Title
Variation in Intraoperative and Postoperative Utilization for 3 Common General Surgery Procedures.

Permalink
https://escholarship.org/uc/item/9mn6t3hc

Authors
Childers, Christopher P
Ettner, Susan L
Hays, Ron D
et al.

Publication Date
2019-08-23

DOI
10.1097/sla.0000000000003571

Peer reviewed
Variation in Intraoperative and Postoperative Utilization for 3 Common General Surgery Procedures

Christopher P. Childers, MD, PhD,* Susan L. Ettnner, PhD,† Ron D. Hays, PhD,‡ Gerald Kominski, PhD,¶ §§ Melinda Maggard-Gibbons, MD, MSHS,* and Rodrigo F. Alban, MD¶

Objective: The aim of this study was to understand variation in intraoperative and postoperative utilization for common general surgery procedures.

Summary Background Data: Reducing surgical costs is paramount to the viability of hospitals.

Methods: Retrospective analysis of electronic health record data for 7762 operations from 2 health systems. Adult patients undergoing laparoscopic cholecystectomy, appendectomy, and inguinal/femoral hernia repair between November 1, 2013 and November 30, 2017 were reviewed for 3 utilization measures: intraoperative disposable supply costs, procedure time, and postoperative length of stay (LOS). Crossed hierarchical regression models were fit to understand case-mixed adjusted variation in utilization across surgeons and locations and to rank surgeons.

Results: The number of surgeons performing each type of operation ranged from 20 to 63. The variation explained by surgeons ranged from 8.9% to 38.2% for supply costs, from 15.1% to 54.6% for procedure time, and from 1.3% to 7.0% for postoperative LOS. The variation explained by location ranged from 12.1% to 26.3% for supply costs, from 0.2% to 2.5% for procedure time, and from 0.0% to 31.8% for postoperative LOS. There was a positive correlation (ρ = 0.49, P = 0.03) between surgeons’ higher supply costs and longer procedure times for hernia repair, but there was no correlation between other utilization measures for hernia repair and no correlation between any of the utilization measures for laparoscopic appendectomy or cholecystectomy.

Conclusions: Surgeons are significant drivers of variation in surgical supply costs and procedure time, but much less so for postoperative LOS. Intraoperative and postoperative utilization profiles can be generated for individual surgeons and may be an important tool for reducing surgical costs.

Keywords: cost, economics, operating room, resource management, surgery, utilization

Anns Surg 2019;xx:xxx–xxx)

From the *Department of Surgery, David Geffen School of Medicine at UCLA, Los Angeles, CA; †Division of General Internal Medicine and Health Services Research, Department of Medicine, David Geffen School of Medicine at UCLA, Los Angeles, CA; ‡Department of Health Policy & Management, UCLA Fielding School of Public Health, Los Angeles, CA; §UCCLA Center for Health Policy Research, Fielding School of Public Health, Los Angeles, California; and ¶Department of Surgery, Cedars Sinai Medical Center, Los Angeles, CA.

©chillders@mednet.ucla.edu.
This manuscript will be presented at the American College of Surgeons 105th Annual Clinical Congress, HSR 1, Scientific Forum, San Francisco, CA, October 2019.

Christopher Childers was funded by AHRQ#F32HS025079. RDH received support from the University of California, Los Angeles (UCLA) and Charles R. Drew University (CDU), Resource Centers for Minority Aging Research Center for Health Improvement of Minority Elderly (RCMAR/CHIME) under NIH/NIA Grant P30-AG021684. The authors report no conflicts of interest.

Supplemental digital content is available for this article. Direct URL citations appear in the HTML and PDF versions of this article on the journal’s Web site (www.annalsofsurgery.com).

Copyright © 2019 Wolters Kluwer Health, Inc. All rights reserved.

