Title
Duty Hour Reform and the Outcomes of Patients Treated by New Surgeons.

Journal
Annals of Surgery, 271(4)

ISSN
0003-4932

Authors
Kelz, Rachel R
Niknam, Bijan A
Sellers, Morgan M
et al.

Publication Date
2020-04-01

DOI
10.1097/sla.0000000000003304

Peer reviewed

HHS Public Access

Author manuscript

Ann Surg. Author manuscript; available in PMC 2021 April 01.

Published in final edited form as:

Ann Surg. 2020 April ; 271(4): 599–605. doi:10.1097/SLA.0000000000003304.

Duty Hour Reform and the Outcomes of Patients Treated by New Surgeons

Rachel R. Kelz, MD[*,†], Bijan A. Niknam, BS[‡], Morgan M. Sellers, MD[*], James E. Sharpe, MS[‡], Paul R. Rosenbaum, PhD[†,§], Alexander S. Hill, BS[‡], Hong Zhou, MS[‡], Lauren L. Hochman, BA[‡], Karl Y. Bilimoria, MD[¶], Kamal Itani, MD[‖,], Patrick S. Romano, MD[††], Jeffrey H. Silber, MD, PhD[†,‡,‡‡,§§,¶¶]**

[*]Department of Surgery, Center for Surgery and Health Economics, The University of Pennsylvania Perelman School of Medicine, Philadelphia, PA

[†]The Leonard Davis Institute of Health Economics, The University of Pennsylvania, Philadelphia, PA

[‡]Center for Outcomes Research, Children's Hospital of Philadelphia, Philadelphia, PA

[§]Department of Statistics, The Wharton School, The University of Pennsylvania, Philadelphia, PA

[¶]Department of Surgery and Center for Healthcare Studies, Northwestern University, Chicago IL

[‖]VA Boston Health Care System, Boston, MA

[**]Department of Surgery, Boston University School of Medicine, Boston, MA

[††]Division of General Medicine and Center for Healthcare Policy and Research, University of California Davis School of Medicine, Sacramento, CA

[‡‡]The Department of Pediatrics, The University of Pennsylvania Perelman School of Medicine, Philadelphia, PA

[§§]Department of Anesthesiology and Critical Care, The University of Pennsylvania Perelman School of Medicine, Philadelphia, PA

[¶¶]Department of Health Care Management, The Wharton School, The University of Pennsylvania, Philadelphia, PA.

Abstract

Objective: The aim of the study was to address the controversy surrounding the effects of duty hour reform on new surgeon performance, we analyzed patients treated by new surgeons following the transition to independent practice.

Summary Background Data: In 2003, duty hour reform affected all US surgical training programs. Its impact on the performance of new surgeons remains unstudied.

Reprints: Rachel R. Kelz, MD, MSCE, MBA, FACS, Department of Surgery, Center of Surgery and Health Economics, University of Pennsylvania, Perelman School of Medicine, 3400 Spruce St, 4 Silverstein, Philadelphia, PA 19104. rachel.kelz@uphs.upenn.edu.

The authors have no conflicts of interest to disclose.

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

Methods: We studied 30-day mortality among 1,483,074 Medicare beneficiaries undergoing general and orthopedic operations between 1999 and 2003 ("traditional" era) and 2009 and 2013 ("modern" era). The operations were performed by 2762 new surgeons trained before the reform, 2119 new surgeons trained following reform and 15,041 experienced surgeons. We used a difference-in-differences analysis comparing outcomes in matched patients treated by new versus experienced surgeons within each era, controlling for the hospital, operation, and patient risk factors.

Results: Traditional era odds of 30-day mortality among matched patients treated by new versus experienced surgeons were significantly elevated [odds ratio (OR) 1.13; 95% confidence interval (CI) (1.05, 1.22), $P < 0.001$). The modern era elevated odds of mortality were not significant [OR 1.06; 95% CI (0.97–1.16), $P = 0.239$]. Relative performance of new and experienced surgeons with respect to 30-day mortality did not appear to change from the traditional era to the modern era [OR 0.93; 95% CI (0.83–1.05), $P = 0.233$]. There were statistically significant adverse changes over time in relative performance to experienced surgeons in prolonged length of stay [OR 1.08; 95% CI (1.02–1.15), $P = 0.015$], anesthesia time [9 min; 95% CI (8–10), $P < 0.001$], and costs [255USD; 95% CI (2–508), $P = 0.049$].

Conclusions: Duty hour reform showed no significant effect on 30-day mortality achieved by new surgeons compared to their more experienced colleagues. Patients of new surgeons, however, trained after duty hour reform displayed some increases in the resources needed for their care.

Keywords

duty hour reform; surgical education; surgical outcomes

A major transformation of surgical education commenced in 2003 with the ACGME duty hour reforms[1] heralding a modern training era. The modern era ushered in a restructured resident experience with reduced curricular inefficiencies.[2] Additional support staff hired to assure coverage of patient care needs, changes to the nature of surgical care, and the loss of between 6 months and 1 year of clinical training (as result of reduced work hours), threatened the development of resident skill, judgment, and autonomy.[3]

Program directors, experienced surgeons, and trainees themselves have all expressed concerns regarding the professional development of new surgeons.[4–6] A recent survey of fellowship directors reported that 30% of new surgeons could not independently perform a common, simple procedure at the start of fellowship, and 66% were deemed unable to operate unsupervised for 30 minutes of a major procedure.[7] Furthermore, following duty hour reform, studies have shown decreased resident case volume,[8–10] limited autonomy,[11] increased variability of the resident operative experience,[12] and increased failure rate on board certification exams.[3,13]

Prior studies of duty hour reform have largely examined its impact on residents during training and patients cared for in academic environments affected by the workforce changes.[14–17] The present study assesses the performance of new surgeons trained before and after duty hour reform upon their transition to independent practice. To our knowledge, this is the first study to use the patient outcomes of new surgeons to measure changes in the performance of the education system, an objective of the ACGME's outcomes based educational paradigm.

METHODS

This research protocol was approved by the institutional review board of the University of Pennsylvania.

Data Sources

We acquired administrative claims files for fee-for-service Medicare beneficiaries undergoing general or orthopedic surgery between July 1999 and June 2003, and July 2009 and June 2013. Information on surgeon training and board certification was obtained from the American Medical Association Physician Masterfile.[18]

Dataset Preparation

Index admissions were selected by screening the ICD-9 principal procedure field of the Medicare inpatient claims for a general or orthopedic surgery operation. All operations studied required an incision. We restricted the sample to patients who were at least age 65.5 on admission and had complete part B coverage without HMO in the 6 months before admission, and then assigned patients to their operating physician using the National Provider Identifier or Unique Physician Identification Number fields in con-temporaneous CPT bills found in the Part B file. A random admission was used for patients with multiple qualifying admissions such that each patient was only included once. See Supplemental Digital Content eTable 1, http://links.lww.com/SLA/B628 for ICD-9 codes and procedure categories, eTable2, http://links.lww.com/SLA/B628 for CPT codes and procedure categories, eSection 1 for a description of the patient-to-surgeon assignment algorithm, and eTable 3, http://links.lww.com/SLA/B628 for detailed exclusion steps.

Surgeon Classification

We defined 4 cohorts of surgeons based on training era (traditional or modern) and experience (new or experienced). Surgeons were considered new during the first 3 years of independent billing following residency. Traditional new surgeons completed residency entirely under the prereform system and entered independent practice between 1999 and 2003. Modern new surgeons completed residency following the implementation of duty hour reform and entered independent practice between 2009 and 2013. Experienced surgeons had at least 10 years of independent practice and completed residency in or after 1965. Surgeons were assigned to the hospital in which they performed the plurality of their practice. To be eligible, surgeons must have billed for at least 10 specialty-specific operations on eligible patients in a given era in a single hospital.

