase price of disposable surgical supplies?
> Answer: 5 category scale ranging from *not familiar at all* to *extremely familiar*

Recency of Surgical Training was assessed by asking surgeons when they finished general surgery training and, for those that completed a clinical fellowship, when they finished their final clinical fellowship. Because some surgeons completed a clinical fellowship during their general surgery residency, for example, a surgical critical care fellowship, there were some individuals who completed fellowship before residency. This was not anticipated prior to administering the survey. We therefore operationalized this variable as the latest of the two years provided (end of residency, end of fellowship).

Surgeon's perceived effectiveness of instruments – For two of the instrument comparisons, the surgeons were asked to ascribe the perceived effectiveness of the two instruments for a given clinical scenario (see brief version below and Appendix 7 for details). The clinical scenario was developed in collaboration with the dissertation committee and refined through cognitive interviews. Surgeons were also asked to pick the instrument they would prefer to use in that clinical scenario.

While cost was categorized without an equivalent choice (i.e., A=B), the effectiveness choice did include this as an option. The logic behind this was as follows. For all item comparisons there was a nominal price difference between the two items. Adding an equal comparison would have therefore required a judgement on our part for what constitutes similarly priced items, for which we had no evidence base. It is conceivable that one surgeon may consider $5 a significant difference while the next may consider $1000 the threshold for significance. Further, we do not know if surgeons would have perceived cost difference as relative (i.e., a 50% or 100% difference) or absolute ($5 or $1000). Effectiveness, on the other hand, is a subjective value that may reflect a variety of factors such as efficiency, familiarity, technical aspects etc. We therefore felt it likely and reasonable that to a surgeon, two instruments may have similar efficacy.

Instrument A	**Instrument B**
Single-use 10 mm Endoclip with titanium clips	Single-use 5mm Endoclip with titanium clips

Which instrument is more expensive? (6 category, ranging from A >>> B to A <<< B)
Which instrument is more effective? (7 category, ranging from A >>> B to A <<< B)
Which instrument would you prefer to use in this scenario? (A, B)

<u>Relative importance of cost vs. effectiveness to surgeon</u> – Surgeons were asked one primary question and two sensitivity questions to assess the relative importance of cost versus effectiveness:

(1) Question: When choosing an instrument to be listed on your preference card, how important is the cost/price of a surgical instrument in comparison to the instrument's effectiveness?

Answer: 9 category scale from "only consider cost" to "only consider effectiveness"

(2) Question: Imagine two theoretical instruments - Instrument X costs $250 and can complete an operative step in 10 minutes; Instrument Y has recently come to market which can complete the same step in 3 minutes. The instruments are equally effective. What is the maximum price the hospital should pay for instrument Y?

Answer: Free text

(3) Question: For this question, assume a superficial wound infection is: erythema with no purulence, and a 7-day course of antibiotics as an outpatient results in full resolution of the infection. The patient suffers no additional adverse event related to the infection or the antibiotics.

Now, Imagine two theoretical instruments - Instrument X costs $250 and, when used, the probability of a superficial wound infections is 3 out of 100. Instrument Y has recently come to market which can reduce the probability of a wound infection to 2 out of 100. The instruments are otherwise equally effective. What is the maximum price the hospital should pay for instrument Y?

Answer: Free text

Our primary analysis (described below under *Regression Specification*) used the response to question 1. We also performed sensitivity analyses using the responses to questions 2 and 3. The responses to questions 2 and 3 were converted to the cost of one minute of OR time and the cost of one surgical site infection by taking the respondent's answer, subtracting $250, and then either dividing by 7 (value of one minute) or multiplying by 100 (value of surgical site infection).

5-4C: Electronic Health Record Independent Variables
(Indicated with rounded boxes in the Measurement Model)

Instrument available at facility and quality of postoperative care were operationalized as indicators representing three of the four UCLA sites – RR-OR, RR-ASC, SM-OR, and SM-ASC.

Surgeon's Technical Skill was measured using surgeon volume, operationalized as the number of procedures performed during the study time period by that surgeon (i.e., If surgeon A performed 37 laparoscopic cholecystectomies, their volume was 37).

Resident involvement was a dichotomous variable (1=present, 0=absent).

Patient financial resources were measured as a categorical variable based on the primary expected payer, including (1) Private, (2) Medicare, (3) MediCal/Other Public, and (4) Other/None. The variable was top coded in this order such that a patient with dual Private and Medicare would have been included as Private. Three indicator variables were utilized.

Patient Comorbidities and Case Complexity

Our proxies were operationalized as follows:

- Patient age – continuous, 18+
- Patient sex – dichotomized (1=female, 0=male)
- Patient race/ethnicity – categorized into Hispanic, non-Hispanic white, non-Hispanic black, non-Hispanic other with the use of 3 indicator variables
- Body mass index (BMI) – continuous
- ASA has 5 levels[80]: I (completely fit), II (mild systemic disease), III (severe systemic disease), IV (incapacitating disease that is a constant threat to life), and V (moribund patient not expected to live >24 hours without surgery). ASA also has an emergent component, such that any level can also contain an E suffix (e.g., IIE or IVE) if the case is done on an emergency basis. We included indicators for 4 of the 5 ASA categories (omitting one reference category) as well as an indicator for emergent classification.
- Urgency – the urgency of the case was captured by 5 variables (with 4 indicators): Elective (i.e., scheduled), Inpatient (i.e., unscheduled but no urgency), Urgent, Emergent,

Critically Emergent; the latter 3 assignments are based on how quickly the surgeon would like the operation to be started.

- Elixhauser comorbidity score – the data warehouse provided a list of ICD-10 codes assigned to the encounter. This allowed us to use existing software (STATA *elixhauser* command) to convert these diagnosis codes into a composite comorbidity score that has been shown in large administrative databases to correlate with mortality and readmission rates.[81] The use of an index was especially beneficial in our data set as it only consumed 1 degree of freedom instead of traditional models that require indicator variables for individual comorbidities.
- Indication for operation was extracted from the associated ICD-10 diagnosis codes assigned to the encounter. A few examples are included in Table 7. Indications were categorized into the following: (1) biliary colic / chronic cholecystitis / other elective indications (e.g., gallbladder polyp) (2) obstructive biliary processes (e.g., pancreatitis, choledocholithiasis, cholangitis), and (3) acute cholecystitis.

5-5: Regression Specification, Statistical Considerations, and Missing Data

For each hypothesis we describe: (1) the empirical measures included in the primary regression, (2) measure transformations, when relevant, (3) missing data, (4) statistical considerations, and, (5) sensitivity analyses. A summary of missing data, by hypothesis, is included in Table 8.

Prior to regression modeling, descriptive statistics were generated for all variables. Histograms were created for continuous and count variables. Normality was assessed both graphically by superimposing a normal curve and using the Shapiro-Wilk W test, whereby a statistically significant finding ($p<0.05$) indicates a lack of normality.

5-5A: Question 1

What factors predict a surgeon's accuracy of knowledge of the relative cost of surgical instruments?

Hypothesis 1a: Surgeons with prior exposure to instrument pricing will have more accurate knowledge about the relative cost of surgical instruments

Primary Model

The model included three conceptual variables, each with one measure: (1) accuracy of cost knowledge (outcome), (2) prior exposure to instrument prices (primary regressor), and (3) relative importance of cost versus effectiveness to the surgeon (moderator).

Accuracy of cost knowledge was maintained as a continuous variable, prior exposure was kept as a 4-category variable (operationalized as three indicators with an omitted reference category), and relative importance of cost versus effectiveness was reduced from a 9-category variable to a dichotomous variable with 1 = cost less important than effectiveness and 0 = cost equal or more important than effectiveness when choosing surgical instruments. Missing data for cost knowledge was singly imputed (Table 8).

The unit of analysis for the outcome as well as all covariates was the surgeon (hereafter, subscript j) and each surgeon only contributed one observation; therefore, consideration of clustering or multilevel models was not required.

The following model was fit:

Model 1a) $Y1_j = \beta0 + \beta1$ (Prior exposure$_j$) +$\beta2$ (Relative Importance$_j$) + $\beta3$(Prior Exposure$_j$ * Relative Importance$_j$) + ε_j

Where $Y1_j$ is the accuracy of surgeon's knowledge of the relative cost of instruments and ε_j is a surgeon-specific error term with an expectation of zero.

Sensitivity Analyses:

Several additional analyses were performed:

(1) We analyzed two alternate operationalizations of the original measure for relative cost versus effectiveness to ensure consistency of results. The primary operationalization (described above) was a dichotomous variable. We also tried a continuous version as well as a 5-level categorical version (with 4 indicator variables and an omitted reference category).

(2) Because we had no empiric guidance for generating a measure for the relative importance of cost versus effectiveness, we included two additional measures as a part of the survey – one asking the surgeon to value OR time and the second asking surgeons to value a surgical site infection. Values were left continuous. Missing data were multiply imputed using 20 imputations, predictive mean matching, k =10, and exposure, year finished, institution, and cost knowledge as covariates (Appendix 4). These measures were substituted for the primary measure of relative importance of cost versus effectiveness in model 1a to assess for changes in overall conclusions.

While our primary hypothesis was the relationship between exposure and cost knowledge, we were also interested in understanding the remaining pathways in the causal diagram leading to cost knowledge. As a result, we also performed:

(3) A bivariate assessment of the correlation between familiarity and exposure measures. Familiarity was collapsed from a 5-level to a 3-level categorical variable.

(4) An assessment of the association between years since finished training and exposure after controlling for institutional fixed effects:

Model 1b) $Y2_j = \beta 0 + \beta 1$ (Years since finished training$_j$) $+\alpha_k + \varepsilon_j$

Where Y2j is a 4-level ordered categorical variable for exposure, α_k are institutional fixed effects (UCLA, UCSD, and CSMC), and ε_j is the residual error term with an expectation of zero.

Years since finished training was categorized into approximately 10-year increments including surgeons who graduated before 1990, and between 1991-2000, 2001-2010, and 2011-2017. Ordered logistic regression was performed after testing the proportionality of odds assumption. (5) A reduced-form model assessing the association between years since finished training and cost knowledge, controlling for institutional fixed effects (model 1c). A reduced-form econometric model is one in which mediators on the causal pathway are excluded from the regression model so that the total effect of the antecedent variable can be estimated.

Model 1c) $Y1j = \beta 0 + \beta 1$ (Years since finished training_j) $+\alpha_k + \varepsilon_j$

Where Y1j is the accuracy of surgeon's knowledge of the relative cost of instruments, α_k are institutional fixed effects (UCLA, UCSD, and CSMC), and ε_j is the residual error term with an expectation of zero.

5-5B: Question 2

How is accuracy of knowledge of the relative cost of surgical instruments associated with the cost of a surgeon's instrument preference card for laparoscopic cholecystectomy?

Hypothesis 2a: Surgeons with more accurate knowledge of the relative cost of surgical instruments will prefer a lower-cost combination of surgical instruments.

Hypothesis 2b: The association between a surgeon's accuracy of the relative cost of instruments and the cost of an instrument preference card will be stronger among surgeons who view cost minimization as more important.

Primary model

Four conceptual variables with four measures were included in the model – (1) cost of the a priori preference card (outcome), (2) accuracy of cost knowledge (primary regressor for hypothesis 2a) (3) relative importance of cost versus effectiveness (moderating variable of interest for hypothesis 2b), and (4) facility.

Cost of the a priori preference card and accuracy of cost knowledge were left as continuous variables. Accuracy of cost knowledge was singly imputed in the same manner as question 1 (Table 8). The relative importance of cost versus effectiveness was dichotomized with 1 = cost less important than effectiveness and 0 = cost equal or more important than effectiveness when choosing surgical instruments. Facility was left as a categorical variable operationalized as a set of 3 indicators and an omitted reference group.

There were 3 statistical considerations. First, we ran a reduced-form model between accuracy of knowledge and cost of the a priori preference card. From a conceptual standpoint, this allowed us to test the relationship of interest (i.e., if we can improve a surgeon's knowledge of costs, will that alter the cost of instruments on the a priori preference card?) and second, it allowed us to avoid generating a measure for perceived cost of instruments and adjusting for actual costs, which would have been insurmountably complex.

Second, unlike question 1, question 2 introduced multiple levels of analysis. The primary outcome was at the level of the surgeon/site, whereas variables 2 and 3 were at the level of the surgeon, and variable 4 was at the level of the facility. Observations clustered within both facilities and surgeons. We previously introduced subscript j for surgeon and will now add k for site. We modeled facilities as fixed effects as this allowed generation of consistent estimators by controlling for unmeasured facility-level characteristics. For clustering within surgeons, we included a cluster variance adjustment. Fixed effects at the surgeon level would have been perfectly correlated with our predictor.

Third, we also included an analytic weight for each observation. As discussed earlier, because our data set included multiple observations for some surgeons, ordinary least squares regression may have biased our estimates toward those surgeons with more than one preference card. We therefore assigned a weight of 1.0 to each surgeon, allowing equal weighting across surgeons. Further, the distribution of cases performed by each surgeon was not necessarily even across sites. Consider a surgeon who performs 90% of her operations at site 1 and only 10% at site 2. Assigning equal weight to these preference cards is imprudent as there may be less incentive (and, ultimately, less downstream impact) for the surgeon to maintain their preference card at site 2. In this hypothetical situation, we would have assigned an analytic weight of 0.9 to site 1 and 0.1 to site 2, adding to the total of 1.0 for the surgeon and weighted based on the number of cases actually performed.

The final model was:

Model 2a) $Y_{jk} = \beta 0 + \beta 1 \text{ (cost knowledge}_j) + \beta 2(\text{cost knowledge}_j * \text{relative importance}_j) + \alpha_k + \varepsilon_{jk}$

Where Y_{jk} is the cost of the surgeon and site-specific preference card; α_k is a facility fixed effect; and ε_{jk} is a surgeon-site error term. The fixed effects and surgeon-site level error have expectation 0. Not indicated in the model, but also included, are a cluster variance adjustment and analytic weights.

Sensitivity Analyses:

Two additional analyses were performed:

(1) We analyzed several different variable operationalizations. We dichotomized our primary predictor (cost knowledge) as 0 = score below the mean and 1 = score at or above the mean. We also tested two alternate operationalizations of the original measure for relative cost versus effectiveness including continuous and 5-level categorical versions (with 4 indicator variables and an omitted reference category).

(2) Not included in the primary model is perceived effectiveness of instruments (despite its appearance in the conceptual/measurement models). Even if two surgeons have identical cost knowledge and opinions about the importance of cost, they may still pick different combinations of items based on the perceived effectiveness of those items, leading to different preference card

costs. Unfortunately, there was no complete way to model this as it would require surgeons not only describe the perceived effectiveness of their preference cards but also the relative effectiveness for all possible instrument combinations. However, to assess the significance of this potential bias we conducted the following sensitivity analyses.

Three conceptual variables with four measures were included in the model – (1) choice of instrument A or instrument B (primary outcome) for two instrument comparisons (First comparison: 10mm Endoclip versus 5mm Endoclip; second comparison: Clearify vs. anti-fog), (2) perceived relative effectiveness for each comparison, and (3) accuracy of cost knowledge. Relative importance of cost versus effectiveness was omitted from this model due to lack of significance in the primary regression (see Results).

Choice was a dichotomous variable (Instrument A or Instrument B). Perceived relative effectiveness was collapsed from a 7-level categorical variable to 2-3 categories using the empiric distribution. Accuracy of cost knowledge was operationalized in two ways: (1) whether or not the surgeon knew which item was more expensive of the two being asked about (i.e., did they know the 5mm Endoclip was more expensive than the 10mm?), and (2) their general "cost knowledge" measure as used in Question 1 (i.e., they scored 70% correct on the overall cost assessment).

Because this analysis only relied on survey data, and because we have survey data on a larger sample of surgeons (all 3 institutions), we ran this sensitivity analysis both on the complete survey data and on the subset of surgeons included in the primary regression (from UCLA). All 4 measures were available for the 26 surgeons included in the primary analysis and for 82 of 83 surgeons in the entire sample. No attempt was made to impute the final surgeon's data.

Because surgeons only contributed one observation and all analyses were at the level of the surgeon, multi-level adjustment was not necessary. However, for the analysis of all survey respondents, we did include institutional fixed effects given that we had respondents from more than one institution. The primary goal of this sensitivity analysis was to assess how the coefficient on cost knowledge changed with the inclusion of perceived effectiveness. The final models were therefore:

Model 2b) $Y_j = \beta 0 + \beta 1$ (cost knowledge$_j$) $+ \varepsilon_j + (\alpha_k)$

Model 2c) $Y_j = \beta 0 + \beta 1$ (cost knowledge$_j$) $+ \beta 2$ (perceived effectiveness$_j$) $+ \varepsilon_j + (\alpha_k)$

Where Y_j represents the choice between the two instruments (A and B); $\beta 1$ was operationalized both as the accuracy of their cost knowledge for that comparison and their overall cost knowledge (across all 10 questions); ε_j is a residual error term with an expectation of zero; α_k are institutional fixed effects that are only included when the model was run on survey data from all 3 institutions.

5-5C: Question 3

For laparoscopic cholecystectomy, how much of intraoperative cost variation is explained at the patient, surgeon, and facility levels?

Hypothesis 3a: More of the intraoperative cost variation will be explained at the surgeon compared to the patient or facility level.

The focus of this analysis was the decomposition of variance explained at the patient, surgeon, and facility levels. The primary outcome was the total cost of instruments used in the OR. Nine patient-level measures were included – age, sex, race, BMI, ASA, urgency, Elixhauser score, indication for operation, and presence or absence of a resident physician. Surgeon and facility random effects were included for the second and third levels of the model, respectively.

The operationalization of each variable was articulated earlier in this Methods chapter with the following changes:

- De-meaned and added an exponential term for age.

- Added an exponential term for Elixhauser score (variance inflation was within normal limits without de-meaning).
- Collapsed "critically emergent" and "emergent" categorizes of urgency due to very small sample size in the former.

A statistical complication was that while patients were operated on by only one surgeon at one facility, surgeons are allowed to operate at more than one facility. As a result, the multi-level structure is crossed instead of nested (see Figure 6). "Crossed" refers to combinations existing across levels (facilities, surgeons) instead of in hierarchies.

We introduce subscript i for the patient. To account for the crossed nature of the data, we performed both an additive crossed random-effects model (model 3a) and a crossed random-effects model with a random interaction (model 3b).[82]

Model 3a) $Y_{ijk} = \beta 0 + \beta(\vec{X}i) + \zeta 1_j + \zeta 2_k + \varepsilon_{ijk}$

Model 3b) $Y_{ijk} = \beta 0 + \beta(\vec{X}i) + \zeta 1_j + \zeta 2_k + \zeta 3_{jk} + \varepsilon_{ijk}$

Where $\vec{X}_i$ refers to a vector of all fixed, patient-level covariates including patient comorbidities and case complexity (age, sex, race, BMI, ASA, urgency, Elixhauser score, indication for operation) as well as the presence or absence of a resident physician; $\zeta 1_j + \zeta 2_k$ are additive, uncorrelated random effects for surgeon and facility, respectively, each with expectation 0. The residual error has expectation 0 conditional on both random effects. For model 3b, $\zeta 3_{jk}$ is a random interaction between surgeon j and facility k, allowing correlation between the random effects. In other words, this random interaction allows a random intercept for each combination of surgeon and facility. This random interaction also has expectation 0 and the expectation of the residual error for model 3b has expectation of 0 conditional on all 3 random effects.

The significance of the random interaction was tested by comparing the nested model (additive) against the model with the random interaction using a Likelihood ratio test. After estimating the variance (ψ) at each level, intraclass correlations (ICC) were calculated as follows:

- ICC surgeon = ψ (surgeon) / (ψ (surgeon) + ψ (facility) + ψ (facility * surgeon) + ψ (residual))
- ICC facility = ψ (facility) / (ψ (surgeon) + ψ (facility) + ψ (facility * surgeon) + ψ (residual))
- ICC surgeon, facility = ψ (surgeon) + ψ (facility) + ψ (facility * surgeon) / (ψ (surgeon) + ψ (facility) + ψ (facility * surgeon) + ψ (residual))

Finally, the change in proportional variance from the inclusion of patient-level factors was calculated by comparing the variance estimate at each level (facility, surgeon, residual) before and after the addition of patient-level variables.[83,84]

Four of the patient-level variables had missing data and were multiply imputed. See Table 8 for details.

5-5D: Question 4

How does a surgeon's a priori preference card for laparoscopic cholecystectomy influence intraoperative costs and patient outcomes?

Hypothesis 4a: There will be a significant positive association between the sum of instrument prices listed on the a priori preference card and the sum of actual instrument prices used in the operating room

Primary Model

The dependent variable was the cost of instruments used in the OR and the primary independent variable was the cost of instruments listed on the a priori preference card. We included the same 9 patient-level measures as in question 3 – age, sex, race, BMI, ASA, urgency, Elixhauser score, indication for operation, and presence or absence of a resident physician. The patient-level covariates were operationalized in the same way as question 3.

Because our focus was no longer on decomposing the variance between surgeons and facilities - an analytic approach that required random effects for each level – we instead included facility fixed effects to eliminate bias from unmeasured facility-level effects. We also included surgeon volume as a continuous variable as a proxy for surgeon technical skill. We continued to include surgeon random effects to account for clustering of actual case costs within surgeons. Surgeon fixed effects would not have been possible due to perfect collinearity with preference card costs.

Four of the patient-level variables had missing data and were multiply imputed. See Table 8 for details.

The final model was:

Model 4a) $Y_{ijk} = \beta 0 + \beta 1(\text{preference card cost}_{jk}) + \beta 2(\text{facility}_{k}) + \beta 3\ (\text{surgeon volume}_{j}) + \vec{B}\ (\vec{X}_{i}) + \zeta_j + \varepsilon_{ijk}$

Where Y_{ijk} is the patient-level actual cost of instruments; $\vec{X}_h$ refers to a vector of all fixed, patient-level covariates including patient comorbidities and case complexity (age, sex, race, BMI, ASA, urgency, Elixhauser score, indication for operation) as well as the presence or absence of a resident physician; ζ_j is a surgeon random effect with expectation 0; ε_{ijk} is a patient-level residual error term with expectation 0 conditional on the surgeon random effect.

Sensitivity Analysis

We ran one sensitivity analysis for Hypothesis 4a. As described earlier, surgeons can change their preference cards over time. Our patient-level database included an extended time frame – 2013 to 2018 – however, we only had access to the surgeon's latest preference card. It is possible the preference card that was active for cases early in our database was quite different from the preference card we analyzed. In order to assess the impact of using only the latest preference card, we repeated the regression model on a subset of the observations. For each preference card we knew the date the card was created and the date the card was last reviewed by the nurse administrator. We repeated our analysis on cases that occurred *after* the preference card was created and after the card was last reviewed. Of interest was the change in the magnitude, direction, and significance of the association between preference card cost and actual case cost between the two samples.

Hypothesis 4b: Controlling for patient factors and facility, higher-cost preference cards will be associated with better intraoperative and postoperative outcomes.

<u>Primary Model</u>

We ran a reduced-form model between preference card cost (primary predictor) and patient outcomes (dependent variable). As discussed earlier, a number of factors may confound the relationship between cost of instruments used in the OR and patient outcomes. Most notably, if case complexity is not perfectly captured, this may lead to a bias in the estimate of the causal impact of cost on outcomes. Because preference cards are generated in the absence of a patient, this confounding should be mitigated.

The model was similar to that used for Hypothesis 4a, substituting patient outcomes for case cost. Patient outcomes included procedure and room time (in minutes), postoperative length of stay (in hours), and dichotomous variables for post procedure escalation of care, and return to the ED within 30 and 90 days. For covariates, we included the same 9 patient-level variables and added a 10^{th} – expected payer. All measures were operationalized in the same way as Hypothesis 3 and 4a except that we de-meaned and added a quadratic term for BMI. We continued to include facility fixed effects, surgeon volume, and surgeon random effects. The model was:

Model 4b) $Y_{ijk} = \beta 0 + \beta 1(\text{preference card cost}_{jk}) + \beta 2(\text{facility}_k) + \beta 3 \,(\text{surgeon volume}_j) + \vec{B}\,(\vec{X}_i) + \zeta_j + \varepsilon_{ijk}$

Where Y_{ijk} is the patient-level outcome of interest (e.g., length of stay); $\vec{B}$ ($\vec{X}_h$) represents a vector of all patient level (fixed) variables including (1) patient comorbidities and case complexity (age, sex, race, BMI, ASA, urgency, Elixhauser score and indication for operation), (2) patient financial resources, and (3) presence or absence of a resident physician; ζ_j is a surgeon random effect with expectation 0; ε_{ijk} is a patient-level residual error term with expectation 0 conditional on the surgeon random effect.

Sensitivity Analysis

We ran one sensitivity analysis for Hypothesis 4b, looking at additional clinical outcomes for the subset of inpatient operations with NSQIP data. We first compared the inpatient cases with NSQIP data to those that did not have NSQIP data to ensure the generalizability of this cohort. We then repeated model 4b for four additional dichotomous outcomes – surgical site infection, reoperation, readmission, and related readmission. Because many of these events were rare, we had to modify the operationalization of several variables in order to prevent perfect event prediction – for example, ASA score and urgency were changed from categorical to continuous variables. Full details are included in the footnotes of the Tables.

Chapter 6: Results

6-1: Descriptive Statistics & Psychometrics – Survey

A histogram for accuracy of cost knowledge (Outcome 1) is given in Figure 7. Histograms for continuous/count covariates are presented in Appendix 8. Descriptive statistics are presented in Table 9 for all variables including mean/SD for normally distributed continuous variables, median/IQR for skewed continuous variables, and frequencies/proportions for categorical and dichotomous variables.

Accuracy of cost knowledge ranged from 40% to 100% and was approximately normally distributed (Shapiro Wilk W Test, p=0.62). Psychometrics for the cost knowledge scale are included in Table 10. Overall, the scale had poor internal consistency with an overall Cronbach's alpha of 0.11. Most surgeons (63%, 52/82) reported little or no prior exposure to instrument prices. Slightly over half of surgeons (57%, 48/83) stated effectiveness was more important than cost when choosing surgical instruments, with 35% (29/83) stating cost and effectiveness were equally important, and the remaining 7% (6/83) stating cost was more important than effectiveness.

6-2: Descriptive Statistics – Preference Cards

A histogram of the preference card cost distribution (Outcome 2) is presented in Figure 8. Costs were standardized to protect proprietary cost information and were approximately normally distributed (Shapiro Wilk W Test, p=0.09).

