

Original Investigation

Association of the 2011 ACGME Resident Duty Hour Reform With General Surgery Patient Outcomes and With Resident Examination Performance

Ravi Rajaram, MD; Jeanette W. Chung, PhD; Andrew T. Jones, PhD; Mark E. Cohen, PhD; Allison R. Dahlke, MPH; Clifford Y. Ko, MD, MS, MSHS; John L. Tarpley, MD; Frank R. Lewis, MD; David B. Hoyt, MD; Karl Y. Bilimoria, MD, MS

IMPORTANCE In 2011, the Accreditation Council for Graduate Medical Education (ACGME) restricted resident duty hour requirements beyond those established in 2003, leading to concerns about the effects on patient care and resident training.

OBJECTIVE To determine if the 2011 ACGME duty hour reform was associated with a change in general surgery patient outcomes or in resident examination performance.

DESIGN, SETTING, AND PARTICIPANTS Quasi-experimental study of general surgery patient outcomes 2 years before (academic years 2009-2010) and after (academic years 2012-2013) the 2011 duty hour reform. Teaching and nonteaching hospitals were compared using a difference-in-differences approach adjusted for procedural mix, patient comorbidities, and time trends. Teaching hospitals were defined based on the proportion of cases at which residents were present intraoperatively. Patients were those undergoing surgery at hospitals participating in the American College of Surgeons National Surgical Quality Improvement Program (ACS NSQIP). General surgery resident performance on the annual in-training, written board, and oral board examinations was assessed for this same period.

EXPOSURES National implementation of revised resident duty hour requirements on July 1, 2011, in all ACGME accredited residency programs.

MAIN OUTCOMES AND MEASURES Primary outcome was a composite of death or serious morbidity; secondary outcomes were other postoperative complications and resident examination performance.

RESULTS In the main analysis, 204 641 patients were identified from 23 teaching (n = 102 525) and 31 nonteaching (n = 102 116) hospitals. The unadjusted rate of death or serious morbidity improved during the study period in both teaching (11.6% [95% CI, 11.3%-12.0%] to 9.4% [95% CI, 9.1%-9.8%], $P < .001$) and nonteaching hospitals (8.7% [95% CI, 8.3%-9.0%] to 7.1% [95% CI, 6.8%-7.5%], $P < .001$). In adjusted analyses, the 2011 ACGME duty hour reform was not associated with a significant change in death or serious morbidity in either postreform year 1 (OR, 1.12; 95% CI, 0.98-1.28) or postreform year 2 (OR, 1.00; 95% CI, 0.86-1.17) or when both postreform years were combined (OR, 1.06; 95% CI, 0.93-1.20). There was no association between duty hour reform and any other postoperative adverse outcome. Mean (SD) in-training examination scores did not significantly change from 2010 to 2013 for first-year residents (499.7 [85.2] to 500.5 [84.2], $P = .99$), for residents from other postgraduate years, or for first-time examinees taking the written or oral board examinations during this period.

CONCLUSIONS AND RELEVANCE Implementation of the 2011 ACGME duty hour reform was not associated with a change in general surgery patient outcomes or differences in resident examination performance. The implications of these findings should be considered when evaluating the merit of the 2011 ACGME duty hour reform and revising related policies in the future.

JAMA. 2014;312(22):2374-2384. doi:10.1001/jama.2014.15277

-  Editorial page 2342
-  Related articles pages 2364 and 2401
-  Supplemental content at jama.com

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Karl Y. Bilimoria, MD, MS, Surgical Outcomes and Quality Improvement Center (SOQIC), Department of Surgery and Center for Healthcare Studies, Feinberg School of Medicine, Northwestern University, 633 N St Clair St, 20th Floor, Chicago, IL 60611 (k-bilimoria@northwestern.edu).

Residency duty hour requirements have been under scrutiny for more than 30 years as concerns about medical errors attributable to exhausted residents persist.^{1,2} In 2003, the Accreditation Council for Graduate Medical Education (ACGME) enacted the first national duty hour requirements for all accredited residency programs, which required residents to work 80 hours or less per week averaged over 4 weeks; have 1 day free per week averaged over 4 weeks; have 10 hours off between shifts; not work more than 24 hours of continuous duty with 6 hours allowed for transfer of care activities; and take call no more than every third night averaged over 4 weeks.³

ABS American Board of Surgery

ABSITE American Board of Surgery In-Training Examination

ACGME Accreditation Council for Medical Education

ACS NSQIP American College of Surgeons National Surgical Quality Improvement Program

CPT Current Procedural Terminology

NQF National Quality Forum

PGY Postgraduate Year

SSI surgical-site infection

However, stakeholders continued to express concern about resident fatigue contributing to patient harm.⁴ Effective July 1, 2011, the ACGME further restricted resident duty hours to include the following provisions: first-year trainees limited to 16 hours of continuous in-hospital duty; residents must have at least 8 hours free between shifts; and residents in-house for 24 hours may have up to 4 hours for transfer of care activities and must have at least 14 hours off between shifts.³

These reforms have been met with mixed opinions. Proponents suggest they may reduce avoidable medical errors and improve resident well-being,⁵⁻⁷ whereas opponents note that disruptions to continuity of care may worsen outcomes and compromise resident education.^{8,9} Despite the need for rigorous evaluations of current duty hour policies, most studies to date have been single-institution reports, used administrative data, or focused on the 2003 duty hour reform.

The objectives of this study were to assess whether implementation of the 2011 ACGME duty hour reform was associated with a change in the composite measure of death or serious morbidity in patients undergoing general surgery operations, with other postoperative complications, or in differences in general surgery resident performance on national in-training and board certification examinations.

Methods

This study was deemed exempt by the Northwestern University institutional review board.

Data Source and Study Population

This was a retrospective observational study of the prospectively maintained American College of Surgeons National Surgical Quality Improvement Program (ACS NSQIP) database.^{10,11} In 2013, ACS NSQIP included 452 adult hospitals in the United States, accounting for nearly 10% of all

hospitals and 30% of all surgeries performed throughout the country.¹² Approximately 76% of hospitals in ACS NSQIP are affiliated with a general surgery residency program. The ACS NSQIP database collects surgical patient information from participating hospitals for the purposes of quality improvement. The details and history of this program have been described.^{10,11,13} Patient demographics, comorbidities, intraoperative variables, and detailed clinical outcomes are collected by trained on-site surgical clinical reviewers using data definitions that are standardized across participating institutions. Reviewers participate in frequent in-person training, online educational modules, and annual examinations. Hospitals are subject to detailed data audits to try to ensure high reliability between reviewers.¹⁴ Clinical outcomes are tracked from the date of surgery until 30 days after surgery. Relevant inpatient and outpatient information is captured based on medical record review and on discussions with physicians, contacting patients, or both.

