Association of Hospital Participation in a Quality Reporting Program with Surgical Outcomes and Expenditures for Medicare Beneficiaries

Nicholas H. Osborne, MD, MS, **Lauren H. Nicholas, PhD**, **Andrew M. Ryan, PhD**, **Jyothi R. Thumma, MPH**, and **Justin B. Dimick, MD, MPH**
Center for Healthcare Outcomes and Policy (NO, AR, JT, JD), University of Michigan, Ann Arbor, MI; Center for Clinical Management Research, VA Ann Arbor Healthcare System (NO); Department of Health Policy & Management, School of Public Health (JD, AR), University of Michigan, Ann Arbor, MI; Department of Health Policy and Management (LN), Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

Abstract

Importance—The American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP) provides feedback to hospitals on risk-adjusted outcomes. It is not known if NSQIP participation improves outcomes and reduces costs relative to non-participating hospitals.

Objective—To evaluate the association of enrollment and participation in the ACS-NSQIP with outcomes and Medicare payments compared to control hospitals that did not participate in this program.

Design—Quasi-experimental study using national Medicare data (2003 to 2012) for patients undergoing general and vascular surgery. A difference-in-difference analytic approach was used to evaluate whether participation in ACS-NSQIP was associated with improved outcomes and reduced Medicare payments compared to non-participating hospitals that were otherwise similar. Control hospitals were selected using propensity score matching (2 control hospitals for each ACS-NSQIP hospital).

Setting and Participants—263 hospitals participating in ACS NSQIP and 526 non-participating hospitals and a total of 1,226,479 patients undergoing general and vascular surgical procedures

Main Outcome Measures—30-day mortality, serious complications (e.g. pneumonia, myocardial infarction, or acute renal failure and a length of stay > 75th percentile), reoperation and readmission within 30 days. Hospital costs were assessed using price-standardized Medicare payments during hospitalization and 30 days post-discharge.

Corresponding author: Nicholas Osborne, MD, MS, Section of Vascular Surgery, Cardiovascular Center 5168, 1500 East Medical Center Dr, SPC 5867, Ann Arbor, MI 48109-5867, Phone: (734) 936-5790, Fax: (734) 647-9867, nichosbo@umich.edu.

Results—After accounting for patient factors and preexisting time trends toward improved outcomes, there were no statistically significant improvements in outcomes at 1-, 2- or 3-years after (vs before) enrollment in ACS-NSQIP. For example, in analyses comparing outcomes at 3-years after (vs. before) enrollment, there were no statistically significant differences in risk-adjusted 30-day mortality (4.3% vs. 4.5%, Relative risk [RR] 0.96, 95% CI, 0.89–1.03), serious complications (11.1% vs. 11.0%, RR 0.96, 95% CI, 0.91–1.00), re-operations (0.49% vs. 0.45%, RR 0.97, 95%CI, 0.77–1.16), or readmissions (13.3% vs. 12.8%, RR 0.99, 95% CI, 0.96–1.03). There were also no differences at 3-years after (vs. before) enrollment in mean total Medicare payments ($40, 95% CI −$268–348), or payments for the index admission (−$11, 95% CI, −$278, $257), hospital readmission ($245, 95% CI, −$231, $721), or outliers (−$86, 95% CI, −$1666, $1495).

Conclusions and Relevance—With time, hospitals had progressively better surgical outcomes but enrollment in a national quality reporting program was not associated with the improved outcomes or lower Medicare payments among surgical patients. Feedback on outcomes alone may not be sufficient to improve surgical outcomes.

INTRODUCTION

Increased scrutiny of hospital performance has lead to a proliferation of clinical registries used to benchmark outcomes. One of the most visible national quality reporting programs is the American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP)[1–3]. The cornerstone of this program is an extensive clinical registry, with data abstracted directly from the medical record by trained personnel [1,4,5]. ACS-NSQIP provides hospitals with reports that include a detailed description of their risk-adjusted outcomes (e.g., mortality, specific complications, and length of stay). These reports allow hospitals to benchmark their performance relative to all other ACS-NSQIP hospitals. Participating hospitals are encouraged to focus improvement efforts on areas where they perform poorly.

The extent to which participation in ACS-NSQIP improves outcomes is unclear. Several single-center studies from participating hospitals report improvement in outcomes after targeting an area of poor performance with a quality improvement intervention [6,7]. However, it is uncertain whether these changes represent salutary effects of the ACS-NSQIP program, improvement that would have occurred without enrollment in the program, or simply regression to the mean. The only study evaluating all participating hospitals in the ACS-NSQIP demonstrated that the "majority" of hospitals improved their outcomes over time [8]. This study did not compare ACS-NSQIP hospitals to a control group, making it difficult to conclude whether improvements in outcomes were truly associated with participation in this program, or simply represent background trends towards improved outcomes at all hospitals.

The objective of this study was to evaluate the association of participation in the ACS-NSQIP with outcomes and payments among Medicare patients compared to control hospitals that did not participate in the program over the same time period.

METHODS

Data Source and Study Population

Data from the Medicare Analysis Provider and Review (MEDPAR) files for 2003–12 was used to create the main analysis datasets. This dataset contains hospital discharge abstracts for all fee-for-service acute care hospitalizations of US Medicare recipients, which accounts for approximately 70 percent of such admissions in the Medicare population. The Medicare denominator file was used to assess patient vital status at thirty days. The study was reviewed and approved by the University of Michigan Institutional Review Board and was deemed exempt due to the use of secondary data.