ISSN: 0003-4922/16/XXXX-0001
DOI: 10.1097/SLA.0000000000003571

METHODS

Data Source, Sample, and Ethics Review

This retrospective review utilized data available in the electronic health records (EHRs) of 2 distinct academic health systems. Both systems use the Epic EHR. The bioinformatics infrastructure that allowed patient-level queries has previously been described for one of the health systems.14 We identified adult (aged 18 yrs or older) patients undergoing laparoscopic cholecystectomy, laparoscopic appendectomy, and laparoscopic inguinal/femoral hernia repair between November 1, 2013 and November 30, 2017. Using current procedural terminology (CPT) codes and booking slip information, we excluded cases where an additional procedure was coded except...
those that represented a coding phenomenon (ie, when diagnostic laparoscopy was always coded in addition to laparoscopic cholecystectomy) or, for laparoscopic cholecystectomy, if a concurrent intraoperative cholangiogram (IOC) was coded. IOC was routinely performed at one institution but not the other and excluding these cases would artificially reduce the variation identified at the location level. Robotic cases were also excluded as relevant costs (ie, acquisition, maintenance, and instrument costs) were not sufficiently itemized.\textsuperscript{15} Finally, we limited our analysis to surgeons who performed the procedure \textgeq 5 times during the study time period. Institutional review board approval was obtained at each health system before starting the research including a waiver of informed consent for patient chart review (IRB#16-001327, Pro00047831).

\textbf{Utilization Measures}

Following the terminology outlined by the Second Panel on Cost Effectiveness in Health and Medicine, the financial perspective of this analysis was that of the healthcare sector—and specifically that of the hospital.\textsuperscript{16} This analysis focused on the actual cost of a hospital delivering a service, an amount that is distinct from that paid by a third party payor (ie, patient, insurer).

Outcomes of interest included the hospital cost of intraoperative disposable supplies, procedure time, and postoperative LOS. A detailed review of assigning and analyzing intraoperative disposable supply costs was reported elsewhere.\textsuperscript{17} In brief, this outcome captured the cost of disposable supplies and implants used or wasted in the OR. This does not include reusable instruments or capital costs. Within each health system, a single price was assigned to each item over the study time period to account for price changes that occur when supplies are reordered or contracts are renegotiated. This adjustment prevents a surgeon for being penalized if they happen to operate the day after a price increase. Price information was not routinely provided to surgeons in this study. Procedure time is measured in minutes and represents skin-to-skin time. This is distinct from room time that may be influenced by factors outside of the surgeons’ control, such as postanesthesia care unit holds. Postoperative LOS was measured in hours from when the patient left the OR to when the nurse closed the patient encounter (ie, when the patient left the room/facility). This is different from when the discharge order was signed. Extreme LOS outliers (>10 d for cholecystectomy and appendectomy, >2 d for hernia repair) were excluded because of their leverage exerted on regression coefficients and also because they were suspected to represent cases that substantially deviate from the norm which may unfairly penalize a surgeon.

\textbf{Covariates}

Covariates were included from 3 levels—location, surgeon, and patient. The data set included operations performed at 6 locations within the 2 health systems—4 locations at UCLA (a main OR and an ambulatory surgery center (ASC) at each of the Ronald Reagan and Santa Monica sites) and 2 (a main OR and an ASC) at Cedars Sinai Medical Center (CSMC). We have previously demonstrated that supply costs vary between facilities and settings even within the same health system.\textsuperscript{19}

Patient-level case mix variables included the following: patient age, body mass index (BMI), the American Association of Anesthesiologists physical status classification score (ASA), sex (male, female), race/ethnicity (Hispanic, non-Hispanic white, non-Hispanic black, non-Hispanic other), and Elixhauser comorbidity index. Race/ethnicity was included because of differences across the various locations as well as potential effects on outcomes because of differences in pathology (eg, gallbladder disease is more common among those of Hispanic ethnicity) and as a potential proxy for socioeconomic status. The elixhauser comorbidity index was generated using ICD-10 codes abstracted from the patient encounter and the STATA command Elixhauser.\textsuperscript{18} For laparoscopic cholecystectomy, we further included a variable for indication [biliary colic/ elective indications, biliary obstruction (ie, cholecloolithiasis, pancreatitis), and acute cholecystitis] using ICD-10 codes.