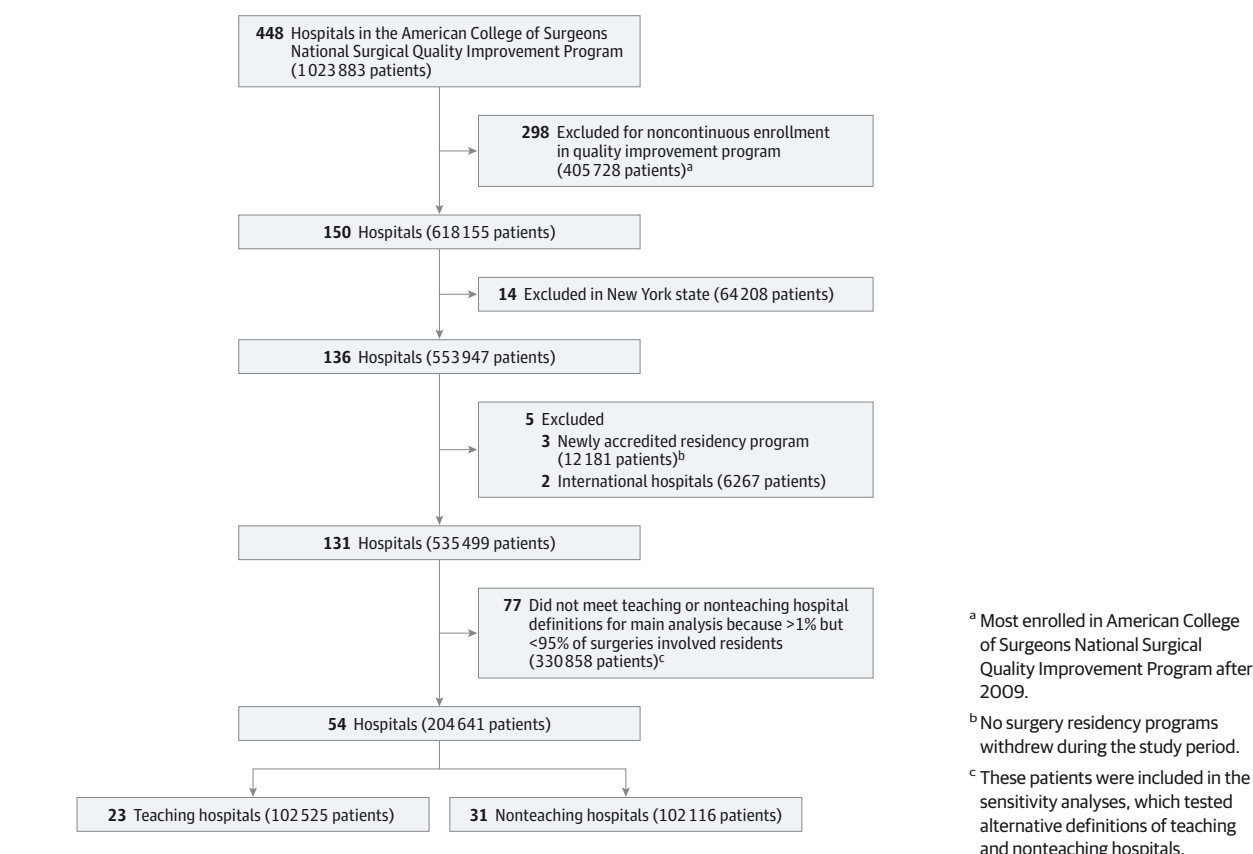
American College of Surgeons National Surgical Quality Improvement Program data were obtained from 2 years before to 2 years after the 2011 ACGME duty hour reform became effective on July 1, 2011. Patients who underwent any general surgery procedure between July 1, 2009, and June 30, 2013, were identified using the standard ACS NSQIP *Current Procedural Terminology* (CPT) codes for general surgery, compiled and vetted by surgeons and coding experts.¹⁵ We excluded hospitals that did not submit at least 25 general surgery cases per quarter to ensure only continuously enrolled sites were included (Figure 1). Patients from hospitals located in New York state were also excluded given preexisting statewide duty hour regulations.¹⁶ We also excluded patients from hospitals that had newly accredited general surgery residency programs, had withdrawn general surgery residence programs during the study period, or had and patients from international ACS NSQIP sites.¹⁷

Teaching vs Nonteaching Hospitals

Teaching status of a hospital was determined based on the proportion of general surgery cases in each hospital in which a resident was present intraoperatively (eTable 1 in the Supplement). Several different cutoffs were examined. To compare clearly contrasting groups for the main analysis, teaching hospitals were defined as those with at least 95% of operations involving a resident. Nonteaching hospitals were defined as having a resident present intraoperatively in 1% or fewer cases (Figure 1). Using these stringent criteria allowed for a comparison between 2 dissimilar groups with respect to surgical resident involvement. Previous studies have similarly compared the extremes of measures reflecting teaching status, such as the resident-to-bed ratio.^{18,19}

For comparison, we also evaluated the following alternative criteria of teaching vs nonteaching hospital status using the percentage of operations with resident involvement: 95% or more vs 5% or less, 90% or more vs 10% or less, 75% or more vs 1% or less, 75% or more vs 10% or less, 50% or more vs 1% or less, 25% or more vs 1% or less, and more than 1% vs 1% or less (eTable 2 in the Supplement).

Figure 1. Inclusion and Exclusion Criteria for Selecting Analytic Study Sample



Outcome Measures

The primary outcome was a composite measure of death or serious morbidity within 30 days of surgery, a National Quality Forum (NQF)-endorsed metric used by the Centers for Medicare & Medicaid Services (NQF, No. 0697) and the main outcome measure for ACS NSQIP.^{20,21} Secondary outcomes included death, serious morbidity, any morbidity, failure to rescue, surgical-site infection (SSI), and sepsis or septic shock within 30 days of surgery. The ACS NSQIP has greater accuracy for detecting postoperative complications when compared with administrative data, and interrater reliability over time has improved for these measures.^{14,22} *Serious morbidity* was defined using the standard ACS NSQIP definition and included one or more of the following complications: deep or organ-space SSI, wound or fascial dehiscence, pneumonia, unplanned intubation, prolonged need for mechanical ventilation (>48 hours), acute renal failure, urinary tract infection, cardiac arrest, myocardial infarction, sepsis or septic shock, or reoperation.²³ The ACS NSQIP definition of *any morbidity* includes all serious morbidity events, excluding reoperation, in addition to superficial SSI, cerebrovascular accident, and bleeding requiring transfusion (≥5 units in first 72 hours).²³ Failure to rescue indicates the proportion of patients who experienced a serious morbidity and subsequently died.²⁴

Statistical Analysis

Risk-adjusted event rates for teaching and nonteaching hospitals were calculated for each outcome using logistic regression models. Models included variables for 21 patient-level demographic factors and comorbidities, with adjustment for surgical case mix using the standard ACS NSQIP methodology (eStatistical Methods in the Supplement).¹¹ This method uses a hierarchical model for each dependent outcome with the CPT code as the random effect. A linear risk score is derived for each procedure and subsequently used to adjust for procedural differences between hospitals.²³ Any missing data are standardly imputed by ACS NSQIP for hospital-level modeling using multivariable regression with single imputation.

To determine the association between the 2011 ACGME duty hour reform and patient outcomes, we used a quasi-experimental study design known as a difference-in-differences analysis (eStatistical Methods in the Supplement).²⁵ This statistical approach is often used to assess changes that occur following natural experiments or policy changes. Using a difference-in-differences model, we compared patient outcomes in teaching hospitals before and after the 2011 ACGME duty hour reform to a contemporaneous control group of nonteaching hospitals during the same period. Because changes to duty hour require-

ments should not affect patient care at nonteaching hospitals, this comparison adjusts for time trends and other unmeasured variables that may affect both hospital groups similarly.