Using procedure codes from the International Classification of Diseases, version 9 (ICD-9-CM), all patients aged 65–99 undergoing any of 11 high-risk general and vascular surgical procedures were identified: esophagectomy, pancreatic resection, colon resection, gastrectomy, liver resection, ventral hernia repair, cholecystectomy, appendectomy, abdominal aortic aneurysm repair, lower extremity bypass, and carotid endarterectomy (see Appendix 1 for complete list of ICD-9-CM codes). These procedures were chosen because they are common, high-risk general and vascular surgical procedures included in the American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP) registry. Because they account for the disproportionate share of morbidity and mortality in ACS-NSQIP, they are the highest priority for quality improvement [9,10]. To enhance the homogeneity of hospital case mix, small patient subgroups with much higher baseline risks were excluded. Also excluded were patients with procedure codes indicating that other operations were simultaneously performed (e.g., coronary artery bypass and carotid endarterectomy) or were performed under extremely high-risk conditions (e.g., ruptured abdominal aortic aneurysm) [11–13]. Missing data were found in only 0.3% of race variable and 0.3% of hospital characteristics. Since these represented less than 1%, those patients were excluded from the analyses [14].

Outcome Variables

Mortality, serious complications, reoperation, and readmission were assessed to determine whether enrollment in ACS-NSQIP was associated with improved outcomes. Mortality was assessed as death within 30-days of the index surgical procedure, which was ascertained from the Medicare beneficiary denominator file. Complications were ascertained from primary and secondary ICD-9-CM diagnostic and procedure codes from the index hospitalization. A subset of codes that have been used in several prior studies of surgical outcomes (See Appendix 2 for full list) was chosen, and these codes have been demonstrated to have high sensitivity and specificity in surgical populations [15–18]. For this analysis, serious complications were defined as the presence of a coded complication and an extended length of stay (greater than or equal to the 75th percentile for each procedure). Since most patients without complications are discharged earlier, the addition of the extended length of stay criterion was intended to increase the specificity of the outcome variable [19,20]. Reoperations were ascertained using ICD-9-CM procedure codes indicating secondary procedures during the index hospitalization (see Appendix 3 for full list of codes). Reoperations are relative common in many of these high risk procedures and are accurately

captured using ICD-9-CM procedure billing codes in administrative datasets [21]. Readmissions were defined as an admission to any hospital within 30-days after discharge from the index procedure using standard methods [22].

Medicare Payments

The association between enrollment in ACS-NSQIP and reduced Medicare payments was also assessed. Quality improvement efforts can potentially decrease costs of care by preventing complications and lowering the intensity of resource use, which would be reflected in lower Medicare payments. Medicare facility payments were therefore used to explore the association of ACS-NSQIP participation and lower resource use among Medicare beneficiaries[23–25]. For this study, Medicare facility payments from MedPAR were used, which include all payments related to the index hospitalization, readmissions, and high-cost outliers. Because Medicare payments vary across hospitals (e.g., payments for disproportionate share of low income patients and graduate medical education) and geographic regions (e.g., payments are indexed to reflect differences in wages), a previously described method to "price-adjust" Medicare payments was used [25,26]. In these analyses, all payments were adjusted for the year of operation by standardizing prices to the most recent year of data available.

Statistical Analysis

The goal of this analysis was to examine whether enrollment in the ACS-NSQIP program was associated with improved outcomes for Medicare patients, compared to similar hospitals that did not participate during the same time period. A difference-in-difference approach was used, which is an econometric method for evaluating changes in outcomes occurring after implementation of a policy. [27–30] This approach isolates the improvement in outcomes related to an intervention (i.e., enrollment in ACS-NSQIP) that exceeds changes over the same time period in a control group that was not exposed to the intervention. Enrollment in the ACS-NSQIP was ascertained from the program's semi-annual reports. Since the University of Michigan was a participating site throughout the study period, regular semi-annual reports were received, which include a list of all currently participating hospitals. Hospitals were assigned an enrollment date based upon when they the first appeared in the semi-annual report. Moreover, these enrollment dates were verified using archival data of the ACS-NSQIP website available to the general public from the Internet Archive (http://archive.org/web/). However, since the hospital list on the website may not be current, this source was only used to confirm (and not rule out) hospital participation.

Because the ACS-NSQIP hospitals may not be representative of all hospitals, two separate strategies to adjust for the potential differences between ACS-NSQIP and control hospitals were employed. First, propensity scores were used to match ACS-NSQIP hospitals and control hospitals on baseline outcomes, surgical volume, and pre-enrollment trends in outcomes. To ensure that hospitals in the study and control groups were on the same trajectory for postoperative outcomes before NSQIP enrollment of the study hospitals, they were matched for risk-adjusted mortality for years 1 and 2 prior to NSQIP enrollment. This matching ensured that our control hospitals and ACS-NSQIP hospitals had "parallel trends" in the pre-enrollment period, which is one of the key assumptions of a difference-in-

difference methodology. [31,32] Second, multivariate adjustment was used to account for all observable hospital and patient characteristics that were not included in the propensity score model.

To create propensity scores for hospital matching, a logistic regression model was created with ACS-NSQIP participation (vs. not participating) as the dependent variable. Annual surgical volume, baseline risk-adjusted outcomes, and pre-enrollment trends in risk-adjusted outcomes were included as independent variables. This matching creates a matched cohort of ACS-NSQIP and control hospitals with parallel trends in the three years prior to enrollment. Pre-enrollment trends in outcomes, within and between ACS-NSQIP hospitals and control hospitals, were compared using univariate statistics and no significant difference in pre-enrollment mortality was noted over the pre-enrollment period. The c-statistic for the propensity score model was 0.85, indicating excellent discrimination.

For matching, a caliper width of 0.2 times the standard deviation of the propensity score without replacement was used[33]. Using a caliper width of 0.2 times the standard deviation of the propensity score yielded 100% match of all ACS-NSQIP hospitals with excellent reduction in bias (see Appendix 4). A sensitivity analysis was performed narrowing the caliper width to 0.1, which demonstrated similar findings but excluded 60 ACS-NSQIP hospitals so a 0.2 caliper width was chosen. Although there was a very large pool of potential hospitals (i.e. not participating in ACS-NSQIP), it was determined that 1:2 matching (1 ACS-NSQIP hospital to 2 control hospitals) was optimal based on the degree of bias reduction and percent of ACS-NSQIP hospitals matched [33]. Further attempts to improve the bias reduction yielded fewer matched hospitals (less than all 263 participating ACS-NSQIP hospitals could be matched to control hospitals). Covariate imbalance before and after matching was checked with t-tests for equality of means, and standardized percentage bias before and after matching (together as the achieved percentage reduction in absolute bias) and using pseudo-R^2. This propensity score matching resulted in an overall 98.4% reduction in bias and excellent overlap in propensity scores for the included variables (Appendix 4 and Figure 1 of Appendix). This reduction in bias reflects that among the 3 variables included in the propensity model, 98% of the imbalance in covariates was removed after matching. The pseudo-R2 was 0.01 following matching. There were no significant differences between pre-enrollment trends in outcomes between matched ACS-NSQIP and control hospitals, thereby satisfying the "parallel trends" assumption[28,32].