Matching Covariates

Patient covariates were defined using the index admission claim and a 6-month look-back in inpatient, outpatient, and carrier/part B files. In addition to the principal procedure, we defined each patient's age at admission, admission year, sex, emergent or transfer status, 30 comorbidities,[19] number of comorbidities, and major operative secondary procedure status.

[20] We also constructed a propensity score for treatment by a new surgeon. Finally, risk scores for 30-day mortality were modeled using patients of surgeons at study hospitals who were not included in the patient match[21] (see Supplemental Digital Content eTables 4a–Supplemental Digital Content eTables 4l, http://links.lww.com/SLA/B628 for risk models).

Outcomes

The primary outcome was 30-day mortality. We also studied 30-day readmissions, anesthesia time,[22] and prolonged length of stay (LOS).[15] A prolonged stay is a procedure- and era-specific LOS that exceeds the point in a hospitalization when rates of discharge begin to decline (see Supplemental Digital Content eTable 5, http://links.lww.com/SLA/B628). Readmissions were defined as any readmission or death within 30 days of discharge. We report readmission or death within 30 days of discharge to avoid giving inappropriate credit in the readmission analysis for an early death. Alternative definitions of readmissions were also examined. Resource utilization was examined using LOS, intensive care unit (ICU) use, and 30-day resource-based costs (see Supplemental Digital Content eSection 2, http://links.lww.com/SLA/B628 for costing algorithm).[23]

The Matching Algorithm

The Surgeon Match—The newest of the new surgeons perform few operations, whereas the most experienced of the new surgeons perform many more; therefore, most patients treated by new surgeons are treated by relatively experienced new surgeons. To prevent the study from focusing on the most experienced of the new surgeons, a random sample of 10 patients was drawn from each new surgeon to represent the typical distribution of a typical new surgeon's practice. Using these sampled patients, each new surgeon was paired with 1 experienced surgeon at the same hospital based on the operations they performed. Matching was performed independently in each era.

The Patient Match—Within each surgeon pair, we matched 10 experienced surgeon patients to the 10 sampled new surgeon patients using the matching covariates previously described. (see Supplemental Digital Content eSection 3, http://links.lww.com/SLA/B628 for a detailed description of the matching covariates and algorithm.)

Matches were considered balanced if all standardized differences after matching were below 0.10. Matches were accomplished using the RCBsubset[24] package in R, version 3.2.1.[25]

Outcome Testing—Outcomes before matching were reported using Fisher exact test for binary outcomes,[26] and Wilcoxon rank sum for continuous outcomes. Outcomes after matching were assessed using paired tests. Within-training era differences in binary outcomes were tested using the McNemar statistic,[27] and expressed as odds ratios. Estimates and tests of LOS, anesthesia time, and costs used M-statistics for matched pairs.[28,29]

The difference-in-differences in binary outcomes between surgeon pairs matched in the traditional and modern eras were tested using the Gart test.[26] Differences-in-differences in continuous outcomes were tested using the Wilcoxon rank sum test. All matching and outcome testing was performed in R, version 3.2.1.[25]

RESULTS

Study Cohort

Across the 2 training eras, 2762 traditional and 2119 modern eligible new surgeons entered practice at 1902 hospitals with 9297 traditional and 8503 modern eligible experienced surgeons. Both new and experienced surgeons practiced in 1478 hospitals in the traditional era and 1432 hospitals in the modern era. Within the same hospitals, new surgeons performed 164,490 operations (Traditional 96,454 and Modern 68,036) and eligible experienced surgeons performed 1,434,264 operations (Traditional 740,099 and Modern 694,165) (see Supplemental Digital Content eTable 6, http://links.lww.com/SLA/B628).

Surgeon and Patient Characteristics

Table 1 describes the characteristics of new and experienced surgeons. Traditional new surgeons had many fewer years in practice at the time of each operation than traditional experienced surgeons (mean: 1.75 vs 18.90 years, $P < 0.0001$). Likewise, modern new surgeons had fewer years in practice than modern experienced surgeons (mean: 1.78 vs 21.60 years, $P < 0.0001$). This pattern persisted within both general and orthopedic surgery subgroups. The differences in the new and experienced surgeon patients can be seen in Table 2.

Quality of Matches: Matching Results

Across the 2 training eras, 4398 new surgeons (90.1%) were matched to experienced surgeons within the same hospitals, with the traditional match comprising 2578 surgeon pairs, and the modern match comprising 1820 surgeon pairs.

The quality of the matches for representative covariates can be seen in Table 2. (see Supplemental Digital Content eTables 7a–Supplemental Digital Content eTables 7f, http://links.lww.com/SLA/B628 for the quality of the match across all covariates). After matching, the standardized differences were <0.10 for all clinical covariates. All operations in each era were exactly balanced.

Clinical and Financial Outcomes

Before matching, the 30-day mortality rates of new surgeons were significantly higher than those of the experienced surgeons in both training eras (traditional 6.5% vs 3.9%, $P <$ 0.0001; modern 5.4% vs 2.7%, $P <$ 0.0001%). There were significant differences between new and experienced surgeons regardless of the training era in many of the unadjusted secondary outcomes (Table 3).

Matched outcome differences between new and experienced surgeons trained in each era are described in Table 4. Outcomes for new surgeons trained in the traditional era are reported on the left; outcomes for new surgeons trained in the modern era are reported on the right. After matching, the 30-day mortality rate for new surgeons was slightly higher than that of the experienced surgeons in both training eras, but significantly higher only in the traditional era (traditional: 7.0% vs 6.3%, $P = 0.0007$; modern: 6.2% vs 5.9%, $P = 0.2391$).

Differences-in-differences in Outcomes Between the Traditional Era (1999–2003) and Modern Era (2009–2013)

The last 2 columns of Table 4 report the difference-in-difference odds ratios or paired differences and *P* value for each outcome, comparing the modern era to the traditional era. After matching, there were no significant differences in the differences between training eras on 30-day mortality, 30-day failure-to-rescue, 30-day readmissions or death, ICU usage, or LOS (see Supplemental Digital Content eTable 8, http://links.lww.com/SLA/B628 for additional readmission definitions). Patients of new surgeons trained in the modern era required more anesthesia time, by an average of 9 minutes, than that needed by the patients of the new surgeons trained in the Traditional era when compared to their experienced surgeon controls ($P < 0.0001$). There were significantly elevated odds of prolonged stay amongst patients treated by new surgeons trained in the modern era when compared to patients treated by new surgeons trained in the Traditional era (Gart's OR = 1.08; $P =$ 0.0149). Notably, prolonged stay declined for both groups from the traditional era to the modern; however, patients treated by experienced surgeons displayed a larger decline in prolonged stay than those treated by new surgeons. Finally, patients of modern new surgeons required slightly more resources than patients of traditional new surgeons when compared to era matched experienced surgeons [255 USD 95% CI (2, 508); $P = 0.0486$].