We estimated logistic regression models with robust cluster-corrected standard errors to account for patient clustering within hospitals. The association between duty hour reform and patient outcomes was identified by the difference between teaching and nonteaching hospitals in pre-post time differences. This required dummy variables indicating (1) whether surgery occurred at a teaching or nonteaching hospital and (2) if the surgery was before or after the duty hour reform was implemented. The interaction term of these 2 variables was the difference-in-differences estimator, and its coefficient reflected the magnitude of association between duty hour reform and the dependent outcome of interest. Models adjusted for patient demographics, comorbidities, and procedure type and also included variables indicating the quarter in which the surgery was performed to control for time trends.

We additionally investigated whether there may have been differential associations with the policy reform in 2 subgroup analyses. First, we tested whether there was a differential policy association in inpatients, defined as having a length of stay of at least 2 days, compared with all other patients. Second, we tested whether there was a differential policy association among high-risk individuals, defined as individuals in the highest quartile of expected probability of risk for each specific adverse outcome, compared with all other patients.²⁶ The expected risk probability for each individual was estimated from logistic regression models adjusted for patient demographic information, comorbidities, and procedure type. To test for subgroup associations, we ran separate logistic regression models for each subgroup analysis that included an interaction term between the subgroup indicator and the differences-in-differences estimator, all lower-order interaction terms, and additional patient-level covariates. This 3-way interaction is the estimator of difference in difference-in-differences, and estimates whether there was a differential association of duty hour reform across inpatients vs outpatients or high-risk vs low-risk patients (eStatistical Methods in the Supplement).²⁷

For each analysis, we pooled data from both postreform year 1 (July 1, 2011-June 30, 2012) and postreform year 2 (July 1, 2012-June 30, 2013) and compared this with the entire prereform period (July 1, 2009-June 30, 2011). Additionally, each postreform year was individually compared with the entire prereform period to test if outcome differences occurred either initially or over time after reform. We report 2-sided *P* values with the level of statistical significance set to .05. We did not adjust our level of significance for multiple comparisons because we hypothesized that duty hour reform was not associated with a change in outcomes; therefore, using uncorrected *P* values presents a more conservative approach. Statistical analyses were performed in SAS version 9.3 (SAS Institute Inc).

Additional Analyses

Difference-in-differences analyses assume that there is no significant change between groups relative to one another prior to the policy change.²⁸ Similar to previous studies, we tested this assumption by performing post hoc analyses for all outcomes comparing teaching and nonteaching hospitals in prereform year 1 with prereform year 2.^{18,19} Using the model previously described, we tested the significance of the interaction term during these periods. If a significant difference was detected, we retested our difference-in-differences model using prereform year 2, the year immediately preceding duty hour reform, as the reference group.

Sensitivity analyses were performed on the primary outcome measure of death or serious morbidity to test the robustness of our findings. This included (1) substituting alternative variables to control for time trends; (2) using resident-to-bed ratios derived from the 2010 American Hospital Association's Annual Survey; and (3) substituting a hierarchical model with the hospital as the random effect (eTable 2 in the Supplement). The resident-to-bed ratio is commonly used in studies evaluating duty hour reform.^{18,19} However, this measure provides only an overall hospital indicator of resident involvement and is not specific to a particular specialty or patient population but was included for comparison with previous research. Hospitals were divided into 1 of 5 groups based on the resident-to-bed ratio: (1) 0.000 (nonteaching); (2) 0.001-0.049 (very minor teaching); (3) 0.050-0.249 (minor teaching); (4) 0.250-0.599 (major teaching); and (5) 0.600 or higher (very major teaching). Teaching hospitals were compared with nonteaching hospitals using the previously described difference-in-differences model.

Resident Examination Performance

Duty hour reform may allow residents increased time to study for examinations that would result in better performance, or conversely duty hour limits may lead to less experienced residents and thus worse examination scores.^{29,30} To assess if the 2011 ACGME duty hour reform was associated with changes in resident education, we evaluated performance on national in-training, written board, and oral board examinations. The American Board of Surgery (ABS) In-Training Examination (ABSITE) is a multiple-choice test taken annually by general surgery residents.³¹ Scores are standardized each year for both the junior-level (postgraduate years [PGY] 1 and 2) and senior-level (PGY 3, 4, and 5) examinations to account for differences in test difficulty from year to year. To become board certified in general surgery, the ABS requires that candidates pass 2 examinations.³¹ The written board examination (Qualifying Examination) is an annual multiple-choice test offered to graduates of accredited residency programs. Successful completion of the written examination is required before individuals may take the oral board examination (Certifying Examination).

Categorical general surgery resident ABSITE scores and performance of first-time US-trained examinees on the written and oral board examinations were obtained from the ABS from 2 years before to 2 years after the 2011 duty hour reform became effective. We tested ABSITE scores for a longitudinal trend

for each postgraduate year using a linear contrast test within an analysis of variance to determine if there were differences in examination performance after duty hour reform. The percentages of individuals passing the written and oral board examinations were compared over time using the Cochran-Armitage trend test. Statistical analyses were performed in SPSS version 22 (IBM Corp).

Results

Patient Outcomes

Of the 1 023 883 patients from 448 ACS NSQIP hospitals, those that were not continuously enrolled from 2009 to 2013 were excluded (298 hospitals, 405 728 patients) (Figure 1). Next, hospitals were excluded if located in New York state (14 hospitals, 64 208 patients), had a new general surgery residency accreditation (3 hospitals, 12 181 patients), or were not in the United States (2 hospitals, 6267 patients). Thus, 535 499 patients from 131 hospitals met the inclusion criteria. Hospitals were then classified into teaching vs nonteaching hospitals using different definitions (eTable 1 in the Supplement). For the main analyses, teaching hospitals were defined as those with a resident involved in at least 95% of general surgery cases and nonteaching hospitals as those with a resident involved in 1% or fewer cases (54 hospitals, 204 641 patients). Patients were nearly evenly divided between the 23 teaching ($n = 102\,525$) and 31 nonteaching ($n = 102\,116$) hospitals. Patient characteristics are reported in Table 1 and the most frequently performed operations are summarized in eTable 3 in the Supplement.

The unadjusted rate of the death or serious morbidity composite outcome improved from prereform year 1 to postreform year 2 in both teaching (11.6% [95% CI, 11.3%-12.0%] to 9.4% [95% CI, 9.1%-9.8%]; $P < .001$) and nonteaching hospitals (8.7% [95% CI, 8.3%-9.0%] to 7.1% [95% CI, 6.8%-7.5%]; $P < .001$; Table 2). Unadjusted rates of any morbidity, failure to rescue, SSI, and sepsis or septic shock remained relatively stable in both hospital groups over time. Adverse event rates adjusted for patient demographics, comorbidities, and procedural case mix for teaching and nonteaching hospitals are shown in Figure 2.

On adjusted analyses comparing outcomes before and after the 2011 ACGME duty hour reform at teaching and nonteaching hospitals, reform was not associated with a change in the likelihood of death or serious morbidity (odds ratio [OR], 1.06; 95% CI, 0.93-1.20; Table 3). Duty hour reform was also not associated with early or delayed differences in death or serious morbidity when evaluating postreform year 1 (OR, 1.12; 95% CI, 0.98-1.28) and postreform year 2 (OR, 1.00; 95% CI, 0.86-1.17) separately. Moreover, duty hour reform was not associated with a significant change in any of the other adverse outcomes examined (Table 3).