6-3: Descriptive Statistics – Medical Record Data

A histogram of the actual case cost distribution (Outcome 3) is presented in Figure 9. Histograms of continuous/count patient outcomes (Outcome 4) are presented in Figure 10. Histograms are presented for continuous and count covariates in Appendix 9. Descriptive statistics are presented in Table 11 for all variables including mean/SD for normally distributed continuous variables, median/IQR for skewed continuous variables, and frequencies/proportions for categorical and dichotomous variables.

As expected, actual case costs, procedure times, and postoperative length of stay were all skewed to the right (i.e., positively skewed), most significantly for postoperative length of stay. The additional outcomes – post procedure escalation of care and return to the ED within 30 and 90 days – ranged in frequency from 2% (16/993 for transfer from floor to the ICU) to 19% (154/824 for transfer from outpatient to inpatient care). Median surgeon volume in our sample was 60 cases over the study time period. The patient profile was relatively diverse, including a mix of patient demographic characteristics (race/ethnicity, primary expected payer) as well as clinical severity (ASA, urgency, and indication for operation). For example, 64% of patients in the sample were female, 29% were Hispanic, and 52% and 38% were undergoing surgery for biliary colic and acute cholecystitis, respectively.

6-4: Question 1

Hypothesis 1a: Surgeons with prior exposure to instrument pricing will have more accurate knowledge about the relative cost of surgical instruments

6-4A: Regression Analysis

Results are included in Table 12. No level of exposure added significant predictive power to the model, either as an individual indicator or when tested together. Relative importance of cost versus effectiveness also did not add to the model, either as a main effect or as a moderator.

6-4B: Sensitivity Analyses

(1) *Analyzing different operationalizations of the original measure for the relative importance of cost versus effectiveness*: Results of the primary regression did not change substantively with different operationalizations of this variable.

(2) *Two alternative measures for the relative importance of cost versus effectiveness (value of surgical site infection and value of OR time):* Regression tables are presented in Tables 13 & 14. There was no moderating effect using the value of a surgical site infection as the moderator. While there was a significant interaction term when using the value of OR time as a measure for the relative importance of cost versus effectiveness, the effect across levels of the interaction term was not monotonic, nor did we have any significant main effects, rendering the interpretation of this effect unclear.

(3) *A bivariate assessment of the correlation between familiarity and exposure measures*: The correlation between exposure to instrument pricing and a surgeon's perceived familiarity with instrument prices was 0.72 (Figure 11), supporting the purported relationship in the measurement/conceptual model ($p<0.001$).

(4) *An assessment of the association between years since finished training and exposure to instrument pricing after controlling for institutional fixed effects:* $Y_j = \beta 0 + \beta 1$ *(Years since finished training*$_j$*)* $+ \alpha_y + \varepsilon_j$. Results of the ordered logistic regression are shown in Table 15. Tests failed to reject the assumption of the proportionality of odds. After adjustment for multiple comparisons (Sidak's method), those surgeons who finished training between 1991 and 2010 had higher odds of more exposure than those surgeons who finished training after 2010. Surgeons from the institution (UCSD) with cost report cards had higher odds of more exposure than either of the institutions without the cost report cards.

(5) *A reduced form model assessing the association between years since finished training and cost knowledge, controlling for institutional fixed effects:* $Yj = \beta 0 + \beta 1$ *(Years since finished training*$_j$*)* $+ \alpha_y + \varepsilon_j$. The results of the reduced form model are presented in Table 16. While one of the coefficients for categorical year was significant (1991-2000 compared to the reference group of <=1990), a joint significance test of years since finished training did not add to the model, nor did institution.

Combining the results of the primary regression and the 4th and 5th sensitivity analyses, we conclude that while institution and years since finishing training may affect exposure, because exposure does not affect cost knowledge, institution and years since training do not affect cost knowledge in a reduced-form model.

6-5: Question 2

Hypothesis 2a: Surgeons with more accurate knowledge of the relative cost of surgical instruments will prefer a lower-cost combination of surgical instruments.

Hypothesis 2b: The association between a surgeon's accuracy of the relative cost of instruments and the cost of an instrument preference card will be stronger among surgeons who view cost minimization as more important.

6-5A: Regression Analysis

Results of the regression analyses are presented without (Table 17) and then with (Table 18) the relative importance of cost versus effectiveness as a moderator.

Facility was a significant predictor of preference card cost, with the SM-OR having higher average preference card costs than both the RR-OR and RR-ASC; the SM-ASC also had higher average preference card costs than the RR-ASC. After controlling for facility, we found a negative association between cost knowledge (correct percent) and preference card cost such that for every 1 additional percent correct, the average preference card cost went down by $2.31. However, this association was not significant at an α of 0.05 (p=0.08). A post-hoc power calculation, assuming the same coefficients and standard deviations, estimated power at 0.31; to achieve a power of 0.80 would have required a sample of 132 preference cards. The addition of the relative importance of cost versus effectiveness as a dichotomous variable did not add to the model as either a main effect or as a moderator (likelihood ratio test had a X^2 statistic of 0.02 and p-value of 0.99).

6-5B: Sensitivity Analyses

(1) Dichotomizing cost knowledge above and below the survey mean (66.5%) resulted in largely the same results – the coefficient on cost knowledge in the regression for preference card costs was -$67.95, with a t-statistic of -1.87 and a p-value of 0.07. The alternative operationalizations of the original measure of relative importance of cost versus effectiveness did not add to the model, nor did they change the sign or magnitude of associations in any substantive way.

(2) The overall bias from omitting perceived effectiveness from our primary regression model appeared to be small. There were 8 possible combinations of models – 2 distinct instrument comparisons (Endoclip vs. Endoclip and Clearify vs. anti-fog), each with 2 different operationalizations of the primary predictor (cost knowledge for the individual comparison or "overall cost knowledge" from all 10 questions), with 2 different samples (UCLA-only and all respondents). Because surgeons almost uniformly knew (98%) Clearify was more expensive than anti-fog, two models had no variation in the primary predictor and therefore did not run. The remaining models are included in Tables 19-21.

Table 19 shows the association between cost knowledge for the Endoclip comparison and their choice of Endoclip devices with just cost as a predictor (first row) followed by the addition of perceived effectiveness (second row). The top of the table is limited to UCLA surgeons with laparoscopic cholecystectomy preference cards (same sample as the primary regression) and the bottom shows all survey respondents with complete data (n=82) from all 3 sites. For this comparison, surgeons who knew which item was more expensive were less likely to choose the more expensive item. With the addition of perceived effectiveness, the direction, magnitude, and significance of the cost knowledge coefficient remained almost identical. These same findings were seen when we used general cost knowledge instead of instrument-specific cost knowledge (Table 20) and when we looked at the Clearify vs. anti-fog comparison (Table 21).

6-6: Question 3

Hypothesis 3a: More of the intraoperative cost variation will be explained at the surgeon than the patient or facility level.

Results are presented in Table 22. The likelihood ratio test comparing the additive multi-level model (surgeon and facility random effects only) to the model which also included a random interaction between surgeon and facility was significant with a X^2 of 23.8 and P-value <0.001. All subsequent analyses therefore included the random interaction term.

The relationship between patient characteristics and actual case costs was largely as anticipated, with higher costs associated with older patient age, male sex, higher ASA score, increased comorbidity burden, and for patients undergoing surgery for acute cholecystitis versus biliary colic. There was no statistically significant association between patient race, the urgency of the operation, and the presence or absence of a resident physician and actual case cost.

After controlling for observed patient-level characteristics, the proportion of variance explained by the surgeon and facility was 34% with 14% for the surgeon alone, 13% for facility alone, and 7% attributed to the interaction between surgeon and facility.

Comparing a null model (without patient factors) to the full model (with patient factors), the proportional reduction at the residual level was 7%, at the facility level was 47%, and at the surgeon level was negative 39% (i.e., residual increase). In other words, the inclusion of patient-level factors explained 47% of the cost variation initially identified between facilities but only accounted for 7% of the within-surgeon and within-hospital (i.e., between patient) residual variation. Interestingly, the inclusion of patient-level variables increased the relative variation explained at the surgeon level, i.e., the effect of the surgeon on case cost was actually greater after patient case-mix adjustment.

6-7: Question 4

Hypothesis 4a: There will be a significant positive association between the sum of instrument prices listed on the a priori preference card and the sum of actual instrument prices used in the operating room

6-7A: Regression Analysis for Hypothesis 4a

Results are presented in Table 22. For every one dollar increase in preference card cost, actual case cost increased by 79 cents. This relationship was significant with a p-value <0.001. The relationship between patient-level covariates and actual case cost was mostly as expected – patients who were older, had higher ASA scores (ASA 3 vs. ASA 1), or were undergoing operations for more complex indications (acute cholecystitis vs. biliary colic) had higher mean case costs. Female patients had consistently lower adjusted mean case costs than male patients (coefficient -$44.36, p-value <0.001). We found no association between patient race, BMI, case urgency, or Elixhauser score and mean case cost. We also found no association between surgeon volume or the presence of a resident physician and mean case cost.

6-7B: Sensitivity Analysis for Hypothesis 4a

Of the 1817 cases included in the primary regression analysis, 891 of these occurred after the available preference card was generated and after the last review date. This restriction reduced the number of surgeons from 23 to 21 and the number of preference cards from 36 to 33.

The results are presented in Table 23. In this subsample, for every one dollar increase in preference card cost, actual case cost increased by 84 cents. This relationship continued to be significant with a p-value below 0.001. The small increase in magnitude of the relationship (from $0.79 to $0.84) likely reflects the fact that this subset of cases is more likely to have occurred with the actual preference card analyzed. However, the overall direction and significance remained largely the same, and the significance of the magnitude change appears negligible, reassuring us that the use of all preference card data does not significantly bias our association of interest.

Hypothesis 4b: Controlling for patient factors and facility, higher-cost preference cards will be associated with better technical/clinical outcomes.

<u>6-7C: Regression Analysis for Hypothesis 4b</u>

Results are presented in Table 25. After controlling for patient-level covariates, between facility differences, surgeon volume, and surgeon clustering, we found no significant associations between preference card cost and procedure minutes, room minutes, postoperative length of stay, post procedure escalation of care (either from outpatient to inpatient, or from floor to ICU) or return to the ED within 90 days. We did find a marginally significant effect (p=0.050) of preference card cost and reduced odds (OR 0.83 for every $100 increase in preference card cost) of return to the ED within 30 days. A $100 increase in preference card cost was associated with an 18.3 percent risk reduction or an absolute risk reduction of 1.6 percentage points.

<u>6-7D: Sensitivity Analysis for Hypothesis 4b</u>

Of 1817 patients in the main regression, 992 were for inpatient operations, of which 329 had NSQIP data. The comparison of those with NSQIP data to those without NSQIP data is included in Table 26. Inpatients with NSQIP data were similar to those without NSQIP data across most covariates including facility, patient age, sex, BMI, comorbidity score, ASA, urgency of operation, presence or absence of resident, and indication for operation. The only difference was in the distribution of primary expected payer.

The regression results for the NSQIP outcomes estimated using the subsample with NSQIP data are presented in Table 27. Reoperation was too rare of an event for any model to converge. After adjustment, we found no association between preference card cost and the odds of a surgical site infection, readmission, or related readmission. However, our conclusions from this analysis may be limited due to small samples and rare events.

Chapter 7: Discussion

Several previous evaluations have attempted to reduce variation in intraoperative supply costs, but the efficacy of these efforts has been modest – generally reducing supply costs by less than $100 per case[85]. In part this reflects a "jumping of the gun," as investigators have attempted to reduce costs before understanding what has contributed to variation. If the drivers of variation are not first fully understood, there is little hope that interventions to reduce variation and cost will succeed. Thus, we performed a detailed analysis of intraoperative supply cost variation for the most common general surgery operation – laparoscopic cholecystectomy – at a multi-facility academic health system in Southern California. Combining this data with results from a multi-institutional survey, we explored the chain of events that leads to intraoperative cost variation, starting with surgeons' exposure and knowledge of the price of instruments, the effect of this knowledge on their instrument preferences, and how instrument preferences are related to intraoperative costs and the outcomes of their patients.

7-1: Interpretation of Results & Comparison to Prior Literature

A summary of our hypotheses and regression results are included in Table 28. The discussion below summarizes these findings by question, provides context and potential explanations for the findings, and reviews these findings in relation to previous literature.

7-1A: Question 1
What factors predict a surgeon's accuracy of knowledge of the relative cost of surgical instruments?

Previous evaluations of surgeon cost knowledge were limited to asking surgeons for free text response of the price of an instrument, without providing a relevant comparison.[44-48] Not surprisingly, surgeons (and physicians in general) perform poorly with this task. Our use of comparisons was designed to reflect a more real-world process, emulating the decision that would be made while a surgeon generates their preference card or while they are in the OR. It seemed reasonable that surgeons may perform well on this assessment because these are decisions they face every day and many of the comparisons were quite intuitive. When surgeons in this study were asked a series of questions to identify the more expensive of 2 instruments or instrument groups, they performed better than chance, but only marginally so (the average score was 66.5%). The fact that the overwhelming majority of surgeons (93%) incorrectly answered at least 2 of 10 questions suggests far from perfect knowledge.

Older surgeons had more exposure to instrument prices and that exposure was associated with increased familiarity with instrument prices, but there was no association between exposure and knowledge of relative instrument costs. Further, surgeons at one institution (who had received cost report cards for two years at the time of the survey) had higher self-reported exposure but performed no better on the cost assessment than surgeons from the other two institutions. Taken together, we believe this suggests that the missing link in the path between experience, exposure, and cost knowledge is the lack of an association between exposure and cost knowledge. This may explain, in part, why previous literature has found no relationship between experience and cost knowledge.[45-47] Finally, we found limited evidence of a moderating effect of the relative importance of cost versus effectiveness on the relationship between exposure and cost knowledge.

7-1B: Question 2

How is accuracy of knowledge of the relative cost of surgical instruments associated with the cost of a surgeon's instrument preference card for laparoscopic cholecystectomy?

We found a negative association between surgeon's cost knowledge and the cost of instruments on their preference card for laparoscopic cholecystectomy. While the coefficient was not significant at an α of 0.05, this likely reflects inadequate power from a sample of 38 preference cards. However, if we assume the coefficient is unbiased, the magnitude of effect is noteworthy. If a 1 percent increase in cost knowledge can reduce preference card cost by $2.31, preference card cost may be reduced by $138.60 if cost knowledge increased from the sample minimum to maximum. Alternatively, and perhaps more realistically, moving a surgeon from below the sample average to above average for cost knowledge might reduce preference card cost by $67.95. This cost reduction is similar in size to interventions previously described in the literature.[52,53,55,56] As with Question 1, we found no moderating effect of perceived importance of cost versus effectiveness.

Finally, our sensitivity analyses provided confidence that the omission of perceived effectiveness did not significantly bias the association between cost knowledge and the cost of the surgeon's preference card. We were unable to model perceived effectiveness for an entire preference card, but we were able to ask surgeons about the influence of cost and perceived effectiveness on two instrument choices. In none of these analyses did the direction or statistical significance of the coefficient between cost knowledge and instrument choice change.

7-1C: Question 3

For laparoscopic cholecystectomy, how much of intraoperative cost variation is explained at the patient, surgeon, and facility levels?

In our evaluation of 1817 laparoscopic cholecystectomies, we found that the surgeon and facility combined explained 34% of the variation in intraoperative disposable supply cost. While the surgeon alone explained slightly more (14%) than the facility alone (13%), it is clear that both are important when considering supply costs. Worth repeating is the fact that our analysis included only one institution, but this health system included two distinct geographic locations, each with two different settings (Main OR vs. ASC), for a total of four facilities. Previous studies from one "medical center" may have included multiple facilities, using our definition, but did not include facility as a covariate.[61-63,66] The few studies that explicitly included multiple hospitals/facilities have been limited to bivariate analysis[64] or single-level linear regression.[65] While these previous studies identified variation in cost by hospital and surgeon, the lack of multi-level modeling prevented the decomposition of variance.

We identified a few patient-level variables that were predictive of intraoperative case cost. For example, operating on a patient with ASA 3 vs. ASA 1 increased average case cost by $47.98; or operating for acute cholecystitis vs. biliary colic increased average case cost by $65.03. Previous literature generally supports the associations we found between patient-level variables and case costs, with the most robust of these prior analyses also finding that females, older patients, and those with more complexity had higher case cost.[66] Nevertheless, it is important to remember that these variables only reduced residual variation by 7%, despite the inclusion of 9 conceptually-driven measures. Finally, the proportion of variation explained at the surgeon level increased with the inclusion of patient-level variables. This is the opposite of what we usually find – where case-mix adjustment narrows difference in outcomes between providers

– but may suggest that costs, unlike other outcomes (e.g., quality measures) do not follow the same rules.

7-1D: Question 4

How does a surgeon's a priori preference card for laparoscopic cholecystectomy influence intraoperative costs and patient outcomes?

Preference card cost was a strong positive predictor of case cost, with every dollar increase in preference card cost associated with a 79-cent increase in case cost. Hypotheses 3a and 4a utilized similar regression models with the main differences being that the model for 4a included preference card cost and included facility fixed effects instead of random effects. The coefficients for the patient-level variables (e.g., age, sex, race etc.) were almost identical in the two models. This has two implications. First, the inclusion of facility fixed effects did not significantly bias the findings we identified in Hypothesis 3a when we used facility random effects. Second, and more importantly, that preference card cost is likely unrelated to these patient-level variables, confirming our proposition that preference cards are generated without a particular patient population in mind. When we repeated this analysis on only the subset of cases after the creation date and last reviewed date of the preference cards, the magnitude and significance of preference card cost coefficient remained largely the same. This suggests that surgeon preference cards only change nominally with each review and that surgeon preferences are stable over time.

Finally, we found no significant associations between preference card cost and most patient outcomes, such that surgeons with lower preference card costs do not necessarily risk patient safety. The only notable exception was a marginally significant value (p=0.050) for return to the ED within 30 days. While there are reasons to believe this finding may be spurious – such as lack of multiple comparisons adjustment and the fact that this includes any reason to visit the ED – if true, the finding does have significant implications. From a purely economic perspective, for every $100 increase in preference card cost the absolute risk reduction in 30-day return to the ED was 1.6%. This equates to a number needed to treat of 62.5. Estimating the cost of an ED visit is beyond the scope of this dissertation, but if this cost is greater than $6250 (the added cost of the preference card multiplied by the number needed to treat), attempts to reduce preference card cost may be counterproductive. Evidence that this finding may not be spurious can be found in the sensitivity analyses where higher cost preference cards were associated with reduced readmissions, including related readmissions, although these did not reach statistical significance. The association between preference card cost and patient outcomes stands in stark contrast to the previous literature, none of which found a significant association between cost and outcomes.[59,62,63,65,66] However, all of these previous studies related case costs to outcomes instead of preference card costs to outcomes. We know, both conceptually and empirically (from this analysis), that case costs and patient outcomes are both influenced by patient complexity, and therefore the associations identified in these previous studies may be biased by unmeasured patient complexity.

7-2: Limitations

Before we discuss the implications of these findings, we must acknowledge a number of limitations. We will first discuss threats to internal validity, followed by threats to external validity.

7-2A: Internal Validity

We partition threats to internal validity into those related to the data source and those related to our analytic approach.

(1) Data source – survey

We had limited empiric data to guide the generation of the survey instrument. We relied heavily on our conceptual model to specify the survey components and to create the individual questions. We revised these questions iteratively with the help of the dissertation committee and also through six cognitive interviews divided over two institutions. Despite these efforts, there are a number of potential issues with our survey instrument.

Our questions may not have accurately captured the underlying concept. An example is our measure of the surgeon's perceived importance of cost versus effectiveness. We ultimately generated three questions, each with a different approach, and with different response scales. We generally found no moderating effect of this perceived importance on surgeon cost knowledge or on instrument preferences. Conceptually it still seems likely that surgeons who prioritize cost more than other surgeons may seek price information, retain price information, and ultimately choose cheaper instruments, therefore the lack of association in this study is surprising.

The cost assessment was the primary outcome for Question 1 and a primary predictor for Question 2, making it an integral component of this analysis, however, our psychometric evaluation suggested that our scale had poor internal consistency reliability. Our scale is therefore difficult to interpret and may explain why we found no association between any of our covariates and cost knowledge. While we cannot say with certainty the reason for the poor internal consistency reliability there are several possible contributors. First, we forced all three institutions to have the same instrument comparisons. While we took steps to make comparisons institution-specific – by using manufacturer-specific (e.g., Covidien vs. Ethicon) versions and using the most common version of each instrument (e.g., Green vs. Blue staple loads) – the comparisons themselves may not be as relevant for some institutions than for others. For example, the survey included two clip appliers – a 5mm and 10mm version – however, at one of the institutions, a number of surgeons use an entirely different system (hem-o-lock) which was not included. Second, our assignment of prices to these items (to consider the answer "correct" or "incorrect") was taken at one point in time. It is conceivable that by the time the surgeon took the survey, the price comparison was no longer valid. Third, it is possible that surgeons correctly identified the more expensive item for the "easy" questions (e.g., Clearify vs. Anti-fog, where 98% of surgeons recorded the correct answer), and simply guessed at the remainder. All three of these problems would add significant noise to our results which may explain the low Cronbach's alpha.

(2) Data source – preference cards

There are two primary limitations to our use of preference cards. First, in an ideal world, for each operation (in the 1817 case database), we would have had access to the surgeon's preference card *at the time* of the operation. Instead, we only had access to the surgeon's latest preference card. Therefore, we were retrospectively assigning a preference card that may have been different from the one used at the time of the operation. If a surgeon systematically increased or decreased the price of their preference card over time, our estimate would be biased towards the null for Hypothesis 4a. Evidence of this bias is the fact that our coefficient for the association between preference card cost and case cost increased when we looked at only cases that occurred after preference card creation and last reviewed date. Second, and perhaps more importantly, these preference cards may not reflect the surgeon's *actual* preferences. It is possible that much of a surgeon's preference is known to the nurses that work with him/her on a regular basis.

Alternatively, the nurse in the OR may call the surgeon prior to an operation to identify the items they want instead of referring to the preference card. If these interactions and relationships are common, there would be little incentive for surgeons to update their preference cards to reflect their actual preferences. The direction of bias introduced by this is unclear, but would likely introduce noise in our estimate of preference card costs biasing our association towards the null hypothesis.

(3) Data source – electronic medical record data

There are a number of potential limitations to using electronic medical record data. First, as with all database analyses, the data generating process relies on billing codes and computer algorithms which are susceptible to typographic errors, misclassifications, and missing data. For missing data, we conducted occasional manual chart review and performed multiple imputation, but we have no systematic way of identifying errors and misclassifications. For example, we initially intended to analyze readmission rates for the entire cohort of patients. As we manually reviewed charts to ensure the data were accurately capturing readmissions, we found that some cases were being identified as readmissions when the patient was coming in for an outpatient x-ray or colonoscopy (which are not readmissions), and conversely, some patients who were actually readmitted were not being captured. As a result, this variable was unusable, and we had to instead analyze the subset of patients with NSQIP data. It is possible there were additional misclassifications that we did not identify. Second, our identification of cases may have resulted in the inclusion of non-comparable cases and exclusion of cases that should have been included. In an ideal world, a clinical abstractor would have reviewed the history and physical and operative note for each operation to ascertain what the operation was and the indication. Instead, we made a series of judgments based on procedure name, primary DRG, and assigned CPT codes to infer the operation and we used ICD-10 codes to infer the diagnosis. Data errors and misclassification of cases would generally bias our estimates towards the null hypothesis. Third, errors in the assignment of instrument prices may have inadvertently over- or under-estimated instrument prices. If one surgeon or one facility used these items more frequently than other surgeons/facilities, this may have biased the variability we identified in each level of the model. Fourth, for the entire cohort, we only had access to all-cause reasons for ED visit within 30 and 90 days. It is likely the probability of a patient coming back for reasons related to their operation (ED visits of interest) is lower, and therefore our estimate was biased away from the null hypothesis. Evidence for this can be seen in the sensitivity analysis where we looked at readmissions. Because related readmissions are rarer than all readmissions, the standard error for the estimate is larger, and we are less likely to reject the null hypothesis. Fifth, and finally, we were limited in the outcomes that were available to us. A comprehensive analysis of outcomes would have included additional intraoperative outcomes such as estimated blood loss and rates of transfusion, as well as postoperative outcomes such as bile duct injury. Outcomes would have ideally been available for the entire cohort, validated by a clinical abstractor, and captured over a longer time horizon.

(4) Analytic approach – question 1

The primary limitation for question 1 has already been addressed. Because our cost knowledge scale was not internally consistent our finding that there was no association between cost exposure and cost knowledge may be expected.

(5) Analytic approach – question 2

The two primary limitations for our analysis of question 2 were sample size and the omission of perceived effectiveness. Because we only had access to preference cards at one of the three survey sites, and because some of the surveyed surgeons do not have laparoscopic cholecystectomy preference cards, our sample for this analysis was small, limiting power. The magnitude and sign of the association we identified were as anticipated but our analysis was under-powered to detect this difference. Second, our conceptual model indicates that the choice of items surgeons list on their preference card should reflect the combination of perceived costs and perceived effectiveness of those items. We were not able to measure the perceived effectiveness of the preference card (and all possible alternatives), so instead performed a sensitivity analysis assessing the effect of adding perceived effectiveness into a simplified equation looking at individual instrument comparisons. While these analyses concluded the magnitude of the bias was likely small, it is unlikely that surgeons make decisions on instruments in a vacuum – but instead in the context of all other instruments that are already open and could be opened – therefore this sensitivity analysis was artificial and may not accurately capture the bias of perceived effectiveness on the relationship between cost knowledge and preference card cost.

(6) Analytic approach – question 3

The primary limitation of our decomposition of variance is the sample size at the facility level. Random effects models perform better with a large number of clusters. However, we only had access to data from four facilities. Further, because surgeons can operate at more than one facility, we were forced to run a crossed random effects model and to include a random interaction term. While this was the correct decision analytically, it significantly limits interpretability. In our analysis, 7% of the variation we identified was related to an interaction between the surgeon and facility. Given that we found similar variance explanations at the surgeon and facility levels, it would have been nice to parse this variation to know definitively whether the surgeon or facility explain more variation. Finally, we had a significant residual variance, even after including nine patient-level covariates, suggesting either substantial noise in the model or that there are a number of unobserved patient-level variables.