The outcomes by specialty can be seen in Table 4. The results were consistent with the overall results, with 2 notable exceptions in general surgery. First, the patients of Modern new surgeons showed a trend toward lower 30-day all location Mortality than patients of traditional new surgeons when compared to contemporaneous experienced surgeons [0.86 95% CI (0.74, 1.00); $P = 0.0513$]. Second, the patients of new surgeons in the modern era showed relatively lower 30-day readmissions than patients of new surgeons in the Traditional era when compared to contemporaneous experienced surgeons [0.86 95% CI (0.79, 0.95); $P = 0.0025$]. The difference in readmissions reflected the slight increase in the rate of 30-day readmissions among the experienced surgeons across the 2 training eras rather than a change in the rates of readmission among the patients of the new surgeons.

DISCUSSION

This nationwide difference-in-differences analysis of Medicare beneficiaries compared patients treated by new surgeons trained before and after duty hour reform. New surgeons and their patients were matched to experienced surgeons (practicing in the same hospitals) and their patients. Although new surgeons had slightly worse 30-day mortality rates than their matched experienced surgeons, there was no evidence to suggest that duty hour reform negatively affected the performance of new surgeons. There were also no significant differences in the differences between new surgeons trained before and after duty hour reform on ICU usage rate, LOS, failure-to-rescue, and 30-day resource cost when compared to matched experienced surgeons practicing in the same era. Significant differences between training eras were noted for a number of secondary outcomes concerned with resource utilization. Patients of new surgeons trained after duty hour reform experienced a greater odds of prolonged LOS, required more anesthesia time and incurred slightly higher resource costs than those of traditionally trained new surgeons (all relative to their contemporaneous experienced surgeons). Furthermore, within general surgery only, patients of modern new surgeons experienced a trend toward a reduced rate of all location 30-day death and a relatively reduced rate of all location 30-day readmission or death compared to traditional new surgeon patients relative to the contemporaneously matched experienced surgeons' patients.

Our finding that the modern training paradigm prepares new surgeons for practice with only slight limitations when compared to the traditional training paradigm supports the perception of graduating residents and new surgeons that they are well prepared for the transition to practice.[30] The finding that new surgeons in the modern era use slightly more resources for patient care is in keeping with the perceptions of fellowship directors; however, with respect to clinical outcomes of 30-day mortality and 30-day readmission or death, our results suggest that the gap in new surgeon performance is more a function of their inexperience and less a product of their training era as previously suggested by the survey of program directors and experienced surgeons.[7,30,31] Moreover, there were only slight significant differences between the modern and traditional eras in outcomes for the subgroups by specialty that might be more sensitive to the differences in duty hour policies.

Within general surgery, the difference in difference analysis almost reached significance with a clinically meaningful improvement in 30-day all location mortality suggesting that new surgeon performance may be better in the new training paradigm when compared to the traditional era. Furthermore, there is evidence that new surgeon performance in the modern era is slightly better than that of the traditional era (when compared to contemporaneous experienced surgeons) on 30-day readmissions or death and slightly worse in prolonged stay. For both of the latter measures, it is important to consider that the modern training paradigm has a greater emphasis on systems-based practice[32] than the traditional training. This may encourage new surgeons to preferentially extend LOS to mitigate the risk of readmission, whereas the experienced surgeons, who were trained in the traditional era, may be more inclined to aim for a shorter LOS without as much concern for the measurement of readmission.

Other studies examining the effects of duty hour reform on patient outcomes have focused on care delivered in the academic medical centers without attention to the performance of the new surgeons. These studies reported minimal changes following the implementation of the 2003 duty hours standards across a broad array of measures.[16,17,33–35] These studies showed that hospitals were able to adapt to the dramatic reduction in the service provided by surgical residents without compromising the care provided under the direction of traditionally trained surgeons. These studies, however, did not assess the impact of these reforms on the performance of the affected trainees as new surgeons after their transition to practice.

Although it is thought that established relationships between experienced surgeons and hospital staff might result in extra attention for their patients, the relationship between surgeon experience and hospital quality has not been explored previously; therefore, another important finding in this study was that patients of new surgeons, regardless of their training era, achieved near equivalent rates of failure-to-rescue, when compared to patients of experienced surgeons. This suggests that the working environment in which new surgeons enter practice is robust to their relative lack of familiarity with coworkers and inexperience. Moreover, it would appear that changes to surgical practice that accompanied duty hour reform were not associated with a penalty to the modern new surgeons' patients.

There are several limitations of this study. First, we only examined inpatient operations performed within 2 surgical specialties; as such the findings may not generalize to surgeons who practice predominantly in the outpatient setting or within specialties that rely on different training paradigms. Second, we could not study the effects of duty hour reform on the new surgeons' mental and physical wellbeing. Third, we deliberately did not consider the effects of fellowship on the outcomes achieved by new surgeons upon the transition to practice. While surgical residents today often enter into fellowships for additional training rather than entering practice directly after residency,[36–39]cross-sectional analyses of surgical practices repeatedly show that the majority of new surgeons perform a broad range of operations,[40‘41] many of which would only have been covered during residency training.[42] Surgical fellows are permitted to independently perform operations within the scope of their residency training (which may be outside of their subspecialty) during the fellowship period. Thus, our results reflect the outcomes of new surgeons who independently care for patients. Fourth, additional changes to the nature of surgical care and further hours' reform for junior residents occurred in 2011 during the study time frame. For these reasons, we used a difference-in-differences framework to benchmark our findings and we restricted the study time frame such that all modern trainees were subject to the same duty hour exposures. Finally, we only studied the outcomes of new and experienced surgeons who practiced in the same hospitals. Although there were very few new surgeons who practiced in hospitals without an eligible experienced surgeon (eTable 3, http://links.lww.com/SLA/B628), there were many experienced surgeons who practiced at hospitals without a new surgeon, and our experienced surgeon comparators might not be reflective of full breadth of experienced surgeons.

By benchmarking new surgeons to experienced surgeons in the same hospital, we were able to separate the effect of training reform from patient, hospital, or time period effects. Given the overall similarities in performance between new surgeons who trained before and after duty hour reform, the findings of this study should provide reassurance of the effectiveness of the new training paradigm. The methods employed in this study may provide a useful template for the evaluation of future changes to training across medical specialties.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The authors thank Traci Frank, AA, Kathryn Yucha, MSN, RN, Joseph G. Reiter, MS, and Orit Even-Shoshan, MS (Center for Outcomes Research, The Children's Hospital of Philadelphia, Philadelphia, PA) for their assistance with this research.

All phases of this study were supported by National Institute on Aging/National Institutes of Health grant R01 AG049757. NIA/NIH had no role in the design or conduct of the study; collection, management, analysis, or interpretation of the data; or preparation, review, or approval of the manuscript to submit for publication.