As a sensitivity test, we compared less restrictive definitions of teaching and nonteaching hospital status using the percentage of general surgery cases in which a resident was involved. This resulted in more hospitals and patients being included in the models (eTable 2 in the Supplement). Using

these alternative definitions of teaching status, there were no significant associations between duty hour reform and death or serious morbidity in postreform year 1 or postreform year 2 or when both postreform years were combined (eTable 2 in the Supplement). In additional analyses, there was no significant association between duty hour reform and death or serious morbidity when comparing inpatients with outpatients for the entire postreform period (coefficient, 0.10; 95% CI, -0.12 to 0.31; $P = .38$) or when evaluating each postreform year separately (eTables 4, 5, and 8 in the Supplement). When high-risk patients were compared with low-risk patients, there was also no significant differential association of duty hour reform with death or serious morbidity or any other adverse outcome examined for any postreform period (eTable 6, 7, and 8 in the Supplement).

On post hoc analysis, when adverse outcomes at teaching and nonteaching hospitals were compared between prereform year 1 and 2, a significant preexisting difference was found only for any morbidity ($P = .02$). However, on subsequent analyses using prereform year 2 as the reference group, there was no association between duty hour reform and any morbidity when comparing the entire postreform period, postreform year 1, or postreform year 2.

Additional sensitivity analyses were performed. When controlling for time trends with the academic year rather than the quarter in which surgery was performed, our results were unchanged (eTable 2 in the Supplement). We also evaluated teaching and nonteaching hospitals using the hospital-level resident-to-bed ratio. On adjusted analyses comparing very minor (0.001-0.049), minor (0.050-0.249), and major teaching (0.250-0.599) hospitals with nonteaching (0.000) hospitals, there was no association between duty hour reform and death or serious morbidity. When very major teaching (≥ 0.600) hospitals were compared with nonteaching hospitals, duty hour reform was associated with a significant increase in death or serious morbidity in postreform year 1 (OR, 1.13; 95% CI, 1.00-1.26), but not in the other postreform periods examined. When a hierarchical model with the hospital as the random effect was used, duty hour reform was again associated with an increased likelihood of death or serious morbidity in the first year after reform (OR, 1.13; 95% CI, 1.02-1.29), but not in postreform year 2 or when pooling both postreform years together.

Resident Examination Performance

Mean (SD) ABSITE scores for PGY-1 residents did not significantly change from 2010 to 2013 (499.7 [85.2] to 500.5 [84.2], $P = .99$) when examining all programs. Nonsignificant changes in ABSITE scores were also found in other PGY classes during this period (Table 4).

The percentage of first-time examinees passing the written board examination significantly increased from 2010 to 2013 (83.1% [95% CI, 80.8%-85.3%] to 88.1% [95% CI, 86.1%-90.0%], $P < .001$). When only the 2011 to 2013 period was examined, the findings were no longer significant (87.5% [95% CI, 85.5%-89.5%] to 88.1% [95% CI, 86.1%-90.0%], $P = .41$). There was no change in oral board examination pass rates over time (81.7% [95% CI, 79.2%-84.2%] to 80.9% [95% CI, 78.7%-83.2%], $P = .21$) (Table 4).

Table 1. Patient Characteristics at Teaching and Nonteaching Hospitals ^a

	Duty Hour Reform, No. (%) of Patients							
	Before Reform (Academic Year)				After Reform (Academic Year)			
	Year 1 (2009-2010)		Year 2 (2010-2011)		Year 1 (2011-2012)		Year 2 (2012-2013)	
	Teaching (n = 25 949)	Nonteaching (n = 26 579)	Teaching (n = 25 968)	Nonteaching (n = 25 830)	Teaching (n = 25 587)	Nonteaching (n = 24 763)	Teaching (n = 25 021)	Nonteaching (n = 24 944)
Age, mean (SD), y	52.9 (16.4)	54.6 (17.1)	53.1 (16.3)	55.0 (17.1)	53.4 (16.3)	54.7 (17.1)	53.7 (16.2)	54.9 (17.0)
Women	15 419 (59.4)	16 009 (60.2)	15 695 (60.4)	15 724 (60.9)	14 923 (58.3)	14 688 (59.3)	14 471 (57.8)	14 882 (59.7)
BMI								
<18.5	596 (2.3)	475 (1.8)	592 (2.3)	412 (1.6)	525 (2.1)	385 (1.6)	509 (2.0)	396 (1.6)
18.5-24.9	6786 (26.2)	6304 (23.7)	6952 (26.8)	6208 (24.0)	6795 (26.6)	5981 (24.2)	6560 (26.2)	5927 (23.8)
25.0-29.9	7882 (30.4)	7904 (29.7)	7779 (30.0)	7685 (29.8)	7707 (30.1)	7323 (29.6)	7553 (30.2)	7373 (29.6)
30.0-34.9	5089 (19.6)	5243 (19.7)	4928 (19.0)	5151 (19.9)	4929 (19.3)	4959 (20.0)	5028 (20.1)	5036 (20.2)
>35.0	5596 (21.6)	6653 (25.0)	5717 (22.0)	6374 (24.7)	5631 (22.0)	6115 (24.7)	5371 (21.5)	6212 (24.9)
Smoker in past year	5127 (19.8)	4847 (18.2)	4877 (18.8)	4736 (18.3)	4751 (18.6)	4474 (18.1)	4501 (18.0)	4504 (18.1)
ASA class ^b								
I/II	15 850 (61.1)	16 309 (61.4)	15 073 (58.0)	15 415 (59.7)	14 219 (55.6)	14 847 (60.0)	13 824 (55.2)	14 827 (59.4)
III	8869 (34.2)	9184 (34.6)	9642 (37.1)	9233 (35.7)	10 100 (39.5)	8871 (35.8)	10 063 (40.2)	8989 (36.0)
IV/V	1230 (4.7)	1086 (4.1)	1253 (4.8)	1182 (4.6)	1268 (5.0)	1045 (4.2)	1134 (4.5)	1128 (4.5)
Preoperative functional status								
Independent	24 608 (94.8)	25 536 (96.1)	24 710 (95.2)	24 971 (96.7)	24 973 (97.6)	24 093 (97.3)	24 529 (98.0)	24 334 (97.6)
Partially dependent	808 (3.1)	745 (2.8)	849 (3.3)	653 (2.5)	458 (1.8)	538 (2.2)	396 (1.6)	476 (1.9)
Totally dependent	533 (2.1)	298 (1.1)	409 (1.6)	206 (0.8)	156 (0.6)	132 (0.5)	96 (0.4)	134 (0.5)
Preoperative history								
Hypertension	10 774 (41.5)	11 727 (44.1)	10 852 (41.8)	11 404 (44.2)	10 924 (42.7)	10 779 (43.5)	10 509 (42.0)	10 775 (43.2)
CHF	141 (0.5)	133 (0.5)	153 (0.6)	114 (0.4)	124 (0.5)	136 (0.5)	114 (0.5)	135 (0.5)
Dyspnea	2231 (8.6)	2504 (9.4)	2099 (8.1)	2519 (9.8)	1678 (6.6)	2065 (8.3)	1348 (5.4)	1674 (6.7)
COPD	993 (3.8)	1053 (4.0)	1062 (4.1)	1135 (4.4)	1011 (4.0)	1074 (4.3)	903 (3.6)	999 (4.0)
Ventilator dependent	351 (1.4)	147 (0.6)	307 (1.2)	134 (0.5)	241 (0.9)	114 (0.5)	150 (0.6)	89 (0.4)
Diabetes								
Oral	2119 (8.2)	2440 (9.2)	2174 (8.4)	2508 (9.7)	2215 (8.7)	2365 (9.6)	2142 (8.6)	2429 (9.7)
Insulin	1320 (5.1)	1261 (4.7)	1419 (5.5)	1264 (4.9)	1520 (5.9)	1305 (5.3)	1500 (6.0)	1338 (5.4)
Corticosteroid use	942 (3.6)	619 (2.3)	1073 (4.1)	556 (2.2)	1197 (4.7)	716 (2.9)	1263 (5.0)	753 (3.0)
Albumin <3.0 g/dL	2009 (7.7)	1443 (5.4)	2130 (8.2)	1521 (5.9)	1991 (7.8)	1472 (5.9)	1766 (7.1)	1385 (5.6)
Ascites	286 (1.1)	202 (0.8)	260 (1.0)	168 (0.7)	235 (0.9)	150 (0.6)	217 (0.9)	147 (0.6)
Renal failure/dialysis	429 (1.7)	376 (1.4)	421 (1.6)	379 (1.5)	465 (1.8)	385 (1.6)	402 (1.6)	441 (1.8)
Bleeding disorder	833 (3.2)	940 (3.5)	917 (3.5)	943 (3.7)	753 (2.9)	897 (3.6)	691 (2.8)	869 (3.5)
SIRS/sepsis/septic shock	2506 (9.7)	2601 (9.8)	2491 (9.6)	2419 (9.4)	2172 (8.5)	2169 (8.8)	1811 (7.2)	2236 (9.0)
Disseminated cancer	767 (3.0)	377 (1.4)	791 (3.0)	450 (1.7)	920 (3.6)	427 (1.7)	976 (3.9)	471 (1.9)
Wound class								
Clean	11 522 (44.4)	11 379 (42.8)	11 274 (43.4)	11 045 (42.8)	10 818 (42.3)	10 689 (43.2)	10 849 (43.4)	10 449 (41.9)
Clean-contaminated	9035 (34.8)	10 504 (39.5)	9214 (35.5)	10 048 (38.9)	9404 (36.8)	9418 (38.0)	9245 (36.9)	9654 (38.7)
Contaminated	3032 (11.7)	2749 (10.3)	3182 (12.3)	2885 (11.2)	3094 (12.1)	2866 (11.6)	2978 (11.9)	2824 (11.3)
Dirty-infected	2360 (9.1)	1947 (7.3)	2298 (8.8)	1852 (7.2)	2271 (8.9)	1790 (7.2)	1949 (7.8)	2017 (8.1)
Emergency operations	4033 (15.5)	3868 (14.6)	3986 (15.3)	3599 (13.9)	3780 (14.8)	3460 (14.0)	2979 (11.9)	3458 (13.9)