(7) Analytic approach – question 4

The most important limitation to this analysis is a lack of measure for surgeon technical skill. The analysis associating cost with patient outcomes improves upon the existing literature by utilizing a reduced-form model between preference card cost and patient outcomes. This eliminates the confounding from unmeasured patient complexity. However, the remaining issue is that a surgeon's inherent "skill" may affect both the instruments they list on their preference card and the outcomes of the patients they operate on. Specifically, it is conceivable that a highly talented surgeon may use cheap instruments and also have excellent patient outcomes, biasing our association towards the null hypothesis. While we attempted to control for this through the addition of surgeon volume, this may not be sufficient. The inclusion of surgeon fixed effects would be ideal, but, because preference card cost is a surgeon-level covariate, the predictor would be collinear with the fixed effect. Unfortunately, measures of a surgeon's technical skill are still very much in their infancy, and to our knowledge, there is no readily available measure that we can add to our model that would improve our capture of this concept.

7:2B: External Validity

We identify two primary limitations as it relates to external validity – the hypothetical nature of the survey and the generalizability of our study cohorts.

First, our survey was a hypothetical exercise. The survey introduced implausible assumptions; for example, when surgeons were asked about the value of OR time, they were told that two theoretical instruments had identical efficacy. For our perceived effectiveness sensitivity analysis, surgeons were only given the choice of two instruments and were provided a single, narrow, clinical vignette about the situation with which they could use those items. How surgeons weigh cost and effectiveness will likely vary by the severity of the situation such that for a routine, low-risk, operation they may be more willing to accept lower costs and lower efficacy, but in a high-risk, life-and-death situation, this trade-off may be less tenable. Finally, our cost assessment included only ten instrument comparisons – a number chosen to be manageable to a group of respondents with limited time – but with so few items, it is possible that if the same surgeons were given a different set of questions, their cost knowledge score may have been different.

Second, the survey was administered to general and subspecialist surgeons at three academic hospitals in southern California. The results of the survey may not generalize to non-academic hospitals, facilities outside of southern California, or other surgical specialties. Orthopedic operations, for example, have been the initial target for emerging value-based payment methods – such as bundled payments – and therefore this specialty may be more attuned to the cost of the instruments they use[86,87]. Teaching hospitals are unique in their educational mission and are generally not-for-profit; non-teaching and for-profit hospitals may have different institutional policies, may recruit surgeons with different opinions on the importance of cost, and may place higher emphasis on cost control. The analysis of preference cards was further limited to a single institution and a single operation. It is possible that the negative association we found between cost knowledge and preference card cost may not apply to other procedures (e.g., laparoscopic appendectomy, hernia repair) even within the same institution. Finally, our medical record analysis was also restricted to a single operation at a single institution, and further, our sub analysis of NSQIP data only applies to inpatient operations.

7-3: Implications of Findings and Directions for Future Research

Our analysis found that, even after adjustment for patient case-mix, there is significant variation in disposable supply costs for laparoscopic cholecystectomy between surgeons and facilities. The cornerstone of intraoperative supply costs appears to be the surgeon's preference card. Surgeons who are able to accurately discriminate the cost of common general surgery items may choose cheaper preference cards but passive exposure to instrument costs, such as through cost report cards, do not appear to increase cost knowledge. Finally, and perhaps most importantly, surgeons with more expensive preference cards do not appear to improve their patients' short-term clinical outcomes.

With rising healthcare costs, there is increasing pressure on the part of payers to implement value-based programs. One example of value-based payment is bundled payments – whereby a payer, such as Medicare, pays a set amount for the 90-day perioperative period beginning with hospital admission and extending through the OR, postoperative care, skilled nursing, home health etc. Bundled payments for surgery were introduced as a part of the 2015 Medicare Access and CHIP Reauthorization Act (MACRA). One of the major pathways introduced in this legislation were Alternative Payment Models, of which, bundled payments for surgery was one example. The merits and critiques of this program are beyond the scope of this dissertation, but ultimately the focus of this program (and many other value-based payment programs) is capitating reimbursement and shifting the financial risk away from the payer and onto the provider and hospital. To compete in these programs, hospitals must therefore identify and eliminate excessive or wasteful costs. This study, along with others, suggests that identical patients operated on by two different surgeons may result in entirely different intraoperative supply costs – and yet their outcomes may not be any different.

This analysis suggests that preference cards may be one promising focus of efforts to identify and eliminate supply cost waste. Asking surgeons to refrain from using an expensive item while a patient is on the operating table will, beyond being met with understandable hostility, likely be ineffective. The preference card, on the other hand, is generated during a non-life-or-death moment, can influence the cost of every subsequent case, and appears to stay relatively constant over time. Interventions that target the moment a preference card is generated – perhaps by having an OR nurse walk them through their various options (and their costs) - may be effective. An unexpected finding in this study was the variability in preference cards from one facility to the next even for the same surgeon. Asking surgeons to generate one preference card across facilities may be reasonable. Further, because of the way preference cards are generated at the studied institution (copied from an existing surgeon's card), the difference in supply costs across facilities may reflect a cyclic process where a senior surgeon's preference card cost (high or low) subsequently dictates the cost of future generations of surgeons at that facility.

Promoting cost awareness among surgeons is not an easy, nor particularly effective, task. Passive exposure through retrospective cost reports does not appear to increase cost knowledge and therefore cannot be expected to reduce costs downstream. This does make intuitive sense;

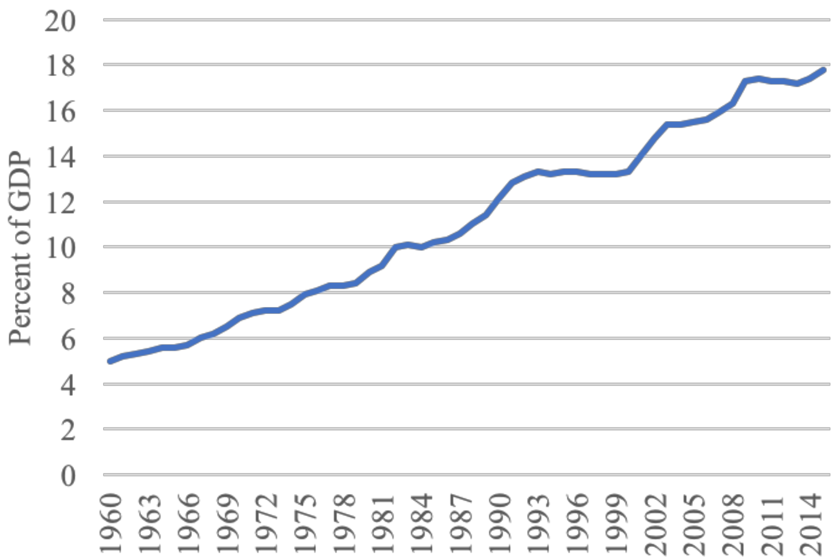
translating a passive cost report card to actual cost changes would require significant effort on the part of surgeons – to first open, read and study the report card; to then seek out information about alternative instruments and their prices; and finally, to identify the nurse or individual that can change their preference card. Eliminating these steps and integrating cost knowledge in an active, not passive, fashion will likely be more effective.

This field of study is still very much in its infancy. While we believe this analysis has helped explain the complex interaction between exposure, cost knowledge, preference cards, actual costs, and patient outcomes, there is still much work to be done. First, broadening these analyses is paramount – to other operations, institutions, and settings - in order to ascertain the generalizability of these findings. There are significant challenges that will be faced in this process. First, hospitals are unlikely to share proprietary cost information. In part, this reflects efforts on the part of medical device manufacturers which impose non-disclosure agreements on hospitals to prevent sharing prices. And second, analyzing supply prices is challenging because they constantly change and many prices may be aberrant or missing. Correcting these issues in this study was labor intensive and may not be easily automated. Survey research related to surgical costs needs significant help. Efforts are needed to produce psychometrically sound measures for all the domains addressed here – experience, exposure, knowledge etc. None of the studies identified in our literature review reported psychometric evaluations of their cost knowledge scales. It seems likely their results would parallel our own with poor internal consistency, and therefore this approach for measuring cost knowledge (asking physicians to guess the price of an instrument) should be abandoned. Our modified approach – asking about comparative pricing – also does not appear to be the solution. Finally, with respect to survey research, an interesting future analysis would include a control group (perhaps of non-surgeon physicians, or even a lay audience) to see whether surgeons actually perform better than chance when asked questions about instrument prices. Qualitative studies of how surgeons make instrument decisions would be beneficial. In our few cognitive interviews, we found surgeons had very different opinions of what they considered an "effective" instrument, which may make for an interesting line of inquiry. Finally, studies are needed to test the efficacy of cost-reducing interventions. Direct comparisons of passive and active interventions, such as cost report cards vs. standardization, and especially randomized designs, would inform the utility of these efforts.

The role of the contemporary surgeon is changing. For better or worse, surgeons can no longer singularly focus on improving the outcomes of their patients--they must also consider the resources that go into those efforts. Recognizing that waste exists and generating mechanisms for eliminating that waste will be vital to the financial viability of our practices and healthcare at large.

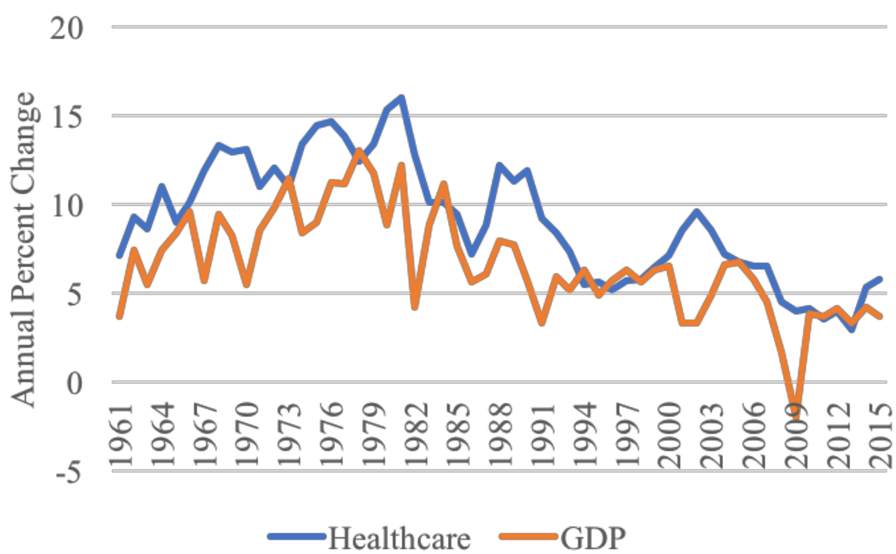
Figures & Tables

Figure 1: Healthcare expenditures as a proportion of United States Gross Domestic Product (GDP)



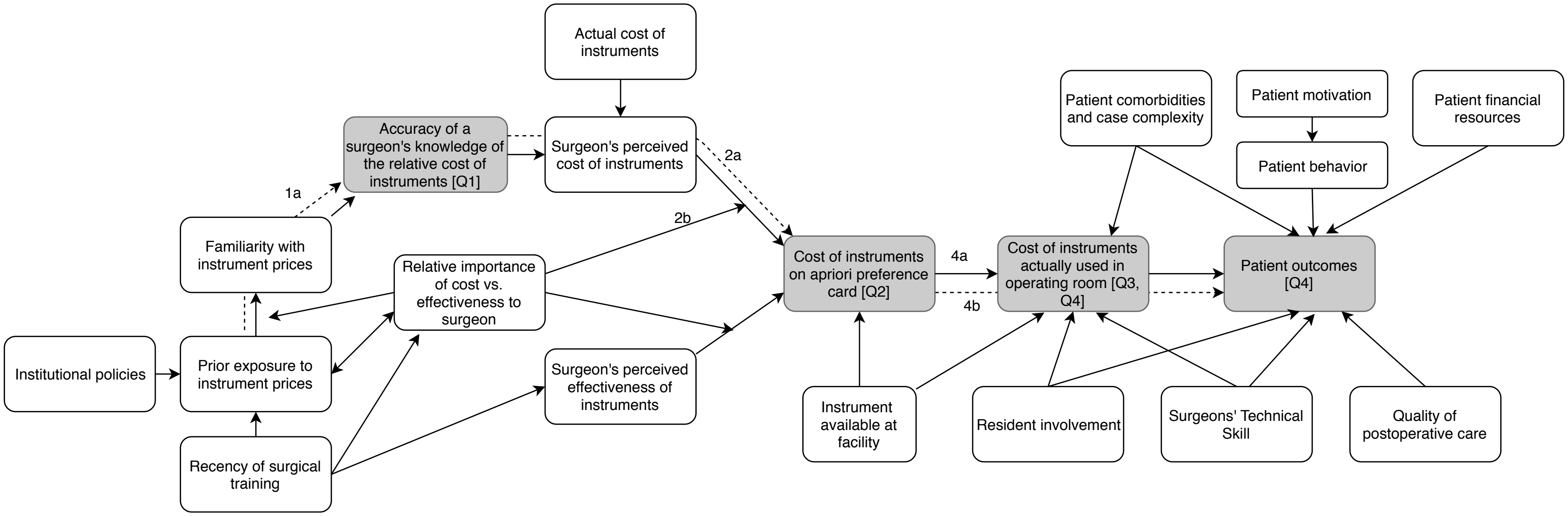
Data Source: CMS, 2016

Figure 2: Growth of healthcare expenditures versus United States GDP



Data Source: CMS, 2016

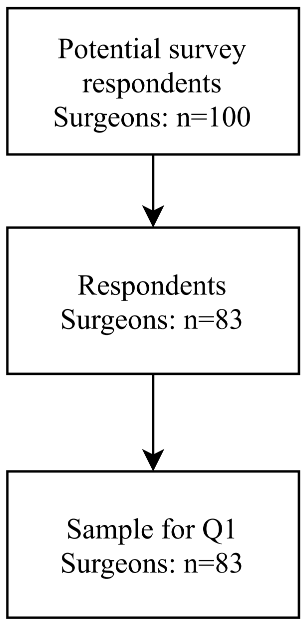
Figure 3: Conceptual model



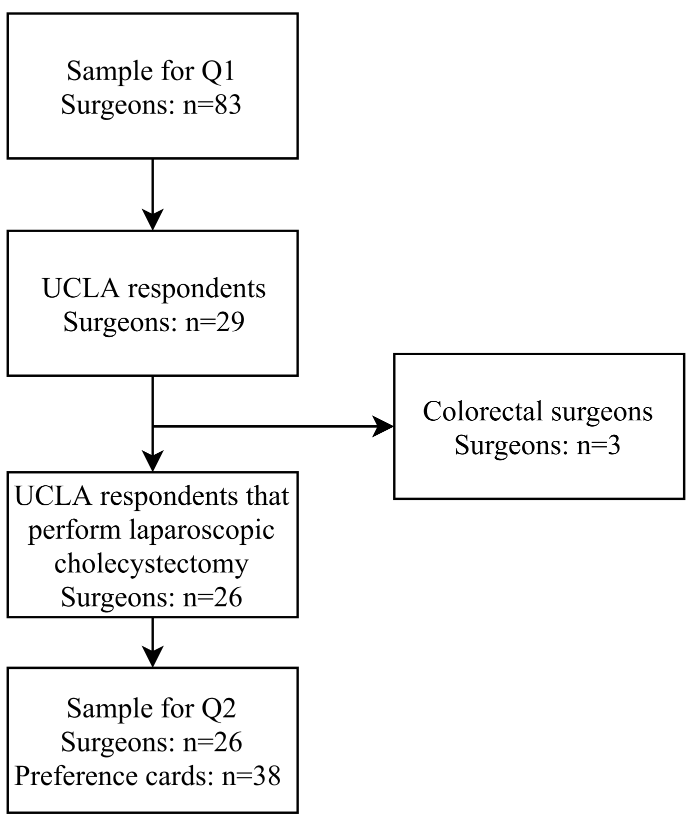
Note: Although most of the conceptual model is generalizable to other surgical procedures, for the purpose of this analysis (and for simplicity) we are focusing on only one operation – laparoscopic cholecystectomy. As such, the "operation" is not included in the model.

Figure 4: Sample size flow diagrams

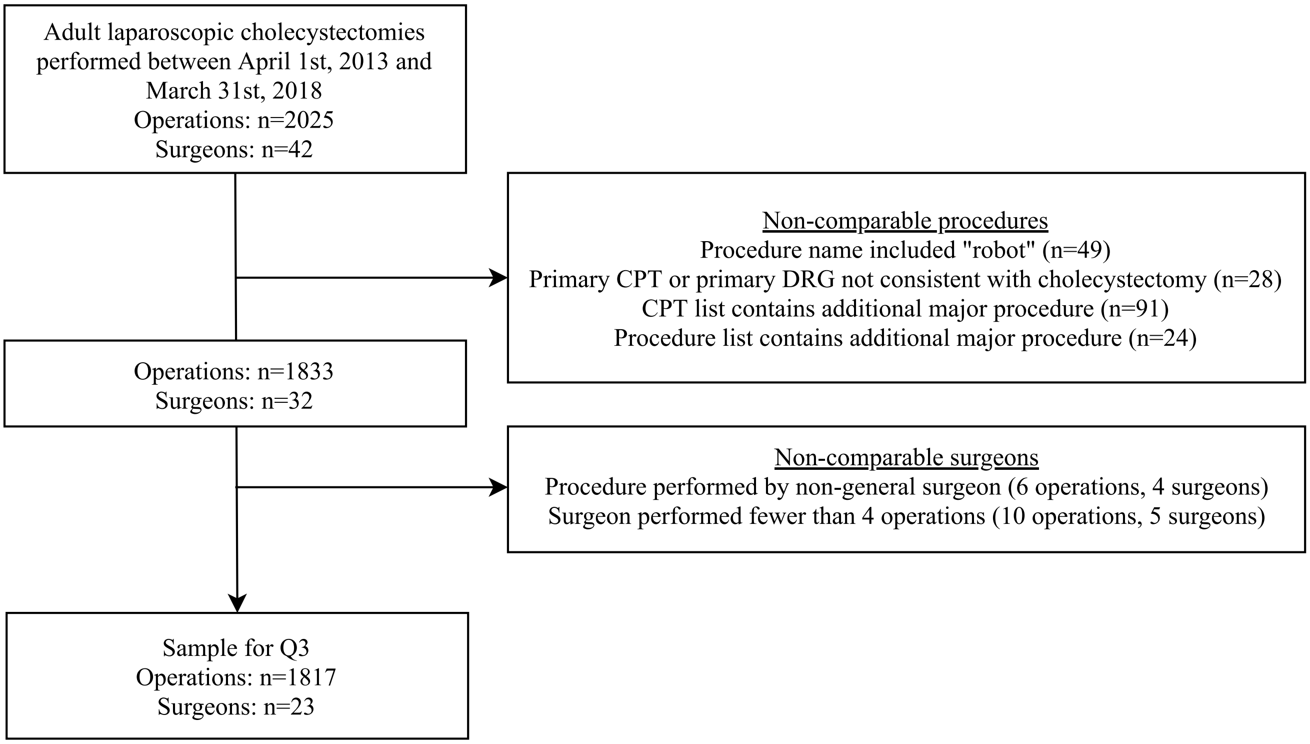
<u>Question 1</u>



<u>Question 2</u>



Question 3



Question 4

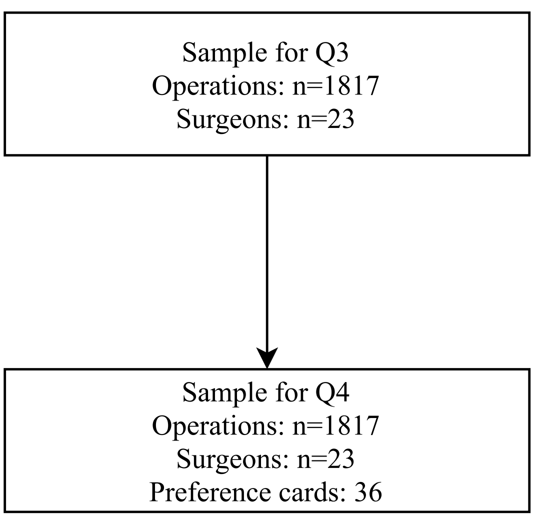
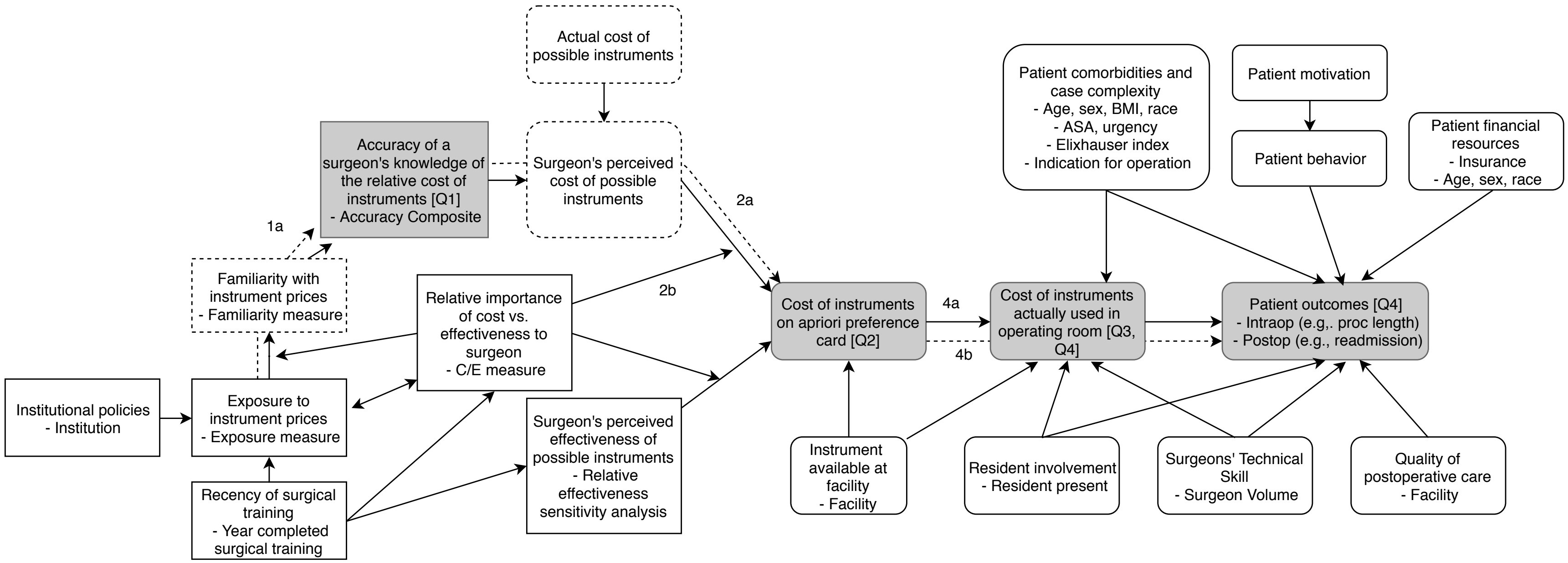
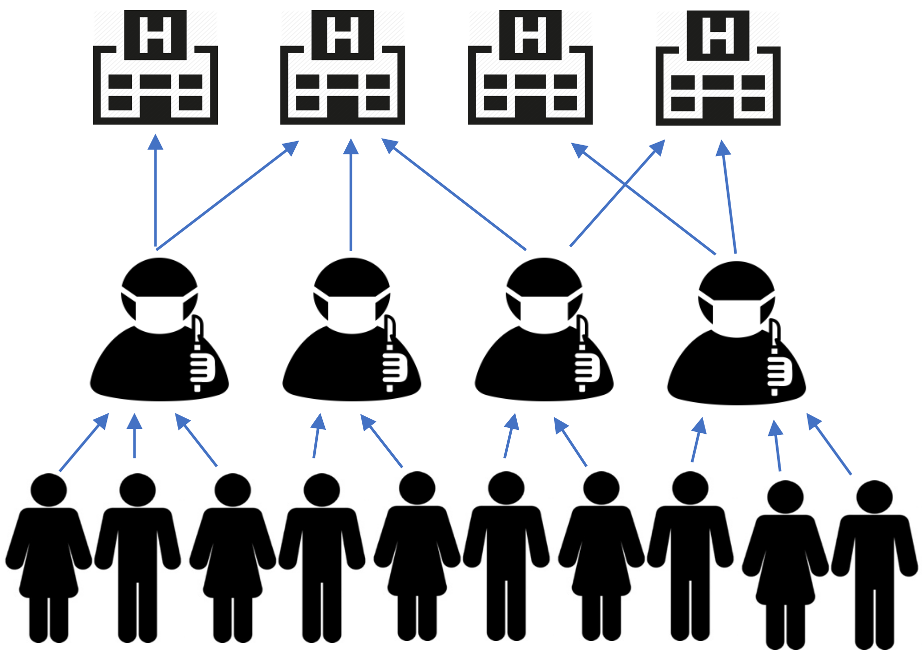


Figure 5: Measurement model



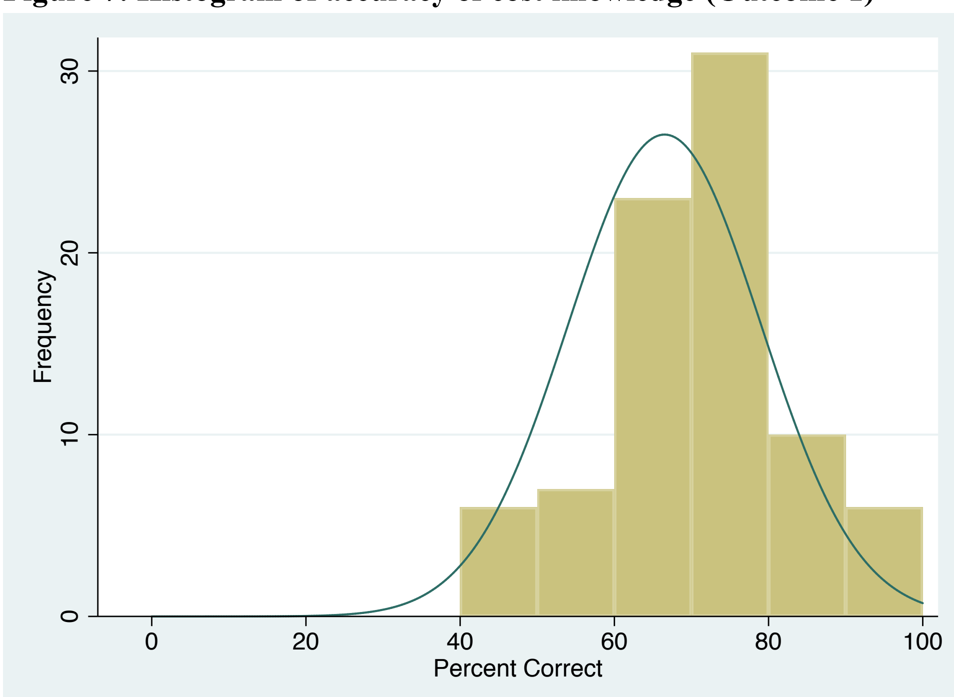
Notes: Square (unrounded) boxes indicate primary data collected via surgeon survey; rounded boxes represent administrative data obtained from the electronic health record; dashed lines and boxes are included when a reduced form model will be tested (e.g., hypothesis 2a will test the reduced form model by regressing cost of the a priori preference card on accuracy of surgeon's knowledge). The third outcome (cost of instruments used in the OR) serves as a dependent variable for hypothesis 3 and 4a but will also be omitted as part of a reduced form model for hypothesis 4b (indicated with the dashed line between cost of the a priori preference card and patient outcomes).