REFERENCES

1. Accreditation Council for Graduate Medical Education. Common Program AQ5 Requirements. Available at: http://acgme.org/acgmeweb. Accessed August 20, 2012.
2. Kellogg KC, Breen E, Ferzoco SJ, et al. Resistance to change in surgical residency: an ethnographic study of work hours reform. J Am Coll Surg. 2006;202:630–636. [PubMed: 16571434]
3. Lewis FR, Klingensmith ME. Issues in general surgery residency training— 2012. Ann Surg. 2012;256:553–559. [PubMed: 22964740]
4. Britt LD, Sachdeva AK, Healy GB, et al. Resident duty hours in surgery for ensuring patient safety, providing optimum resident education and training, and promoting resident well-being: a response from the American College of Surgeons to the Report of the Institute of Medicine, “Resident Duty Hours: Enhancing Sleep, Supervision, and Safety”. Surgery. 2009;146:398–409. [PubMed: 19715795]
5. Drolet BC, Sangisetty S, Tracy TF, et al. Surgical residents’ perceptions of 2011 Accreditation Council for Graduate Medical Education duty hour regulations. JAMA Surg. 2013;148:427–433. [PubMed: 23677406]
6. Mir HR, Cannada LK, Murray JN, et al. Orthopaedic resident and program director opinions of resident duty hours: a national survey. J Bone Joint Surg Am. 2011;93:e1421–1429. [PubMed: 22159864]
7. Mattar SG, Alseidi AA, Jones DB, et al. General surgery residency inadequately prepares trainees for fellowship: results of a survey of fellowship program directors. Ann Surg. 2013;258:440–449. [PubMed: 24022436]
8. Tran J, Lewis R, De Virgilio C. The effect of the 80-hour work week on general surgery resident operative case volume. Am Surg. 2006;72:924–928. [PubMed: 17058737]
9. Bland KI, Stoll DA, Richardson JD, et al. Brief communication of the Residency Review Committee-Surgery (RRC-S) on residents’ surgical volume in general surgery. Am J Surg. 2005;190:345–350. [PubMed: 16105514]
10. Kairys JC, McGuire K, Crawford AG, et al. Cumulative operative experience is decreasing during general surgery residency: a worrisome trend for surgical trainees? J Am Coll Surg. 2008;206:804–811. discussion 811–803. [PubMed: 18471701]
11. Meyerson SL, Teitelbaum EN, George BC, et al. Defining the autonomy gap: when expectations do not meet reality in the operating room. J Surg Educ. 2014;71:e64–e72. [PubMed: 24924583]
12. Quillin RC 3rd, CortezF AR, PrittsF TA. et al. Operative variability among residents has increased since implementation of the 80-hour workweek. J Am Coll Surg. 2016;222:1201–1210. [PubMed: 27068844]
13. Herkowitz HN, Emery SE, Hurwitz SR, et al. Performance of candidates selecting the subspecialty of spine surgery for the Part II American Board of Orthopaedic Surgery oral certification examination. J Bone Joint Surg Am. 2013;95:e135. [PubMed: 24048566]
14. Bilimoria KY, Chung JW, Hedges LV, et al. National cluster-randomized trial of duty-hour flexibility in surgical training. N Engl J Med. 2016;374:713–727. [PubMed: 26836220]
15. Silber JH, Rosenbaum PR, Rosen AK, et al. Prolonged hospital stay and the resident duty hour rules of 2003. Med Care. 2009;47:1191–1200. [PubMed: 19786912]
16. Volpp KG, Rosen AK, Rosenbaum PR, et al. Mortality among patients in VA hospitals in the first 2 years following ACGME resident duty hour reform. JAMA. 2007;298:984–992. [PubMed: 17785643]
17. Baldwin K, Namdari S, Donegan D, et al. Early effects of resident work-hour restrictions on patient safety: a systematic review and plea for improved studies. J Bone Joint Surg Am. 2011;93:e5. [PubMed: 21248206]
18. American Medical Association. About the AMA Physician Masterfile Available at: http://info.commerce.ama-assn.org/ama-physician-masterfile. Accessed February 16, 2019.

19. Center for Outcomes Research, Children's Hospital of Philadelphia. National Quality Forum Outcomes Measures Maintained by the Center for Outcomes Research. 2015 Available at: http://cor.research.chop.edu/node/26. Accessed April 6, 2018.
20. Agency for Healthcare Research and Quality. Clinical Classifications Software (CCS) for ICD-9-CM. Health Care Utilization Project. March 2017. Available at: www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp. Accessed April 20, 2018.
21. Hansen BB. The prognostic analogue of the propensity score. Biometrika. 2008;95:481–488.
22. Silber JH, Rosenbaum PR, Zhang X, et al. Estimating anesthesia and surgical procedure times from Medicare anesthesia claims. Anesthesiology. 2007;106:346–355. [PubMed: 17264730]
23. Silber JH, Rosenbaum PR, Kelz RR, et al. Medical and financial risks associated with surgery in the elderly obese. Ann Surg. 2012;256:79–86. [PubMed: 22566017]
24. Pimentel SD. Optimal Subset Matching with Refined Covariate Balance. Package "rcbsubset". Version 1.1.3. November 12, 2017. Available at: https://cran.r-project.org/web/packages/rcbsubset/rcbsubset.pdf. Accessed April 20, 2018.
25. R Development Core Team. R: A Language and Environment for Statistical Computing. 2018 Available at: http://www.R-project.org. Accessed April 20, 2018.
26. Cox DR, Snell EJ. Analysis of Binary Data. New York, NY: Chapman and Hall; 1970.
27. Bishop YMM, Fienberg SE, Holland PW. Discrete Multivariate Analysis: Theory and Practice. Cambridge, MA: The MIT Press; 1975.
28. Rosenbaum PR. Sensitivity analysis for m-estimates, tests, and confidence intervals in matched observational studies. Biometrics. 2007;63:456–464 (R package sensitivitymv and sensitivitymw). [PubMed: 17688498]
29. Rosenbaum PR. Two R packages for sensitivity analysis in observational studies. Obs Stud. 2015;1:1–17.
30. Napolitano LM, Savarise M, Paramo JC, et al. Are general surgery residents ready to practice? A survey of the American College of Surgeons Board of Governors and Young Fellows Association. J Am Coll Surg. 2014;218:1063.e31–1072.e31.
31. Pellegrini VD Jr, Ferguson PC, Cruess R, et al. Sufficient competence to enter the unsupervised practice of orthopaedics: What is it, when does it occur, and do we know it when we see it? AOA Critical Issues. J Bone Joint Surg Am. 2015;97:1459–1464. [PubMed: 26333742]
32. NEJM Knowledge+ Team. Exploring the ACGME Core Competencies: Systems-Based Practice (Part 4 of 7). 2019 Available at: https://knowledge-plus.nejm.org/blog/acgme-core-competencies-systems-based-practice/. Accessed February 16, 2019.
33. Volpp KG, Rosen AK, Rosenbaum PR, et al. Mortality among hospitalized Medicare beneficiaries in the first 2 years following ACGME resident duty hour reform. JAMA. 2007;298:975–983. [PubMed: 17785642]
34. Rosen AK, Loveland SA, Romano PS, et al. Effects of resident duty hour reform on surgical and procedural patient safety indicators among hospitalized Veterans Health Administration and Medicare patients. Med Care. 2009;47:723–731. [PubMed: 19536029]
35. Shetty KD, Bhattacharya J. Changes in hospital mortality associated with residency work-hour regulations. Ann Intern Med. 2007;147:73–80. [PubMed: 17548403]
36. American College of Surgeons Taskforce on the Resident 80-Hour Work Week. Position of the American College of Surgeons on Restrictions on Resident Work Hours presented to the Institute of Medicine Consensus Committee. 2008 Available at: http://www.facs.org/education/statement.pdf. Accessed August 21, 2012.
37. Borman KR, Vick LR, Biester TW, et al. Changing demographics of residents choosing fellowships: long-term data from the American Board of Surgery. J Am Coll Surg. 2008;206:782–788. discussion 788–789. [PubMed: 18471695]
38. Ellis MC, Dhungel B, Weerasinghe R, et al. Trends in research time, fellowship training, and practice patterns among general surgery graduates. J Surg Educ. 2011;68:309 [PubMed: 21708369]
39. Samora JB, Bashook P, Jones A, et al. Orthopaedic graduate medical education: a changing paradigm. JBJS Rev. 2014;2:pii: 01874474–201411000-00004.