Abbreviations: ASA, American Society of Anesthesiology; BMI, body mass index, calculated as weight in kilograms divided by height in meters squared; CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; SIRS, systemic inflammatory response syndrome.

^a See Methods section for hospital definitions. There were a total of 54

hospitals, 23 teaching and 31 nonteaching, included for all 4 years of the study period in the main analysis.

^b The ASA class I indicates healthy; II, mild systemic disease; III, severe systemic disease; IV, severe life-threatening systemic disease; and V, moribund not expected to survive without the operation.¹⁵

Table 2. Unadjusted Adverse Outcome Rates in Patients Before and After the 2011 Duty Hour Reform at Teaching and Nonteaching Hospitals

	Duty Hour Reform, No. of Events/No. of Patients (%)			
	Before Reform (Academic Year)		After Reform (Academic Year)	
	Year 1 (2009-2010)	Year 2 (2010-2011)	Year 1 (2011-2012)	Year 2 (2012-2013)
Death or serious morbidity				
Teaching	3021/25 949 (11.6)	2839/25 968 (10.9)	2669/25 587 (10.4)	2364/25 021 (9.4)
Nonteaching	2305/26 579 (8.7)	2152/25 830 (8.3)	1734/24 763 (7.0)	1778/24 944 (7.1)
Death				
Teaching	380/25 949 (1.5)	357/25 968 (1.4)	353/25 587 (1.4)	260/25 021 (1.0)
Nonteaching	349/26 579 (1.3)	322/25 830 (1.2)	281/24 763 (1.1)	261/24 944 (1.0)
Serious morbidity				
Teaching	2905/25 949 (11.2)	2720/25 968 (10.5)	2560/25 587 (10.0)	2300/25 021 (9.2)
Nonteaching	2191/26 579 (8.2)	2046/25 830 (7.9)	1617/24 763 (6.5)	1694/24 944 (6.8)
Any morbidity				
Teaching	3514/25 949 (13.5)	3763/25 968 (14.5)	3798/25 587 (14.8)	3484/25 021 (13.9)
Nonteaching	2295/26 579 (8.6)	2579/25 830 (10.0)	2384/24 763 (9.6)	2269/24 944 (9.1)
Failure to rescue				
Teaching	264/2905 (9.1)	238/2720 (8.8)	244/2560 (9.5)	196/2300 (8.5)
Nonteaching	235/2191 (10.7)	216/2046 (10.6)	164/1617 (10.1)	177/1694 (10.4)
Surgical site infection				
Teaching	1426/25 949 (5.5)	1470/25 968 (5.7)	1501/25 587 (5.9)	1340/25 021 (5.4)
Nonteaching	959/26 579 (3.6)	931/25 830 (3.6)	851/24 763 (3.4)	872/24 944 (3.5)
Sepsis or septic shock				
Teaching	851/25 949 (3.3)	942/25 968 (3.6)	1032/25 587 (4.0)	905/25 021 (3.6)
Nonteaching	539/26 579 (2.0)	561/25 830 (2.2)	484/24 763 (2.0)	626/24 944 (2.5)

Discussion

In this study, the 2011 ACGME duty hour reform was not associated with a change in general surgery patient outcomes or differences in resident in-training, written board, or oral board examination performance. To our knowledge, this is one of the first national empirical evaluations of the 2011 ACGME duty hour reform. Our results may be considered when examining the merit of this reform and to inform future duty hour policies.

Duty Hour Reform and Clinical Outcomes

Previous studies evaluating patient outcomes before and after the 2003 ACGME duty hour reform found that reform was not associated with differences in surgical mortality.^{18,19} Several recent meta-analyses and systematic reviews have also concluded that duty hour reforms were not associated with improvements in surgical care, with some supporting possibly worsened outcomes.^{2,32,33} Using clinically collected data, we found that there was no association between the 2011 ACGME duty hour reform and outcomes in general surgery patients.