To perform the difference-in-difference analysis, regression models were used to evaluate the relationship between each dependent variable (mortality, serious complications, reoperations, readmissions, and Medicare payments) and enrollment in ACS-NSQIP. Non-participating control hospitals were assigned the same enrollment year as their corresponding matched ACS-NSQIP hospitals. For the dichotomous outcome variables, logistic regression was used and for the continuous Medicare payment variables, generalized linear models with a log link were used. A dummy variable was included, indicating whether the patient had surgery before enrollment or after enrollment in ACS-NSQIP, defined at post-enrollment year 1, post-enrollment year 2 and post-enrollment year 3. To adjust for linear time trends, a yearly time variable was included. Finally, 3 interaction terms of the ACS-NSQIP (vs. non-ACS-NSQIP control hospital) variable and the pre-post policy

implementation variable (ACS-NSQIP*post-year1, ACS-NSQIP*post-year2, ACS-NSQIP*post-year3) were added. The coefficient from these interaction terms, i.e., the difference-in-difference estimators, can be interpreted as the independent relationship of enrollment in ACS-NSQIP and outcomes for Medicare patients at those time periods [29,34,35]. In all models evaluating outcomes and Medicare payments, patient characteristics were adjusted for by entering the 29 Elixhauser comorbid diseases as individual covariates, a widely used and previously validated approach for risk-adjustment in administrative data.[36,37] These comorbidities were obtained from the ICD-9-CM coding during the same hospital admission. All models were adjusted for the type of surgery by including a categorical variable for each procedure. The difference-in-difference analyses were performed adjusting for all hospital covariates not included in the propensity score model (for-profit status, geographic region, bed size, teaching hospital status and urban location). Additionally all difference-indifference analyses were performed adjusting for clustering at the hospital level with robust standard errors.

In addition to the main analysis, several sensitivity analyses were performed. First, to assess the effect of including the highest risk patients, a difference-in-difference analysis including the previously excluded high-risk patient subgroups (e.g., emergency surgery and ruptured abdominal aortic aneurysm repair) was performed. Second, to assess the effect of using hierarchical modeling rather than robust standard errors to account for clustering of similar patients within hospitals, a sensitivity analysis was performed using a multi-level model with hospital-level random effects.

All odds ratios were converted to relative risk because the former may not be an accurate representation of the risk ratio when an outcome variable is relatively common [38] and all 95% confidence intervals were calculated using robust variance estimates. A $p<.05$ was used as the threshold for statistical significance and all reported P-values are two-sided. Model fit was assessed using goodness of fit and model discrimination was assessed using ROC curves. All statistical analyses were conducted using STATA 12.0 (College Station, Texas).

RESULTS

A total of 294 hospitals enrolled in the ACS-NSQIP during the study period. Of these, 20 hospitals were excluded because they performed only pediatric surgery or had no Medicare identifier. Among the 274 remaining hospitals, 5 hospitals were excluded due to incomplete participation (i.e. hospitals that joined and dropped ACS-NSQIP during the study period) and 6 hospitals were excluded because they only had 1 year of participation during our study period. In total, 263 ACS-NSQIP hospitals were each matched with 2 control hospitals to yield 526 non-ACS-NSQIP control hospitals. The 263 ACS-NSQIP hospitals had a median follow-up after enrollment in the program of 3.8 years and a minimum of 2 years. Table 1 shows the hospital characteristics before and after propensity score matching for participating and non-participating hospitals. ACS-NSQIP and control hospitals were well matched for the variables used in the propensity score matching, including surgical volume and baseline outcomes for the 2 years prior to NSQIP enrollment. ACS-NSQIP and control hospitals' baseline (during the year prior to enrollment) 30-day mortality was 4.9% vs. 5.0%, p=0.55, serious complications were 11.3% vs. 10.2%, p=0.005, reoperation was 0.5%

vs. 0.5%, p=0.66, and readmission was 13.0% vs. 12.6%, p=0.28 (Table 1). While many other hospital characteristics were clinically similar (nurse to patient ratio, % Medicaid, and urban location,), even after propensity matching, the ACS-NSQIP hospitals were slightly larger with more admissions, higher total surgical operations, more employees, more operating rooms and were more likely to have not-for-profit status and be teaching hospitals (Table 1).

Patient characteristics were generally clinically similar at ACS-NSQIP and control hospitals despite statistically significant differences (Table 2). Patients were clinically similar with respect to average age (75.71 vs. 76.10 years old, p <0.0001) and the proportion that were female (49.02% vs. 49.96%, p <0.0001) and non-white race (11.51% vs. 9.22%, p <0.0001). Approximately two-thirds of the included surgical cases at both ACS-NSQIP hospitals (65.90%) and control hospitals (64.60%) were general surgery with the remaining cases representing major vascular procedures (Table 2). The procedure mix was comparable for both general and vascular surgery cases, although ACS-NSQIP hospitals tended to perform more complex gastrointestinal cancer resections than control hospitals (esophagectomy 1.44% vs. 0.64%, p <0.0001; pancreatectomy 2.00% vs. 0.78%, p <0.0001; gastrectomy 2.57% vs. 1.74%, p <0.0001). Although statistically significant differences were noted, patients at participating and non-participating hospitals were generally similar in terms of comorbid diseases with no clinically important differences apparent (Table 2).