40. Sheikh F, Gray RJ, Ferrara J, et al. Disparity between actual case volume and the perceptions of case volume needed to train competent general surgeons. J Surg Educ. 2010;67:371–375. [PubMed: 21156293]
41. Black KP, Alman BA, Levine WN, et al. Orthopaedic resident education-it's a whole new game: "If I'm going to be a spine surgeon, why do I need to learn how to reconstruct an anterior cruciate ligament?": AOA critical issues. J Bone Joint Surg Am. 2012;94:e96. [PubMed: 22760397]
42. Miller BJ, Rajani R, Leddy L, et al. How much tumor surgery do early-career orthopaedic oncologists perform? Clin Orthop Relat Res. 2015;473:695–702. [PubMed: 25224820]

TABLE 1.

Characteristics of New and Experienced Surgeons Practicing Within the Same Hospitals

	Traditional Era: 1999–2003 Cohort						Modern Era: 2009–2013 Cohort					
	Surgeon-patients before Sampling or Matching			Surgeon-patients after Sampling and Matching			Surgeon-patients before Sampling or Matching			Surgeon-patients after Sampling and Matching		
Covariate (% Unless Noted)	**New**	**Experienced**	***P***	**New**	**Experienced**	***P***	**New**	**Experienced**	***P***	**New**	**Experienced**	***P***
All surgery												
No of surgeons	2762	9297		2578	2578		2119	8503		1,820	1,820	
N Patients	96,454	740,099		25,780	25,780		68,036	694,165		18,200	18,200	
Surgeon characteristics												
Mean years of practice at date of surgery	1.75	18.90	<0.0001	1.64	18.80	<0.0001	1.78	21.60	<0.0001	1.62	21.30	<0.0001
% Board-certified‡	79.0	89.4	<0.0001	74.4	87.0	<0.0001	70.3	92.5	<0.0001	68.6	90.8	<0.0001
Mean operations per surgeon, N	34.9	79.6	<0.0001	34.8	100.5	<0.0001	32.1	81.6	<0.0001	31.2	111.6	<0.0001
Operations analyzed per surgeon, N	NA	NA	NA	10.0	10.0	1.0000	NA	NA	NA	10.0	10.0	1.0000
General surgery												
N of surgeons	1401	4783		1377	1377		1027	4324		994	994	
N of patients	47,819	313,255		13,770	13,770		28,536	220,137		9940	9940	
Surgeon characteristics												
Mean Years of Practice at Date of Surgery	1.66	18.80	<0.0001	1.58	18.70	<0.0001	1.60	21.30	<0.0001	1.52	21.00	<0.0001
% Board-certified in general surgery‡	86.7	88.7	<0.0001	82.9	86.6	<0.0001	76.1	92.2	<0.0001	74.8	91.5	<0.0001
Mean operations per surgeon, N	34.1	65.4	<0.0001	34.2	86.1	<0.0001	27.7	50.9	<0.0001	28.1	72.5	<0.0001
Operations analyzed per surgeon, N	NA	NA	NA	10.0	10.0	1.0000	NA	NA	NA	10.0	10.0	1.0000
Orthopedics												
No of surgeons	1361	4514		1201	1201		1092	4179		826	826	
No of patients	48,635	426,844		12,010	12,010		39,500	474,028		8260	8260	
Surgeon characteristics												
Mean Years of Practice at Date of Surgery	1.83	18.90	<0.0001	1.72	18.90	<0.0001	1.91	21.70	<0.0001	1.74	21.80	<0.0001
% Board-certified in orthopedic surgery‡	71.5	90.0	<0.0001	64.6	87.4	<0.0001	66.1	92.6	<0.0001	61.1	90.1	<0.0001
Mean operations per surgeon, N	35.7	94.5	<0.0001	35.6	117.1	<0.0001	36.1	113.0	<0.0001	34.9	158.7	<0.0001
Operations analyzed per surgeon, N	NA	NA	NA	10.0	10.0	1.0000	NA	NA	NA	10.0	10.0	1.0000

There were 10,496 traditional and 9318 modern new general and orthopedic surgeons identified in the AMA Masterfile based on years of training and experience level. The data presented in this table reflect the characteristics of the surgeons who also met eligibility for study inclusion based on the performance of qualifying operations and practice composition and location.

[*] Board certification status was assessed at the end of each study era.

[†] The left set of columns in each era represents the practice distribution of all eligible new and experienced surgeons who had at least 10 cases and were practicing in the same hospitals. The right set of columns in each era represents the random selection of 10 patients per new surgeon, and the experienced surgeon patients they were matched to. One hundred seventy-five (6.3%) new surgeons in the traditional era and 299 (14.1%) new surgeons in the modern era could not be satisfactorily matched and were dropped from the analysis. All covariates are weighted at the patient level.

TABLE 2.

The Quality of the Matches for (A) All Surgery, (B) General Surgery, and (C) Orthopedic Surgery

A. All Surgery	**Traditional Era: 1999–2003 Cohort**						**Modern Era: 2009–2013 Cohort**					
	Before Match			**After Match**			**Before Match**			**After Match**		
Covariate (% Unless Noted)	**New Surgeon-Patients**	**Exp. Surgeon-Patients**	**Std. Diff.**	**New Surgeon-Patients**	**Exp. Surgeon-Patients**	**Std. Diff**	**New Surgeon-Patients**	**Exp. Surgeon-Patients**	**Std. Diff.**	**New Surgeon-Patients**	**Exp. Surgeon-Patients**	**Std. Diff**
N of surgeons	2762	9297		2578	2578		2119	8503		1.820	1.820	
No of patients	27,620	740,099		25,780	25,780		21,190	694,165		18,200	18,200	
Selected principal procedure types												
Appendectomy	3.9	2.5	0.08	3.9	3.9	0.00	5.4	3.7	0.09	5.3	5.3	0.00
Bariatric procedure	0.1	0.1	0.00	0.1	0.1	0.00	0.5	1.2	−0.08	0.5	0.5	0.00
Cholecystectomy	29.9	25.3	0.10	29.9	29.9	0.00	29.9	23.1	0.15	29.8	29.8	0.00
Colectomy, partial, Lap	NA	NA	NA	NA	NA	NA	9.0	9.0	0.00	9.1	9.1	0.00
Colectomy, partial, open	25.6	26.2	−0.01	25.6	25.6	0.00	18.1	17.2	0.02	18.0	18.0	0.00
Pancreatectomy	0.5	1.1	−0.07	0.5	0.5	0.00	0.7	2.5	−0.14	0.7	0.7	0.00
Small bowel resection	5.7	4.2	0.07	5.7	5.7	0.00	8.2	6.1	0.08	8.3	8.3	0.00
Orthopedic surgery												
Hip fracture repair	38.4	17.4	0.48	39.1	39.1	0.00	29.5	9.0	0.54	29.7	29.7	0.00
Hip replacement	42.4	35.5	0.14	42.9	42.9	0.00	48.9	33.4	0.32	50.1	50.1	0.00
Hip replacement revision	2.1	3.9	−0.10	1.7	1.7	0.00	2.7	3.2	−0.03	2.0	2.0	0.00
Knee replacement	15.7	40.1	−0.57	15.3	15.3	0.00	17.0	51.0	−0.77	16.8	16.8	0.00
Knee replacement revision	1.5	3.1	−0.11	1.0	1.0	0.00	2.0	3.5	−0.09	1.5	1.5	0.00
Selected admission covariates												
Age at admission (y, mean)	79.3	77.6	0.23	79.3	79.3	−0.01	79.1	76.7	0.32	79.0	79.1	−0.02
Sex (% female)	67.5	65.7	0.04	67.1	67.3	−0.01	63.7	63.0	0.01	63.0	63.1	0.00
Emergent admission	47.7	26.9	0.44	48.0	47.9	0.00	53.4	23.0	0.66	53.9	53.9	0.00
Transfer-in	2.7	1.6	0.08	2.7	2.7	0.00	2.5	1.2	0.09	2.5	2.5	0.00
Probability of 30-day mortality	6.1	4.0	0.28	6.2	6.1	0.01	5.4	2.7	0.38	5.6	5.5	0.01
Secondary procedure	23.5	20.8	0.07	24.3	24.7	−0.01	24.5	18.3	0.15	26.7	25.7	0.03
Selected comorbidities (from 30)												
CHF	22.6	17.0	0.14	22.6	22.3	0.01	22.1	14.0	0.21	22.2	21.4	0.02