Figure 6: Crossed relationship between surgeons and facilities



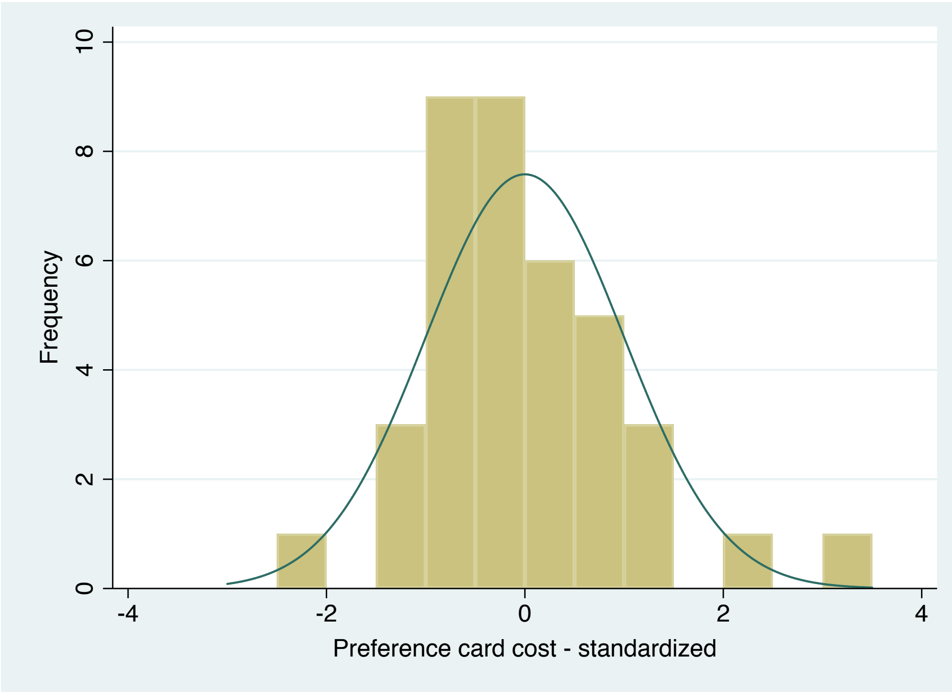
Note: While patients are nested one-to-one within surgeons, surgeons may operate at multiple facilities and therefore the relationship is not strictly hierarchical.

Figure 7: Histogram of accuracy of cost knowledge (Outcome 1)



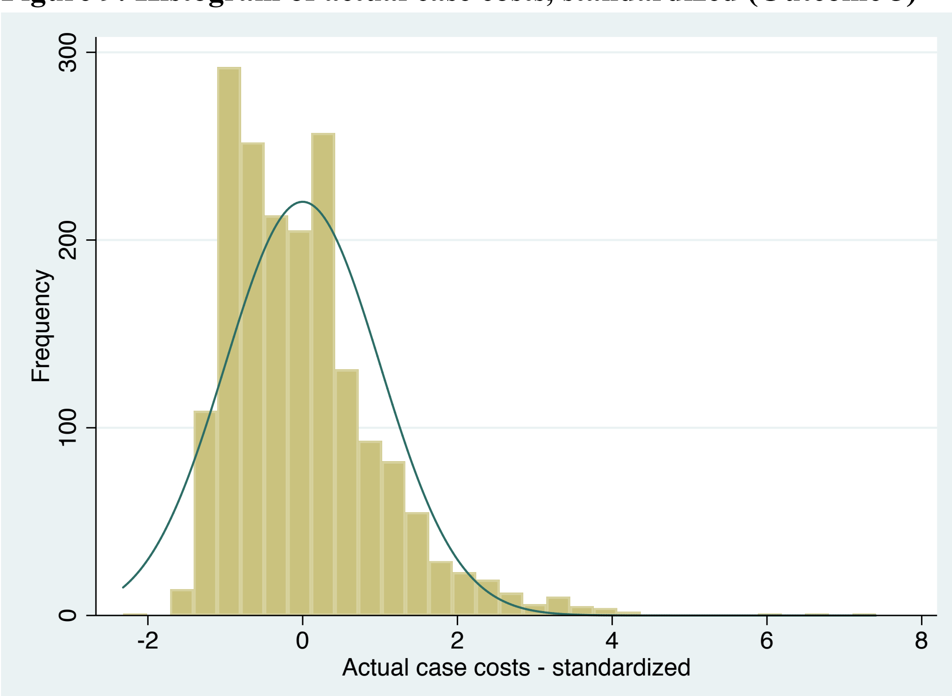
Note: N=83; single imputation was used for 4 surgeons with partial cost knowledge information

Figure 8: Histogram of preference card costs, standardized (Outcome 2)



Note: N=38

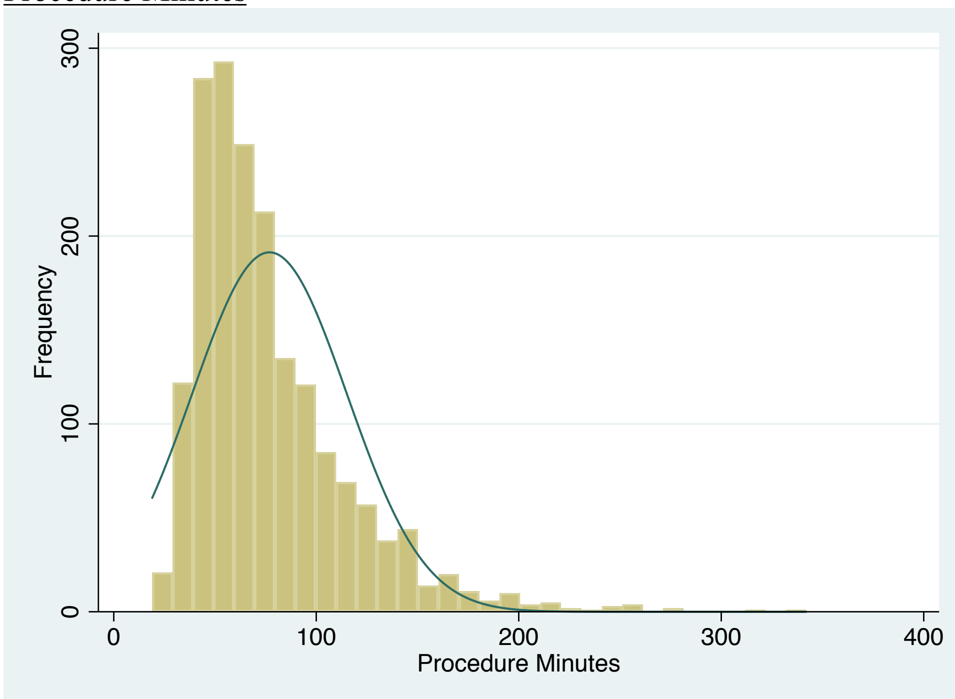
Figure 9: Histogram of actual case costs, standardized (Outcome 3)



Note: N=1817

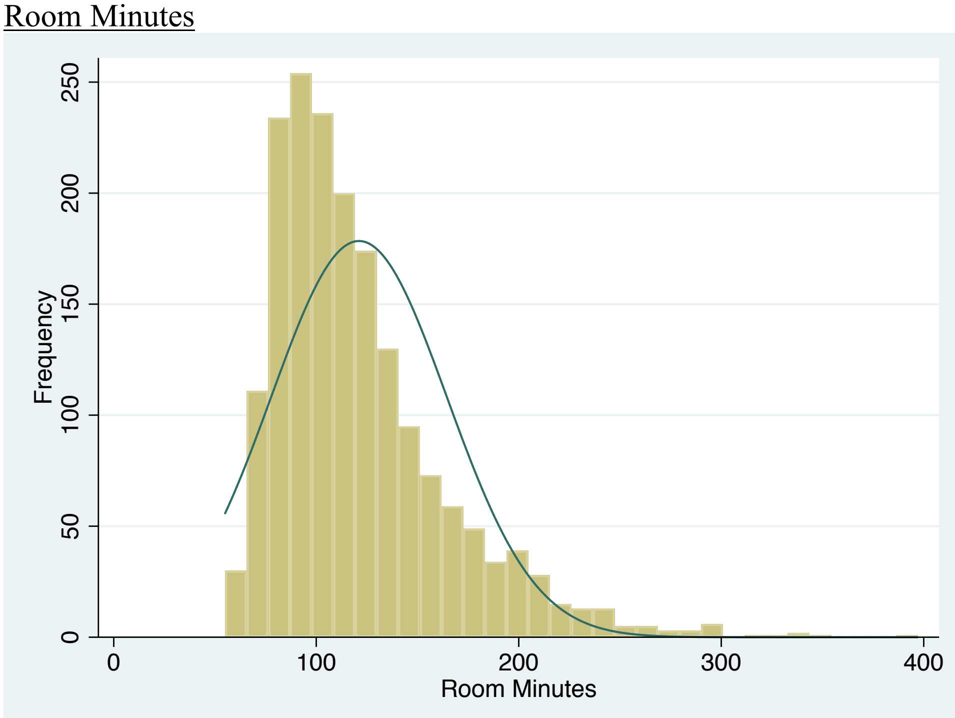
Figure 10: Histograms of continuous/count patient outcomes (Outcome 4)

Procedure Minutes



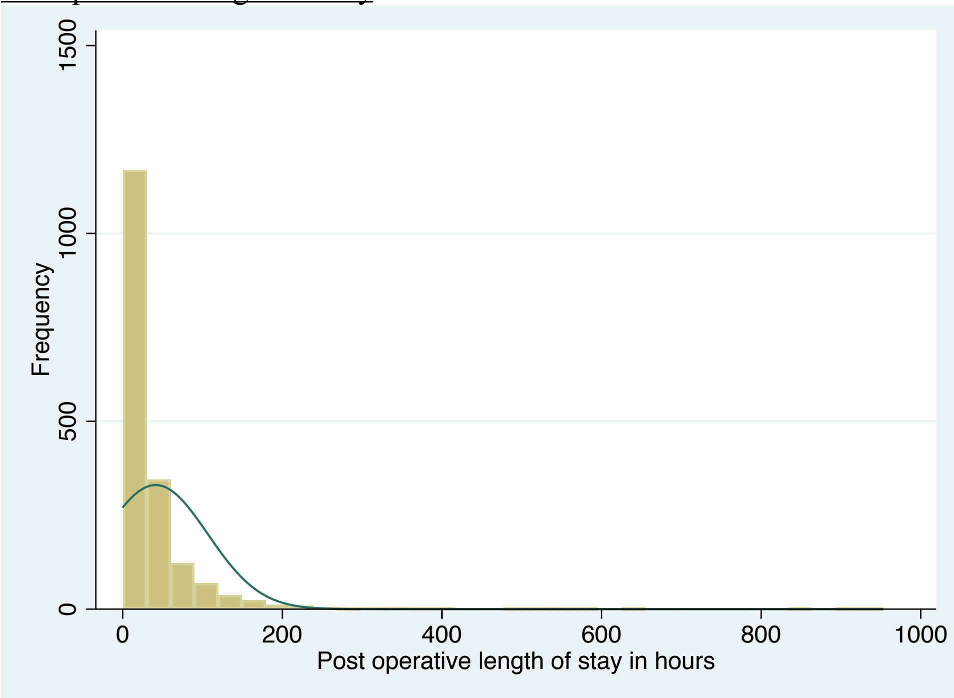
Note: N=1815

Room Minutes



Note: N=1815

Postoperative Length of Stay



Note: N=1817

Figure 11: Weighted scatterplot comparing familiarity and exposure to instrument prices

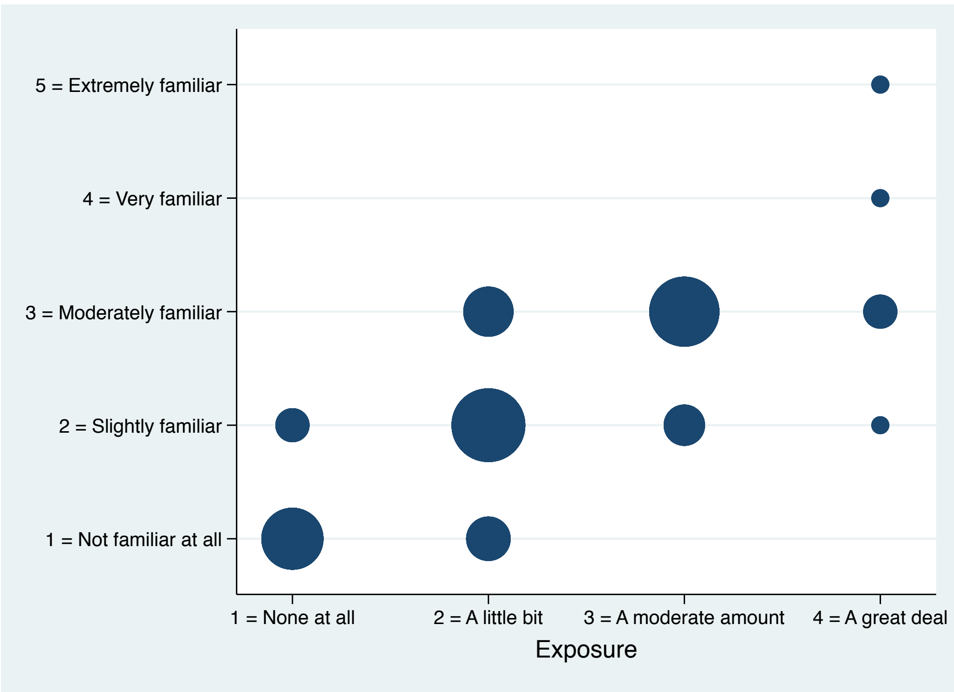


Table 1: Literature review table for question 1

Author, Year	Population, Setting	Primary Research Question & Hypotheses	Study Design	Primary Outcome & Operationalization	Covariates	Sample Size	Key Findings
Okike, 2014[46]	Orthopedic surgery; residents and faculty at 7 U.S. academic medical centers	(1) To what extent are orthopedic surgeons able to estimate the cost of commonly used orthopedic devices? (2) What factors are associated with knowledge about device costs?	Cross-sectional online survey	Estimate the cost of 13 orthopedic devices; prices were derived locally at each institution; some items were paired (i.e., items that are used for the same task but have different price points); accurate if estimate +/- 20%	Resident vs. Attending, years of experience, participation on value analysis committee, frequency of use of instrument, own knowledge of costs, the role cost should play in selection of devices	503 (96% RR), 217 attendings, 286 residents	19% overall accuracy Residents lower than attendings (17 vs. 21%) For attendings: No difference across sites or years in practice; prior experience on value analysis committee and frequent use of device positively associated with accuracy. For residents: No difference across site or prior experience with value analysis committee; however, more senior residents and more frequent use of items positively correlated with accuracy.
Parnes, 2015[47]	ENT, residents and faculty at 2 Canadian institutions	(1) To what extent are otolaryngologists able to estimate the cost of commonly used disposable instruments? (2) How interested are residents and attendings interested in having access to cost information and what is the perceived role of cost information in patient care?	Cross-sectional online survey	Estimate the cost of 23 common disposables; accurate if +/- 50%, or if cost <$1 if the estimate was <$2 Questions about interest in cost, sense of transparency of cost information in the hospital, and sense of how increased knowledge would affect practice	Resident vs. Attending, years of experience, subspecialty	43 respondents (Unclear RR), 26 attendings and 17 residents	30% overall accuracy; no difference between residents and attendings; no difference by site or years of experience; unable to assess relationship between confidence in estimates and accuracy due to small sample sizes. Most participants felt they had poor knowledge and poor confidence in estimates; most felt more knowledge would be good and may decrease their expenses.
Jackson, 2014[45]	ENT, residents and faculty at 1 U.S. academic medical center	(1) To what extent are otolaryngologists able to estimate the cost of OR supply and implant costs? (2) What factors affect accuracy of predictions and how does accuracy relate to the surgeon's self-assessment of cost knowledge?	Cross-sectional online survey	59 items identified via Pareto analysis; attendings asked to estimate cost of 20 items (10 general, 10 specialty specific) and residents asked to estimate cost of 25 items; accurate if +/- 20% Questions about knowledge of costs of supplies, previous attempts to obtain cost information, ease of obtaining information, ideal format of obtaining cost information; how often they review preference card; importance of costs, system emphasis on cost, leaders in cost, and who should lead discussion on cost	Resident vs. Attending, sub-specialty, frequency of item use, years in practice, confidence in cost knowledge	51 respondents (100% RR), 24 Attendings, 27 Residents/fellows	Faculty 25%, trainees 12%. Increased chance of accurate guess with specialty specific items (for faculty); no difference based on frequency of item use, years in practice, self-rated knowledge of cost.

Author, Year	Population, Setting	Primary Research Question & Hypotheses	Study Design	Primary Outcome & Operationalization	Covariates	Sample Size	Key Findings
Bade, 2015[44]	General surgery, residents and faculty at 3 public hospitals in New Zealand	(1) To what extent are general surgeons able to estimate the cost of common diagnostic tests, procedures, and hospital resources? How does position and setting affect accuracy?	Cross-sectional mailed survey	Estimate the cost of 14 common diagnostic tests, procedures, and hospital resources; costs included staffing, supplies, infrastructure and outsourced costs; accurate if +/- 25%	Consultants vs. House Officers, Public vs. Private	57 respondents (100% RR) --> excluded 4 for missing demographics; 25 house officers and 28 consultants/registrars	Overall 19% accurate; no difference between positions or public vs. private practices.
		(2) How important is cost information to surgical practice?		Questions about own accuracy and perception of importance of cost to surgical practice			Most thought cost information was important but rated own knowledge as poor.
Allan, 2008[48]	Studies of surveys assessing physician or trainee knowledge of costs for diagnostics/therapeutics including medical supplies	(1) To what extent are physicians and trainees aware of the cost of diagnostic/therapeutics?	Meta-analysis	Average cost accuracy, average percent of estimates over and under true cost, average percent error	Year of publication, location of study, training level of participants, specialty, and cost of item (expensive vs. inexpensive); also: study quality (response rate, sampling method, distribution)	14 studies; 5 in U.S., 4 U.K, 3 Canada, 2 other; samples ranged from 20 to 506 with RR from 48% to 100%	33% overall accuracy (when using +/- 20-25%), 50% (when using +/- 50-100%).
		(2) What factors influence accuracy of predictions?					No effect of country, year of publication, specialty (GPs vs. specialists) (lack of relationship holds regardless of criteria used); percent accuracy slightly worse for higher quality studies vs. moderate/low (27 vs. 33%).

ENT = Ear, Nose & Throat (Otolaryngology); GP = General practitioner; RR = Response rate; U.K = United Kingdom; U.S. = United States

Table 2: Literature review table for question 3

Author, Year	Population, Setting	Primary Research Question & Hypotheses	Study Design	Outcome(s) & Operationalization	Covariates	Sample Size	Analysis	Key Findings
Adkins, 2016[57]	Elective laparoscopic cholecystectomy at one U.S. medical center	(1) What is the variation in equipment and OR costs for laparoscopic cholecystectomy between surgeons?	Retrospective medical chart review	Variable supply cost, total cost, morbidity, mortality	Age, BMI, Sex, Race	272 patients, 3 surgeons	Descriptive, univariate	Variable supply cost ranged from $412 to $924; there was "no difference" in patient demographics between the 3 surgeons, no difference in morbidity and no mortalities.
		(2) How does resident involvement impact OR cost?			Resident involvement		Descriptive, univariate	Cases with residents were longer, on average, and cost more overall but no analysis of supply use.
Benoit, 2001[58]	Radical prostatectomy at one U.S. medical center	What factors influence the hospital costs of radical prostatectomy?	Retrospective medical chart review	Total hospital costs and costs broken into 11 cost centers	Age, PSA, Gleason score, clinical stage, ASA, smoking history, DM, intraoperative variables (OR time, surgical time, EBL, transfusions)	104 patients	Bivariate analysis between each covariate and total hospital cost	Patient factors (age, ASA, DM, smoking) were not correlated with total hospital costs.
Brauer, 2015[59]	Laparoscopic cholecystectomy (+/- IOC) at 7 U.S. medical centers	What are the intraoperative and surgeon-specific sources of cost variation in laparoscopic cholecystectomy?	Retrospective review (administrative data)	Intraoperative supply cost ("estimated" cost)	Case volume	2178 patients, 55 surgeons	Descriptive, graphical	Median cost ~ $500 with variation (not statistically defined) at the hospital and surgeon level; no graphical relationship between case volume and median cost; largest costs related to trocars, packs and clip appliers.
Chung, 2010[60]	CABG at 2 Taiwanese medical centers	What is the relationship between hospital, surgeon, and patient characteristics and cost?	Retrospective medical chart review	Multiple; primary outcome was total OR cost; one subcomponent was direct instrument costs	Hospital volume, surgeon, OR time, LOS, patient age and gender	238 patients, 3 surgeons, 2 facilities	Bivariate, Multivariate	Low volume hospitals higher surgical instrument costs than high volume hospital (bivariate); multivariate analysis only on total OR cost and showed inverse relationship between LOS and hospital volume and OR cost and direct relationship between OR time and OR cost; no relationship between gender and cost; positive relationship between patient age and cost.

Author, Year	Population, Setting	Primary Research Question & Hypotheses	Study Design	Outcome(s) & Operationalization	Covariates	Sample Size	Analysis	Key Findings
Kazberouk, 2016[61]	Spine surgery at 1 U.S. medical center	(1) What is the intersurgeon variation in total procedure cost after adjusting for patient case-mix?	Retrospective medical chart review	Mean adjusted cost per procedure (direct costs for items charged to patients)	Physician, procedure, patient age, CHF, ethnicity, CCI	1241 patients, 7 surgeons	Bivariate, multivariate	In unadjusted analysis: cost per procedure varied by factor of 1.32 to 1.81 depending on operation; in adjusted analysis: CCI increased by 27% per point, ethnicity other than white increased by 11%, CHF increased by 82%; age did not influence; after removing these patient characteristics, variation by factor of 1.31 still existed between surgeons.
		(2) What cost categories drive this variation?		Mean adjusted cost per procedure per surgeon	Cost categories (supply, OR cost, anesthesia cost, other, inpatient stay)			Surgeons with lower supply costs didn't necessarily have lower total costs; for example, 4 surgeons had below average supply costs but only 2 of the 4 had below average total costs.
Meier, 2014[62]	Pediatric adenotonscillectomy at 1 U.S. medical center	How much variation exists between surgeons and which buckets drive this variation?	Retrospective medical chart review	Total encounter cost including OR, supplies (only those >$25), PACU, same-day services, anesthesia, Rx, Other; TDABC accounting	Surgeon	4824 patients, 14 surgeons	Bivariate	Significant variation in total cost and individual buckets between surgeons; main drivers of total variation were supply cost and "OR cost".
Rosenbaum, 2017[63]	Lumbar spine surgery at 1 U.S. medical center	How much variation exists between surgeons and what are the primary drivers of variation?	Retrospective medical chart review	Supply and implant costs, artificially manipulated to relative cost unit	Surgeon, inpatient/outpatient, pt age, pt sex, operation duration, mean difference in HSM (QOL metric)	652 patients, 9 surgeons	Bivariate, Multivariate	Significant variation between surgeons; on multivariate regression only surgeon and location were associated with surgical costs.
Sjogren, 2016[64]	Pediatric tympanostomy tube placement; 15 hospitals within 1 U.S. healthcare system	What are the major expenses and where does variation lie between surgeons and hospitals?	Retrospective medical chart review	Total cost per case; also looked at subcategories - of interest, OR and OR supplies; TDABC accounting	Surgeon, facility	5,623 patients, 67 surgeons	Bivariate	For total costs: hospitals ranged from $509-1212; surgeons ranged from $660-1330 (no multivariable analysis to ascertain which is the dominant driver); For supplies range $18-44 by surgeon and for OR costs range from $200-$349.

