Improved Surgical Outcomes for ACS NSQIP Hospitals Over Time

Evaluation of Hospital Cohorts With up to 8 Years of Participation

Mark E. Cohen, PhD,* Yaoming Liu, PhD,* Clifford Y. Ko, MD, MS, MSHS, FACS,* ‡ and Bruce L. Hall, MD, PhD, MBA, FACS* ‡

Background: The American College of Surgeons, National Surgical Quality Improvement Program (ACS NSQIP) surgical quality feedback models are recalibrated every 6 months, and each hospital is given risk-adjusted, hierarchical model, odds ratios that permit comparisons to an estimated average NSQIP hospital at a particular point in time. This approach is appropriate for “relative” benchmarking, and for targeting quality improvement efforts, but does not permit evaluation of hospital or program-wide changes in quality over time. We report on long-term improvement in surgical outcomes associated with participation in ACS NSQIP.

Study Design: ACS NSQIP data (2006-2013) were used to create prediction models for mortality, morbidity, and surgical site infection (SSI), respectively. For each model, for each hospital, and for year of first participation (hospital cohort), hierarchical model observed/expected (O/E) ratios were computed. The primary performance metric was the within-hospital trend in logged O/E ratios over time (slope) for mortality, morbidity, and SSI, respectively. For these hospitals, we estimate 0.8%, 3.1%, and 2.6% annual reductions (with respect to prior year’s rates) for mortality, morbidity, and SSI, respectively. For hospitals currently in the program for at least 3 years, 69%, 65%, and 60% of hospitals had negative slopes for mortality, morbidity, and any SSI, respectively.

Results: Hospital-averaged log O/E ratio slopes were generally negative, indicating improving performance over time. For all hospitals, 62%, 70%, and 65% of hospitals had negative slopes for mortality, morbidity, and any SSI, respectively. For hospitals currently in the program for at least 3 years, 69%, 79%, and 71% showed improvement in mortality, morbidity, and SSI, respectively. For these hospitals, we estimate 0.8%, 3.1%, and 2.6% annual reductions (with respect to prior year’s rates) for mortality, morbidity, and SSI, respectively.

Conclusions: Participation in ACS NSQIP is associated with reductions in adverse events after surgery. The magnitude of quality improvement increases with time in the program.

Keywords: ACS NSQIP, profiling, risk adjustment, surgical quality improvement, time trends

(Ann Surg 2015;00:1–7)

The American College of Surgeons, National Surgical Quality Improvement Program’s (ACS NSQIP) general approach to data collection and statistical modeling for purposes of risk adjustment has been described elsewhere. A central programmatic feature is that evaluations are recalibrated on the basis of incorporation of new data (and removal of old data) on a semiannual basis, and results for overlapping 12 months of data are reported in a “Semi-annual report” (SAR). Hospitals are provided with this report, which benchmarks their performance in comparison to an estimated average NSQIP hospital. Hospitals with this report, which benchmarks their performance in comparison to an estimated average NSQIP hospital would perform if doing the same procedures on the same patient. A consequence of this remodeling is that performance is continuously re-standardized to the current timeframe. Hospitals will know how they are doing in comparison to other hospitals currently in the program but cannot determine if they are improving over time relative to some fixed standard. This is a rational choice for providing feedback that will continuously drive quality improvement.

Consideration of changes in raw mortality and complication rates could be informative on evaluating performance over time, but without risk adjustment this can be misleading. ACS NSQIP hospitals frequently shift sampling protocols to focus on different groups of operations, frequently from those with lesser risk to those with greater risk (eg, when transitioning to the “procedure-targeted” option). Also potentially informative regarding longitudinal improvement is the growing literature that describes how participation in ACS NSQIP helps to successfully target quality improvement efforts. Improvement in surgical quality tide for all. The magnitude of this rising tide has been infrequently evaluated in the literature. In an important study on this topic, logistic modeling of ACS NSQIP data from 2005 to 2007 was used to estimate observed/expected O/E ratios derived from a time-constant equation or from a constant hospital cohort. Improvement in surgical quality over the time period was observed for the majority of participating hospitals. The objectives of this study were to evaluate ACS NSQIP quality improvement effect over a longer time frame—from 2006 to 2013 and use hierarchical modeling methods to permit stabilized estimates of hospital performance nested within years (using a 3-level hierarchical model—patients nested within years, nested within hospitals). Definitions of certain ACS NSQIP predictive and outcome variables have changed over time and this precludes generating overall models of number and scope similar to those included in standard SARs. However, this is not relevant to this study as the proposed overtime information is not intended for detailed benchmarking or for drilling down to identify quality problems and to direct solutions. Rather, it is intended to provide a broad programmatic evaluation of
the value of participation in ACS NSQIP. Thus, we focus on modeling 3 important outcomes, which have been consistently defined, using a small, fixed set of stable predictors.