\textbf{Analysis}

All statistical analyses were performed in STATA v15.1 using 2-sided tests and an alpha of 0.05. Multiple mixed-effects regression models were fit for each procedure and each utilization measure. All models included covariates as described above. Nonlinear relationships and collinearity were addressed through a combination of quadratic terms and de-meaning (or centering). To address our aims, models varied in their hierarchical structure as described below.

To estimate how much variation of each outcome was explained by the location and surgeon, we fit crossed random-effects models. Crossed refers to combinations existing across levels (locations, surgeons) instead of in hierarchies. An additive crossed-effects model and a crossed-effects model with random interaction were run to assess the need for a random intercept for each combination of surgeon and location. Intraclass correlations were calculated by dividing the variance of the random intercept at a given level (location, surgeon) by the sum of all variances across levels including the residual error. To provide context to the magnitude of the intraclass correlations, we also estimated the amount of variation explained by patient-level covariates. We compared a null model that included random effects for the surgeon and location only to a model that included random effects for the surgeon and location along with all patient-level covariates.

After calculating the proportion of variance at each level, we then focused on understanding variation at the surgeon level. We repeated our full model using location fixed effects instead of random effects. This simplified the model while also controlling for unmeasured facility-level factors (eg, institutional policies, purchasing contracts, nurse staffing) that may influence each outcome. Following estimation, best linear unbiased predictions of the surgeon effects and standard errors were calculated. Estimates and 95% confidence intervals that were entirely above or below 0 represented “outlier” surgeons with risk-adjusted averages higher or lower than the overall average for that outcome (eg, LOS).

Most variables had complete or near-complete (<2% missing) data; the only exception was BMI for appendectomy, which was missing in 5.5% of cases. Given the low rates of missingness, all analyses were therefore done on a complete case basis.

\textbf{RESULTS}

\textbf{Sample Characteristics}

The sample included 7762 operations, including 4089 laparoscopic cholecystectomies, 2489 laparoscopic appendectomies, and 1184 laparoscopic hernia repairs (Table 1). The number of surgeons performing each operation ranged from 20 for laparoscopic hernia repair to 63 for laparoscopic cholecystectomy.

Patient demographics and case mix factors varied across procedures as expected. For example, patients undergoing appendectomy were typically younger than those undergoing cholecystectomy and most hernia repairs were performed in males, whereas cholecystectomies were predominantly performed in females. For cholecystectomy, 50.5% (2043/4042, 47 missing) of operations were for elective indications, 39.2% (1583/4042) were for acute cholecystitis, and 10.3% (416/4042) were for biliary obstruction (ie, pancreatitis, cholecloolithiasis).
TABLE 1. Descriptive Data