Author, Year	Population, Setting	Primary Research Question & Hypotheses	Study Design	Outcome(s) & Operationalization	Covariates	Sample Size	Analysis	Key Findings
Thomas, 2016[65]	ENT; 21 facilities within 1 U.S. healthcare system	How much variation exists between surgeons and what are the primary drivers of variation?	Retrospective medical chart review	Total cost per case; subcategories included OR time, OR supplies, same-day services, Rx, PACU, anesthesia supplies, other, lab	Surgeon, facility, surgeon specialty; patient characteristics included age and gender	4007 patients, 72 surgeons	Multivariate	For total costs: hospitals ranged from $2,073 to $3,991 and surgeons varied from $1735 to $6,940. For supplies: Supplies accounted for 25% of variable costs, significant variation between surgeons for OR supply costs (raw data not shown, just p-value <0.01); also, ENT had higher supply costs than OMFS or plastic surgeons
Zygourakis, 2016[66]	All inpatient NSG procedures at 1 U.S. teaching hospital	What is the effect of patient, procedural, and provider-specific factors on NSG supply cost?	Retrospective medical chart review	Total direct cost of surgical supplies (including disposable, implantable and nonimplantable); not instrument sets	Fixed: age, gender, payer, CMI, BMI, admission source, procedure length, procedure date, surgeon case volume; Random: Surgeon	4,904 patients, 24 surgeons	Mixed effects	Females, older patients, and more complex patients (CMI) had higher supply costs; patients with Medicare had higher supply costs compared to commercial; longer procedures had higher supply costs; surgeons with higher volumes had higher supply costs

ASA = American Society of Anesthesiologists physical status scale; BMI = Body Mass Index; CABG = Coronary Artery Bypass Graft; CCI = Charlson Comorbidity Index; CHF = Congestive Heart Failure; CMI = Case Mix Index; DM = Diabetes Mellitus; IOC = Intraoperative Cholangiogram; LOS = Length of Stay; NSG = Neurosurgery; OR = Operating Room; PACU = Post Anesthesia Care Unit; PSA = Prostate Specific Antigen; QOL = Quality of Life; Rx = Pharmacy; TDABC = Time-Driven Activity-Based Costing

Table 3: Question, unit of analysis, restrictions, and sample size for each question

<table>
<tr><th>Question</th><th>Unit of Analysis</th><th>Patient/Surgeon/Site Restriction</th><th>Sample for primary analysis (N)</th></tr>
<tr><td>1</td><td>Surgeon</td><td>UCLA: Current surgeons in the divisions of general, colorectal, surgical oncology, minimally invasive, and pediatric surgery + surgeons with relevant operations during the study time period (who have subsequently left).
CSMC: Current surgeons in the divisions of general, colorectal, minimally invasive, hepatobiliary, and pediatric surgery.
UCSD: Current surgeons in the divisions of general, colorectal, surgical oncology, minimally invasive, and transplant surgery.</td><td>83 surgeons</td></tr>
<tr><td>2</td><td>Surgeon/Site</td><td>Surgeons at UCLA with laparoscopic cholecystectomy preference cards</td><td>26 surgeons, 38 preference cards</td></tr>
<tr><td>3</td><td>Patient</td><td rowspan="2">Surgeons at UCLA with >=4 relevant laparoscopic cholecystectomies during study time period</td><td>23 surgeons, 1817 operations</td></tr>
<tr><td>4</td><td>Patient</td><td>23 surgeons, 36 preference cards, 1817 operations</td></tr>
</table>

Notes: flow diagrams are included in Figure 4; the sample sizes described in this Table refer to only the primary analysis. A more detailed Table and flow diagrams for the various sensitivity analyses are included in Appendix 4.

Table 4: Distribution of survey respondents by institution

Institution	Potential Respondents, N	Started survey	Completed >= 50% of eligible questions	Response Rate (%)
UCLA	31	30	29	93.5
UCSD	25	25	24	96.0
CSMC	44	33	30	68.2
Totals	100	88	83	83.0

Table 5: Case inclusion/exclusion criteria

	Inclusion	Exclusion
Population	Adults (18 years or older) at time of operation	
Facility/Time	Operation performed April 1st, 2013 – March 31st, 2018 Any UCLA Site (RR-OR, RR-ASC, SM-OR, SM-ASC)	
Operation	Booking slip with laparoscopic cholecystectomy	Booking slip contains: "Robot" Additional major operation (e.g., colorectal resection, herniorrhaphy)
	Post procedure primary CPT consistent with cholecystectomy or primary DRG consistent with cholecystectomy (Table 6)	CPT list contains additional major operation
Surgeon	Surgeon performed at least 4 relevant operations during study time period	Non-general surgeon (e.g., liver transplant)

ASC = Ambulatory Surgery Center; CPT = Current Procedural Terminology; DRG = Diagnosis Related Group

Table 6: Included cases based on post procedure assigned primary CPT code or primary DRG

CPT Code	Description	Primary DRG	Description
47562	Laparoscopic cholecystectomy	414	Cholecystectomy Except By Laparoscope W/O C.D.E. W Mcc
47563	Laparoscopic cholecystectomy with cholangiography	415	Cholecystectomy Except By Laparoscope W/O C.D.E. W Cc
47564	Laparoscopic cholecystectomy with exploration of common duct	416	Cholecystectomy Except By Laparoscope W/O C.D.E. W/O Cc/Mcc
47600	Cholecystectomy	417	Laparoscopic Cholecystectomy W/O C.D.E. W Mcc
47605	Cholecystectomy with cholangiography	418	Laparoscopic Cholecystectomy W/O C.D.E. W Cc
47610	Cholecystectomy with exploration of common duct	419	Laparoscopic Cholecystectomy W/O C.D.E. W/O Cc/Mcc

C.D.E = Common duct exploration; Cc = Comorbidity or complication; CPT = Current Procedural Terminology; DRG = Diagnosis Related Group; Mcc = Major comorbidity or complication.

Table 7: Example ICD-10 codes used to identify case indication

ICD-10 Code	ICD-10 Description	Category
K80.20	Calculus of gallbladder	(1) biliary colic / chronic cholecystitis / other elective
K80.44	Calculus of gallbladder with acute cholecystitis	(2) calculus of bile duct with chronic cholecystitis without obstruction
K81.0	Acute cholecystitis	(3) acute cholecystitis

Table 8: Missing data by question/hypothesis

Variable	Unit	Complete Data, N	Missing Data, N	Mechanism of Dealing with Missingness
Question 1				
Accuracy of cost knowledge	Surgeon	79	4[1]	Single (regression) imputation
Prior exposure to instrument prices	Surgeon	82	1	Multiple imputation not possible due to perfect prediction – complete case analysis performed when exposure is included in any model
Relative importance of cost vs. effectiveness	Surgeon	83	0	NA
Question 2				
Preference card cost	Pref Card	38	0	NA
Accuracy of cost knowledge	Surgeon	23	3	Single (regression) imputation using all survey data (all 79 respondents with complete data)
Facility	Pref Card	38	0	NA
Relative importance of cost vs. effectiveness	Surgeon	26	0	NA
Questions 3 & 4				
Cost of instruments used in operation	Patient	1817	0	NA
Preference card cost	Pref Card	36	0	NA
Patient Outcomes	Patient			
Procedure time		1815	2	Complete case analysis
Room time		1815	2	
Length of stay		1817	0	
Escalation of care		1817	0	
Return to ED (30 / 90 days)		1817	0	
Surgeon	Surgeon	23	0	NA
Facility[2]	Facility	4	0	NA
Patient-level variables	Patient			
Age		1817	0	
Race/Ethnicity		1803	14	Race/ethnicity, BMI, ASA, and urgency were multiply imputed using multinomial (race), linear (BMI), and ordered logit (ASA/urgency) imputation with 20 imputation sets with the outcome (e.g., cost of instruments), age, sex, Elixhauser score, ASA emergent status, indication, and resident as predictors
BMI		1794	23	
ASA		1815	2	
ASA Emergent		1817	0	
Urgency		1782	35	
Elixhauser Score		1817	0	
Indication[3]		1817	0	
Resident		1817	0	
Primary Payer		1817	0	

Notes: Missing data for the sensitivity analyses are included in appendix 4.
[1] Three respondents answered 9/10 and one respondent answered 8/10 questions
[2] Nine operations were missing facility in the administrative database; these were corrected via manual chart review
[3] For 11 operations, the identified ICD-10 codes did not provide an indication; these were corrected via manual chart review

Table 9: Descriptive statistics for survey variables

Variable		Frequency (%) unless otherwise indicated
N = 83 unless otherwise indicated		
Outcomes		
Accuracy of cost knowledge % (Mean, SD)[1]		66.5 +/- 12.5
Choice of Instruments		
Instrument Choice 1[2]	A, 10mm Endoclip	29 (35.4)
	B, 5mm Endoclip	53 (64.6)
Instrument Choice 2[2]	A, Clearify	35 (42.7)
	B, Anti-fog	47 (57.3)
Covariates		
Institution	UCLA	29 (34.9)
	UCSD	24 (28.9)
	CSMC	30 (36.1)
Year finished training, categorical	<=1990	8 (9.6)
	1991-2000	12 (14.5)
	2001-2010	30 (36.1)
	2011-2017	33 (39.8)
Exposure to instrument prices [2]	None at all	17 (20.7)
	A little bit	35 (42.7)
	A moderate amount	23 (28.0)
	A great deal	7 (8.5)
Familiarity with instrument prices	Not at all	20 (24.1)
	Slightly	30 (36.1)
	Moderately	30 (36.1)
	Very	1 (1.2)
	Extremely	2 (2.4)
Relative importance of cost vs. effectiveness		
Measure 1 - Direct measure	<=4	6 (7.2)
	5	29 (35.0)
	6	10 (12.0)
	7	26 (31.3)
	8-9	12 (14.5)
Measure 2 - Value of OR time (Median, IQR)[3]		$21.43 [7.14-35.71]
Measure 3 - Value of surgical site infection (Median, IQR)[4]		$5000 [500-15,000]
Perceived effectiveness		
Instrument Choice 1[2]	A > B	20 (24.4)
	A = B	53 (64.6)
	A < B	9 (11.0)
Instrument Choice 2[2]	A > B	48 (58.5)
	A = B	33 (40.2)
	A < B	1 (1.2)

Freq = Frequency; IQR = Interquartile Region; SD = Standard Deviation

[1] Singly imputed income for the 4 surgeons with partial data

[2] Complete case, n=82

[3] Converted from native answer by subtracting $250 and dividing by 7; complete case, n=75

[4] Converted from native answer by subtracting $250 and multiplying by 100; complete case, n=78

Table 10: Cost knowledge scale psychometrics

Question Number	Observations (Number)	Item-Rest Correlation	Alpha
1	83	0.0346	0.096
2	83	0.1023	0.0545
3	81	-0.0678	0.172
4	82	-0.0272	0.1404
5	83	0.1007	0.0618
6	83	0.0875	0.095
7	83	0.0057	0.1279
8	83	0.0781	0.063
9	82	-0.0287	0.1602
10	82	0.222	0.06
Overall Scale			0.1142

Table 11: Descriptive statistics for medical record data

Variable		Frequency (%), Mean +/-, or Median [IQR]
N = 1817 unless otherwise indicated		
Outcomes		
Cost of instruments used in the operating room		See Histogram[1]
Procedure Time (Min, n=1815)		76.9 +/- 38.2
Room Time (Min, n=1815)		121.1 +/- 43.4
Length of Stay (Hours)		23.5 [16.3-45.0]
Post procedure escalation of care	Outpatient --> Inpatient (n=824)	154 (18.7)
	Floor --> ICU (n=993)	16 (1.6)
Return to ED, 30 days		186 (10.2)
Return to ED, 90 days		268 (14.7)
Covariates		
Surgeon Procedure Volume (n=23)		60 [14-99]
Age		51.5 +/- 18.0
Body Mass Index		27.5 [24.2-32]
Elixhauser Comorbidity Score		1 [0-2]
Facility	RR-OR	572 (31.5)
	RR-ASC	193 (10.6)
	SM-OR	827 (45.5)
	SM-ASC	225 (12.4)
Resident Present		1731 (95.3)
Primary Expected Payer	Private	1213 (66.8)
	Medicare	361 (19.9)
	Medical/CHIP/Other Public	204 (11.2)
	Other/Missing	39 (2.1)
Female		1162 (64.0)
Race/Ethnicity (n=1803)	Non-Hispanic White	822 (45.6)
	Non-Hispanic Black	117 (6.5)
	Non-Hispanic Other	342 (19.0)
	Hispanic	522 (29.0)
ASA score[2]	I	235 (12.9)
	II	1013 (55.8)
	III	537 (29.6)
	IV	30 (1.7)
	ASA Emergent	120 (6.6)
Urgency (n=1782)	Elective	956 (53.6)
	Inpatient	201 (11.3)
	Urgent	447 (25.1)
	Emergent/Critically Emergent	178 (10.0)
Indication for Operation	Biliary colic / Chronic Cholecystitis / Other elective	940 (51.7)
	Obstructive Biliary	193 (10.6)
	Acute Cholecystitis	684 (37.6)

ASC = Ambulatory Surgery Center; CHIP = Children's health insurance program; OR = Main Operating Room; RR = Ronald Reagan; SM = Santa Monica

All values based on complete case analysis, sample sizes are indicated throughout

[1] To protect hospital financial information, cost data were standardized and presented in histograms

[2] ASA score was missing for n=2; ASA emergent status was not missing for any patient

Table 12: Effect of exposure to instrument prices on cost knowledge, including assessment of the moderating effect of the relative importance of cost versus effectiveness

	Change in accuracy of cost knowledge[1]	Standard Error	Test statistic	P-Value
		N=82		
Exposure to instrument prices				
None		Reference		
A little bit	7.50	6.03	1.24	0.22
A moderate amount	7.27	6.13	1.19	0.24
A great deal	5.53	7.43	0.74	0.46
Relative importance of cost versus effectiveness				
Cost equivalent to or more important than effectiveness		Reference		
Cost less important than effectiveness	6.00	6.25	0.96	0.34
Interactions				
A little bit * Cost less important than effectiveness	-6.38	7.71	-0.83	0.41
A moderate amount * Cost less important than effectiveness	-6.66	8.19	-0.81	0.42
A great deal * Cost less important than effectiveness	-6.53	12.32	-0.53	0.60
Constant	60.00	4.80		

[1] Cost knowledge was on a percentage scale, ranging from 0 to 100%

Notes: The likelihood ratio test comparing regression with and without the relative importance of cost versus effectiveness (as main effect and moderator) had a $\chi2$ statistic of 1.0 and p-value of 0.90. The results remained substantively the same using a continuous and categorical operationalization of the relative importance of cost versus effectiveness.

Table 13: Sensitivity analysis - moderating effect of the value of a surgical site infection on the impact of exposure to instrument prices on cost knowledge

	Change in accuracy of cost knowledge[1]	Standard Error	Test statistic	P-Value
		N=82		
Exposure to instrument prices				
None		Reference		
A little bit	-2.23	5.31	-0.42	0.68
A moderate amount	-2.10	5.42	-0.39	0.70
A great deal	-1.98	7.22	-0.27	0.78
USD value assigned for one SSI	-0.0004	0.0003	-1.32	0.20
Interaction Terms				
SSI * A little bit	0.0006	0.0004	1.54	0.13
SSI * A moderate amount	0.0006	0.0004	1.49	0.14
SSI * A great deal	0.0003	0.0007	0.46	0.64
Constant	67.79	4.72		

[1] Cost knowledge was on a percentage scale, ranging from 0 to 100%

SSI = Surgical site infection; USD = US dollars

Notes: Complete data for SSI values were available for 78 surgeons; 20 imputations were used to fill values for the remaining surgeons using predictive mean matching, k=10, with exposure, year finished, institution and cost knowledge as covariates.

Table 14: Sensitivity analysis - moderating effect of the value of one minute of OR time on the impact of exposure to instrument prices on cost knowledge

	Change in accuracy of cost knowledge[1]	Standard Error	Test statistic	P-Value
		N=82		
Exposure to instrument prices				
None		Reference		
A little bit	-4.13	5.77	-0.72	0.48
A moderate amount	-10.07	6.69	-1.50	0.14
A great deal	-6.96	7.37	-0.94	0.35
USD value assigned for one minute of OR time	-0.25	0.13	-1.87	0.07
Interaction Terms				
OR Time * A little bit	0.25	0.15	1.69	0.10
OR Time * A moderate amount	0.46	0.19	2.42	0.019*
OR Time * A great deal	0.34	0.35	0.97	0.34
Constant	70.45	5.41		

[1] Cost knowledge was on a percentage scale, ranging from 0 to 100%

USD = US dollars, * = p<0.05

Notes: Complete data for OR cost values were available for 75 surgeons; 20 imputations were used to fill values for the remaining surgeons using predictive mean matching, k=10, with exposure, year finished, institution and cost knowledge as covariates.

Table 15: Sensitivity analysis - ordered logistic regression assessing the effect of recency of surgical training on exposure to instrument prices

	Odds ratio	**Standard Error**	**Test statistic**	**P-Value**
		N=82		
Year, Categorical				
<1990		Reference		
1991-2000	8.01	7.88	2.11	0.04*
2001-2010	3.01	2.54	1.31	0.19
2011-2017	0.81	0.68	-0.25	0.80
Institution				
CSMC		Reference		
UCLA	2.07	1.03	1.45	0.15
UCSD	5.67	3.19	3.09	<0.001*
	Coefficient	Standard Error		
Cut 1	-0.32	0.79		
Cut 2	1.95	0.82		
Cut 3	4.10	0.92		

* $p<0.05$

Notes: Likelihood ratio test of proportionality of odds across response categories had a $\chi 2$ statistic of 8.21 and p-value of 0.08; the likelihood ratio test comparing the nested model (categorical year only) versus the model with year and institution had a $\chi 2$ statistic of 10.1 and a p-value of 0.006. Pairwise comparisons between year categories suggested significantly lower exposure in 2010-2017 compared to both 1991-2000 and 2001-2010, after Sidak's adjustment.

Table 16: Sensitivity analysis - reduced form model of effect of years since finished training on cost knowledge

	Change in accuracy of cost knowledge[1]	Standard Error	Test statistic	P-Value
		N=82		
Year, Categorical				
<1990		Reference		
1991-2000	12.01	5.71	2.10	0.04*
2001-2010	9.29	4.95	1.88	0.07
2011-2017	6.03	4.95	1.22	0.23
Institution				
CSMC		Reference		
UCLA	1.56	3.19	0.49	0.63
UCSD	1.24	3.47	0.36	0.72
Constant	57.89	4.61		

[1] Cost knowledge was on a percentage scale, ranging from 0 to 100%

* = p<0.05

Notes: The omnibus test of the null hypothesis that all coefficients for categorical year = 0 had an F statistic of 1.9 and p-value of 0.14; the likelihood ratio test comparing the naive model (institution only) versus institution + year had a $\chi 2$ statistic of 5.8 and p-value of 0.12.

Table 17: Effect of cost knowledge on preference card costs

	Change in preference card cost (USD)	Standard Error	Test statistic	P-Value
		N=38		
Cost knowledge[1]	-2.31	1.25	-1.84	0.08
Facility				
RR-ASC		Reference		
RR-OR	65.13	45.25	1.44	0.16
SM-OR	304.71	53.32	5.71	<0.001*
SM-ASC	113.17	40.20	2.82	0.009*
Constant	506.90	93.31		

[1] Cost knowledge was on a percentage scale, ranging from 0 to 100%

* = p<0.05

ASC = Ambulatory Surgery Center, OR = Main Operating Room; SM = Santa Monica; RR = Ronald Reagan; USD = US dollars

Notes: Model includes analytic weights based on actual utilization of each preference card by surgeon and standard errors adjusted for surgeon clustering. Pairwise comparisons of facility, including Sidak's adjustment, indicated that the SM-OR had higher mean preference card costs than both RR-ASC and RR-OR.

Table 18: Effect of cost knowledge on preference card cost, including assessment of the moderating effect of relative importance of cost versus effectiveness

	Change in preference card cost (USD)	Standard Error	Test statistic	P-Value
		N=38		
Cost knowledge[1]	-1.84	2.00	-0.92	0.37
Relative importance of cost versus effectiveness				
Cost equivalent to or more important than effectiveness		Reference		
Cost less important than effectiveness	37.62	192.84	0.20	0.85
Interaction				
Cost less important than effectiveness * Correct percent	-0.58	2.72	-0.21	0.83
Constant	480.17	139.38		

USD = US dollars

[1] Cost knowledge was on a percentage scale, ranging from 0 to 100%

Note: Model also includes institutional fixed effects, analytic weights based on actual utilization of each preference card by surgeon, and standard errors adjusted for surgeon clustering. The likelihood ratio test comparing the nested model with institution and correct percent only compared to institution, correct percent, and a main and moderating effect of relative importance of cost had a $\chi 2$ statistic of 0.02 and p-value of 0.99.

Table 19: Sensitivity analysis – effect of cost knowledge for an individual instrument comparison (5mm Endoclip versus 10mm Endoclip) on instrument choice, including assessment of the bias introduced by omitting perceived effectiveness

	% change in choice of more expensive instrument (5mm)	Standard Error	Test Statistic	P-Value
	UCLA only (n=26)			
Correct, knew 5mm > 10mm (Simple)	-56.9%	0.18	-3.17	0.002*
Correct, knew 5mm > 10mm (+Perceived effectiveness)[1]	-49.8%	0.17	-2.99	0.003*
	All 3 institutions (n=82)[2]			
Correct, knew 5mm > 10mm (Simple)	-25.5%	0.09	-2.98	0.003*
Correct, knew 5mm > 10mm (+Perceived effectiveness)[3]	-24.1%	0.07	-3.30	0.001*

* = p<0.05

[1] Perceived effectiveness was dichotomized based on the empiric distribution, with n=13 stating the 10mm was better than the 5mm (A > B) and n=13 stating the 10mm was the same or worse than the 5mm (A < B)

[2] Models included institutional fixed effects

[3] Perceived effectiveness was categorized based on the empiric distribution, with n=20 stating the 10mm was better than the 5mm (A>B), n=53 stating the 10mm was equal to the 5mm (A=B), and n=9 stating the 10mm was worse than the 5mm (A < B)

Table 20: Sensitivity analysis - effect of general cost knowledge on instrument choice (5mm Endoclip versus 10mm Endoclip), including assessment of the bias introduced by omitting perceived effectiveness

	% change in choice of more expensive instrument (5mm)	Standard Error	Test Statistic	P-Value
		UCLA only (n=26)		
Correct percent (Simple)	-1.7%	0.003	-5.47	<0.001*
Correct percent (+Perceived effectiveness)[1]	-2.0%	0.003	-6.53	<0.001*
		All 3 institutions (n=82)[2]		
Correct percent (Simple)	-0.4%	0.004	-1.24	0.22
Correct percent (+Perceived effectiveness)[3]	-0.5%	0.003	-1.54	0.12

* = p<0.05

[1] Perceived effectiveness was dichotomized based on the empiric distribution, with n=13 stating the 10mm was better than the 5mm (A > B) and n=13 stating the 10mm was the same or worse than the 5mm (A < B)

[2] Models included institutional fixed effects

[3] Perceived effectiveness was categorized based on the empiric distribution, with n=20 stating the 10mm was better than the 5mm (A>B), n=53 stating the 10mm was equal to the 5mm (A=B), and n=9 stating the 10mm was worse than the 5mm (A < B)

Table 21: Sensitivity analysis - effect of general cost knowledge on instrument choice (Clearify vs. anti-fog), including assessment of the bias introduced by omitting perceived effectiveness

	% change in choice of more expensive instrument (Clearify)	Standard Error	Test Statistic	P-Value
	UCLA only (n=26)			
Correct percent (Simple)	-0.8%	0.01	-1.17	0.24
Correct percent (+Perceived effectiveness)[1]	-0.5%	0.01	-0.67	0.50
	All 3 institutions (n=82)[2]			
Correct percent (Simple)	-0.2%	0.004	-0.42	0.68
Correct percent (+Perceived effectiveness)[3]	-0.2%	0.004	-0.45	0.65

[1] Perceived effectiveness was dichotomized based on the empiric distribution, with n=10 stating Clearify was significantly better than anti-fog and n=16 stating Clearify was slightly better or equal to anti-fog
[2] Models included institutional fixed effects
[3] Perceived effectiveness was categorized based on the empiric distribution with n=13 stating Clearify was significantly better than anti-fog, n=25 stating Clearify was slightly better than anti-fog, and 34 stating Clearify was the same or worse than anti-fog

Table 22: Decomposition of variance for the cost of supplies used during an operation

Fixed Effects Parameters		Change in mean case cost (USD)	Standard Error	Test statistic	P-Value
		N=1817			
Age		1.18	0.37	3.18	0.001*
Age ^ 2		0.03	0.02	1.95	0.05
Sex	Male		Reference		
	Female	-43.77	11.53	-3.80	<0.001*
Race	Non-Hispanic White		Reference		
	Non-Hispanic Black	10.74	22.53	0.48	0.63
	Non-Hispanic Other	-0.87	14.65	-0.06	0.95
	Hispanic	2.53	13.12	0.19	0.85
BMI		1.25	0.91	1.38	0.17
ASA	1		Reference		
	2	12.33	17.46	0.71	0.48
	3	47.98	22.08	2.17	0.03*
	4	67.51	47.87	1.41	0.16
	Non -Emergent		Reference		
	Emergent[1]	41.22	22.99	1.79	0.07
Urgency	Elective		Reference		
	Inpatient	49.36	23.05	2.14	0.03*
	Urgent	11.60	18.89	0.61	0.54
	Emergent[1]	-14.72	23.45	-0.63	0.53
Elixhauser Score		18.88	9.59	1.97	0.049*
Elixhauser Score ^ 2		-3.59	1.76	-2.04	0.041*
Indication	Biliary colic		Reference		
	Biliary Obstruction	32.02	19.55	1.64	0.10
	Acute Cholecystitis	65.03	15.20	4.28	<0.001*
Resident	No		Reference		
	Yes	20.97	26.52	0.79	0.43
Constant		528.56	68.11		

Random Effects	Variance Estimate	Standard Error	Confidence Interval
Facility	9681.02	1851.18	(1743.76-53747.05)
Surgeon	10175.10	781.74	(3432.95-30158.52)
Facility * Surgeon	5265.27	528.68	(1520.46-18233.32)
Residual	49438.37	13.92	(46291.07-52799.63)
Intraclass Correlation			
Surgeon	14%		
Facility	13%		
Surgeon & Facility	34%		
Proportional change in variance			
	Null model	Full model	Proportional Change in Variance
Facility	18382.64	9681.02	0.47
Surgeon	7339.36	10175.10	-0.39
Facility * Surgeon	5477.80	5265.27	0.04
Residual	53149.89	49438.37	0.07

* = $p<0.05$

ASA = American Society of Anesthesiologists; BMI = Body Mass Index

Note: The joint significance of ASA score had a P-value of 0.07 and the joint significance of urgency had a P-value of 0.053

[1] ASA emergent status is a dichotomous variable provided by the anesthesiologist based on their perception of the timeliness of the operation; urgency is determined based on the booking slip and the surgeon's indication of how quickly the operation needs to get done

Table 23: Effect of preference card cost on intraoperative disposable supply cost