**METHODS**

We used ACS NSQIP data from 2006 through 2013, which are described in Table 1. Mortality and morbidity events (a binary outcome, no events vs 1 or more events) are based on 30-day follow-up regardless of discharge status. Because it was necessary to disregard morbidity events which were not defined consistently from 2006 through 2013, morbidity included only 7 of ACS NSQIP's standard morbidity outcomes: superficial, deep, or organ space surgical site infection (SSI); failure to wean; pneumonia; renal complications; urinary tract infection; cardiac complications; and vein thrombosis requiring therapy/pulmonary embolism. In standard risk adjustment for SAR morbidity models, we account for present-at-time-of-surgery (P ATOS) conditions. Specifically, a patient would not be counted as having had a postoperative event if the associated preoperative condition was recorded as present. For example, postoperative pneumonia would not be attributed to a patient, if pneumonia were present at the time of surgery. However, as P ATOS conditions and definitions have not been consistent over time, none of the morbidity events evaluated in this study were adjusted for P ATOS. Although this strategy will result in generally higher observed event rates, it permits fairer comparisons of those rates over time. In addition to mortality and composite morbidity, we also evaluated SSI.

Among all standard ACS NSQIP predictors, 14 had definitions that were relatively stable across the time period and these were used as predictors in modeling: Relative Value unit, age, gender, ventilator dependence, ascites, history of chronic obstructive pulmonary disease, American Society of Anesthesiologists class, history of congestive heart failure, body mass index, hypertension, smoking, diabetes, dialysis, and Current Procedural Terminology (CPT) code linear risk. CPT linear risk (a continuous variable representing endogenous risk associated with each CPT code) for each outcome was estimated in preliminary models as described elsewhere.1

A 3-level hierarchical model (patients nested within years, nested within hospitals) was applied to each of the 3 outcomes (mortality, morbidity, and SSI) using SAS version 9.4, PROC GLIMMIX (SAS Institute Inc, Cary, NC). Hierarchical models have several advantages including more stabilized estimates resulting from shrinkage (toward the grand mean) imposed with magnitude inversely proportional to sample size.1 The hospital hierarchical O/E ratio was constructed as the quotient of BLUP/NOBLUP (best linear unbiased estimates).
predictor, BLUP; probabilities are based on the random hospital and year effects and the fixed, patient-level effects whereas NOBLUP probabilities are based only on the fixed effects) probabilities summed for all patients within each combination of hospital and year.

The primary results of interest were trends in O/E ratios, by hospital and over time. For purposes of comparison and to provide a measure of the average effect across all hospitals, we computed hospital-averaged log O/E ratio slopes (hospitals carried equal weight) for each hospital cohort for each year, where slopes were estimated by ordinary linear regression, except when there were only 2 years of data. (Logged values are technically superior as they are symmetrically distributed around 1.0, rather than bound by 0 for values below 1 and by positive infinity for values above 1.) Each hospital was assigned to a cohort based on the year in which its data first appeared (8 cohorts corresponding to each year, 2006 through 2013). Separately, we evaluated only hospitals, from each cohort, that are currently in the program (provided data for 2013) and have been in the program for at least 3 years.

In addition, we evaluated time trends in estimated expected risk for patients, or “E.” It could be argued that longer hospital experience in ACS NSQIP might be associated with greater facility by surgical clinical reviewers in identifying preoperative risk factors in patients (although accumulated experience could also affect detection of outcomes). As a consequence, overtime improvement in O/E ratios might therefore be attributed to an upward drift in E, in the absence of any true reduction in risk-adjusted O. This hypothesis is less likely if values of E fail to rise, especially given that many hospitals have recently shifted to higher risk-targeted procedures.