Although there were slight trends towards improved outcomes in ACS-NSQIP hospitals before vs. after enrollment (1-, 2-, and 3- years), there were similar trends in control hospitals (Table 3). For example, 30-day mortality in ACS-NSQIP hospitals declined from 4.6% (95% CI 4.6%–4.7%) to 4.2% (95% CI 4.2%–4.3%) during the study period (p-value <0.001), as compared to 4.9% (95% CI 4.8%–4.9%) to 4.6% (95% CI 4.5%–4.6%) among control hospitals (p-value <0.001). In difference-in-difference analyses, there was no statistically significant reduction in any measured outcome after enrollment in ACS-NSQIP (Table 4). For example, there was no significant difference in risk-adjusted 30-day mortality in the three years following enrollment: post-enrollment year 1 (Relative Risk [RR], 0.96, 95% CI, 0.90–1.02); post-enrollment year 2 (RR 0.94, 95% CI, 0.88–1.00); post-enrollment year 3 (RR 0.96, 95% CI, 0.89–1.03) (Table 4). Even at 3-years following enrollment, there remained no significant differences in the rates of serious complications (RR 0.96, 95% CI, 0.91–1.00), reoperations (RR 0.97, 95% CI, 0.77–1.16), and readmissions (RR 1.01, 95% CI, 0.98–1.03) (Table 4).

There were no statistically significant differences in 30-day Medicare payments in difference-indifference analyses, even when facility payments were separated into payments for the index hospital stay, payments for readmissions, and payments for outliers (Table 5). For example, at 3-years following enrollment, there were no significant differences in mean total Medicare payments ($40, 95% CI −$268–348), payments for index admission (−$11, 95% CI, −$278–$257), payment for readmission ($245, 95% CI, −$231–$721), payment for outliers (−$86, 95% CI, −$1666, 1495).

In a sensitivity analysis including patients that were previously excluded high-risk patient subgroups, there were also no significant difference in the rates of 30-day mortality, serious

complications, reoperations, or readmission following enrollment in the ACS-NSQIP. In a second sensitivity analysis using hierarchical modeling, there were also no significant differences in the rates of 30-day mortality, serious complications, reoperations, or readmissions following enrollment in the ACS-NSQIP.

DISCUSSION

In this study, there was a slight time trend toward improved surgical outcomes in both ACS-NSQIP and control hospitals. To evaluate the extent to which these improved outcomes were independently associated with enrollment in ACS-NSQIP, we matched each ACS-NSQIP hospital with 2 control hospitals that had similar trends in outcomes before enrollment, as well as similar baseline outcomes and surgical volumes. In a comparison between ACS-NSQIP and matched control hospitals, there was no independent association of hospital enrollment in this quality reporting with improved outcomes or decreased Medicare payments at 1-, 2-, or 3-years. Because of this control group of hospitals, the independent association of enrollment in ACS-NSQIP with adverse outcomes and Medicare payments was isolated, removing any confounding background trends towards improved outcomes. These findings imply that participation in hospital quality reporting programs, such as ACS-NSQIP, may not be sufficient to improve outcomes.

Prior studies reported a salutary effect of participation in the ACS-NSQIP program. Several single-center studies have reported improvements in specific complications after a local quality improvement intervention [6,39,40]. Many of these interventions were initiated because the hospital was identified as a poorly performing "outlier" on their ACS-NSQIP report. After implementing best practices at their institution, most studies report improvement in risk-adjusted outcomes [6,39,40]. However, because these studies lack a control group, it is difficult to know whether such changes represent true improvements in outcomes or simply reflect regression to the mean. Regression to the mean is observed when individuals with an extreme value on a measure spontaneously move back toward the average. Establishing differences between true improvement in outcomes and regression to the mean in quality improvement research is difficult. To do this, a control group is necessary. This was achieved by matching ACS-NSQIP hospitals to a larger cohort of non-participating hospitals. Using the ACS-NSQIP clinical registry, Hall and colleagues found that the majority of hospitals improved their risk-adjusted outcomes after enrolling in the program during 2005–2007 [8]. However, this study lacked a control group and it is not known if improved outcomes would have occurred in the absence of ACS-NSIP enrollment. By using a control group, secular improvements in mortality were found that were independent of ACS-NSQIP enrollment.

This study has certain limitations. One is the use of administrative data rather than a clinical registry. Clinical registries may have more detailed information on patient risk factors and outcomes. Nonetheless, there is no other source of data that could be used to address this important question since no registry exists that collects data from both participating and non-participating hospitals. Medicare data provides the most comprehensive data available to capture not only the outcomes and payments at the participating hospitals, but also the non-participating hospitals. Further, this study was specifically designed to take advantage of the

strengths and to minimize the weaknesses of administrative data. The first weakness of administrative data is the assessment of patient comorbidities and severity of illness, which are needed for risk-adjustment. In our analysis, the best available comorbid disease index for risk-adjustment was used. Moreover, our study design, difference-in-difference, also mitigates this limitation by adjusting for any unobserved differences in patient case-mix that do not change over time [19,30,34]. There is no reason to believe that changes in patient case-mix differed between participating and non-participating hospitals during the time period of our study. Another limitation of administrative data is the identification of patient outcomes. This was addressed by assessing outcomes reliably coded in billing records, including mortality, reoperations, and readmissions. Identification of complications that rely on ICD-coding was optimized by only using complication codes known to have a high sensitivity and specificity in surgical patients [15,16]. An extended length of stay criterion was added to our assessment of complications (i.e., patients had to have both an ICD-9 code and a prolonged length of stay) to improve the specificity for identification of complication outcomes since patients with a prolonged length of stay likely have had complications [20].

Another potential limitation relates to our ability to only evaluate the association of outcomes and participation in the ACS-NSQIP through 2012. Because there may be a lag between ACS-NSQIP enrollment and improved outcomes, we thought it was important to have at least 2 years of data to evaluate outcomes of participating hospitals. Consequently, all hospitals that enrolled in ACS-NSQIP through 2010 were included and outcomes for at least the following 2 years were assessed. During the time period of our study, outcomes feedback was the primary means by which ACS-NSQIP affected surgical outcomes. Out findings suggest that this was not effective.