Fixed Effects Parameters		Change in mean case cost (USD)	Standard Error	Test statistic	P-Value
		N=1817			
Preference Card Cost		0.79	0.07	11.15	<0.001*
Facility	RR-ASC		Reference		
	RR-OR	70.09	23.34	3.00	0.003*
	SM-OR	179.86	30.73	5.85	<0.001*
	SM-ASC	85.91	27.14	3.17	0.002*
Surgeon Volume		-0.10	0.07	-1.36	0.17
Age		1.19	0.37	3.18	0.001*
Age ^ 2		0.03	0.02	1.98	0.047*
Sex	Male		Reference		
	Female	-44.36	11.60	-3.82	<0.001*
Race	Non-Hispanic White		Reference		
	Non-Hispanic Black	15.60	22.64	0.69	0.49
	Non-Hispanic Other	4.94	14.71	0.34	0.74
	Hispanic	6.35	13.20	0.48	0.63
BMI		1.04	0.91	1.14	0.25
ASA	1		Reference		
	2	11.38	17.57	0.65	0.52
	3	48.13	22.19	2.17	0.03*
	4	66.97	48.16	1.39	0.16
	Non -Emergent		Reference		
	Emergent[1]	36.43	23.17	1.57	0.12
Urgency	Elective		Reference		
	Inpatient	38.98	21.91	1.78	0.08
	Urgent	2.17	17.94	0.12	0.90
	Emergent[1]	-22.74	23.09	-0.99	0.33
Elixhauser Score		16.07	9.67	1.66	0.10
Elixhauser Score ^ 2		-3.11	1.77	-1.75	0.08
Indication	Biliary colic		Reference		
	Biliary Obstruction	30.35	19.61	1.55	0.12
	Acute Cholecystitis	63.84	15.26	4.18	<0.001*
Resident	No		Reference		
	Yes	29.61	25.48	1.16	0.25
Constant		73.08	49.61		
Random Effects	Variance Estimate	Standard Error	95% Confidence Interval		
Surgeon	439.29	262.52	(21.22- 9094.64)		
Residual	50512.29	14.41	(47276.11- 53970.03)		

* = p<0.05

ASA = American Society of Anesthesiologists; ASC = Ambulatory Surgery Center; BMI = Body Mass Index; OR = Main Operating Room; USD = US dollars

Note: The joint significance of ASA score had a P-value of 0.06 and the joint significance of urgency had a P-value of 0.08

[1] ASA emergent status is a dichotomous variable provided by the anesthesiologist based on their perception of the timeliness of the operation; urgency is determined based on the booking slip and the surgeon's indication of how quickly the operation needs to get done

Table 24: Sensitivity analysis – effect of preference card cost on intraoperative disposable supply cost for cases after preference card creation and last review date

Fixed Effects Parameters		Change in mean case cost (USD)	Standard Error	Test statistic	P-Value
		n=891			
Preference Card Cost		0.84	0.11	7.65	<0.001*
Facility	RR-ASC		Reference		
	RR-OR	107.48	50.94	2.11	0.035*
	SM-OR	200.65	60.78	3.30	0.001*
	SM-ASC	98.48	60.65	1.62	0.10
Surgeon Volume		0.03	0.12	0.21	0.83
Age		1.52	0.51	2.96	0.003*
Age ^ 2		0.03	0.02	1.18	0.24
Sex	Male		Reference		
	Female	-31.55	15.67	-2.01	0.044*
Race	Non-Hispanic White		Reference		
	Non-Hispanic Black	-15.67	29.35	-0.53	0.59
	Non-Hispanic Other	0.06	20.28	0.00	1.00
	Hispanic	-8.85	18.29	-0.48	0.63
BMI		2.00	1.18	1.69	0.09
ASA	1		Reference		
	2	16.73	28.02	0.60	0.55
	3	46.48	33.23	1.40	0.16
	4	29.63	55.54	0.53	0.59
	Non -Emergent		Reference		
	Emergent[1]	24.00	27.85	0.86	0.39
Urgency	Elective		Reference		
	Inpatient	7.60	28.96	0.26	0.79
	Urgent	-22.69	25.68	-0.88	0.38
	Emergent[1]	-37.42	30.56	-1.22	0.22
Elixhauser Score		19.32	12.52	1.54	0.12
Elixhauser Score ^ 2		-3.00	2.17	-1.38	0.17
Indication	Biliary colic		Reference		
	Biliary Obstruction	52.97	27.51	1.93	0.054
	Acute Cholecystitis	93.71	21.13	4.44	<0.001*
Resident	No		Reference		
	Yes	27.10	38.22	0.71	0.48
Constant		-29.75	86.39		
Random Effects	Variance Estimate	Standard Error	95% Confidence Interval		
Surgeon	1374.67	184.00	(327.61- 5768.17)		
Residual	45750.47	26.31	(41645.75- 50259.77)		

* = p<0.05

ASA = American Society of Anesthesiologists; ASC = Ambulatory Surgery Center; BMI = Body Mass Index; OR = Main Operating Room; USD = US dollars

Note: The joint significance of ASA score had a P-value of 0.40 and the joint significance of urgency ha a P-value of 0.39

[1] ASA emergent status is a dichotomous variable provided by the anesthesiologist based on their perception of the timeliness of the operation; urgency is determined based on the booking slip and the surgeon's indication of how quickly the operation needs to get done

Table 25: Effect of preference card cost on patient outcomes

	Sample Size	Model, Parameter Estimated	Effect of increasing preference card cost by 100 USD	Confidence Interval	P-Value
Procedure Minutes	1815	Linear, Coefficient	1.22	(-2.36-4.80)	0.51
Room Minutes	1815	Linear, Coefficient	2.54	(-1.34-6.42)	0.20
Length of Stay (Hours)	1817	NB, IRR	0.98	(0.94-1.04)	0.57
Post Procedure Escalation of Care					
Outpatient --> Inpatient	824[1]	Logit, OR	0.79	(0.54-1.16)	0.24
Floor --> ICU	992[2]		0.88	(0.45-1.71)	0.70
Return to ED, 30 days[3]	1817	Logit, OR	0.83	(0.69-1.00)	0.050*
Return to ED, 90 days[3]	1817	Logit, OR	0.94	(0.80-1.10)	0.43

* = p<0.05

ED Emergency Department; ICU = Intensive Care Unit; IRR = Incidence Rate Ratio; NB = Negative Binomial; OR = Odds ratio; USD = US dollars

Unless otherwise indicated, all models controlled for facility, expected payer, urgency and indication of operation, surgeon volume, the presence of absence of a resident and patient age, race, BMI, ASA score, ASA emergent status, and Elixhauser score; quadratic terms were included for age, BMI and Elixhauser score to account for non-linear relationships; a surgeon random effect was included to account for clustering of outcomes within surgeon

[1]ASA emergent status and urgency were removed from the model because of non-applicability to outpatient setting and, because of perfect prediction at one or more levels of a categorical variables, we removed expected payer and converted ASA to a continuous variable

[2] One observation from the SM-ASC was excluded (hence n=992 instead of 993), and, because of perfect prediction at one or more levels of a categorical variable, we removed expected payer and race from the model and converted ASA and urgency from categorical variables into continuous variables

[3] Return to ED includes all causes/reasons

Table 26: Sensitivity analysis – comparison of inpatient observations with and without NSQIP data

	Without NSQIP data	With NSQIP Data	P-Value[1]
	n=663	n=329	
Age	53.56 +/- 19.25	55.27 +/- 18.62	0.18
Body Mass Index	27.4 [24-31.8]	27.6 [24.6-32.5]	0.24
Elixhauser Comorbidity Score	1 [0-2]	1 [0-2]	0.08
Facility			
RR-OR	279 (42.1)	137 (41.6)	0.90
SM-OR	384 (57.9)	192 (58.4)	
Resident Present	634 (95.6)	319 (97.0)	0.31
Primary Expected Payer			
Private	329 (49.6)	186 (56.5)	0.02*
Medicare	166 (25.0)	53 (16.1)	
Medical/CHIP/Other Public	139 (21.0)	85 (25.8)	
Other/Missing	29 (4.4)	5 (1.5)	
Female	408 (61.5)	195 (59.3)	0.21
Race/Ethnicity			
Non-Hispanic White	277 (42.1)	150 (45.7)	0.74
Non-Hispanic Black	49 (7.4)	22 (6.7)	
Non-Hispanic Other	116 (17.6)	56 (17.1)	
Hispanic	216 (32.8)	100 (30.5)	
ASA			
I	67 (10.1)	29 (8.8)	0.91
II	332 (50.2)	165 (50.3)	
III	245 (37.0)	126 (38.4)	
IV	18 (2.7)	8 (2.4)	
ASA Emergent	82 (12.4)	37 (11.2)	0.61
Urgency			
Elective	86 (13.5)	48 (14.9)	0.83
Inpatient	138 (21.7)	63 (19.5)	
Urgent	297 (46.6)	150 (46.4)	
Emergent/Critically Emergent	116 (18.2)	62 (19.2)	
Indication for Operation			
Biliary colic / Chronic Cholecystitis / Other elective	176 (26.5)	70 (21.3)	0.16
Obstructive Biliary	89 (13.4)	52 (15.8)	
Acute Cholecystitis	398 (60.0)	207 (62.9)	

* = $p<0.05$
ASA = American Society of Anesthesiologists; NSQIP = National Surgical Quality Improvement Program; OR = Main Operating Room
Note: Estimates generated using complete case analysis.
[1] Categorical variables (e.g., facility, race) were compared using $\chi 2$ tests of independence; continuous normally distributed variables (e.g., age) were compared using two sample t-tests; continuous non-normally distribute variables (e.g., BMI, elixhauser score) were compared using rank-sum tests.

Table 27: Sensitivity analysis – effect of preference card cost on patient outcomes for NSQIP subsample

	Effect of increasing preference card cost by 100 USD, odds ratio	Confidence Interval	P-Value
	n=329		
SSI (Any)[1]	1.75	(0.75-4.09)	0.20
Reoperation	Too rare of an event for model to converge		
Any Readmission[1,2]	0.58	(0.32-1.06)	0.08
Related Readmission[1,2]	0.54	(0.26-1.13)	0.10

SSI = Surgical site infection
[1]To prevent perfect prediction, expected payer and presence or absence of resident were removed from the model and ASA and indication were included as continuous covariates
[2]To further prevent perfect prediction emergent ASA status was removed from these models

Table 28: Summary of Findings

Question/hypothesis	Findings
Q1: What factors predict a surgeon's accuracy of knowledge of the relative cost of surgical instruments?	
H1a: Surgeons with prior exposure to instrument pricing will have more accurate knowledge about the relative cost of surgical instruments.	No association identified
Q2: How is accuracy of knowledge of the relative cost of surgical instruments associated with the cost of a surgeon's instrument preference card for laparoscopic cholecystectomy?	
H2a: Surgeons with more accurate knowledge of the relative cost of surgical instruments will prefer a lower-cost combination of these devices.	Hypothesis supported, although not significant at α of 0.05
H2b: The association between a surgeon's accuracy of the relative cost of instruments and the cost of an instrument preference card will be stronger among surgeons who view cost minimization as more important.	No association identified
Q3: For laparoscopic cholecystectomy, how much of intraoperative cost variation is explained at the patient, surgeon, and facility levels?	
H3a: More of the intraoperative cost variation will be explained at the surgeon compared to the patient or facility level.	Surgeon and facility combined explain 34% of variation
Q4: How does a surgeon's a priori preference card for laparoscopic cholecystectomy influence intraoperative costs and patient outcomes?	
H4a: There will be a significant positive association between the sum of instrument prices listed on the a priori preference card and the sum of actual instruments prices used in the operating room	Hypothesis supported
H4b: Controlling for patient factors and facility, higher-cost preference cards will be associated with better patient outcomes.	No consistent association identified

Appendices

Appendix 1: Definitions of select financial terminology

Adapted from Childers et al.[3]

Direct Costs	Department specific costs, such as salary, benefits, supplies, and utilities.
Indirect Costs	Costs not directly attributable to a department but are required as part of a larger operation, such as insurance, interest, and land tax.
Fixed Costs	Costs that do not vary based on activity level, e.g., depreciation of a piece of equipment.
Variable Costs	Costs directly proportional to activity level, e.g., surgical supplies.
Cost-Charge Ratio (CCR)	The CCR is the ratio between total patient charges (i.e., billed amounts) and total hospital costs. These ratios can be used to convert a patient's bill to an approximate hospital cost, based on the assumption that CCRs do not vary across service lines or from one patient to the next.
Time-Driven Activity-Based Costing (TDABC)	An accounting system that calculates the resources expended to take care of a patient by aggregating the smallest unit of measurement. For example, one component may be the number of minutes of nursing time multiplied by the cost of one minute of nursing time. TDABC requires extensive process mapping of a patient's route through the hospital to calculate costs, but is considered more accurate and more actionable as it allows identification of waste in the care process.

Appendix 2: Literature indirectly related to question #2

Author, Year	Population, Setting	Primary Research Question & Hypotheses	Study Design	Intervention	Outcome(s) & Operationalization	Covariates	Sample Size	Key Findings
Austin, 2017	Orthopedics, regional healthcare system in U.S.	(1) Does cost disclosure influence surgeons to reduce operating room expenditures?	One-group pre-post	Cost report card	Average disposable cost per case	NA	423 patients	Average supply cost decreased by $269 over study period.
		(2) Do surgeons like the scorecard and does it influence practice?	Cross sectional survey	NA	4 yes/no and 1 Likert-scale question		9 of 11 surgeons responded	89% moderate to extreme interest in scorecard; 89% looked at scorecard; 89% looked at other surgeons' scorecards; 56% say it changed the way performed operation; 33% expressed concern cost may have on referral patterns.
Avansino, 2013	Pediatric surgery, one academic teaching hospital in U.S.	(1) There is significant variation in cost of performing laparoscopic appendectomy	Retrospective review	NA	Average disposable cost per surgeon	NA	145 patients, 10 surgeons	From figure 1: pre-intervention average cost ranged from ~ $650 to $1050.
		(2) Standardizing the operation can reduce costs without affecting quality	One-group pre-post	Standardization of preference card except port selection & mechanism of taking mesoappendix	Average disposable cost per case		145 patients pre, 101 patients post, 10 surgeons	Reduction of $167 per case; no change in OR time, LOS, or adverse events.
		(3) What do operating room staff think of standardized preference card?	Cross sectional survey	NA	7 Likert-scale questions		51 of 65 people responded including 11 surgeons, 25 nurses, 15 scrub techs	General agreement that preference card improved efficiency and reduced supply costs without compromising patient safety; adherence to preference card was mixed with surgeons rating their own adherence much higher than their peers.

Author, Year	Population, Setting	Primary Research Question & Hypotheses	Study Design	Intervention	Outcome(s) & Operationalization	Covariates	Sample Size	Key Findings
Croft, 2017	Obstetrics and gynecology, one community hospital in U.S.	Does surgeon education of costs decrease total instrumentation cost?	One-group pre-post	Meeting with surgeons and individual cost feedback	Average disposable cost per case	NA	13 surgeons, 271 patients pre and 69 patients post intervention	Average decrease of $257 per case; no change in operative time
Gitelis, Guzman, Vigneswarin; 2015	General surgery; one regional healthcare system in the U.S.	Cost reductions will occur if surgeons were educated about their disposable supply usage during common general surgery operations (laparoscopic cholecystectomy, inguinal hernia repair, laparoscopic appendectomy)	One-group pre-post	Educational seminar to surgeons about disposable costs, average cost by surgeon, list of items with costs and strategies to reduce costs	Average disposable cost per case	Age, BMI, ASA	Laparoscopic cholecystectomy: 15 surgeons, 586 patients pre and 428 patients post	Savings of $58 per case; no change in readmissions, reoperations, infections, or transfusions
						Age, Sex, ASA, BMI	Laparoscopic appendectomy: 16 surgeons, 336 patients pre and 357 patients post	Savings of $210 per case; no change in LOS, operative time, readmissions, reoperations, SSI
						Age, ASA	Laparoscopic and open inguinal hernia repairs: 10 surgeons; for lap: 258 patients pre and 274 post; for open: 366 patients pre and 286 post intervention	Savings of $228 for laparoscopic cases, $48 for open cases; no change in LOS, recurrence, OR time
Skarda, 2015	Pediatric surgery, one academic teaching hospital in U.S.	How effective is standardization of technique/preference card at reducing costs and maintaining patient safety?	One-group pre-post	Consensus opinion amongst 6 surgeons based on cost, availability, and utility of instruments	Average disposable cost per case	Age, rupture status	6 surgeons; 346 patients pre and 362 post intervention	Savings of $539 per case; increased procedure and OR time by 2 minutes; no change in LOS, readmissions, ED visits, abscess, reoperation, IR drainage, subQ abscess, c.diff colitis, CT scans

Author, Year	Population, Setting	Primary Research Question & Hypotheses	Study Design	Intervention	Outcome(s) & Operationalization	Covariates	Sample Size	Key Findings
Still, 2017	Plastic surgery, one academic teaching hospital in U.S.	Heat mapping can facilitate standardization of surgeon preference cards and reduce waste	Cross-sectional	Lean method process of categorizing items on preference card into low, medium and high cost using "heat maps"; showing these cost maps to surgeons, meeting with them to discuss preference cards, and eliminating unnecessary equipment	Disposable cost per case	NA	3 Surgeons; No patient sample size	Estimated $18,000 in annual savings by eliminating 1,693 items
Zygourakis, 2017	Multi-department, single academic health system in the U.S.	(1) What is the association between providing surgeons cost feedback and supply costs?	Pre-post controlled study	Monthly supply scorecards with surgeon's baseline and peers' average prices along with lists of expensive items and "bang for your buck" items; intervention arm was orthopedics, ENT and NSG while control groups were cardio/thoracic, general, vascular, pediatric, ob/gyn, ophthalmology and urology; financial incentive of $50k for academic/research for departments that met 5% cost reduction goal (control and intervention departments)	Disposable cost per case, intervention versus control	Surgeon, department, patient demographics, "clinical indicators", CMI	10637 patients pre to 11820 patients post; 63 surgeons in intervention group and 186 surgeons in control group	Median cost decreased by $91 in intervention; compared to increase of $53 in control groups; no difference in 30-day readmissions, but actually improved 30-day mortality and discharge location outcomes in intervention versus control
		(2) How effective was the intervention and what are surgeons' attitudes towards OR costs and healthcare value?	Cross sectional survey	NA	Multiple Likert scales assessing whether surgeons looked at scorecards, if it influenced supply use, and various other questions about healthcare value		91 of 249 attending surgeons responded	86% of surgeons reported looking at scorecards always, often or sometimes and 79% always, often, or sometimes used the data to influence supply use; those receiving the scorecard reported higher values for knowing about their procedure costs compared to their peers and knowing which items contribute to high cost; 77% of scorecard recipients agreed or strongly agreed that the intervention helped learn about cost and efficiency and 80% agreed or strongly agreed program should be continued

ASA = American Society of Anesthesiologists physical status scale; BMI = body mass index; CMI = case mix index; CT = Computed tomography; ENT = Ear, Nose, and Throat; IR = Interventional Radiology; NA = Not Applicable; NSG = neurosurgery LOS = length of stay; OB/Gyn = Obstetrics & Gynecology; OR = Operating room; SSI = Surgical site infection;

Appendix 3: Literature indirectly related to question #4

Author, Year	Population, Setting	Primary Research Question & Hypotheses	Study Design	Outcome(s) & Operationalization	Covariates	Sample Size	Analysis	Key Findings
Brauer, 2015	Laparoscopic cholecystectomy (+/- IOC) at 7 U.S. medical centers	What is the relationship between cost and quality?	Retrospective review (administrative data)	Intraoperative supply cost (estimated); 30-day NSQIP outcomes (subset only)	Case volume	2178 patients, 55 surgeons	Graphical	No graphical relationship between all complication rate and median cost.
Meier 2014	Pediatric adenotonscillectomy at 1 U.S. medical center	What is the relationship between cost and patient outcomes?	Retrospective medical chart review	Total encounter cost; any complication, hospital readmission, observation, ED visit, hemorrhage, hemorrhage requiring procedure	Surgeon	4824 patients, 14 surgeons	Bivariate	Significant variation between surgeons in all complications and ED visits; "Higher costs for the procedure did not correlate with fewer complications".
Rosenbaum, 2017	Lumbar spine surgery at 1 U.S. medical center	What is the correlation between cost and patient reported outcomes?	Retrospective medical chart review	Supply and implant costs; Outcomes were EQ-5D, PDQ, PHQ-9 (patient reported outcome metrics)	Surgeon, inpatient/outpatient, pt age, pt sex, operation duration	652 patients, 9 surgeons; but only 34-52% of patients had pre and postoperative outcomes data	Multivariate	Surgical cost only correlated to mean difference in PDQ ($p=0.03$) with more costly surgeries resulting in worse outcomes.
Thomas, 2016	ENT; 21 facilities within 1 U.S. healthcare system	What is the relationship between total costs and patient outcomes?	Retrospective medical chart review	Total cost per case; 30-day outcomes - return to OR or return to ED	None	4007 patients, 72 surgeons	Bivariate	No difference in total cost between patients with a complication and patients without a complication.
Zygourakis, 2016	All inpatient NSG procedures at 1 U.S. teaching hospital	What is the effect of patient, procedural, and provider-specific factors on patient outcomes?	Retrospective medical chart review	Surgical supply cost; 30-day readmission, 30-day mortality	Patient age, gender, payer, CMI, BMI, admission source, procedure length, procedure date, surgeon case volume, surgeon (random effect)	4,904 patients, 24 surgeons	Mixed effects	No association between supply cost and mortality or readmission rates.

BMI = Body Mass Index; CMI = Case Mix Index; ED = Emergency Department; ENT = Ear, Nose & Throat; IOC = Intraoperative Cholangiogram; NSG = Neurosurgery; NSQIP = National Surgical Quality Improvement Project

Appendix 4: Sensitivity Analyses – Flow diagrams, sample sizes, and missing data

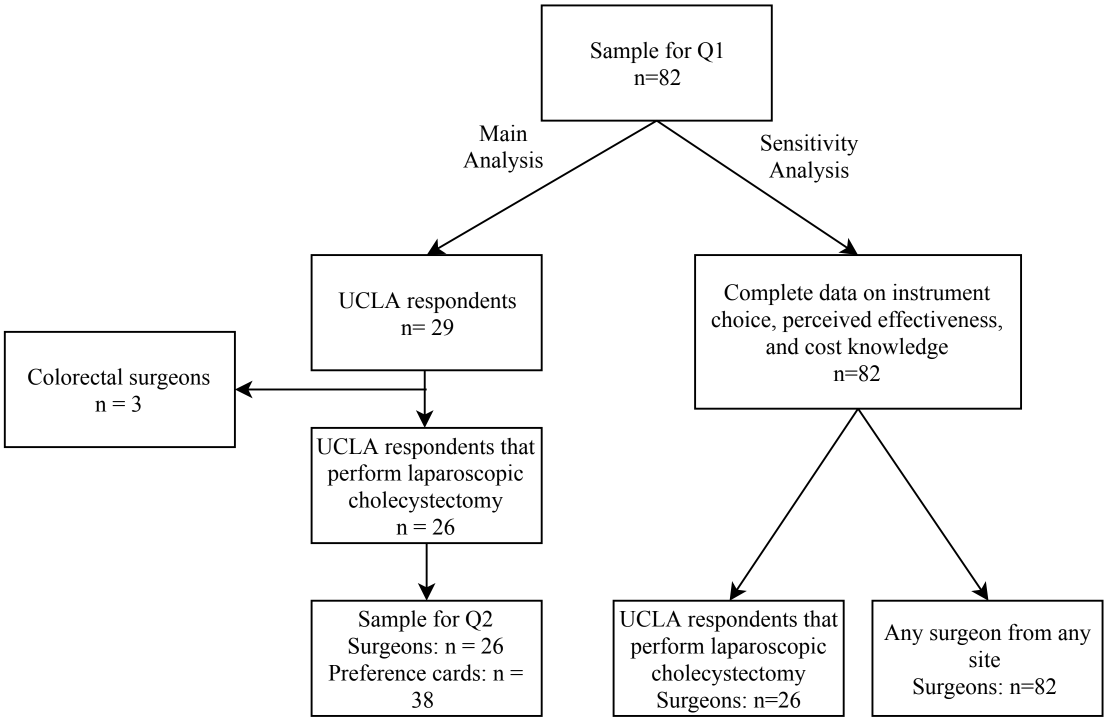
Data source, inclusion/exclusion criteria, and sample size for each sensitivity analysis, by question

	Data Source(s)	Inclusion/Exclusion Criteria	Final Sample Size
Question 1			
(1) Different operationalizations of the moderator (2) Alternative measures for relative importance of cost vs. effectiveness (3) Bivariate association between familiarity and exposure (4) Association between years since finished training and exposure (5) Association between years since finished training and cost knowledge	Survey	All survey respondents with non-missing exposure information	82 surgeons
Question 2			
(1) Different operationalizations of the primary predictor (cost knowledge) and moderator	Survey	All survey respondents	82 surgeons
(2) How does perceived effectiveness bias relationship between cost knowledge and actual preference/choice?	Survey	Respondents from UCLA who perform laparoscopic cholecystectomy	26 surgeons
		Respondents from any institution with non-missing choice information	82 surgeons
Question 3 - NA, No Sensitivity Analyses			
Question 4			
(1) Hypothesis 4a - restricting sample to cases after preference card created and last reviewed	Medical Record Data	Same as main regression + restricted to cases after preference card was created and last reviewed	891 patients, 21 surgeons, 33 preference cards
(2) Hypothesis 4b - restricting sample to inpatient cases with and without NSQIP data	Medical Record Data	Same as main regression + restricted to inpatient operations with NSQIP data	329 patients, 19 surgeons, 22 preference cards

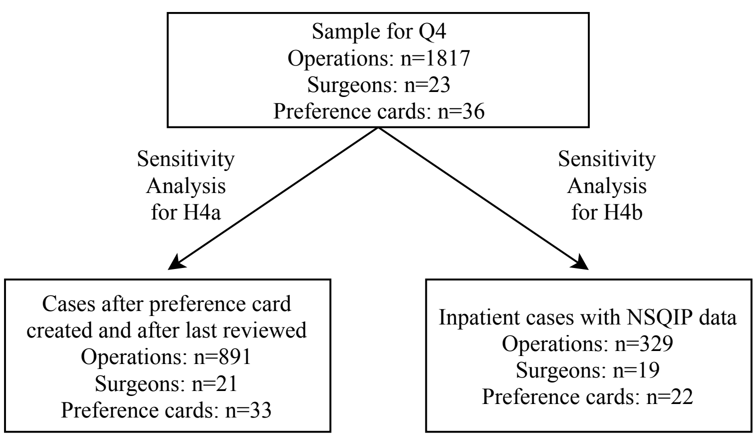
Flow diagrams for sensitivity analysis

Note: Questions 1 & 3 shared the same sampling frame as the main regression; please see the main text for details.