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Appendectomy</th>
<th>Cholecystectomy</th>
<th>Hernia Repair</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Size</td>
<td>Patients</td>
<td>Surgeons</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2489</td>
<td>50 (27–50)</td>
<td>4089</td>
</tr>
<tr>
<td></td>
<td></td>
<td>63 (37–67)</td>
<td>1184</td>
</tr>
<tr>
<td>Age, yrs</td>
<td>36</td>
<td>52 (21.1–28.2)</td>
<td>27.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(24.1–32.1)</td>
<td>24.6 (22.8–27.0)</td>
</tr>
<tr>
<td>ASA score</td>
<td>1023</td>
<td>562 (41.2%)</td>
<td>389 (13.8%)</td>
</tr>
<tr>
<td></td>
<td>(41.2%)</td>
<td>138 (55.2%)</td>
<td>389 (32.9%)</td>
</tr>
<tr>
<td></td>
<td>1214 (48.8%)</td>
<td>3255 (9.7%)</td>
<td>641 (28.5%)</td>
</tr>
<tr>
<td></td>
<td>2355 (28.5%)</td>
<td>154 (54.1%)</td>
<td>13 (13.0%)</td>
</tr>
<tr>
<td></td>
<td>240 (9.7%)</td>
<td>116 (2.2%)</td>
<td>15 (0.0%)</td>
</tr>
<tr>
<td>Elixhauser score</td>
<td>9 (0.4%)</td>
<td>100 (0.4%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td>1199 (48.2%)</td>
<td>389 (10.1%)</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td>1520 (61.4%)</td>
<td>2067 (59.9%)</td>
<td>909 (77.7%)</td>
</tr>
<tr>
<td></td>
<td>Non-Hispanic</td>
<td>1675 (67.3%)</td>
<td>782 (66.0%)</td>
</tr>
<tr>
<td></td>
<td>white</td>
<td>1272 (42.4%)</td>
<td>80 (6.8%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2407 (16.3%)</td>
<td>26 (2.2%)</td>
</tr>
<tr>
<td>Procedure minutes</td>
<td></td>
<td>23 (8.5–46.1)</td>
<td>26 (2.5–12.6)</td>
</tr>
<tr>
<td>Postoperative LOS, h</td>
<td>20.4 (14.0–33.2)</td>
<td>23.9 (8.5–46.1)</td>
<td>4.3 (2.5–12.6)</td>
</tr>
</tbody>
</table>

ASA, American Society of Anesthesiologists Physical Status; BMI, body mass index; LOS, length of stay.

All continuous data are presented as median (IQR).

<1% of data was missing for any variable except for 5.5% of operations were missing BMI for appendectomy; the few missing data points explain why variable samples do not always add to the total sample.

Variation in Utilization Measures Explained by Surgeon and Location

The variation of each utilization measure explained by the surgeon and location are illustrated in Figure 1. Full regression specifications are provided in Appendix 1, http://links.lww.com/SLA/B763. The variation explained by surgeon ranged from 0.49% to 31.8% for postoperative LOS.

Variation in Utilization Measures Explained by Location

The variation explained by location ranged from 12.1% to 26.3% for supply costs, from 0.2% to 7.0% for postoperative LOS. The variation explained by location ranged from 12.1% to 26.3% for supply costs, from 0.2% to 7.0% for postoperative LOS.

Variation in Utilization Measures Explained by Procedure

The variation explained by procedure ranged from 26.3% to 54.6% for procedure time, and from 1.3% to 7.0% for postoperative LOS.

Surgeon Profiles and Association Between Utilization Measures

Surgeon profiles for each procedure and each utilization measure are included in Figure 2 (appendectomy) and Appendices 3 and 4, http://links.lww.com/SLA/B763. The variation explained by surgeon ranged from 0.49% to 31.8% for postoperative LOS. The variation explained by location ranged from 12.1% to 26.3% for supply costs, and from 0.2% to 5.3% for procedure time, and from 0.0% to 31.8% for postoperative LOS.

DISCUSSION

In this multihealth system evaluation of cost drivers for 3 common general surgery operations, surgeons accounted for a significant portion of variation in OR supply costs and procedure time, with a much smaller influence on postoperative LOS. Surgeons with outlier utilization for one component (eg, supply costs) were generally not outliers on other components, suggesting that increased utilization in one cost component does not “make up” for other components.

This analysis provides a number of insights for understanding ways to control surgical costs. First, location matters when evaluating supply costs. Not just the health system or the facility within a health system, but also the setting (ie, main OR vs ASC) in which a procedure is performed is associated with costs. Differences in supply costs between facilities and settings have been shown even when the same distributors and purchasing department are involved. For example, one location may use bundled packs, whereas the other locations require supplies to be assembled individually, preference cards can vary—even for the same surgeon—across settings, and certain items may be stocked (or simply more conveniently located) in one setting but not the other. Analyses that evaluate supply costs must stratify by facility and setting to create a complete picture.