There are several potential reasons why improved outcomes among participating hospitals were not found. Conceivably, participating hospitals may not have initiated quality improvement efforts after receiving ACS-NSQIP reports. The ACS-NSQIP provides non-publicly reported performance feedback, which may not adequately motivate participating hospitals to make changes. Other strategies have much stronger incentives. For example, the accountability of publicly reporting of hospital performance can motivate improvement [41,42]. Other strategies, such as value-based purchasing, including pay-for-performance and non-payment for adverse events, directly incentivize hospitals financially [43,44]. It is also possible that hospitals participating in ACS-NSQIP implemented quality improvement efforts but they did not improve outcomes [45]. Clinical quality improvement is challenging for hospitals. Changing physician practice requires complex, sustained, multifaceted interventions, and most hospitals may not have the expertise or resources to launch effective quality improvement interventions.

CONCLUSIONS

Enrollment in a national surgical quality reporting program was not associated with improved outcomes or lower payments among Medicare patients. Feedback of outcomes alone may not be sufficient to improve surgical outcomes.

Acknowledgments

References

1. 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. [PubMed: 20919525]

2. Fink AS, Campbell DA Jr, Mentzer RM Jr, et al. The National Surgical Quality Improvement Program in non-veterans administration hospitals: initial demonstration of feasibility. Ann Surg. Sep; 2002 236(3):344–353. discussion 353–344. [PubMed: 12192321]

3. Khuri SF, Henderson WG, Daley J, et al. 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. Aug; 2008 248(2):329–336. [PubMed: 18650645]

4. Best WR, Khuri SF, Phelan M, et al. Identifying patient preoperative risk factors and postoperative adverse events in administrative databases: results from the Department of Veterans Affairs National Surgical Quality Improvement Program. J Am Coll Surg. Mar; 2002 194(3):257–266. [PubMed: 11893128]

5. 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. Dec; 2012 256(6):973–981. [PubMed: 23095667]

6. Ellner SJ. Hospital puts ACS NSQIP to the test and improves patient safety. Bull Am Coll Surg. Sep; 2011 96(9):9–11. [PubMed: 22259928]

7. Glickson J. ACS NSQIP National conference: speakers promote professionalism, collaboration to achieve quality improvement. Bull Am Coll Surg. Oct; 2013 98(10):66–71. [PubMed: 24266125]

8. Hall BL, Hamilton BH, Richards K, Bilimoria KY, Cohen ME, Ko CY. Does surgical quality improve in the American College of Surgeons National Surgical Quality Improvement Program: an evaluation of all participating hospitals. Ann Surg. Sep; 2009 250(3):363–376. [PubMed: 19644350]

9. Schilling PL, Dimick JB, Birkmeyer JD. Prioritizing quality improvement in general surgery. J Am Coll Surg. Nov; 2008 207(5):698–704. [PubMed: 18954782]

10. Schilling PL, Dimick JB, Birkmeyer JD. Prioritizing quality improvement in vascular surgery. Surg Innov. Jun; 2010 17(2):127–131. [PubMed: 20504789]

11. Finks JF, Osborne NH, Birkmeyer JD. Trends in hospital volume and operative mortality for high-risk surgery. N Engl J Med. Jun 2; 2011 364(22):2128–2137. [PubMed: 21631325]

12. Ghaferi AA, Birkmeyer JD, Dimick JB. Variation in hospital mortality associated with inpatient surgery. N Engl J Med. Oct 1; 2009 361(14):1368–1375. [PubMed: 19797283]

13. Birkmeyer JD, Stukel TA, Siewers AE, Goodney PP, Wennberg DE, Lucas FL. Surgeon volume and operative mortality in the United States. N Engl J Med. Nov 27; 2003 349(22):2117–2127. [PubMed: 14645640]

14. Little, RJA.; Rubin, DB. Statistical analysis with missing data. 2. Hoboken, N.J: Wiley; 2002.

15. Iezzoni LI, Daley J, Heeren T, et al. Using administrative data to screen hospitals for high complication rates. Inquiry. Spring;1994 31(1):40–55. [PubMed: 8168908]

16. Weingart SN, Iezzoni LI, Davis RB, et al. Use of administrative data to find substandard care: validation of the complications screening program. Medical care. Aug; 2000 38(8):796–806. [PubMed: 10929992]

17. Iezzoni LI, Daley J, Heeren T, et al. Identifying complications of care using administrative data. Medical care. Jul; 1994 32(7):700–715. [PubMed: 8028405]

18. Lawthers AG, McCarthy EP, Davis RB, Peterson LE, Palmer RH, Iezzoni LI. Identification of inhospital complications from claims data. Is it valid? Medical care. Aug; 2000 38(8):785–795. [PubMed: 10929991]

19. Dimick JB, Nicholas LH, Ryan AM, Thumma JR, Birkmeyer JD. Bariatric surgery complications before vs after implementation of a national policy restricting coverage to centers of excellence. JAMA. Feb 27; 2013 309(8):792–799. [PubMed: 23443442]

20. Livingston EH. Procedure incidence and in-hospital complication rates of bariatric surgery in the United States. Am J Surg. Aug; 2004 188(2):105–110. [PubMed: 15249233]

21. Morris AM, Baldwin LM, Matthews B, et al. Reoperation as a quality indicator in colorectal surgery: a population-based analysis. Ann Surg. Jan; 2007 245(1):73–79. [PubMed: 17197968]

22. Tsai TC, Joynt KE, Orav EJ, Gawande AA, Jha AK. Variation in surgical-readmission rates and quality of hospital care. N Engl J Med. Sep 19; 2013 369(12):1134–1142. [PubMed: 24047062]

23. Birkmeyer JD, Gust C, Baser O, Dimick JB, Sutherland JM, Skinner JS. Medicare payments for common inpatient procedures: implications for episode-based payment bundling. Health services research. Dec; 2010 45(6 Pt 1):1783–1795. [PubMed: 20698899]