Question 2:



Question 4:



Missing data for variables included in sensitivity analyses, by question

Variable	Complete Data, n	Missing Data, n	Mechanism of Dealing with Missingness
Question 1			
Relative importance of cost vs. effectiveness, alternative operationalizations			
Value of one minute of OR time	75	8	Multiple imputation using predictive mean matching, k=10, with exposure, year finished, institution, and cost knowledge as covariates.
Value of surgical site infection	78	5	
Question 2			
Choice of instruments			
Comparison 1, Endoclip Comparison 2, Clearify vs. Anti-fog	26 (UCLA) 82 (All respondents)	0 (UCLA) 1 (All respondents)	NA for UCLA data; complete case analysis for all respondents
Effectiveness of instruments			
Comparison 1, Endoclip Comparison 2, Clearify vs. Anti-fog	26 (UCLA) 82 (All respondents)	0 (UCLA) 1 (All respondents)	NA for UCLA data; complete case analysis for all respondents
Cost of instruments			
Comparison 1, Endoclip Comparison 2, Clearify vs. Anti-fog	26 (UCLA) 82 (All respondents)	0 (UCLA) 1 (All respondents)	NA for UCLA data; complete case analysis for all respondents
Accuracy of cost knowledge	23 (UCLA) 79 (All respondents)	3 (UCLA) 4 (All respondents)	Single (regression) imputation using all survey day (all 79 respondents with complete data)
Question 3 - Not applicable, no sensitivity analyses			
Question 4			
Hypothesis 4a			
Race/Ethnicity	883	8	Multiple imputation using multinomial (race), linear (BMI), and ordered logit (ASA/urgency) imputation with 20 imputation sets with the outcome (e.g., cost of instruments), age, sex, Elixhauser score, ASA emergent status, indication, and resident as predictors
BMI	880	11	
ASA	890	1	
Urgency	872	19	
Hypothesis 4b			
BMI	323	6	Multiple imputation using linear (BMI), and ordered logit (ASA/urgency) imputation with 20 imputation sets with the outcome (e.g., readmission), age, sex, payer, Elixhauser score, ASA emergent status, indication, and resident as predictors
ASA	328	1	
Urgency	323	6	

Appendix 5: Cognitive interviewing script

Introduce self, and if applicable, introduce additional interviewer(s) by name:

Hi my name is ___________________.

Thank you for taking the time to meet with us today. Before we begin, let me review some general information. This interview is part of a research study about surgeon knowledge and opinions of instrument costs, and is expected to last no more than 45 minutes. The purpose of these interviews is to refine a survey instrument that will be distributed to surgeons around southern California. The de-identified, aggregated results of the survey will be shared with you, once complete.

This interview is entirely voluntary. You are free to skip questions, or stop or postpone this interview at any time. To protect privacy, throughout this interview I will not refer to you by name. It would be helpful if you could also refer to colleagues by title rather than name. If you forget and do so, that is okay; we will just redact their names from the transcript. Because we value everything you have to say and want to keep a good record of the conversation, I would like to request your permission to audio record this interview. Only project staff will hear the audio recording and it will be stored in a HIPAA compliant system. The recordings will be transcribed, analyzed, and summarized. Everything you say today will be kept confidential in all summary reports, meaning your name will not be used in the interview paperwork or in summaries. Instead, each participant will receive a unique ID number that will be used in place of your name or other identifying information. If you are not comfortable being recorded, I can take written notes so I can accurately record your information.

Do you have any questions about the study or about your participation?

We're ready to begin. May I record the interview? [PAUSE] Now that the recorder is on, I wanted to confirm again that I have your permission to record this interview. If you want to pause or turn off the recorder at any time during the interview, just let me know. Are you okay with this? Okay, I will say the code number, and then we can get started. This is code number XXXXXX, and the date is XXXXXXX.

Code numbers will be assigned based on institution and subject number

In a couple of minutes, I'm going to open the online survey. When I do, I would like you to talk out loud about your reactions to the form as you read questions and fill it out. I would like to know everything you think about. Talking out loud about these sorts of things may seem a little unusual, so before I give you the official survey, I have a short practice question:

Show practice questions (number of windows)

Possible probes:
- *Did you count sliding glass door?*
- *What did you do with multi-pane windows?*

Now we are going to move onto the survey itself. Please fill it out at your own pace and talk out loud about your impressions of it. Anything that you would have read to yourself, please read out loud.

Notes & possible probes

<u>Intro Page</u>

<u>Recency of Surgical Training</u>
Were the choices of finishing general surgery residency adequate?
Is it clear what is meant by clinical fellowship?

<u>Exposure</u>
What comes to mind when you think of "disposable surgical supplies"?
For this study, what is meant when we refer to supply costs?
Why did you pick the number you did for familiarity?
What types of exposures have you had to instrument prices?
How often have these exposures occurred?
Have these exposures changed over the course of your career?

<u>Cost vs. Effectiveness</u>
Are there enough answer choices?
When we think of an instrument's effectiveness – what factors do you consider?
If you were to order the factors above, which would be more important? Which would be less important?
How do the above factors change when you are preparing for an operation (i.e. generating a preference card) versus in the middle of an operation?
(Opportunity cost question): When coming up with a number, what are you considering?
(Wound infection question): Do you have enough clinical details to make a decision? When you think about "costs" for this question – whose perspective are you considering? (examples include the patient, hospital, payer, etc.)

<u>Accuracy of Cost Knowledge</u>
Is it clear what we are asking you to do?
What do you think we mean when we say "ignore any capital expenditures that would be required to use the instrument"?
Are the pictures large enough?
Is it clear what is being compared?
Are the answer choices sufficient?
When two pictures are shown is it clear why?
Is it clear when there are two or more items being included in a comparison group?

<u>Perceived Effectiveness</u>
Is it clear what we are asking you to do?
Is the clinical scenario adequately detailed?

Are the answer choices sufficient?
How do the effectiveness/cost answers influence your final choice?

Debriefing Questions
How easy was the instrument?
Was the instrument the appropriate length?
What do you think are the barriers to getting surgeon's to reduce costs? To standardize the instruments they use?
Anything else you want to add?

That is all I have for your today. Thank you very much for your time. We anticipate administering this survey in the coming months. We look forward to being able to share the findings with you.

Appendix 6: Sample surgeon/site preference card

Surgeon Information

Surgeon	Surgeon ID
[redacted]	

Supplies

Bin Loc.	Item Name	EMPAC	Mfg. Num.	Open	PRN	Latex
Totes						
909320	**TOTE LAP CHOLE SURGITRACK**	319139	SMLC0026	1	0	
Suture						
1610	**SUT MONOCRYL 4-0 PS-2 27IN J&J**	116551	Y426H	2	0	
2120	**SUT VICRYL 0 CT2 27IN**	071849	J270H	4	0	
SCC6M55	**SUT SILK 0 PSL 18IN J&J**	075385	580H	1	0	
Other						
160	**NEEDLE INSUFFLATION 14GAX120MM 10EA/BX**	319383	S100000	1	0	
410	**APPLIER ENDO CLIP 5MM**	135637	176630	1	0	
690	**TROCAR VERSASTEP 5MM 3EA/BX**	320157	VS101005	3	0	
700	**TROCAR VERSASTEP PLUS 12MM**	501638	VS101012P	1	0	
1495	**ELECTRODE SLR/DIV LIGASURE 5MM**	521138	LF1637	0	1	
1695	**DHELP DEFOG HEAT ENDO LENS PRO TECTOR**	635324	21345	1	0	
11140	**KIT INF CONTROL TURNOVER RR**	521998	UCLA-06	1	0	
AN405	**PACK LAP CHOLE SM**	542199	DYNJ0780078N	1	0	
BA210	**SLEEVE CALF 18IN RPR 2EA/PR**	519076	ALP1UCLA	1	0	
CB705	**DRAIN EVACUATOR 100CC DAVL** Comment: HOLD	001866	0070900	0	1	
EC205	**APPLICATOR CHLORAPREP 26ML**	500250	260815	2	0	
EC525	**DRAIN WOUND FULL 19FR .25IN**	314078	DYNJWE2190	0	1	
SCC5D05	**TUBING IRRIG LAP SUCT 5MMX33CM**	315065	5552002	1	0	
SCC6M30	**BLADE SURG 11 CARBON STRL ***DROP SHIP ONLY***** Comment: inside the pack	020182	73-0411	0	1	
Dressing						
EC315	**BALL COTTON LG STRL** Comment: FOR UMBILICUS - AS DRESSING	000525	08168110100	1	0	
FA205	**DRESSING COVERLET 3.25 X 3.375**	312310	66000709	4	0	

Drugs

Item Name	Open	PRN	Amount
bupivacaine 0.5%-epinephrine (PF) injection 1:200,000	0	0	0 mL
sodium chloride (GU Irrigant) 0.9% Soln	0	0	1000 mL

Instruments

Bin Location	Item Name	Open	PRN
Cart 01	**SMH GRASPER** [redacted]	2	0
Cart 11	**SMH LONG VASCULAR CLAMPS 19901**	0	1
SMBS01D/E	**SMH Basic Major Tray Gen 18583**	1	0

SMBS02C/D	**SMH Clickline Set 5mm 18566**	1	0
SMBS03B	**SMH MIS Telescopes Set 5mm/10mm Storz 18564**	1	0
SPCt1B3	**SMH CS CLAMP TOWEL LG 15925**	0	1
	SMH Handle Camera Light Cover 104505	1	0
	SMH Image One HD Camera Storz 104289	1	0

Equipment

Bin Location	Item Name	Open	PRN
BACK HALLWAY	**SMH EQ SUCTION NEPTUNE**	1	0
OR ROOMS	**SMH EQ OR TABLE SKYTRON 3501B X 4**	1	0
OR'S	**SMH EQ WARMER BAIR HUGGER**	2	0
	SMH EQ DONUT FOAM 9IN	1	0
	SMH EQ ESU	1	0
	SMH EQ IRRIGATION CONTROLLER DAVOL	1	0
	SMH EQ PROTECTOR HEEL	1	0
	SMH EQ PROTECTOR ULNAR NERVE	1	0
	SMH EQ PULSATILE SCD MACHINE	1	0

Implant Trays

None

Instructions

Scheduling Instructions

Patient Instructions

Nursing Instructions

Comments:

[redacted]

FOOT PEDAL TO DR`S RIGHT FOOT.

CUT = 40; COAG = 40.

MD STANDS ON PT`S RIGHT SIDE.
HAVE PT. URINATE JUST BEFORE GOING TO SURGERY

ON THE MAYO STAND:
MAYO CURVE SCISSOR, CRILE, DEBAKEY X2
[redacted] ENDO PEANUTS, LAP. RIGHT ANGLE, MARYLAND, SPATULA WITH SHEATH

FOR CHOLANGIOGRAM:
SYRINGE 20ML X 1 FOR SALINE
SYRINGE 10ML X 1 FOR CONRAY
EXTENSION SET 34 IN. #2C6227
DRAPE C-ARM 27X70IN. #4951
STOPCOCK 3WAY #MX53IIL
CONRAY FOR CONTRAST
INJECTABLE NORMAL SALINE

Pre-proc Prep Instructions

Drapes:

USES BLUE TOWELS TO SQUARE OFF (NOT PAPER) ; LAP CHOLE DRAPE.

Prep:
CHLOROPREP X 2

Pharmacy:
(as of 01.11.2016) - MARCAINE 0.5% WITH EPINEPHRINE
0.9% NACL IV BAGS 1000CC (NOT 3,000ML, PLEASE);

HOLD: CONRAY 50% + 0.9% NACL INJECTABLE 100CC;
ANCEF 1GM TO ANESTHESIA.

Dressing:
BENZOIN; STERISTRIPS; OPSITE ; COTTONBALL X 1 ON THE UMBILICUS.

Position Instructions

SUPINE. ARMS EXTENDED ON ARMBOARD WITH ULNAR PADS;
PILLOW UNDER KNEES; HEEL PADS; SAFETY STRAP

Positioning Information

Position #1

Laterality: Default

Body	Left Arm	Right Arm
Supine Sheet Draw Strap Safety	Extended Armboard Pad Foam Arm Strap Safety	Extended Armboard Pad Foam Arm Strap Safety
Head	**Left Leg**	**Right Leg**
Aligned Pillow Donut Foam 9 Inch	Flexed Slightly Pillow Pad Foam Feet SCD / Pulsatile Stockings	Flexed Slightly Pillow Pad Foam Feet SCD / Pulsatile Stockings

Preference Cards Using Settings From This Preference Card

None

Pick Lists Using Settings From This Pick List

None

Review History

User	Date	Time	Comments
Seroje, Gladys Radaza, RN	10/27/2016	0829	
Seroje, Gladys Radaza, RN	7/13/2015	1442	
Seroje, Gladys Radaza, RN	3/16/2015	1742	
Seroje, Gladys Radaza, RN	10/30/2014	1245	
Seroje, Gladys Radaza, RN	10/6/2014	1158	
Seroje, Gladys Radaza, RN	3/25/2014	0946	
Seroje, Gladys Radaza, RN	6/18/2013	1625	
Seroje, Gladys	12/13/2012	1924	
Seroje, Gladys	12/5/2012	1456	

Appendix 7: Survey Instrument

Welcome

Hello, and welcome -

You have been selected to participate in this survey because you are a surgeon. Your participation in this survey is voluntary. If you choose to participate, the survey will inquire about the instruments you use in the operating room and the factors you consider in this process. The survey will take approximately 10 minutes to complete. Your answers will be kept confidential and any presentations (oral or written) of data will be in the aggregate.

Thank you very much for your time.

If you have questions please contact any of the following study investigators:

UCLA
Melinda Maggard-Gibbons (mmaggard@mednet.ucla.edu)
Chris Childers (cchilders@mednet.ucla.edu)

UCSD
Bryan Clary (bclary@ucsd.edu)
Mark Zhao (markzhao@ucsd.edu)

CSMC
Rodrigo Alban (rodrigo.alban@cshs.org)
Joshua Tseng (joshua.tseng@cshs.org)

Recency of surgical training

At which institution do you primarily operate?

- ○ UCSD
- ○ CSMC
- ○ UCLA

What year did you finish general surgery residency?

[▾]

Did you complete a **clinical** fellowship? (eg, minimally invasive, colorectal)
Do not include fellowships dedicated primarily to research

- ○ Yes
- ○ No

What year did you finish your final clinical fellowship?

[▾]

Prior exposure to instrument prices

For the remainder of the survey we will be discussing disposable surgical instruments/supplies. That is, single use devices such as staplers and dressings. These questions ask about the price the hospital pays to a vendor to acquire the supply ("cost").

	Not familiar at all	Slightly familiar	Moderately familiar	Very familiar	Extremely familiar
How familiar are you with the purchase price of surgical instruments?	○	○	○	○	○

Surgeons vary in their exposure to instrument prices. They can be exposed to prices by participating on supply purchasing committees, seeing cost report cards, or asking about the price of instruments.

How much exposure have you had to the prices of instruments you use in surgery?

○ None at all
○ A little bit
○ A moderate amount
○ A great deal

Relative importance of cost vs. effectiveness to the surgeon

When choosing an instrument to be listed on your preference card, how important is the cost of a surgical instrument in comparison to the instrument's effectiveness?

Only consider cost				Consider cost and effectiveness equally					Only consider effectiveness
	○	○	○	○	○	○	○	○	○

When choosing an instrument to be listed on your preference card, how important is the cost of a surgical instrument in comparison to the instrument's effectiveness?

Only consider cost				Consider cost and effectiveness equally				Only consider effectiveness
○	○	○	○	○	○	○	○	○

Imagine two instruments - Instrument X costs $250 and can complete an operative step in 10 minutes; Instrument Y has recently come to market and completes the same step in 3 minutes. The instruments are equally effective. What is the maximum price the hospital should pay for Instrument Y?

In other words:
Instrument X: 10 minutes -- $250

Instrument Y: 3 minutes -- $??

For this question, assume a superficial wound infection is: erythema with no purulence, and a 7-day course of antibiotics as an outpatient results in full resolution of the infection. The patient suffers no additional adverse event related to the infection or the antibiotics.

Now, imagine two instruments - Instrument X costs $250 and, when used, the probability of a superficial wound infections is 3 out of 100. Instrument Y has recently come to market and reduces the probability of a superficial wound infection to 2 out of 100. The instruments are otherwise equally effective. What is the maximum price the hospital should pay for Instrument Y?

In other words:
Instrument X: SSI rate 3/100 -- $250
Instrument Y: SSI rate 2/100 -- $??

Accuracy of cost knowledge

For each of the following comparisons, please select which item is more expensive and by how much. Ignore any capital expenditures required to use the instrument (eg, ignore the cost of the electrical source for a Ligasure, only consider the cost of the handpiece).

Comparison 1 of 8

INSTRUMENT A

CHLORAPREP

INSTRUMENT B

IODINE "WET" PREP

	A is more expensive			B is more expensive		
	A>>>B	A>>B	A>B	A<B	A<<B	A<<<B
Which item is more expensive and by how much?	○	○	○	○	○	○

Comparison 2 of 8

INSTRUMENT A

ENDOCLOSE

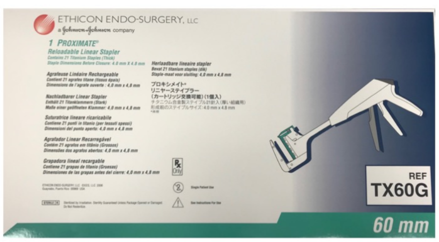
INSTRUMENT B

CARTER-THOMASON

Comparison 3 of 8

INSTRUMENT A	**INSTRUMENT B**

60MM TA STAPLER **80MM GIA <u>RELOAD</u>**

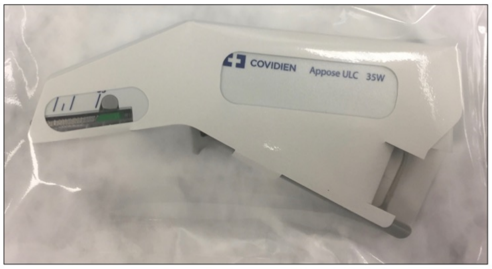


60MM TA STAPLER **75MM GIA <u>RELOAD</u>**

60MM TA STAPLER **75MM GIA <u>RELOAD</u>**

INSTRUMENT A

INSTRUMENT B



+

SKIN STAPLER

4-0 MONOCRYL x 2

+

SKIN STAPLER

4-0 MONOCRYL x 2

Comparison 5 of 8

INSTRUMENT A	INSTRUMENT B

ARGON BEAM HANDPIECE

2 PACKS OF 4” X 4” FIBRILLAR

ARGON BEAM HANDPIECE

2 PACKS OF 4” X 8” SURGICEL

ARGON BEAM HANDPIECE

2 PACKS OF 2” X 14” SURGICEL

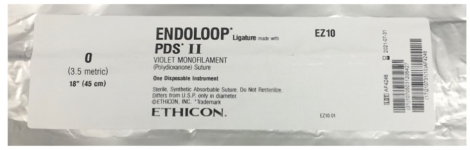
Comparison 6 of 8

INSTRUMENT A

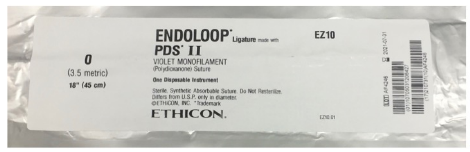
+

ENDO-STAPLER +
BLUE LOAD

INSTRUMENT B

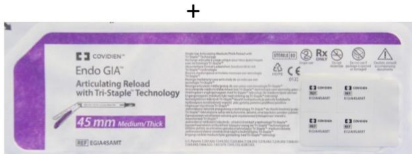


+

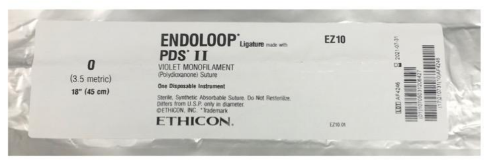


ENDOLOOPS X 2

+



ENDO-STAPLER +
PURPLE LOAD

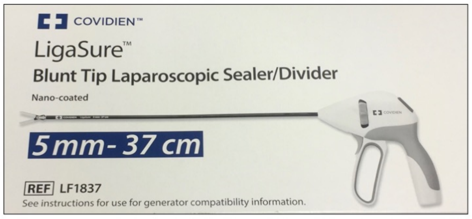


+

ENDOLOOPS X 2

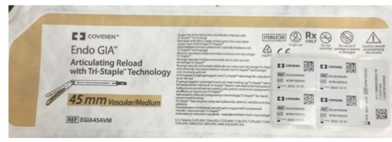
Comparison 7 of 8

INSTRUMENT A	INSTRUMENT B

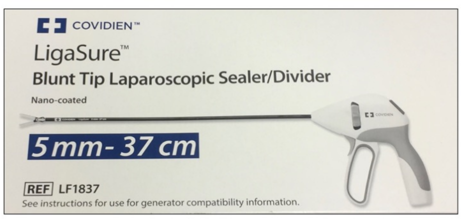


LIGASURE

+



ENDO-STAPLER +
VASCULAR LOAD



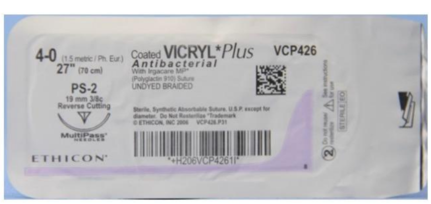
LIGASURE

+

ENDO-STAPLER +
VASCULAR LOAD

Comparison 8 of 8

INSTRUMENT A	INSTRUMENT B



4-0 VICRYL x 1

4-0 MONOCRYL x 1

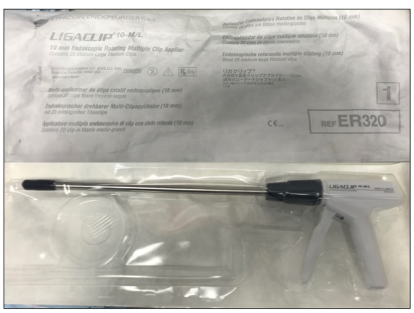
Perceived Effectiveness Block

For the final 2 comparisons, assume the following scenario when making your choices:

You are performing an elective outpatient cholecystectomy on a 20 y/o, BMI 30, female for biliary colic. She has no prior history of surgery. The gallbladder has no active inflammation and the duct is of a normal caliber.

Comparison 1:

INSTRUMENT A



LIGACLIP
10MM DIAMETER
MEDIUM/LARGE CLIPS

INSTRUMENT B

LIGAMAX
5MM DIAMETER
MEDIUM/LARGE CLIPS

ENDOCLIP
10MM DIAMETER
MEDIUM/LARGE CLIPS

ENDOCLIP
5MM DIAMETER
MEDIUM/LARGE CLIPS

In this clinical scenario, which of these two instruments **would you prefer to use**?

(Reminder of the scenario: You are performing an elective outpatient cholecystectomy on a 20 y/o, BMI 30, female for biliary colic. She has no prior history of surgery. The gallbladder has no active inflammation and the duct is of a normal caliber.)

- ○ Instrument A
- ○ Instrument B

Which instrument is more **expensive**?

(Reminder of the scenario: You are performing an elective outpatient cholecystectomy on a 20 y/o, BMI 30, female for biliary colic. She has no prior history of surgery. The gallbladder has no active inflammation and the duct is of a normal caliber.)

A is more expensive			B is more expensive		
A>>>B	A>>B	A>B	A<B	A<<B	A<<<B
○	○	○	○	○	○

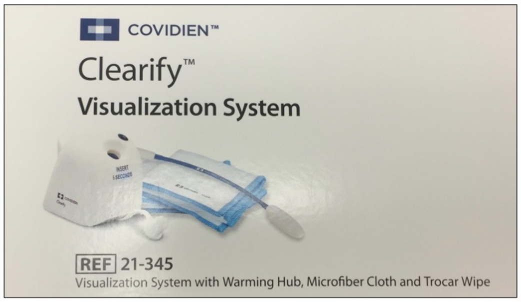
Which instrument is more **effective**?

(Reminder of the scenario: You are performing an elective outpatient cholecystectomy on a 20 y/o, BMI 30, female for biliary colic. She has no prior history of surgery. The gallbladder has no active inflammation and the duct is of a normal caliber.)

A is more effective				B is more effective		
A>>>B	A>>B	A>B	A=B	A<B	A<<B	A<<<B
○	○	○	○	○	○	○

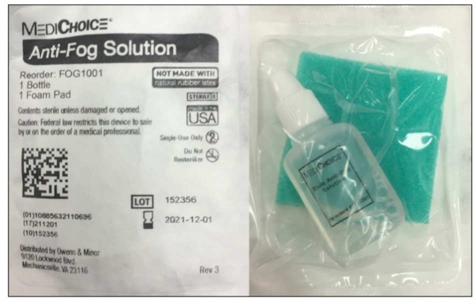
Comparison 2 (Same Questions):

INSTRUMENT A



CLEARIFY VISUALIZATION SYSTEM

INSTRUMENT B



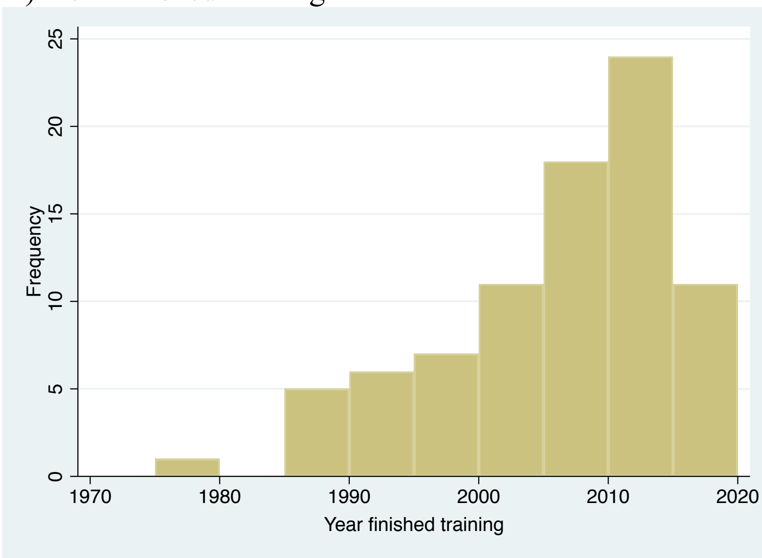
ANTI-FOG SOLUTION & PAD

Block 6

We would appreciate any feedback or comments you may have.
When you are finished, please click the **SUBMIT** button to record your answers.