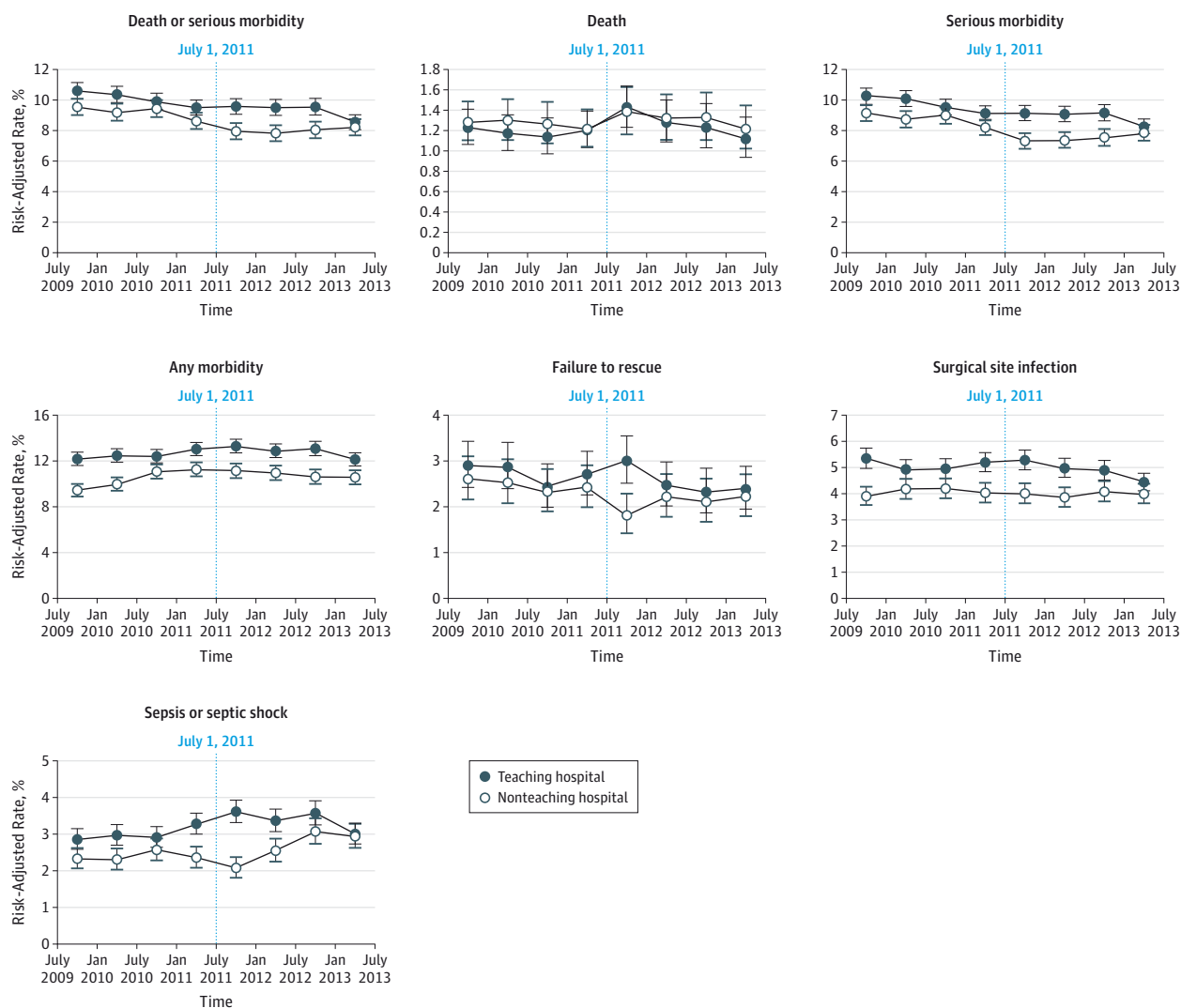
Some have suggested that outcomes for high-risk patients may better reflect resident involvement and be more sensitive to duty hour reforms.^{26,34} Volpp et al²⁶ examined Medicare and Veterans Affairs patients in the highest quartile of risk and found no consistent association between the 2003 reform and mortality or failure to rescue. In our study, we

also found that duty hour reform was not associated with a differential association in high-risk vs low-risk patients. Although continuity of care may be negatively affected by duty hour policies, this may be mitigated by hospitals anticipating the effects of these regulations and responding accordingly (eg, hiring midlevel support staff, formalized training in patient handoffs, etc).

Duty Hour Reform and Resident Examination Performance

The few single-institution studies comparing surgical resident in-training examination performance before and after the 2003 ACGME duty hour reform have demonstrated conflicting results.^{29,30} We found that there were no changes in in-training examination performance for all general surgery programs in the United States during this period. Moreover, first-year trainees who were most directly affected by the 2011 reform did not improve their ABSITE scores, despite presumably more free time to prepare.

Prior evaluations of board certification examinations demonstrated that written examination failure rates were relatively stable over time whereas oral board failure rates had increased.^{35,36} We found no significant change over time for written board examination and oral board examination pass rates. Although there was a significant increase in written board examination scores over the entire period, this was predominantly due to the anomalous poor performance of 2010 examinees. When comparing written examination performance the year before reform (2011) with those after the reform (2012-2013), pass rates appeared to be stable. A significant differ-

Figure 2. Risk-Adjusted Rates of Adverse Outcomes Before and After the 2011 Duty Hour Reform at Teaching vs Nonteaching Hospitals^a

^a Each data point represents 6 months of data during the 4-year period. The error bars indicate the 95% confidence intervals. The dotted blue line represents the period in which the 2011 duty hour reform was implemented,

beginning July 1, 2011. Models are adjusted for patient demographics, comorbidities, and differences in procedure mix (see eStatistical Methods in the Supplement).

ence in board examination performance was not expected because these examinees had only trained for 1 or 2 years under the 2011 duty hour reform. Nevertheless, the 2011 duty hour reform was not associated with an improvement in these standard measures of resident education.

Limitations

There are several possible limitations of our study. First, we were only able to assess the first 2 years following duty hour reform. There may be differences in patient care or resident examination performance that are evident only several years after implementation and adoption of new duty hour requirements. Second, to be able to examine outcomes before and after duty hour reform using clinical data, we were limited to hospitals participating in ACS NSQIP. These hospitals likely have more resources and possibly greater

interest in quality improvement than other institutions. As such, more financially-strained hospitals may be less able to mitigate the effects of duty hour reform through the use of midlevel practitioners and ancillary staff and thus may be more sensitive to duty hour reforms. Third, although we were able to risk-adjust using detailed clinical data and control for unmeasured changes using a difference-in-differences approach, our study was still observational. Any unobserved confounders that influenced patient outcomes differentially at teaching or nonteaching hospitals may have biased our findings. Fourth, although we found no statistically significant association between the 2011 ACGME duty hour reform and patient outcomes, some confidence intervals are fairly wide and encompass ranges of values that may be clinically important. Nevertheless, our results were unchanged in sensitivity analyses that were based on a larger

Table 3. Adjusted Analyses of Association Between Adverse Outcomes in Teaching Hospitals vs Nonteaching Hospitals After the 2011 Duty Hour Reform

	Postreform Years 1 and 2 Combined			Postreform Year 1			Postreform Year 2		
	No. of Patients	OR (95% CI) ^a	P Value	No. of Patients	OR (95% CI) ^a	P Value	No. of Patients	OR (95% CI) ^a	P Value
Death or serious morbidity	204 641	1.06 (0.93-1.20)	.37	154 676	1.12 (0.98-1.28)	.09	154 291	1.00 (0.86-1.17)	.96
Death	204 641	1.05 (0.86-1.28)	.62	154 676	1.11 (0.88-1.41)	.38	154 291	0.99 (0.75-1.29)	.92
Serious morbidity	204 641	1.07 (0.94-1.21)	.31	154 676	1.13 (0.99-1.30)	.07	154 291	1.01 (0.86-1.17)	.93
Any morbidity ^b	204 641	0.98 (0.88-1.10)	.72	154 676	0.98 (0.88-1.11)	.79	154 291	0.98 (0.85-1.13)	.78
Failure to rescue ^c	18 033	1.05 (0.85-1.30)	.65	14 039	1.13 (0.88-1.45)	.35	13 856	0.98 (0.72-1.32)	.88
Surgical-site infection	204 641	0.97 (0.84-1.14)	.74	154 676	1.05 (0.89-1.23)	.58	154 291	0.91 (0.77-1.08)	.29
Sepsis or septic shock	204 641	1.02 (0.79-1.31)	.89	154 676	1.23 (0.97-1.56)	.09	154 291	0.85 (0.62-1.17)	.32

^a Odds ratios (ORs) are determined from coefficient of interaction term (teaching hospital status × postreform period) with 3 postreform periods evaluated relative to the prereform period: postreform year 1 and 2 combined, postreform year 1 only, and postreform year 2 only.