24. Dimick JB, Weeks WB, Karia RJ, Das S, Campbell DA Jr. Who pays for poor surgical quality? Building a business case for quality improvement. J Am Coll Surg. Jun; 2006 202(6):933–937. [PubMed: 16735208]

25. Miller DC, Gust C, Dimick JB, Birkmeyer N, Skinner J, Birkmeyer JD. Large variations in Medicare payments for surgery highlight savings potential from bundled payment programs. Health Aff (Millwood). Nov; 2011 30(11):2107–2115. [PubMed: 22068403]

26. Gottlieb DJ, Zhou W, Song Y, Andrews KG, Skinner JS, Sutherland JM. Prices don't drive regional Medicare spending variations. Health Aff (Millwood). Mar-Apr;2010 29(3):537–543. [PubMed: 20110290]

27. Colla CH, Wennberg DE, Meara E, et al. Spending differences associated with the Medicare Physician Group Practice Demonstration. JAMA. Sep 12; 2012 308(10):1015–1023. [PubMed: 22968890]

28. Dimick J, Ryan A. Methods for Evaluating Changes in Health Care Policy: The Difference-inDifference Approach. JAMA. 2014; 312(22)

29. 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. Sep 5; 2007 298(9):975–983. [PubMed: 17785642]

30. Wooldridge, JM. Introductory econometrics: a modern approach. 4. Mason, OH: South Western, Cengage Learning; 2009.

31. Bertrand M, Duflo E, Mullainathan S. How much should we trust differences-in-differences estimates? Q J Econ. Feb; 2004 119(1):249–275.

32. Ryan A, JB, Dimick J. Why we shouldn't be indifferent to specification in difference-in-differences analysis. Health services research. 2014

33. Austin PC. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. Pharmaceutical statistics. Mar-Apr; 2011 10(2):150–161. [PubMed: 20925139]

34. Donald SG, Lang K. Inference with difference-in-differences and other panel data. Review of Economics and Statistics. May; 2007 89(2):221–233.

35. Ryan AM. Effects of the Premier Hospital Quality Incentive Demonstration on Medicare Patient Mortality and Cost. Health Services Research. Jun; 2009 44(3):821–842. [PubMed: 19674427]

36. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. Medical care. Jan; 1998 36(1):8–27. [PubMed: 9431328]

37. Southern DA, Quan H, Ghali WA. Comparison of the Elixhauser and Charlson/Deyo methods of comorbidity measurement in administrative data. Medical care. Apr; 2004 42(4):355–360. [PubMed: 15076812]

38. Zhang J, Yu KF. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. JAMA. Nov 18; 1998 280(19):1690–1691. [PubMed: 9832001]

39. Neumayer L, Mastin M, Vanderhoof L, Hinson D. Using the Veterans Administration National Surgical Quality Improvement Program to improve patient outcomes. J Surg Res. Jan; 2000 88(1): 58–61. [PubMed: 10644468]

40. Rowell KS, Turrentine FE, Hutter MM, Khuri SF, Henderson WG. Use of national surgical quality improvement program data as a catalyst for quality improvement. J Am Coll Surg. Jun; 2007 204(6):1293–1300. [PubMed: 17544087]

41. Hannan EL, Sarrazin MS, Doran DR, Rosenthal GE. Provider profiling and quality improvement efforts in coronary artery bypass graft surgery: the effect on short-term mortality among Medicare beneficiaries. Medical care. Oct; 2003 41(10):1164–1172. [PubMed: 14515112]

42. Hibbard JH, Stockard J, Tusler M. Hospital performance reports: impact on quality, market share, and reputation. Health Aff (Millwood). Jul-Aug;2005 24(4):1150–1160. [PubMed: 16012155]

43. Rosenthal MB. Nonpayment for performance? Medicare's new reimbursement rule. N Engl J Med. Oct 18; 2007 357(16):1573–1575. [PubMed: 17942869]

44. Rosenthal MB. Beyond pay for performance--emerging models of provider-payment reform. N Engl J Med. Sep 18; 2008 359(12):1197–1200. [PubMed: 18799554]

45. Walshe K, Freeman T. Effectiveness of quality improvement: learning from evaluations. Qual Saf Health Care. Mar; 2002 11(1):85–87. [PubMed: 12078378]

Table 1

Characteristics of hospitals participating in the American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP) compared to non-participating hospitals before and after propensity score matching.

	ACS-NSQIP Hospitals	Non-ACS-NSQIP Hospitals		P-value ACS-NSQIP vs non-ACS-NSQIP Hospitals after matching
Hospital Characteristics		Before matching	After matching	
# of hospitals	263	4,789	526	
# of patients	430,179	3,918,956	796,318	
Annual Surgical volume, median (25th – 75th %)	442 (298–615)	230 (123–381)	411 (280–566)	<.0001
Risk-adjusted mortality pre-enrollment year1 (%)	5.0	5.3	5.0	0.5762
Risk-adjusted mortality pre-enrollment year2 (%)	5.1	5.3	5.1	0.6310
Baseline risk-adjusted outcomes[*]				
Mortality N (%)	4.9	5.7	5.0	0.5527
Serious complications N (%)	11.3	9.2	10.2	0.0005
Reoperation N (%)	0.5	0.4	0.5	0.6592
Readmissions N (%)	13.0	13.4	12.6	0.2845
Geographic region N (%)				
Northeast	66 (25.1)	498 (12.2)	83 (15.8)	0.0029
Midwest	81 (30.8)	1,177 (28.9)	124 (23.6)	0.0342
South	64 (24.3)	1,635 (40.1)	221 (42.0)	<.0001
West	52 (19.8)	769 (18.9)	98 (18.6)	0.703
Bed size N (%)				
<200	34 (12.9)	3,286 (80.6)	152 (28.9)	<.0001
200–349	67 (25.5)	558 (13.7)	161 (30.6)	0.1272
350–499	57 (21.7)	166 (4.1)	115 (21.9)	0.9514
500	105 (39.9)	69 (1.7)	98 (18.6)	<.0001
Profit status N (%)				
For-profit	15 (5.7)	1,052 (25.8)	83 (15.8)	<.0001
Non-profit	214 (81.4)	2,206 (54.1)	382 (72.6)	0.0049
Other	34 (12.9)	821 (20.1)	61 (11.6)	0.5947
Other characteristics				
Nurse ratio, median (25th –75th %)	8.2 (6.8–9.7)	7.4 (4.7–10.6)	7.4 (6.2–8.9)	0.0034
% Medicaid days, median (25th –75th %)	18 (12–26)	15 (8–22)	17 (11–22)	0.015
Teaching hospital (%)	213 (81.0)	837 (20.5)	276 (52.5)	<.0001