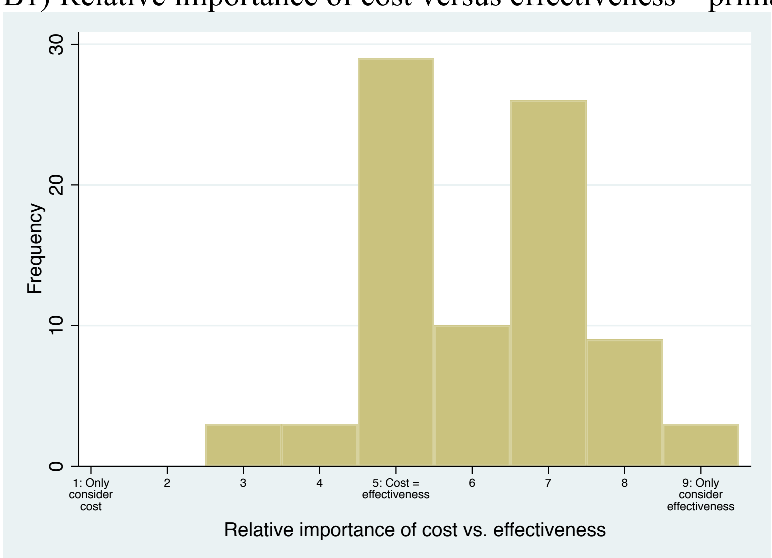
Appendix 8: Histograms of continuous/count covariates - survey

A) Year finished training



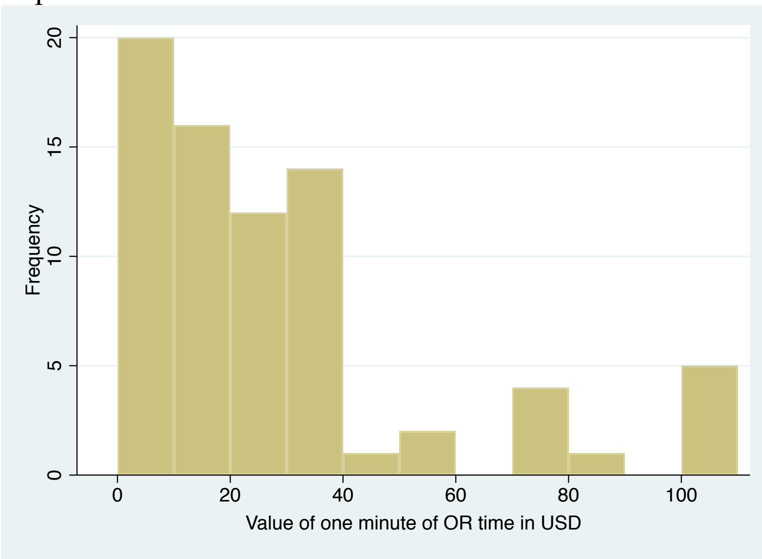
Note: N=83

B1) Relative importance of cost versus effectiveness – primary measure



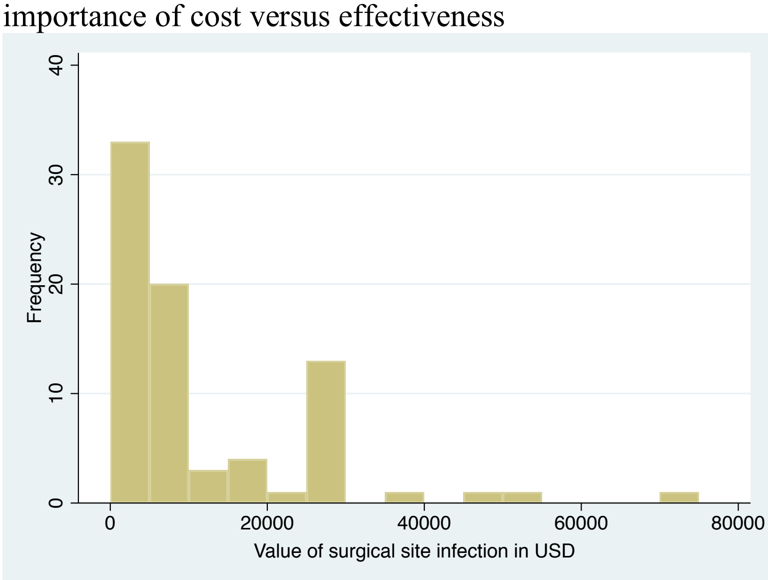
Note: N=83

B2) Value surgeon assigned to one minute of OR time – alternative measure of relative importance of cost versus effectiveness



Note: Histogram generated using complete data (N=75). Raw answers were converted to one minute of OR time by subtracting $250 and dividing by 7

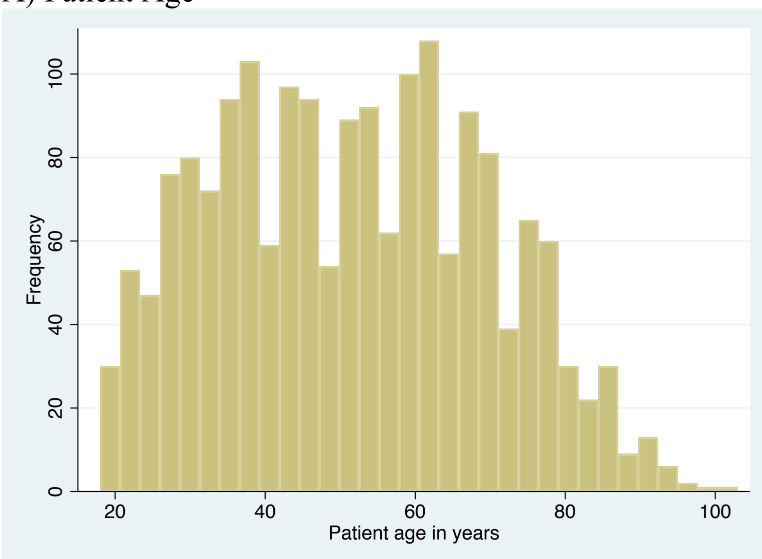
B3) Value surgeon assigned to surgical site infection – alternative measure of relative importance of cost versus effectiveness



Note: Histogram generated using complete data (N=78). Raw answers were converted to the value of one surgical site infection by subtracting $250 and multiplying by 100.

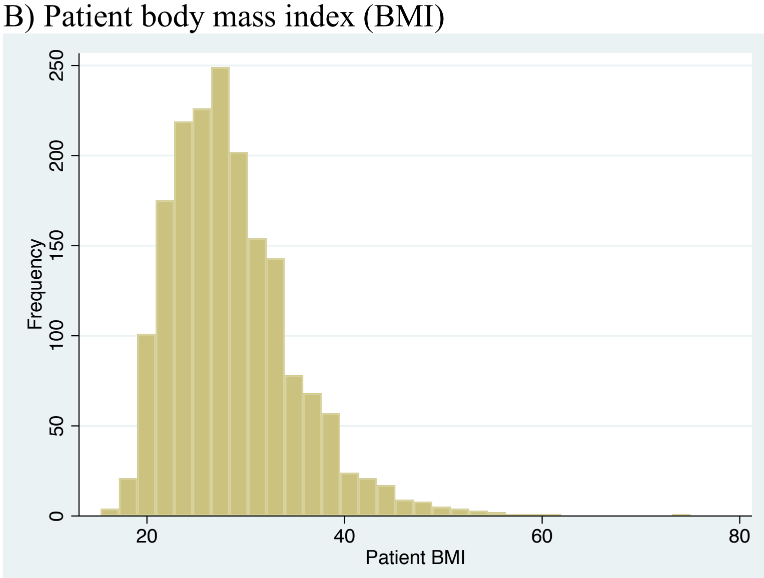
Appendix 9: Histograms of continuous/count covariates – medical record data

A) Patient Age



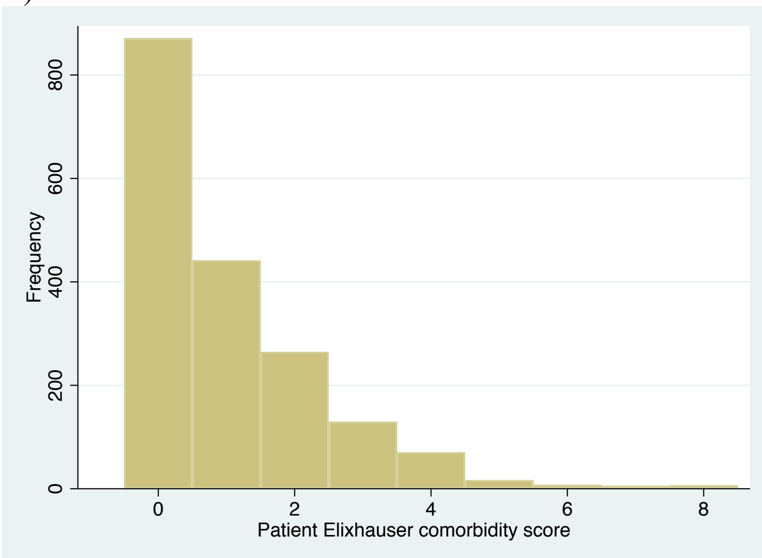
Note: N=1,817

B) Patient body mass index (BMI)



Note: N=1,794

C) Elixhauser Score



Note: N=1,817

Bibliography

1. Centers for Medicare & Medicaid Services. 2016. NHE Summary including share of GDP, CY 1960-2015 [Data File]. https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical.html. Accessed August 7th, 2017.

2. Munoz E, Munoz W, 3rd, Wise L. National and surgical health care expenditures, 2005-2025. *Annals of surgery.* 2010;251(2):195-200.

3. Childers CP, Maggard-Gibbons M. Understanding Costs of Care in the Operating Room. *JAMA surgery.* 2018;153(4):e176233.

4. Stey AM, Brook RH, Needleman J, et al. Hospital costs by cost center of inpatient hospitalization for medicare patients undergoing major abdominal surgery. *Journal of the American College of Surgeons.* 2015;220(2):207-217 e211.

5. Souhrada L. OR and materials: the yin and yang. *Mater Manag Health Care.* 1999;8(3):18-20.

6. Nathan H, Dimick JB. Medicare's Shift to Mandatory Alternative Payment Models: Why Surgeons Should Care. *JAMA surgery.* 2017;152(2):125-126.

7. Squires D, Anderson C. U.S. Health Care from a Global Perspective: Spending, Use of Services, Prices, and Health in 13 Countries, The Commonwealth Fund, October 2015.

8. Moore FD. Surgical streams in the flow of health care financing. The role of surgery in national expenditures: what costs are controllable? *Annals of surgery.* 1985;201(2):132-141.

9. Ginsburg P HM, Adler L, Burke S, Hoagland G W, Jennings C, and Lieberman S. *What is Driving U.S. Health Care Spending? America's Unsustainable Health Care Cost Growth.* Bipartisan Policy Center's Health Care Cost Containment Series September 2012 2012.

10. Cutler DM, McClellan M. Is Technological Change In Medicine Worth It? *Health Affairs.* 2001;20(5):11-29.

11. Skinner JS, Staiger DO, Fisher ES. Is technological change in medicine always worth it? The case of acute myocardial infarction. *Health Aff (Millwood).* 2006;25(2):w34-47.

12. Dieleman JL, Squires E, Bui AL, et al. Factors Associated With Increases in US Health Care Spending, 1996-2013. *JAMA.* 2017;318(17):1668-1678.

13. Steiner CA, Karaca Z, Moore BJ, Imshaug MC, Pickens G. Surgeries in Hospital-Based Ambulatory Surgery and Hospital Inpatient Settings, 2014. HCUP Statistical Brief #223. May 2017. Agency for Healthcare Research and Quality, Rockville, MD.

14. Sabiston DC, Townsend CM. *Sabiston textbook of surgery : the biological basis of modern surgical practice.* 19th ed. Philadelphia, PA: Elsevier Saunders; 2012.

15. Stinton LM, Myers RP, Shaffer EA. Epidemiology of gallstones. *Gastroenterol Clin North Am.* 2010;39(2):157-169, vii.

16. Zakko SF. Uncomplicated gallstone disease in adults. In: Post TW, ed. *UpToDate.* Waltham, MA: UpToDate; 2016.

17. McSherry CK. Cholecystectomy: the gold standard. *American journal of surgery.* 1989;158(3):174-178.

18. National Institutes of Health Consensus Development Conference Statement on Gallstones and Laparoscopic Cholecystectomy. *American journal of surgery.* 1993;165(4):390-398.

19. Fendrick AM, Escarce JJ, McLane C, Shea JA, Schwartz JS. Hospital adoption of laparoscopic cholecystectomy. *Med Care.* 1994;32(10):1058-1063.

20. Steiner CA, Bass EB, Talamini MA, Pitt HA, Steinberg EP. Surgical rates and operative mortality for open and laparoscopic cholecystectomy in Maryland. *N Engl J Med.* 1994;330(6):403-408.

21. Escarce JJ, Chen W, Schwartz JS. Falling cholecystectomy thresholds since the introduction of laparoscopic cholecystectomy. *JAMA.* 1995;273(20):1581-1585.

22. Huang Y, Chua TC, Maddern GJ, Samra JS. Robotic cholecystectomy versus conventional laparoscopic cholecystectomy: A meta-analysis. *Surgery.* 2017;161(3):628-636.

23. Omar MA, Redwan AA, Mahmoud AG. Single-incision versus 3-port laparoscopic cholecystectomy in symptomatic gallstones: A prospective randomized study. *Surgery.* 2017;162(1):96-103.

24. Bernhardt J, Sasse S, Ludwig K, Meier PN. Update in Natural Orifice Translumenal Endoscopic Surgery (NOTES). *Curr Opin Gastroenterol.* 2017;33(5):346-351.

25. Johnson & Johnson. 2016 Annual Report. https://jnj.brightspotcdn.com/88/3f/b666368546bcab9fd520594a6016/2017-0310-ar-bookmarked.pdf. Accessed August 7th, 2017.

26. Ethicon. Wound Closure Overview. http://www.ethicon.com/healthcare-professionals/products/wound-closure. Accessed August 7th, 2017.

27. US Food and Drug Administration. Overview of Device Regulation. https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/Overview/default.htm#mdr. Accessed August 7th, 2017.

28. Shapiro JK. Substantial Equivalence Premarket Review: the Right Approach for Most Medical Devices. *Food Drug Law J.* 2014;69(3):365-397, i.

29. Ibrahim AM, Dimick JB. Monitoring Medical Devices: Missed Warning Signs Within Existing Data. *JAMA.* 2017;318(4):327-328.

30. Institute of Medicine (U.S.). Committee on the Public Health Effectiveness of the FDA 510(k) Clearance Process. *Report Brief: Medical devices and the public's health : the FDA 510(k) clearance process at 35 years.* Washington, D.C.: National Academies Press; 2011.

31. Pauly MV, Burns LR. Price transparency for medical devices. *Health Aff (Millwood).* 2008;27(6):1544-1553.

32. Sanders GD, Neumann PJ, Basu A, et al. Recommendations for Conduct, Methodological Practices, and Reporting of Cost-effectiveness Analyses: Second Panel on Cost-Effectiveness in Health and Medicine. *JAMA.* 2016;316(10):1093-1103.

33. Najjar PA, Strickland M, Kaplan RS. Time-Driven Activity-Based Costing for Surgical Episodes. *JAMA surgery.* 2017;152(1):96-97.

34. Kaplan RS, Porter ME. How to solve the cost crisis in health care. *Harv Bus Rev.* 2011;89(9):46-52, 54, 56-61 passim.

35. DiGioia AM, 3rd, Greenhouse PK, Giarrusso ML, Kress JM. Determining the True Cost to Deliver Total Hip and Knee Arthroplasty Over the Full Cycle of Care: Preparing for Bundling and Reference-Based Pricing. *J Arthroplasty.* 2016;31(1):1-6.

36. McLaughlin N, Burke MA, Setlur NP, et al. Time-driven activity-based costing: a driver for provider engagement in costing activities and redesign initiatives. *Neurosurg Focus.* 2014;37(5):E3.

37. Farrokhi FR, Gunther M, Williams B, Blackmore CC. Application of Lean Methodology for Improved Quality and Efficiency in Operating Room Instrument Availability. *Journal for healthcare quality : official publication of the National Association for Healthcare Quality.* 2015;37(5):277-286.

38. Holloran-Schwartz MB, Gavard JA, Martin JC, Blaskiewicz RJ, Yeung PP, Jr. Single-Use Energy Sources and Operating Room Time for Laparoscopic Hysterectomy: A Randomized Controlled Trial. *J Minim Invasive Gynecol.* 2016;23(1):72-77.

39. Hsu AR, Gross CE, Bhatia S, Levine BR. Template-directed instrumentation in total knee arthroplasty: cost savings analysis. *Orthopedics.* 2012;35(11):e1596-1600.

40. Seligson D, Douglas L. Experience with a large-frame, disposable external fixator. *Orthopedics.* 2010;33(3).

41. Siegel GW, Patel NN, Milshteyn MA, Buzas D, Lombardo DJ, Morawa LG. Cost Analysis and Surgical Site Infection Rates in Total Knee Arthroplasty Comparing Traditional vs. Single-Use Instrumentation. *The Journal of arthroplasty.* 2015;30(12):2271-2274.

42. Wannemuehler TJ, Elghouche AN, Kokoska MS, Deig CR, Matt BH. Impact of Lean on surgical instrument reduction: Less is more. *The Laryngoscope.* 2015;125(12):2810-2815.

43. Yung E, Gagner M, Pomp A, Dakin G, Milone L, Strain G. Cost comparison of reusable and single-use ultrasonic shears for laparoscopic bariatric surgery. *Obesity surgery.* 2010;20(4):512-518.

44. Bade K, Hoogerbrug J. Awareness of surgical costs: a multicenter cross-sectional survey. *J Surg Educ.* 2015;72(1):23-27.

45. Jackson CR, Eavey RD, Francis DO. Surgeon Awareness of Operating Room Supply Costs. *Ann Otol Rhinol Laryngol.* 2016;125(5):369-377.

46. Okike K, O'Toole RV, Pollak AN, et al. Survey finds few orthopedic surgeons know the costs of the devices they implant. *Health Aff (Millwood).* 2014;33(1):103-109.

47. Parnes JA, Yeh DH, Glicksman JT, Tewfik MA, Sowerby LJ. Surgeon-estimated costs of common consumables in otolaryngology. *The Laryngoscope.* 2015;125(12):2690-2694.

48. Allan GM, Lexchin J. Physician awareness of diagnostic and nondrug therapeutic costs: a systematic review. *Int J Technol Assess Health Care.* 2008;24(2):158-165.

49. Austin LS, Tjoumakaris FP, Ong AC, Lombardi NJ, Wowkanech CD, Mehnert MJ. Surgical Cost Disclosure May Reduce Operating Room Expenditures. *Orthopedics.* 2017;40(2):e269-e274.

50. Avansino JR, Goldin AB, Risley R, Waldhausen JH, Sawin RS. Standardization of operative equipment reduces cost. *Journal of pediatric surgery.* 2013;48(9):1843-1849.

51. Croft K, Mattingly PJ, Bosse P, Naumann RW. Physician Education on Controllable Costs Significantly Reduces Cost of Laparoscopic Hysterectomy. *J Minim Invasive Gynecol.* 2017;24(1):62-66.

52. Gitelis M, Vigneswaran Y, Ujiki MB, et al. Educating surgeons on intraoperative disposable supply costs during laparoscopic cholecystectomy: a regional health system's experience. *American journal of surgery.* 2015;209(3):488-492.

53. Guzman MJ, Gitelis ME, Linn JG, et al. A Model of Cost Reduction and Standardization: Improved Cost Savings While Maintaining the Quality of Care. *Diseases of the colon and rectum.* 2015;58(11):1104-1107.

54. Skarda DE, Rollins M, Andrews S, et al. One hospital, one appendectomy: The cost effectiveness of a standardized doctor's preference card. *Journal of pediatric surgery.* 2015;50(6):919-922.

55. Vigneswaran Y, Linn JG, Gitelis M, et al. Educating surgeons may allow for reduced intraoperative costs for inguinal herniorrhaphy. *Journal of the American College of Surgeons.* 2015;220(6):1107-1112.

56. Zygourakis CC, Valencia V, Moriates C, et al. Association Between Surgeon Scorecard Use and Operating Room Costs. *JAMA surgery.* 2017;152(3):284-291.

57. Adkins HH, Hardacker TJ, Ceppa EP. Examining variation in cost based on surgeon choices for elective laparoscopic cholecystectomy. *Surgical endoscopy.* 2016;30(7):2679-2684.

58. Benoit RM, Cohen JK. The relationship between quality and costs: factors that affect the hospital costs of radical prostatectomy. *Prostate cancer and prostatic diseases.* 2001;4(4):213-216.

59. Brauer DG, Hawkins WG, Strasberg SM, et al. Cost variation in a laparoscopic cholecystectomy and the association with outcomes across a single health system: implications for standardization and improved resource utilization. *HPB : the official journal of the International Hepato Pancreato Biliary Association.* 2015;17(12):1113-1118.

60. Chung WC, Fan PL, Chiu HC, Yang CY, Huang KL, Tzeng DS. Operating room cost for coronary artery bypass graft procedures: does experience or severity of illness matter? *Journal of evaluation in clinical practice.* 2010;16(6):1063-1070.

61. Kazberouk A, Sagy I, Novack V, McGuire K. Understanding the Extent and Drivers of Interphysician Cost Variation for Spine Procedures. *Spine.* 2016;41(13):1111-1117.

62. Meier JD, Duval M, Wilkes J, et al. Surgeon dependent variation in adenotonsillectomy costs in children. *Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery.* 2014;150(5):887-892.

63. Rosenbaum BP, Modic MT, Krishnaney AA. Value in Single-level Lumbar Discectomy: Surgical Disposable Item Cost and Relationship to Patient-reported Outcomes. *Clin Spine Surg.* 2017;30(9):E1227-E1232.

64. Sjogren PP, Gale C, Henrichsen J, et al. Variation in costs among surgeons and hospitals in Pediatric tympanostomy tube placement. *The Laryngoscope.* 2016;126(8):1935-1939.

65. Thomas A, Alt J, Gale C, et al. Surgeon and hospital cost variability for septoplasty and inferior turbinate reduction. *International forum of allergy & rhinology.* 2016;6(10):1069-1074.

66. Zygourakis CC, Valencia V, Boscardin C, et al. Predictors of Variation in Neurosurgical Supply Costs and Outcomes Across 4904 Surgeries at a Single Institution. *World neurosurgery.* 2016;96:177-183.

67. Ting NT, Moric MM, Della Valle CJ, Levine BR. Use of knotless suture for closure of total hip and knee arthroplasties: a prospective, randomized clinical trial. *The Journal of arthroplasty.* 2012;27(10):1783-1788.

68. Eggers MD, Fang L, Lionberger DR. A comparison of wound closure techniques for total knee arthroplasty. *The Journal of arthroplasty.* 2011;26(8):1251-1258 e1251-1254.

69. Lukish J, Powell D, Morrow S, Cruess D, Guzzetta P. Laparoscopic appendectomy in children: use of the endoloop vs the endostapler. *Archives of surgery (Chicago, Ill : 1960).* 2007;142(1):58-61; discussion 62.

70. Brown JK, Campbell BT, Drongowski RA, et al. A prospective, randomized comparison of skin adhesive and subcuticular suture for closure of pediatric hernia incisions: cost and cosmetic considerations. *Journal of pediatric surgery.* 2009;44(7):1418-1422.

71. Tillmanns TD, Mabe A, Ulm MA, Lee D, Lowe P, Kumar S. Vaginal Cuff Closure in Robotic Hysterectomy: A Randomized Controlled Trial Comparing Barbed Versus Standard Suture. *Journal of Gynecologic Surgery.* 2016;32(4):215-219.

72. Cohen DJ, Reynolds MR. Interpreting the results of cost-effectiveness studies. *J Am Coll Cardiol.* 2008;52(25):2119-2126.

73. Dillman DA, Smyth JD, Christian LM, Dillman DA. *Internet, mail, and mixed-mode surveys : the tailored design method.* 3rd ed. Hoboken, N.J.: Wiley & Sons; 2009.

74. The American Association for Public Opinion Research. 2016. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 9th edition. AAPOR.

75. Hofer IS, Gabel E, Pfeffer M, Mahbouba M, Mahajan A. A Systematic Approach to Creation of a Perioperative Data Warehouse. *Anesth Analg.* 2016;122(6):1880-1884.

76. Childers CP, Maggard-Gibbons M. Estimation of the Acquisition and Operating Costs for Robotic Surgery. *JAMA.* 2018;320(8):835-836.

77. Giger UF, Michel JM, Opitz I, et al. Risk factors for perioperative complications in patients undergoing laparoscopic cholecystectomy: analysis of 22,953 consecutive cases from the Swiss Association of Laparoscopic and Thoracoscopic Surgery database. *Journal of the American College of Surgeons.* 2006;203(5):723-728.

78. Birkmeyer JD, Stukel TA, Siewers AE, Goodney PP, Wennberg DE, Lucas FL. Surgeon volume and operative mortality in the United States. *N Engl J Med.* 2003;349(22):2117-2127.

79. Childers CP, Hofer IS, Cheng DS, Maggard-Gibbons M. Evaluating Surgeons on Intraoperative Disposable Supply Costs: Details Matter. *J Gastrointest Surg.* 2018.

80. Daabiss M. American Society of Anaesthesiologists physical status classification. *Indian J Anaesth.* 2011;55(2):111-115.

81. Moore BJ, White S, Washington R, Coenen N, Elixhauser A. Identifying Increased Risk of Readmission and In-hospital Mortality Using Hospital Administrative Data: The AHRQ Elixhauser Comorbidity Index. *Med Care.* 2017;55(7):698-705.

82. Rabe-Hesketh S, Skrondal A. *Multilevel and longitudinal modeling using Stata.* 3rd ed. College Station, Tex.: Stata Press Publication; 2012.

83. Merlo J, Yang M, Chaix B, Lynch J, Rastam L. A brief conceptual tutorial on multilevel analysis in social epidemiology: investigating contextual phenomena in different groups of people. *J Epidemiol Community Health.* 2005;59(9):729-736.

84. Aquina CT, Becerra AZ, Justiniano CF, et al. Surgeon, Hospital, and Geographic Variation in Minimally Invasive Colectomy. *Annals of surgery.* 2018.

85. Childers CP, Showen A, Nuckols T, Maggard-Gibbons M. Interventions to Reduce Intraoperative Costs: A Systematic Review. *Annals of surgery.* 2018;268(1):48-57.

86. Navathe AS, Liao JM, Dykstra SE, et al. Association of Hospital Participation in a Medicare Bundled Payment Program With Volume and Case Mix of Lower Extremity Joint Replacement Episodes. *JAMA.* 2018;320(9):901-910.

87. Finkelstein A, Ji Y, Mahoney N, Skinner J. Mandatory Medicare Bundled Payment Program for Lower Extremity Joint Replacement and Discharge to Institutional Postacute Care: Interim Analysis of the First Year of a 5-Year Randomized Trial. *JAMA.* 2018;320(9):892-900.