Although we believe that a 3-level hierarchical model is appropriate for these purposes, we confirmed these results using 2 alternative methods. First, a 2-level hierarchical model, with patients nested in (random) years, was constructed for each hospital with 2 or more years in ACS NSQIP during the time period. For these models, parameters (except CPT linear risk) were estimated separately for each hospital model. Second, a logistic model approach (where there were neither random effects for hospital nor year) was applied to all data and logistic O/E ratios were constructed for each quasi-hospital formed by the combination of hospital and year.

**RESULTS**

Results using the 3 different approaches to statistical modeling were very similar so only results for the 3-level models are presented, since these models are most conceptually appealing.

Table 1 shows the number of hospitals and number of cases in each cohort for each year, along with raw event rates for all patients within cohort-year. Hospital dropout rates ranged from 25% [(121-91)/121] for the 2006 cohort to 3% for the 2012 cohort.

Figures 1A to 1C show hospital-averaged raw event rates for each cohort over time for mortality, morbidity, and SSI, respectively. In general, raw event rates are decreasing over time within hospital cohorts and across hospital cohorts—later cohorts have lower raw event rates. The latter effect might be the result of each succeeding cohort having larger percentages of smaller, nonresearch hospitals that are less likely to undertake more complex procedures on higher risk patients.

Figures 2A to 2C show hospital-averaged model-based expected rates for each cohort over time for mortality, morbidity, and SSI, respectively. Expected rates are also shown to decrease over time within hospitals and across hospital cohorts. Thus, later cohorts have both lower event rates and lower predicted event rates. Comparing the patterns in Figures 1 and 2, it does not appear that declines for raw event rates and expected probabilities differ dramatically.

Figures 3A to 3C show hospital-averaged O/E ratios for each cohort over time for mortality, morbidity, and SSI, respectively. O/E ratios are seen to decline over time within a hospital cohort but the orderly segregation of cohorts seen for raw rates and expected rates does not appear; risk adjustment has presumably compensated for differences in patient characteristics and the risk profile of procedures performed that might be associated with cohort.

Figure 4 shows mean hospital slopes, median slope, range of hospital slopes, and percent of hospitals with negative slopes based on O/E ratios and log O/E ratios over time, for mortality, morbidity, and SSI. Mean slopes are generally negative, indicating improving performance for all outcomes, for all cohorts.

Table 2 also shows results for changes on the more appropriate log O/E scale. Depending on whether the original or log scale is used or whether only hospitals that appear in 2013 are considered, Table 2 shows that the percentage of (All) hospitals that improve was 62% for mortality, ranging from 70% to 71% for morbidity, and ranging from 64 to 65% for SSI. Performance was most improved for hospitals currently in the program (they have data for 2013) for at least 3 years (2006-2011 cohorts). The percentage of improving hospitals for this subgroup was 69% for mortality, 79% for morbidity, and 71% for SSI.

Although slopes on the original and log-transformed scale are very close, slopes of logged data have the advantage of being directly interpretable as annual percentage change. Furthermore, because the risk-adjusted rate is the O/E times the reference population rate

![Figure 1](image-url)
FIGURE 2. Hospital-averaged model-based expected rates for each cohort over time for mortality (A), morbidity (B), and SSI (C).

(which is a constant), the magnitude of annual percentage change in the O/E equals the magnitude of annual percentage change in the rate. Thus, on the basis of the observed mean hospital slopes of log O/E ratios we estimate, for all hospitals combined (and in the program as of 2013), a 0.8% annual reduction from the prior year’s mortality rate \[ \text{current year’s rate} = \text{prior year’s rate} \times (1–0.008) \], a 3.1% annual reduction in the rate of patients with 1 or more morbidity events, and a 2.2% annual reduction in the rate of patients with an SSI event. Limiting consideration only to hospitals currently in the program for at least 3 years, we estimate a 0.8% annual reduction in the mortality rate, a 3.1% annual reduction in the rate of patients with 1 or more morbidity events, and a 2.6% annual reduction in the rate of patients with an SSI event.

DISCUSSION

These results confirm that participation in ACS NSQIP, for up to 8 years, is associated with declining O/E ratios (improving performance). We confirmed the effect using 3 different statistical methodologies, which incorporate hospital and year random effects, only the hospital random effect, or no random effects (though only 1 set of statistical model results are presented). The effect is also observed when all hospitals are considered or when analysis is restricted to hospitals that have currently been in the program for at least 3 years. This effect is not due to an upward shift in estimates of risk (values of “E”) over time, as values of E are in fact observed to decline, despite the likely inclusion of more high risk cases as hospitals have transitioned to the ACS NSQIP “procedure-targeted” surgery program in the recent past.