A. All Surgery	Traditional Era: 1999–2003 Cohort						Modern Era: 2009–2013 Cohort					
	Before Match			After Match			Before Match			After Match		
Covariate (% Unless Noted)	New Surgeon-Patients	Exp. Surgeon-Patients	Std. Diff.	New Surgeon-Patients	Exp. Surgeon-Patients	Std. Diff	New Surgeon-Patients	Exp. Surgeon-Patients	Std. Diff.	New Surgeon-Patients	Exp. Surgeon-Patients	Std. Diff
Past myocardial infarction	9.3	8.2	0.04	9.4	8.2	0.04	10.1	8.0	0.07	10.4	9.1	0.04
Diabetes	24.2	22.0	0.05	24.4	24.8	−0.01	32.3	29.8	0.05	32.5	32.9	−0.01
Renal dysfunction	8.1	5.6	0.10	8.2	7.6	0.02	24.8	15.5	0.23	25.2	24.8	0.01
Dementia	19.4	11.3	0.23	19.2	18.5	0.02	21.8	10.7	0.31	21.2	20.5	0.02
COPD	26.4	21.9	0.10	26.4	25.9	0.01	27.0	20.4	0.16	27.3	26.4	0.02
Number of comorbidities (mean)	3.4	3.0	0.17	3.4	3.3	0.03	4.1	3.5	0.28	4.1	4.0	0.04

	Traditional Era: 1999–2003 Cohort						Modern Era: 2009–2013 Cohort					
	Before Match			After Match			Before Match			After Match		
B. General Surgery	New Surgeon-Patients	Exp. Surgeon-Patients	Std. Diff.	New Surgeon-Patients	Exp. Surgeon-Patients	Std. Diff	New Surgeon-Patients	Exp. Surgeon-Patients	Std. Diff.	New Surgeon-Patients	Exp. Surgeon-Patients	Std. Diff
No of surgeons	1401	4783		1377	1377		1027	4324		994	994	
No of patients	14,010	313,255		13,770	13,770		10,270	220,137		9940	9940	
Selected principal procedure types												
Appendectomy	3.9	2.5	0.08	3.9	3.9	0.00	5.4	3.7	0.09	5.3	5.3	0.00
Bariatric procedure	0.1	0.1	0.00	0.1	0.1	0.00	0.5	1.2	−0.08	0.5	0.5	0.00
Cholecystectomy	29.9	25.3	0.10	29.9	29.9	0.00	29.9	23.1	0.15	29.8	29.8	0.00
Colectomy, partial, Lap	NA	NA	NA	NA	NA	NA	9.0	9.0	0.00	9.1	9.1	0.00
Colectomy, partial, open	25.6	26.2	−0.01	25.6	25.6	0.00	18.1	17.2	0.02	18.0	18.0	0.00
Pancreatectomy	0.5	1.1	−0.07	0.5	0.5	0.00	0.7	2.5	−0.14	0.7	0.7	0.00
Small bowel resection	5.7	4.2	0.07	5.7	5.7	0.00	8.2	6.1	0.08	8.3	8.3	0.00
Selected admission covariates												
Age at admission (y, mean)	77.1	76.8	0.05	77.2	77.3	−0.02	77.0	76.6	0.06	77.0	77.0	0.00
Sex (% female)	61.0	61.3	−0.01	60.9	61.2	−0.01	56.4	58.1	−0.03	56.5	56.8	−0.01
Emergent admission	41.3	30.1	0.24	41.7	41.7	0.00	52.4	37.9	0.29	52.7	52.7	0.00
Transfer-in	3.4	2.5	0.05	3.4	3.5	0.00	3.5	2.6	0.05	3.5	3.5	0.00
Probability of 30-day mortality	6.6	5.4	0.13	6.7	6.4	0.02	6.2	5.1	0.13	6.2	6.0	0.03

A. All Surgery	Traditional Era: 1999–2003 Cohort						Modern Era: 2009–2013 Cohort					
	Before Match			After Match			Before Match			After Match		
Covariate (% Unless Noted)	New Surgeon-Patients	Exp. Surgeon-Patients	Std. Diff.	New Surgeon-Patients	Exp. Surgeon-Patients	Std. Diff	New Surgeon-Patients	Exp. Surgeon-Patients	Std. Diff.	New Surgeon-Patients	Exp. Surgeon-Patients	Std. Diff
Secondary procedure	40.6	39.0	0.03	40.6	41.1	−0.01	43.4	42.4	0.02	43.5	42.0	0.03
Selected comorbidities (from 30)												
CHF	22.2	19.7	0.06	22.3	21.5	0.02	22.7	19.2	0.09	22.8	21.1	0.04
Past myocardial infarction	10.2	9.4	0.03	10.3	8.3	0.07	11.2	10.3	0.03	11.2	9.6	0.05
Diabetes	26.0	24.6	0.03	26.0	26.1	0.00	35.4	35.0	0.01	35.4	35.7	0.00
Renal dysfunction	10.1	8.0	0.07	10.1	9.2	0.03	28.2	22.6	0.13	28.2	27.2	0.02
Dementia	10.4	8.1	0.08	10.4	9.3	0.04	13.3	10.5	0.09	13.3	11.6	0.05
COPD	26.8	25.1	0.04	26.9	26.0	0.02	28.7	26.5	0.05	28.6	27.6	0.02
Number of comorbidities (mean)	3.5	3.4	0.04	3.5	3.4	0.05	4.3	4.2	0.03	4.3	4.2	0.04

C. Orthopedic Surgery	Traditional Era: 1999–2003 Cohort						Traditional Era: 2009–2013 Cohort					
	Before Match			After Match			Before Match			After Match		
Covariate (% Unless Noted)	New Surgeon-Patients	Exp. Surgeon-Patients	Std. Diff.	New Surgeon-Patients	Exp. Surgeon-Patients	Std. Diff	New Surgeon-Patients	Exp. Surgeon-Patients	Std. Diff.	New Surgeon-Patients	Exp. Surgeon-Patients	Std. Diff
No of surgeons	1361	4514		1201	1201		1092	4179		826	826	
No of patients	13,610	426,844		12,010	12,010		10,920	474,028		8260	8260	
Principal procedure types												
Hip fracture repair	38.4	17.4	0.48	39.1	39.1	0.00	29.5	9.0	0.54	29.7	29.7	0.00
Hip replacement	42.4	35.5	0.14	42.9	42.9	0.00	48.9	33.4	0.32	50.1	50.1	0.00
Hip replacement revision	2.1	3.9	−0.10	1.7	1.7	0.00	2.7	3.2	−0.03	2.0	2.0	0.00
Knee replacement	15.7	40.1	−0.57	15.3	15.3	0.00	17.0	51.0	−0.77	16.8	16.8	0.00
Knee replacement revision	1.5	3.1	−0.11	1.0	1.0	0.00	2.0	3.5	−0.09	1.5	1.5	0.00
Selected admission covariates												
Age at admission (y, mean)	81.5	78.2	0.44	81.7	81.7	−0.01	81.1	76.7	0.56	81.3	81.7	−0.05
Sex (% female)	74.1	68.8	0.12	74.1	74.4	−0.01	70.7	65.3	0.11	70.9	70.7	0.00
Emergent admission	54.2	24.5	0.64	55.2	55.2	0.00	54.2	16.1	0.87	55.3	55.3	0.00
Transfer-in	1.9	0.9	0.09	1.9	1.9	0.00	1.5	0.6	0.09	1.3	1.3	0.00