	ACS-NSQIP Hospitals	Non-ACS-NSQIP Hospitals		P-value ACS-NSQIP vs non-ACS-NSQIP Hospitals after matching
Hospital Characteristics		**Before matching**	**After matching**	
Urban location (%)	246 (94.3)	3,311 (93.8)	473 (91.0)	0.0862
Licensed Beds, median (25th –75th %)	496 (307–699)	100 (44–204)	365 (210–519)	<.0001
Total Admissions, median (25th –75th %)	21,207 (13,689–32,008)	2,426 (936–7,111)	15,059 (8,101–22,620)	<.0001
Full-Time Equivalent Employees, median (25th –75th %)	2,905 (1,765–5,000	371 (178–793)	1,644 (948–2,587)	<.0001
Total Annual Surgical Operations, median (25th –75th %)	15,181 (9,772–24,816)	2,634 (931–5,537)	1,0161 (6,052–15,843)	<.0001
Number of Operating Rooms, median (25th – 75th %)	21 (13–33)	4 (2–8)	15 (9–23)	0.0154

*
Risk-adjusted rates are adjusted for patient characteristics, comorbidities and procedure type.

Table 2

Patient characteristics at hospitals participating in the American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP) compared to non-participating hospitals before and after propensity score matching.

	ACS-NSQIP Hospitals	Non-ACS-NSQIP Hospitals Before Propensity Score Matching	Non-ACS-NSQIP Control Hospitals After Propensity Score Matching	P-value (ACS-NSQIP vs. Non-ACS-NSQIP Control hospitals)
No. of hospitals	263	4,789	526	
No. of patients	430,179	3,918,956	796,318	
Patient characteristics				
Mean (SD) age (yrs)	75.7 (6.8)	75.9 (7.1)	76.1 (6.9)	<0.0001
Female (%)	49.02	51.71	49.96	<0.0001
Non-white race (%)	11.51	11.27	9.22	<0.0001
General surgery cases (%)	65.9	72.46	64.6	<0.0001
Esophagectomy	1.44	0.49	0.64	<0.0001
Pancreatic resection	2	0.53	0.78	<0.0001
Colon resection	22.82	24.6	22.46	<0.0001
Gastrectomy	2.57	1.69	1.74	<0.0001
Liver resection	1.67	0.51	0.71	<0.0001
Hernia repair	11.98	12.02	11.12	<0.0001
Cholecystectomy	18.57	26.09	21.88	<0.0001
Appendectomy	4.85	6.51	5.28	<0.0001
Vascular surgery cases (%)	34.1	27.54	35.4	<0.0001
Abdominal aortic aneurysm repair	8.33	5.23	6.99	<0.0001
Lower extremity bypass	9.23	6.86	8.67	<0.0001
Carotid endarterectomy	16.54	15.46	19.74	<0.0001
Comorbid diseases				
Hypertension (%)	58.07	57.37	58.78	<.0001
Diabetes w/o complications (%)	18.35	18.95	18.92	<.0001
Chronic pulmonary disease (%)	17.83	18.92	19.59	<.0001
Electrolyte disorders (%)	16.46	18.1	16.51	0.3267
Peripheral vascular disease (%)	12.41	10.54	12.6	<.0001
Hypothyroidism (%)	8.77	8.79	8.84	0.1057
Congestive heart failure (%)	8.57	9.44	9.14	<.0001
Metastatic cancer (%)	7.96	6.38	6.3	<.0001
Deficiency Anemias (%)	7.58	9.04	8.37	<.0001
Renal failure (%)	6.73	6.46	6.76	0.332
Valvular disease (%)	4.88	4.65	5.18	<.0001

	ACS-NSQIP Hospitals	Non-ACS-NSQIP Hospitals Before Propensity Score Matching	Non-ACS-NSQIP Control Hospitals After Propensity Score Matching	P-value (ACS-NSQIP vs. Non-ACS-NSQIP Control hospitals)
Weight loss (%)	4.8	5.23	4.91	0.0002
Obesity (%)	4.5	4.76	4.56	0.0437
Depression (%)	3.37	3.35	3.28	0.0005
Coagulopthy (%)	3.24	2.76	2.9	<.0001
Other neurological disorders (%)	2.71	3.1	2.81	<.0001
Solid tumor w/o metastasis (%)	2.56	2.31	2.29	<.0001
Diabetes w/complications (%)	2.38	2.26	2.19	<.0001
Rheumatologic disorder (%)	1.82	1.73	1.79	0.1326
Liver disease (%)	1.8	1.61	1.51	<.0001
Chronic blood loss anemia (%)	1.6	2	1.83	<.0001
Pulmonary hypertension (%)	1.46	1.28	1.26	<.0001

Table 3

Risk-adjusted patient outcomes before *vs.* after enrolling in the American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP) as compared to Non-ACS-NSQIP matched control hospitals. P-value for trends over time are displayed for ACS-NSQIP and control hospitals.