For mortality and morbidity, we observed 62% and 71% of all hospitals improving, compared to 66% and 82% that were reported in the study by Hall et al (Table 3 in this reference—118 hospitals evaluated from 2006 to 2007). The differences in findings from the
### TABLE 2. Hospital-averaged Slope, Median Slope, Range of Hospital Slopes, and Percent of Hospitals With Negative Slopes for O/E Ratios Over Time, for Mortality, Morbidity, and SSI on the Original Scale and When O/E Ratios Are Log Transformed. All Results Are Based on a 3-level Hierarchical Model

<table>
<thead>
<tr>
<th>Cohort*</th>
<th>Mortality</th>
<th>Morbidity</th>
<th>SSI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Median</td>
<td>Range</td>
</tr>
<tr>
<td><strong>O/E Scale</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>-0.008 (0.035)</td>
<td>-0.006</td>
<td>(-0.289, 0.111)</td>
</tr>
<tr>
<td>2007</td>
<td>-0.009 (0.028)</td>
<td>-0.006</td>
<td>(-0.163, 0.042)</td>
</tr>
<tr>
<td>2008</td>
<td>-0.013 (0.032)</td>
<td>-0.010</td>
<td>(-0.120, 0.098)</td>
</tr>
<tr>
<td>2009</td>
<td>-0.015 (0.032)</td>
<td>-0.009</td>
<td>(-0.126, 0.037)</td>
</tr>
<tr>
<td>2010</td>
<td>0.002 (0.031)</td>
<td>-0.000</td>
<td>(-0.066, 0.072)</td>
</tr>
<tr>
<td>2011</td>
<td>-0.005 (0.040)</td>
<td>-0.009</td>
<td>(-0.119, 0.072)</td>
</tr>
<tr>
<td>2012</td>
<td>-0.008 (0.080)</td>
<td>-0.006</td>
<td>(-0.220, 0.152)</td>
</tr>
<tr>
<td>All</td>
<td>-0.008 (0.045)</td>
<td>-0.007</td>
<td>(-0.289, 0.152)</td>
</tr>
<tr>
<td><strong>Log O/E Scale</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>-0.008 (0.028)</td>
<td>-0.006</td>
<td>(-0.215, 0.111)</td>
</tr>
<tr>
<td>2007</td>
<td>-0.009 (0.025)</td>
<td>-0.005</td>
<td>(-0.157, 0.030)</td>
</tr>
<tr>
<td>2008</td>
<td>-0.011 (0.030)</td>
<td>-0.011</td>
<td>(-0.101, 0.101)</td>
</tr>
<tr>
<td>2009</td>
<td>-0.014 (0.029)</td>
<td>-0.009</td>
<td>(-0.106, 0.036)</td>
</tr>
<tr>
<td>2010</td>
<td>0.000 (0.031)</td>
<td>-0.001</td>
<td>(-0.074, 0.024)</td>
</tr>
<tr>
<td>2011</td>
<td>-0.007 (0.038)</td>
<td>-0.010</td>
<td>(-0.094, 0.074)</td>
</tr>
<tr>
<td>2012</td>
<td>-0.009 (0.076)</td>
<td>-0.006</td>
<td>(-0.176, 0.140)</td>
</tr>
<tr>
<td>All</td>
<td>-0.008 (0.042)</td>
<td>-0.007</td>
<td>(-0.215, 0.140)</td>
</tr>
</tbody>
</table>

For computation of slopes, hospitals had to have O/E ratios computed for at least 2 years between their initial year and 2013. For 2006 through 2012, for hospitals in the program in 2013, 91, 39, 29, 30, 26, 70, 73, and 358 hospitals met this criterion, respectively. For 2006 through 2012, for hospitals in the program in 2013, 91, 39, 29, 30, 26, 70, 73, and 358 hospitals met this criterion. These sample sizes were used to generate O/E ratios and percent hospitals improving as reported in the text when attention was restricted to hospitals in the program for at least 3 years.

Percentages of hospitals with negative slopes in the “All” cohort were evaluated using a 1-sample, 2-tailed test against the null of 50. In all cases, the test was rejected at P < 0.001.