A. All Surgery	Traditional Era: 1999–2003 Cohort						Modern Era: 2009–2013 Cohort					
	Before Match			After Match			Before Match			After Match		
Covariate (% Unless Noted)	New Surgeon-Patients	Exp. Surgeon-Patients	Std. Diff.	New Surgeon-Patients	Exp. Surgeon-Patients	Std. Diff	New Surgeon-Patients	Exp. Surgeon-Patients	Std. Diff.	New Surgeon-Patients	Exp. Surgeon-Patients	Std. Diff
Probability of 30-day mortality	5.6	2.9	0.45	5.7	5.7	−0.01	4.7	1.6	0.57	4.8	4.9	−0.03
Secondary procedure	5.9	7.4	−0.06	5.7	5.8	0.00	6.8	7.1	−0.01	6.6	5.9	0.02
Selected comorbidities (from 30)												
CHF	23.0	15.1	0.20	22.9	23.3	−0.01	21.5	11.5	0.27	21.5	21.9	−0.01
Past myocardial infarction	8.4	7.3	0.04	8.3	8.0	0.01	9.0	6.9	0.08	9.4	8.6	0.03
Diabetes	22.5	20.2	0.06	22.6	23.2	−0.01	29.4	27.4	0.04	29.0	29.7	−0.02
Renal dysfunction	6.0	3.9	0.10	5.9	5.7	0.01	21.7	12.2	0.26	21.6	21.9	−0.01
Dementia	28.7	13.6	0.38	29.2	29.0	0.00	29.9	10.8	0.49	30.8	31.3	−0.01
COPD	25.9	19.6	0.15	25.8	25.9	0.00	25.5	17.6	0.19	25.8	25.0	0.02
Number of comorbidities (mean)	3.3	2.8	0.27	3.3	3.3	0.00	4.0	3.1	0.38	4.0	3.9	0.04

* Describes the new surgeon practice distribution after randomly sampling 10 patients per surgeon. These patients were then matched to the patients of experienced surgeons. Note the practice patterns of new and experienced surgeons are dissimilar before matching, and similar after matching.

TABLE 3.

Pre-matching Outcomes

Outcome (% Unless Noted)	(1) Traditional Era: 1999–2003 Match New Surgeon-Patients	Exp. Surgeon-Patients	Diff	P^*	(2) Modern Era: 2009–2013 Match New Surgeon-Patients	Exp. Surgeon-Patients	Diff	P^*	(2) vs (1) D-in-D D-in-D	$P^\dagger$
All surgery										
No of surgeons	2762	9297			2119	8503				
No of patients	96,454	740,099			68,036	694,165				
30-Day mortality	6.5	3.9	2.6	<0.0001	5.4	2.7	2.7	<0.0001	0.1	<0.0001
30-Day failure-to-rescue	10.3	7.5	2.9	<0.0001	8.6	5.3	3.3	<0.0001	0.4	<0.0001
30-Day readmission or death (all location)	18.7	13.7	5.1	<0.0001	17.7	12.3	5.4	<0.0001	0.3	0.0003
ICU usage	21.1	14.4	6.7	<0.0001	18.6	11.1	7.5	<0.0001	0.8	<0.0001
Prolonged stay	61.1	50.4	10.8	<0.0001	49.8	32.3	17.5	<0.0001	6.7	<0.0001
Anesthesia time (min, m-est.)	149.2	147.9	1.3	<0.0001	149.8	135.6	14.2	<0.0001	12.9	<0.0001
Length of stay (days, m-est.)	7.5	6.6	0.9	<0.0001	6.7	5.1	1.6	<0.0001	0.7	<0.0001
30-Day resource cost ($, m-est.)	19,657	16,737	2920	<0.0001	20,041	15,094	4947	<0.0001	2027	<0.0001
General surgery										
No of surgeons	1401	4783			1027	4324				
No of patients	47,819	313,255			28,536	220,137				
30-Day mortality	7.4	5.3	2.1	<0.0001	6.8	5.2	1.6	<0.0001	−0.5	<0.0001
30-Day failure-to-rescue	11.1	9.1	2.0	<0.0001	10.1	9.0	1.2	<0.0001	−0.8	<0.0001
30-Day readmission or death (all location)	20.9	17.3	3.6	<0.0001	21.3	18.5	2.9	<0.0001	−0.8	0.0003
ICU usage	32.8	26.3	6.5	<0.0001	31.9	26.2	5.7	<0.0001	−0.8	<0.0001
Prolonged stay	70.2	63.6	6.6	<0.0001	69.4	62.3	7.1	<0.0001	0.5	<0.0001
Anesthesia time (min, m-est)	153.9	152.2	1.7	<0.0001	156.2	149.2	7.0	<0.0001	5.3	<0.0001
Length of stay (days, m-est)	9.1	8.5	0.6	<0.0001	8.7	8.0	0.7	<0.0001	0.1	0.5025
30-Day resource cost ($, m-est.)	24,792	22,049	2743	<0.0001	27,564	24,368	3196	<0.0001	453	0.0317
Orthopedics										
No of surgeons	1361	4514			1092	4179				
No of patients	48,635	426,844			39,500	474,028				
30-Day mortality	5.5	2.8	2.7	<0.0001	4.4	1.6	2.8	<0.0001	0.1	<0.0001

Outcome (% Unless Noted)	(1) Traditional Era: 1999–2003 Match				(2) Modern Era: 2009–2013 Match				(2) vs (1) D-in-D	
	New Surgeon-Patients	Exp. Surgeon-Patients	Diff	P^*	New Surgeon-Patients	Exp. Surgeon-Patients	Diff	P^*	D-in-D	$P^†$
30-Day failure-to-rescue	9.4	5.9	3.5	<0.0001	7.4	3.3	4.1	<0.0001	0.6	<0.0001
30-Day readmission or death (all location)	16.6	11.0	5.5	<0.0001	15.0	9.4	5.6	<0.0001	0.0	0.0003
ICU usage	9.6	5.7	3.9	<0.0001	9.0	4.1	4.9	<0.0001	1.0	<0.0001
Prolonged stay	52.3	40.7	11.5	<0.0001	35.6	18.4	17.2	<0.0001	5.7	<0.0001
Anesthesia time (min, m-est.)	145.4	145.7	−0.3	0.0328	148.9	131.7	17.2	<0.0001	17.5	<0.0001
Length of stay (days, m-est.)	6.4	5.7	0.7	<0.0001	5.5	4.4	1.1	<0.0001	0.4	0.9578
30-Day resource cost ($, m-est.)	15,793	13,917	1876	<0.0001	15,879	12,520	3359	<0.0001	1483	<0.0001

Displays the overall outcome differences between new and experienced surgeons practicing in the same hospitals before sampling 10 new surgeon patients or matching.