Risk-adjusted outcomes	ACS-NSQIP Hospitals							Non-ACS-NSQIP Control Hospitals						
	Pre-enrollment			Post-enrollment			p-value	Pre-enrollment			Post-enrollment			p-value
	Year 3	Year 2	Year 1	Year 1	Year 2	Year 3		Year 3	Year 2	Year 1	Year 1	Year 2	Year 3	
Hospitals, N	204	263	263	263	263	261		413	526	526	525	519	515	
Patients, N	60,640	79,651	77,033	73,566	70,816	68,473		113,107	148,198	143,358	137,616	130,082	123,957	
Mortality, N	2739	3637	3391	3225	3057	3014		5315	6829	6359	6392	6059	5697	
% (95% CI) *	4.6 (4.6–4.7)	4.6 (4.6–4.7)	4.4 (4.4–4.5)	4.3 (4.2–4.3)	4.1 (4.0–4.1)	4.2 (4.2–4.3)	<.0001	4.9 (4.8–4.9)	4.8 (4.8–4.8)	4.6 (4.5–4.6)	4.6 (4.6–4.7)	4.5 (4.4–4.5)	4.6 (4.5–4.6)	<.0001
Serious complications, N	6555	9003	8830	8313	7790	7669		11321	14880	14859	14460	14311	13735	
% (95% CI) *	11.0 (11.0–11.1)	11.3 (11.2–11.4)	11.5 (11.4–11.6)	11.3 (11.2–11.4)	11.0 (10.9–11.1)	11.2 (11.1–11.3)	0.0102	10.4 (10.4–10.5)	10.4 (10.4–10.5)	10.6 (10.6–10.7)	10.4 (10.4–10.5)	10.6 (10.5–10.7)	10.9 (10.8–11.0)	<.0001
Reoperations, N	252	352	366	411	399	382		432	595	610	588	622	648	
% (95% CI) *	0.4 (0.4–0.4)	0.4 (0.44–0.44)	0.5 (0.46–0.47)	0.5 (0.52–0.53)	0.5 (0.5–0.5)	0.5 (0.5–0.5)	0.0156	0.4 (0.4–0.4)	0.4 (0.4–0.4)	0.5 (0.4–0.5)	0.4 (0.4–0.4)	0.5 (0.5–0.5)	0.5 (0.5–0.5)	0.0176
Readmissions, N	7617	10250	10228	10188	9879	9720		13025	17374	17202	16770	16478	15903	
% (95% CI) *	12.5 (12.5–12.6)	12.7 (12.7–12.8)	13.1 (13.1–13.2)	13.6 (13.6–13.6)	13.6 (13.6–13.7)	13.9 (13.9–14.0)	<.0001	11.7 (11.6–11.7)	11.9 (11.9–11.9)	12.1 (12.1–12.2)	12.2 (12.2–12.2)	12.7 (12.6–12.7)	12.9 (12.9–13.0)	<.0001

*
All models are adjusted for patient characteristics, comorbidities and procedure type.

Table 4

Relative risk of risk-adjusted outcomes in ACS-NSQIP hospitals and Non-ACS-NSQIP matched control hospitals assessed at 1-, 2- and 3- years following enrollment in simple pre-post (ACS-NSQIP hospitals after enrollment vs. before enrollment) and difference-indifference analysis

Adverse Outcome	Relative Risk of Adverse Outcome in ACS-NSQIP Hospitals After vs. Before Enrollment (95% CI) N=430,179 Post-enrollment			Relative Risk of Adverse Outcomes After vs. Before Enrollment in ACS-NSQIP Compared to Control Hospitals (95% CI)* N=1,226,497 Post-enrollment		
	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Mortality	0.92 (0.85, 0.98)	0.76 (0.62, 0.9)	0.79 (0.59, 0.99)	0.96 (0.90, 1.02)	0.94 (0.88, 1.00)	0.96 (0.89, 1.03)
Serious complications	1.01 (0.97, 1.05)	0.99 (0.88, 1.09)	1.00 (0.86, 1.14)	1.00 (0.96, 1.03)	0.96 (0.93, 1.00)	0.96 (0.91, 1.00)
Reoperations	1.07 (0.91, 1.22)	1.15 (0.77, 1.53)	1.29 (0.86, 1.73)	1.14 (0.98, 1.31)	1.06 (0.90, 1.22)	0.97 (0.77, 1.16)
Readmissions	1.06 (1.02, 1.09)	1.12 (1.03, 1.20)	1.16 (1.05, 1.27)	1.04 (1.01, 1.07)	1.00 (0.97, 1.03)	0.99 (0.96, 1.03)

*
All models are adjusted for patient characteristics, comorbidities, year of surgery, procedure type, and hospital characteristics.

Table 5

Medicare payments before vs. after enrolling in the American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP) compared to Non-ACS-NSQIP matched control hospitals. Pre-enrollment payments represent mean payments over the three years prior to enrollment.

	ACS-NSQIP Hospitals	Non-ACS-NSQIP Control Hospitals	Difference in Average Medicare Payments After Enrollment vs. Before in ACS-NSQIP Hospitals (95% CI)			Difference in Average Medicare Payments in NSQIP Hospitals vs. Non-ACS-NSQIP Control Hospitals After Enrollment (95% CI)[*]		
	N=217,324	N=404,663	N=430,179			N=1,226,497		
	Pre-enrollment		Post-enrollment			Post-enrollment		
			Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Total payments, mean	$20,529	$18,740	$642 (280,1003)	$879 (377,1382)	$1058 (414,1703)	$166 (−78,410)	$180 (−88,449)	$40 (−268,348)
Payments for index admission, mean	$18,746	$17,168	$534 (219,848)	$815 (378, 253)	$957 (396,518)	$54 (−163,272)	$123 (−103,349)	−$11 (−278,257)
Payments for readmissions								
% with readmission	14.7	13.2	15.2	15.0	15.0			
Payment, mean	$12,124	$11,889	−$12 (−497, 473)	−$154 (−738, 431)	−$268 (−897, 360)	$95 (−405,596)	$201 (−339, 740)	$245 (−231, 721)
Payments for outliers								
% with outlier payment	6.0	5.1	7.0	7.2	7.4			
Payment, mean	$23,507	$20,153	$1625 (215, 3035)	$3391 (1662, 5120)	$2872 (740, 5004)	$454 (−961,1868)	$1459 (−48, 2965)	−$86 (−1666, 1495)

[*] All models are adjusted for patient characteristics, comorbidities, year of surgery, procedure type and hospital characteristics.