*For computation of slopes, hospitals had to have O/E ratios computed for at least 2 years between their initial year and 2013. For 2006 through 2012 and for “All,” 119, 64, 36, 28, 73, and 428 hospitals met this criterion, respectively. For 2006 through 2012, for hospitals in the program in 2013, 91, 39, 29, 30, 26, 70, 73, and 358 hospitals met this criterion. These sample sizes were used to generate O/E ratios and percent hospitals improving as reported in the text when attention was restricted to hospitals in the program for at least 3 years.

Log O/E Scale for hospitals in the program for their initial year and 2013
current work and the prior work might be due to a larger sample size that would exhibit less variation due to chance, a longer observation period (8 years vs 2 years, where the “easy” quality improvement targets might already have been addressed), differences in cohorts (such as any bias toward early participation of high performing hospitals), or differences in statistical methods. In addition, very importantly, there could be a critical difference in that this study employs a more limited list of events constituting morbidity, which was necessary for stability in definitions over time, but which reduces the ability to detect improvements on multiple axes. However, overall, the current and prior studies are reasonably consistent.

Although the percentage of all hospitals that are improving is slightly smaller than that reported by Hall et al, hospitals that are currently in the program for at least 3 years are estimated to actually improve at about the same rate as the historical figure. Mortality rate is improving for 69% of these hospitals (vs 66% in Hall et al) and morbidity is improving for 79% of these hospitals (vs 82%). It is reasonable to conclude that participation in ACS NSQIP is more likely to lead to improvement when hospitals are committed to the program (are current as of 2013) and have had time to collect data, receive risk-adjusted reports, implement quality improvement projects, and evaluate resultant change. Mere “participation” in ACS NSQIP is not a panacea; hospitals must make a long-term commitment to using the information provided to guide quality improvement efforts.

Using rate reductions estimated for hospitals currently in the program for at least 3 years and initial rates from the 2006 cohort, we observe annual relative reductions of 0.8% from the baseline overall mortality rate of 1.81%, annual relative reductions of 3.1% from the baseline overall “any” morbidity rate of 10.32%, and annual relative reductions of 2.6% from the baseline overall SSI rate of 5.34%. These are cumulative annual reductions, so changes that might at first seem small in fact become substantial. Thus, in 5 years’ time, a starting mortality rate of 1.81% would be reduced to 1.74% [= 1.81 × (1 − 0.008)5], for morbidity a starting rate of 10.32% would be reduced to 8.82% [= 10.32 × (1 − 0.031)5], and for SSI a starting rate of 5.34% would be reduced to 4.68% [= 5.34 × (1 − 0.026)5]. In the fifth year, for every 10,000 surgical procedures, similar to this data set, the improving hospital would have avoided 7 deaths [(0.0181 − 0.0174) × 10,000], converted 150 [(0.1032 − 0.0882) × 10,000] patients from having 1 or more complications to having none, and converted 66 [(0.0534 − 0.0468) × 10,000] patients from having 1 or more SSI to having none. We emphasize, this is the annual improvement, observed in the fifth year. A large hospital (800 − 1000 beds) might perform twice this many procedures (reflecting the risk levels captured in this data set) annually and thus save twice as many adverse events (in raw number) annually by the fifth year (14 deaths, 300 morbidities, and 132 surgical infections). Furthermore, because these outcomes (for morbidity and SSI) are “any” in structure, these reductions are underestimate of true improvement—these estimates are for patients experiencing “one or more” events, and transitioning to no events. Patients who would otherwise have 3 events and only suffer 2, or would have 2 events and only suffer 1, are not fully reflected as improvement by this current method. In future work, modeling by counts could be applied to refine these estimates. In addition, as noted earlier, because this work included only outcome events that were consistently defined over this time period, these results also underestimate improvement that might be observed across all adverse events.

Although there is a general downward trend in mean hospital O/E slope, cohorts do perform somewhat differently. A determination of factors contributing to these differences is beyond the scope of this article but, as risk adjustment presumably accounts for patient and procedure risk, this may involve differences in hospital characteristics associated with cohort, including initial quality, quality culture, and surveillance mechanisms. Differences in cohorts are clearly evidenced by patterns observed in raw event rates and predicted rates. Using results from the 3-level model, the ACS NSQIP would be able to provide individual hospitals with a summary of their performance during their participation in ACS NSQIP. Regardless of shifts over time in their performance relative to other hospitals at any particular time (viz, as reported in sequential SARs), a determination of absolute change across time could be generated from this approach. Most hospitals are improving in surgical quality and this is important information that can be used in many contexts including cost/benefit analyses of participation in ACS NSQIP. “Over time” results would also provide an alternative criterion (compared to the existing SAR reporting) for identifying hospitals most in need of surgical quality improvement and other hospitals that could share their methodologies for achieving quality improvement goals.