Second, the surgeon “effect” on procedure time is as strong as, if not stronger than, the effect of surgeons on supply costs. The total cost of running an OR as of 2014 was $37/min. The marginal cost—the value of adding or removing a minute of OR time—is likely smaller, with activity-based costing estimates around $10/
As a result, surgeons with procedure times within a couple minutes of average will likely have a negligible effect on costs, but some of the outliers in this study were more substantial. At the extremes, for appendectomy, one surgeon had risk-adjusted procedure times 23 minutes shorter than average (n = 108 operations), and another had procedure times 21 minutes longer than average (n = 49 operations). Bringing this second surgeon to the average would have cumulatively saved 1029 minutes in OR time. This difference may have a real impact on hospital finances—not just the added labor costs, but also potentially loss of revenue from operations that could have been performed (ie, opportunity costs). There are certainly justifiable reasons for high outliers—such as the surgeon’s learning curve or resident teaching—but these profiles can quickly identify surgeons that may require further investigation. Conversely, what

**FIGURE 1.** Variation in resource utilization explained by surgeon and location. LOS, length of stay; variation explained at each level (surgeon, location) was calculated as the variance of the random intercept for that level (surgeon, location) divided by the sum of variance for the surgeon, location, and residual, after patient case-mix adjustment.
strategies or techniques does the low outlier surgeon employ to achieve such short operative times?

Third, targeting surgeons to reduce LOS is likely not a high-yield strategy for cost containment in these 3 common, but relatively simple, procedures. There was certainly variation in LOS, but this variation was randomly distributed across surgeons after controlling for patient case mix. Fundamentally, this suggests a different driving force. Although supplies and procedure time are inherently under the surgeon’s control, postoperative stay may be more influenced by patient and disease characteristics than by a particular surgeon’s management strategy. Further evidence for this is the relatively larger fraction of variance explained by patient factors for postoperative LOS compared with the other outcomes.

Fourth, providing surgeons with a utilization profile across all 3 measures simultaneously may be a much more powerful tool to motivate change behavior than looking at one in isolation. A number of commercial software systems have come to market profiling surgeons on supply costs alone. Indeed, when we started this collaboration, surgical supplies were our focus, but surgeons quickly noted that supply costs may be inextricably linked to other utilization measures such as procedure time and LOS. This analysis shows that, in general, this is not true—being a high outlier in one domain does not portend a low outlier in another—and also provides a graphical way of illustrating this. Surgeons may be more amenable to change if they see that they are a high outlier of supply utilization and yet do not “compensate” for this utilization in other areas.

Fifth, and finally, using procedure time and LOS—rather than estimating a cost for these utilization measures—may be more accurate and approachable for clinicians. Commercial software systems can generate a total “cost” for a patient’s stay and can identify surgeons that have overall costs higher than others. However, the validity of these costs is only as good as the underlying accounting methodology. Many of these systems rely on cost-to-charge ratios.
or other blunt accounting methods that may produce inaccurate estimates for surgical patients. Procedure time and LOS are measures that are in units the clinician can easily appreciate, are valid, and can be weighed against each other based on finances that are specific to that institution and that procedure.

This study has a number of limitations. First, although this is the first study to evaluate intraoperative and postoperative utilization at more than one health system, we were still limited to 2 organizations in southern California and 3 different procedures. The variation at each level will undoubtedly vary by institution, specialty, and procedure. Second, the risk-adjustment variables available to us were limited, relying largely on ICD-10 codes. If a clinical registry, such as NSQIP, would add disposable supply costs, it is likely these utilization profiles could be generated across hospitals and surgeons using more robust risk-adjustment methods. Third, our analysis was limited to common laparoscopic operations with short postoperative stays that may have limited our ability to detect differences in this utilization measure across surgeons. Lastly, we only had information for the inpatient stay. Readmissions are an important driver of cost in many surgical procedures, but they are also quite rare and detecting differences across surgeons would require large surgeon-specific samples which may not be possible.

CONCLUSION

Surgeons are significant drivers of variation in surgical supplies and OR time, and much less so for postoperative LOS. Intraoperative and postoperative utilization profiles can be generated for individual surgeons and are important tools for those interested in reducing surgical costs.

REFERENCES