^b Significant difference in trend between teaching and nonteaching hospitals in the 2 years before duty hour reform. Analysis shown herein presents

prereform year 1 and year 2 combined. In post hoc analyses using prereform year 2 as the reference year, there were also no significant associations between duty hour reform and any morbidity.

^c Fewer patients compared with other outcomes as the denominator only includes individuals who experienced a serious morbidity.

Table 4. Performance of Categorical General Surgery Residents on the American Board of Surgery In-Training Examination, Written Board Examination, and Oral Board Examination from All Programs in the United States

	Before Reform				After Reform				P Value ^a
	2010		2011		2012		2013		
	No. of Examinees	Mean (SD)	No. of Examinees	Mean (SD)	No. of Examinees	Mean (SD)	No. of Examinees	Mean (SD)	
ABSITE Scores ^b									
Postgraduate year									
1	1151	499.7 (85.2)	1180	500.7 (86.3)	1224	498.5 (85.8)	1259	500.5 (84.2)	.99
2	1400	545.7 (74.2)	1177	538.3 (77.2)	1191	538.5 (76.7)	1253	542.7 (74.7)	.36
3	1163	480.2 (104.5)	1149	477.6 (103.1)	1194	468.4 (103.3)	1669	475.7 (103.1)	.07
4	1060	509.5 (93.1)	1099	518.3 (92.2)	1094	513.2 (89.6)	1154	510.4 (91.9)	.84
5	1014	527.7 (88.6)	1040	518.1 (90.7)	1080	537.5 (84.6)	1113	528.2 (87.1)	.08
General Surgery Board Certification-Pass Rate									
	No. of Examinees	No. (%) Passing, [95% CI]	No. of Examinees	No. (%) Passing, [95% CI]	No. of Examinees	No. (%) Passing, [95% CI]	No. of Examinees	No. (%) Passing, [95% CI]	
Written examination ^c	1063	883 (83.1), [80.8-85.3]	1038	908 (87.5), [85.5-89.5]	1082	959 (88.6), [86.7-90.5]	1097	966 (88.1), [86.1-90.0]	<.001
Oral examination	906	740 (81.7), [79.2-84.2]	1086	811 (74.7), [72.1-77.3]	1046	863 (82.5), [80.2-84.8]	1154	934 (80.9), [78.7-83.2]	.21

Abbreviation: ABSITE, American Board of Surgery In-Training Examination.

^a ABSITE scores were compared for each postgraduate year class from 2010 through 2013 using a linear contrast test within an analysis of variance. The percentage of individuals who passed the written and oral board examination were compared from 2010 through 2013 using the Cochran-Armitage trend test.

^b ABSITE scores are nationally standardized each year to an overall mean of 500

and standard deviation of 100 for both the junior-level (postgraduate years 1 and 2) and senior-level (postgraduate years 3, 4, and 5) examinations.

^c Although the trend over time for written board examination pass rates was significant, this was due to unusually poor performance in 2010. When the postreform years are compared with 2011 alone, there was not a significant difference in written board examination pass rates ($P = .41$).

sample size (131 hospitals and 535 499 patients) than the main analysis.

Implications

The study findings could be interpreted in at least 2 ways. First, there is no evidence of worsened patient care or resident education, and given assumed improvements to resi-

dent well-being, this could indicate that current policies should continue forward as they are. Conversely, the potential harm from poor continuity of care, increased handoffs, trainees feeling unprepared to practice, and concern regarding residents developing a shift-work mentality engendered by these policies could suggest that the duty hour reform may require significant revision or reconsideration.^{8,37-39}

Although many of these concerns have not been substantiated by consistent evidence, they reflect the intense interest duty hour reform has generated from the clinical and educational community.

The issues surrounding duty hour reform may have implications for patient care and resident training, and the lack of high-level evidence to guide policy decisions in this area needs to be addressed with randomized trials. To that end, a national multicenter cluster randomized trial is being conducted (the Flexibility In Duty Hour Requirements for Surgical Trainees (FIRST) Trial, <http://www.Clinicaltrials.gov/NCT02050789>) comparing current duty hour requirements vs flexible duty hours to assess the effects of this interven-

tion on patient outcomes and resident well-being.⁴⁰ This trial may further inform the debate of how to optimally structure postgraduate training.

Conclusions

Implementation of the 2011 ACGME duty hour reform was not associated with a change in surgical patient outcomes or resident examination performance. The implications of these findings should be considered when evaluating the merit of the 2011 ACGME duty hour reform and revising related policies in the future.

ARTICLE INFORMATION

Author Affiliations: Division of Research and Optimal Patient Care, American College of Surgeons, Chicago, Illinois (Rajaram, Cohen, Ko, Hoyt, Bilimoria); Surgical Outcomes and Quality Improvement Center (SOQIC), Department of Surgery and Center for Healthcare Studies in the Institute for Public Health and Medicine, Feinberg School of Medicine, Northwestern University, Chicago, Illinois (Rajaram, Chung, Dahlke, Bilimoria); American Board of Surgery, Philadelphia, Pennsylvania (Jones, Lewis); Department of Surgery, University of California, Los Angeles (Ko); VA Greater Los Angeles Healthcare System, Los Angeles, California (Ko); Department of Surgery, Vanderbilt University Medical Center, Nashville, Tennessee (Tarpley).

Author Contributions: Dr Rajaram had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: All authors.

Acquisition, analysis, or interpretation of data: All authors.

Drafting of the manuscript: Rajaram, Chung, Jones, Bilimoria.

Critical revision of the manuscript for important intellectual content: All authors.

Statistical analysis: Rajaram, Chung, Jones, Cohen, Bilimoria.

Obtained funding: Bilimoria.

Administrative, technical, or material support: Bilimoria.

Study supervision: Ko, Tarpley, Lewis, Hoyt, Bilimoria.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Bilimoria reported support from the National Institutes of Health, Agency for Healthcare Research and Quality, American Board of Surgery, American College of Surgeons, Accreditation Council for Graduate Medical Education, National Comprehensive Cancer Network, American Cancer Society, Health Care Services Corp, California Health Care Foundation, Northwestern University, the Robert H. Lurie Comprehensive Cancer Center, Northwestern Memorial Foundation, and Northwestern Memorial Hospital. Dr Bilimoria has received honoraria from hospitals, professional societies, and continuing medical education companies for clinical care and quality improvement research presentations.

Funding/Support: Dr Rajaram is supported by grant T32HS000078 from the Agency for Healthcare Research and Quality, the American College of Surgeons Clinical Scholars in Residence Program, and an unrestricted educational grant from Merck.

Role of the Funder/Sponsor: The funding sponsors had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Correction: This article was corrected on December 18, 2014, to reorder footnotes in Table 4.