* *P* values were calculated using Fisher exact test for binary variables and Wilcoxon Rank Sum for continuous variable.

† *P* values were calculated using logistic regression and m-estimation for binary and continuous variables, respectively.

‡ Resource costs were calculated using 2013 dollars.

TABLE 4.

Outcome Differences Between Patients Treated by New or Experienced Surgeons, After Matching

	(1) Traditional Era: 1999–2003 Match				(2) Modern Era: 2009–2013 Match				(2) vs. (1) D-in-D	
Outcome (% Unless Noted)	**New Surgeon-Patients**	**Exp. Surgeon-Patients**	**Odds Ratio or Paired Diff (95% CI)**	***P***	**New Surgeon-Patients**	**Exp. Surgeon-Patients**	**Odds Ratio or Paired Diff (95% CI)**	***P***	**Odds Ratio or Difference (95% CI)**	***P***
All surgery										
No of surgeons	2578	2578			1820	1820				
No of patients	25,780	25,780			18,200	18,200				
30-Day mortality	7.0	6.3	1.13 (1.05, 1.22)	0.0007	6.2	5.9	1.06 (0.97, 1.16)	0.2391	0.93 (0.83, 1.05)	0.2333
30-Day failure-to-rescue	10.8	10.0	1.07 (0.99, 1.17)	0.1059	9.5	9.3	0.99 (0.89, 1.10)	0.8725	0.92 (0.80, 1.06)	0.2428
30-Day readmission or Death	19.8	18.1	1.13 (1.08, 1.18)	<0.0001	19.4	18.6	1.06 (1.01, 1.12)	0.0305	0.94 (0.88, 1.01)	0.1008
ICU usage	22.9	20.3	1.23 (1.17, 1.29)	<0.0001	22.8	20.5	1.19 (1.13, 1.27)	<0.0001	0.98 (0.90, 1.05)	0.5225
Prolonged stay	63.7	62.8	1.05 (1.01, 1.10)	0.0109	56.9	54.5	1.14 (1.08, 1.19)	<0.0001	1.08 (1.02, 1.15)	0.0149
Anesthesia time (min, m-est.)	150.6	141.3	7.7 (6.4, 9.1)	<0.0001	155.4	137.6	18.7 (17.4, 20.1)	<0.0001	9.0 (7.5, 10.5)	<0.0001
Length of stay (days, m-est.)	7.9	7.7	0.1 (0.0, 0.2)	0.0006	7.3	7.2	0.2 (0.1, 0.2)	0.0002	0.0 (0.0, 0.0)	0.4046
30-Day resource cost ($, m-est.)	20,713	19,751	860 (620, 1,103)	<0.0001	22,404	20,998	1257 (946, 1568)	<0.0001	255 (2, 508)	0.0486
General surgery										
No of surgeons	1377	1377			994	994				
No of patients	13,770	13,770			9940	9940				
30-Day mortality	7.4	6.4	1.18 (1.07, 1.30)	0.0010	6.9	6.8	1.01 (0.90, 1.14)	0.8380	0.86 (0.74, 1.00)	0.0513
30-Day failure-to-rescue	10.9	10.0	1.08 (0.96, 1.21)	0.1938	10.3	10.5	0.97 (0.85, 1.11)	0.6609	0.90 (0.75, 1.07)	0.2347
30-Day readmission or death	21.5	18.9	1.19 (1.12, 1.26)	<0.0001	21.3	20.8	1.03 (0.96, 1.10)	0.4707	0.86 (0.79, 0.95)	0.0025
ICU usage	33.6	30.3	1.23 (1.16, 1.30)	<0.0001	33.2	30.6	1.17 (1.09, 1.26)	<0.0001	0.95 (0.87, 1.05)	0.3128
Prolonged stay	70.3	70.1	1.01 (0.95, 1.07)	0.7507	70.1	68.2	1.10 (1.03, 1.18)	0.0028	1.09 (1.00, 1.19)	0.0437
Anesthesia time (min, m-est.)	156.4	147.0	7.6 (5.5, 9.7)	<0.0001	165.0	145.6	18.3 (16.0, 20.6)	<0.0001	9.0 (6.0, 10.5)	<.0001
Length of stay (days, m-est.)	9.2	9.0	0.1 (0.0, 0.3)	0.0145	8.8	8.5	0.2 (0.0, 0.4)	0.0277	0.0 (0.0, 0.0)	0.5333
30-Day resource cost ($, m-est.)	25,271	23,703	1,363 (895, 1,833)	<0.0001	27,996	26,203	1650 (1035, 2268)	<0.0001	244 (−274, 762)	0.3565
Orthopedics										
No of surgeons	1201	1201			826	826				
No of patients	12,010	12,010			8260	8260				
30-Day mortality	6.5	6.0	1.08 (0.97, 1.20)	0.1599	5.3	4.7	1.12 (0.97, 1.30)	0.1149	1.04 (0.87, 1.24)	0.6861

Outcome (% Unless Noted)	(1) Traditional Era: 1999–2003 Match				(2) Modern Era: 2009–2013 Match				(2) vs. (1) D-in-D	
	New Surgeon-Patients	Exp. Surgeon-Patients	Odds Ratio or Paired Diff (95% CI)	*P*	New Surgeon-Patients	Exp. Surgeon-Patients	Odds Ratio or Paired Diff (95% CI)	*P*	Odds Ratio or Difference (95% CI)	*P*
30-Day failure-to-rescue	10.7	10.1	1.07 (0.93, 1.22)	0.3504	8.5	7.8	1.03 (0.86, 1.23)	0.7923	0.96 (0.77, 1.20)	0.7410
30-Day readmission or death	18.0	17.2	1.05 (0.98, 1.13)	0.1293	17.3	15.8	1.11 (1.02, 1.21)	0.0132	1.05 (0.95, 1.18)	0.3342
ICU usage	10.5	8.9	1.22 (1.11, 1.33)	<0.0001	10.2	8.5	1.26 (1.12, 1.40)	0.0001	1.03 (0.89, 1.19)	0.6901
Prolonged stay	56.2	54.3	1.10 (1.04, 1.16)	0.0012	41.1	37.9	1.18 (1.10, 1.27)	<0.0001	1.08 (0.98, 1.18)	0.1046
Anesthesia time (min, m-est.)	145.3	136.1	7.9 (6.3, 9.5)	<0.0001	147.7	127.7	19.2 (17.7, 20.8)	<0.0001	9.0 (7.5, 10.5)	<0.0001
Length of stay (days, m-est.)	6.8	6.7	0.1 (0.0, 0.2)	0.0068	5.9	5.8	0.1 (0.0, 0.2)	0.0017	0.0 (0.0, 0.0)	0.7436
30-Day resource cost ($, m-est.)	16,565	16,053	535 (328, 741)	<0.0001	17,044	16,042	916 (654, 1,177)	<0.0001	243 (7, 479)	0.0435

[*] Differences in binary outcomes within each era were assessed using McNemar test. Estimates of and differences in continuous outcomes within each era used weighted M-statistics. Difference-in-differences were tested using Gart test for binary outcomes, and the Wilcoxon rank sum test for continuous ones.

[†] Binary outcomes are reported using paired odds ratios and continuous outcomes are reported using paired differences.

[‡] Resource costs were calculated using 2013 dollars.

[§] See eTable 8 for additional secondary outcomes.