The effects of ACS NSQIP observed here cannot be evaluated as if hospital participation/nonparticipation was randomly assigned and all other factors were controlled. As has been suggested earlier, hospitals that chose to participate early (2006) versus late (2013) may not be the same in terms of institutional culture, motivation to change, or initial quality. In addition, ACS NSQIP has evolved so that the program’s ability to influence quality has probably increased over time, and there have been other important influences on surgical quality over the time period, apart from participation in ACS NSQIP.

The potential influence of secular trends, unrelated to ACS NSQIP, cannot be isolated based on this work, though there are some data to suggest that the effects observed coincident with NSQIP cannot be entirely due to unrelated trends. This is an important issue as the period from 2006 through 2013 saw exponential growth in quality and performance improvement efforts in medicine and in surgery, with respect to both processes and outcomes, motivated in part by intrinsic provider culture, patient advocates, payors, and private and governmental agencies. Clearly, these other factors cannot be excluded from contributing to the observed trends among ACS NSQIP hospitals.

Although our methodology cannot segregate the influence of ACS NSQIP participation from “everything else,” a perfect methodological solution may not be available. Although a control group of non-ACS NSQIP hospitals exists, there is no structure for randomization, nor can non-ACS NSQIP hospitals be expected to provide the type of clinical data needed for an analysis of scale and duration undertaken here. It is also true that contributions from “everything else” may not be free of the influence of ACS NSQIP. ACS NSQIP is capable of improving performance among participating hospitals, but, as an etiological factor in the evolution of quality improvement in surgery generally, it has also contributed to the quality culture that now motivates other hospitals and agencies to improve. Thus, although ACS NSQIP can improve quality in participating hospitals directly, it has also changed the quality improvement landscape for both participating and nonparticipating hospitals indirectly. Although the absence of a non-ACS NSQIP participating control group is methodologically problematic, the beneficial cultural carryover effect of ACS NSQIP to nonparticipating hospitals is also difficult to isolate. Thus, although results from this and other studies lead us to believe that participation in ACS NSQIP improves quality to a greater extent than secular trend (or is a critical component of observed trends), the current data cannot demonstrate this unequivocally. The issue of quantifying causative attribution remains open to continuing research.

Although these results indicate improved surgical outcomes, we have not estimated the magnitude of the effect precisely in terms of “counts” of adverse events avoided. Cases that hospitals submit to ACS NSQIP are, in substantial part, determined so as to meet sample...
size requirements for self-identified areas of interest (eg, general/vascular surgery, any of 8 other surgical specialties, or any of 34 surgical targets), which may change over time. Because raw event rates are different across surgical groupings and because different groupings might present greater or lesser opportunities to improve surgical outcomes, it is problematic to translate changes in O/E ratios to changes in event rates. Although it is not uncommon to multiply an O/E ratio by a global event rate to arrive at a risk-adjusted rate, for the reasons described, this could be inexact. It also needs to be understood that (excluding postoperative death) morbidity and SSI are treated as simply present or absent—‘all or none.’ This work does not model the number of events, only whether there were 1 or more events. For these reasons, a percentage decline in O/E ratios over time will underestimate the percentage decline in total adverse events. Finally, as mentioned, we only considered adverse outcomes whose definitions had not changed from 2006 through 2013. Thus, we are not considering reductions in other outcomes. A further consequence of this is that if hospitals had chosen to target the unexamined outcomes for quality improvement, and not others, we might not detect a negative slope for our limited morbidity measure in which that outcome was not included, again potentially underestimating the true improvement effect.

CONCLUSIONS
Commitment to ACS NSQIP participation is associated with improved surgical outcomes. The magnitude of that improvement depends on the duration of participation and presumably also efforts devoted toward using the information provided to direct quality improvement efforts. By 5 years’ time, a large hospital could conservatively be avoiding 14 deaths annually, and annually converting 300 patients from any complication to none, and 132 patients from any SSI to none. These figures are likely to be underestimates of the overall impact and, furthermore, hospitals are likely to continue to improve.

REFERENCES