REFERENCES

1. Bell BM. The new hospital code and the supervision of residents. *N Y State J Med*. 1988;88(12):617-619.
2. Philibert I, Nasca T, Brigham T, Shapiro J. Duty-hour limits and patient care and resident outcomes: can high-quality studies offer insight into complex relationships? *Annu Rev Med*. 2013; 64:467-483.
3. Resident duty hours in the learning and working environment. <https://www.acgme.org/acgme/Portals/0/PDFs/dh-ComparisonTable2003v2011.pdf>. Accessed July 18, 2014.
4. *Resident Duty Hours: Enhancing Sleep, Supervision, and Safety*. Washington, DC: Institute of Medicine; 2008.
5. Landrigan CP, Rothschild JM, Cronin JW, et al. Effect of reducing interns' work hours on serious medical errors in intensive care units. *N Engl J Med*. 2004;351(18):1838-1848.
6. Lockley SW, Cronin JW, Evans EE, et al; Harvard Work Hours, Health and Safety Group. Effect of reducing interns' weekly work hours on sleep and attentional failures. *N Engl J Med*. 2004;351(18): 1829-1837.
7. Fletcher KE, Reed DA, Arora VM. Patient safety, resident education and resident well-being following implementation of the 2003 ACGME duty hour rules. *J Gen Intern Med*. 2011;26(8):907-919.
8. Antiel RM, Thompson SM, Hafferty FW, et al. Duty hour recommendations and implications for meeting the ACGME core competencies: views of residency directors. *Mayo Clin Proc*. 2011;86(3):185-191.
9. Desai SV, Feldman L, Brown L, et al. Effect of the 2011 vs 2003 duty hour regulation-compliant models on sleep duration, trainee education, and continuity of patient care among internal medicine house staff: a randomized trial. *JAMA Intern Med*. 2013;173(8):649-655.
10. Ingraham AM, Richards KE, Hall BL, Ko CY. Quality improvement in surgery: the American College of Surgeons National Surgical Quality Improvement Program approach. *Adv Surg*. 2010; 44:251-267.
11. Cohen ME, Ko CY, Bilimoria KY, et al. Optimizing ACS NSQIP modeling for evaluation of surgical quality and risk: patient risk adjustment, procedure mix adjustment, shrinkage adjustment, and surgical focus. *J Am Coll Surg*. 2013;217(2):336-346.
12. American Hospital Association. Fast FACTS on US hospitals. <http://www.aha.org/research/rc/stat-studies/fast-facts.shtml>. Accessed September 11, 2014.
13. Khuri SF, Henderson WG, Daley J, et al; Principal Investigators of the Patient Safety in Surgery Study. Successful implementation of the Department of Veterans Affairs' National Surgical Quality Improvement Program in the private sector: the Patient Safety in Surgery study. *Ann Surg*. 2008; 248(2):329-336.
14. Shiloach M, Frencher SK Jr, Steeger JE, et al. Toward robust information: data quality and inter-rater reliability in the American College of Surgeons National Surgical Quality Improvement Program. *J Am Coll Surg*. 2010;210(1):6-16.
15. *Chapter 4. ACS NSQIP Variables & Definitions*. Chicago, Illinois: American College of Surgeons National Surgical Quality Improvement Program; 2013.
16. Ritterband AB. The New York State Department of Health's Part 405 regulations. *N Y State J Med*. 1990;90(11):565-566.
17. Accreditation Council for Graduate Medical Education (ACGME)—Public. <http://www.acgme.org/ads/public>. Accessed May 15, 2014.
18. 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(9):984-992.
19. 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(9): 975-983.
20. National Quality Forum. Measuring performance. http://www.qualityforum.org/Measuring_Performance/Measuring_Performance.aspx. Accessed July 10, 2014.

21. Hospital Compare. Medicare.gov. <http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalCompare.html>. Accessed July 12, 2014.
22. Lawson EH, Louie R, Zingmond DS, et al. A comparison of clinical registry versus administrative claims data for reporting of 30-day surgical complications. *Ann Surg*. 2012;256(6):973-981.
23. ACS NSQIP Semiannual Report: January 1, 2013 through December 31, 2013. American College of Surgeons National Surgical Quality Improvement Program; 2014. Working report.
24. Silber JH, Williams SV, Krakauer H, Schwartz JS. Hospital and patient characteristics associated with death after surgery: a study of adverse occurrence and failure to rescue. *Med Care*. 1992;30(7):615-629.
25. Woolridge JM. *Introductory Econometrics: A Modern Approach*. 3rd ed. Mason, OH: Thomson/South-Western; 2006.
26. Volpp KG, Rosen AK, Rosenbaum PR, et al. Did duty hour reform lead to better outcomes among the highest risk patients? *J Gen Intern Med*. 2009;24(10):1149-1155.
27. Wang R, Lagakos SW, Ware JH, Hunter DJ, Drazen JM. Statistics in medicine—reporting of subgroup analyses in clinical trials. *N Engl J Med*. 2007;357(21):2189-2194.
28. Angrist JD, Krueger AB. Empirical strategies in labor economics. In: Card D, Ashenfelter O, eds. Amsterdam, the Netherlands: *Handbook of Labor Economics*. Vol 3. Elsevier Science B.V; 1999.
29. Durkin ET, McDonald R, Munoz A, Mahvi D. The impact of work hour restrictions on surgical resident education. *J Surg Educ*. 2008;65(1):54-60.
30. de Virgilio C, Yaghoobian A, Lewis RJ, Stabile BE, Putnam BA. The 80-hour resident workweek does not adversely affect patient outcomes or resident education. *Curr Surg*. 2006;63(6):435-439.
31. The American Board of Surgery. General surgery examinations. <http://www.absurgery.org/default.jsp?examoffered.gs>. Accessed July 10, 2014.
32. Ahmed N, Devitt KS, Keshet I, et al. A systematic review of the effects of resident duty hour restrictions in surgery: impact on resident wellness, training, and patient outcomes. *Ann Surg*. 2014;259(6):1041-1053.
33. Jamal MH, Doi SA, Rousseau M, et al. Systematic review and meta-analysis of the effect of North American working hours restrictions on mortality and morbidity in surgical patients. *Br J Surg*. 2012;99(3):336-344.
34. Shetty KD, Bhattacharya J. Changes in hospital mortality associated with residency work-hour regulations. *Ann Intern Med*. 2007;147(2):73-80.
35. Lewis FR, Klingensmith ME. Issues in general surgery residency training—2012. *Ann Surg*. 2012;256(4):553-559.
36. Moffatt-Bruce SD, Ross P, Williams TE Jr. American Board of Thoracic Surgery examination: fewer graduates, more failures. *J Thorac Cardiovasc Surg*. 2014;147(5):1464-1469.
37. Antiel RM, Reed DA, Van Arendonk KJ, et al. Effects of duty hour restrictions on core competencies, education, quality of life, and burnout among general surgery interns. *JAMA Surg*. 2013;148(5):448-455.
38. Coleman JJ, Esposito TJ, Rozycki GS, Feliciano DV. Early subspecialization and perceived competence in surgical training: are residents ready? *J Am Coll Surg*. 2013;216(4):764-771.
39. Fischer JE, Healy GB, Britt LD. Surgery is different: a response to the IOM report. *Am J Surg*. 2009;197(2):135-136.
40. Flexibility in duty hour requirements for surgical trainees trial. The FIRST Trial. <http://www.thefirsttrial.org>. Accessed September 15, 2014.