Effectiveness of Nonpublic Report Cards for Reducing Trauma Mortality

Laurent G. Glance, MD; Turner M. Osler, MD; Dana B. Mukamel, PhD; J. Wayne Meredith, MD; Andrew W. Dick, PhD

T
he release of the influential Institute of Medicine report on patient safety, 1 citing medical errors as the 8th leading cause of death fueled demand to use quality measurement as a catalyst for improving health care quality. Efforts by the Veterans Administration (VA), 2 New York state, 3 and hospitals in Northern New England 4 showed that nonpublic reporting was associated with significant reductions in mortality and morbidity in patients undergoing cardiac and noncardiac surgery. More recently, the American College of Surgeons 5 and the Society of Thoracic Surgeons 6 have spearheaded national efforts to improve patient outcomes using performance benchmarking. The Centers for Medicare and Medicaid Services publicly reports mortality rates for patients hospitalized with acute myocardial infarctions, heart failure, and pneumonia, 7 and it is expanding these efforts to include many other areas of health care. 8 The need to control runaway health care spending has further magnified the need for quality measurement to ensure that health care quality is not sacrificed to save health care dollars.

Twenty-five years ago, the American College of Surgeons (ACS) created the National Trauma Databank (NTDB) as a “foundation for evidence-based practice, performance improvement, and research.” 9 At its inception, this registry was not used to provide participating hospitals with information on their risk-adjusted outcomes. There is now mounting evidence that patient outcomes following traumatic injury are determined not only by the dose of trauma, but also by the hospital where the patient is treated. 10-12 There is remarkable variability in trauma mortality outcomes across hospitals, with up to 4-fold differences in risk-adjusted mortality rates between the best...

IMPORTANCE
An Institute of Medicine report on patient safety that cited medical errors as the 8th leading cause of death fueled demand to use quality measurement as a catalyst for improving health care quality.

OBJECTIVE
To determine whether providing hospitals with benchmarking information on their risk-adjusted trauma mortality outcomes will decrease mortality in trauma patients.

DESIGN, SETTING, AND PARTICIPANTS
Hospitals were provided confidential reports of their trauma risk-adjusted mortality rates using data from the National Trauma Data Bank. Regression discontinuity modeling was used to examine the impact of nonpublic reporting on in-hospital mortality in a cohort of 326,206 trauma patients admitted to 44 hospitals, controlling for injury severity, patient case mix, hospital effects, and preexisting time trends.

MAIN OUTCOMES AND MEASURES
In-hospital mortality rates.

RESULTS
Performance benchmarking was not significantly associated with lower in-hospital mortality (adjusted odds ratio [AOR], 0.89; 95% CI, 0.68-1.16; P = .39). Similar results were obtained in secondary analyses after stratifying patients by mechanism of trauma: blunt trauma (AOR, 0.91; 95% CI, 0.69-1.20; P = .51) and penetrating trauma (AOR, 0.75; 95% CI, 0.44-1.28; P = .29). We also did not find a significant association between nonpublic reporting and in-hospital mortality in either low-risk (AOR, 0.84; 95% CI, 0.57-1.25; P = .40) or high-risk (AOR, 0.88; 95% CI, 0.67-1.17; P = .38) patients.

CONCLUSIONS AND RELEVANCE
Nonpublic reporting of hospital risk-adjusted mortality rates does not lead to improved trauma mortality outcomes. The findings of this study may prove useful to the American College of Surgeons as it moves ahead to further develop and expand its national trauma benchmarking program.

Published online December 11, 2013.
and worst-performing hospitals. This quality gap presents an opportunity to improve trauma outcomes using performance feedback to bridge the divide between lower-performance and higher-performance hospitals.

With funding from the Agency for Healthcare Research and Quality, and in collaboration with the ACS, we conducted a prospective trial to test whether nonpublic reporting leads to lower trauma mortality. Participating hospitals were provided with detailed reports of their risk-adjusted mortality outcomes. We designed this study to examine the feasibility and impact of using the data infrastructure in the NTDB to improve trauma outcomes using nonpublic report cards. In our analysis examining the impact of nonpublic reporting, we controlled for temporal trends and hospital effects, in addition to controlling for patient case mix and injury severity. The objective of this article was to report the findings of this trial.

Methods

Data Source and Study Population

The University of Rochester School of Medicine institutional review board approved this study; the need for written informed patient consent was waived. This study was designed to determine whether nonpublic reporting leads to a reduction in in-hospital mortality in injured patients using data from the NTDB. The NTDB was created by the ACS to serve as a national repository for trauma center data. The NTDB includes the following data elements: patient demographics, hospital demographics, International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) diagnostic and injury codes, encrypted hospital identifiers, physiology values, and in-hospital mortality.

National Trauma Databank coding practices dictate how missing data, invalid data, and inconsistent data are handled once the data have been transmitted to the NTDB. Data reports submitted by individual hospitals are checked by the NTDB using software edit tools. Internal consistency is assessed by comparing the values for related data elements. For example, the intensive care unit length of stay must be less than the total hospital stay. Each hospital submitting data to the NTDB is given a screening report and has the opportunity to resubmit their data after correcting their errors.

At the inception of this trial, we identified 2 separate hospital cohorts depending on whether hospitals were primarily using either ICD-9-CM or Abbreviated Injury Scale (AIS) codes to code patient injuries. We developed and validated 2 versions of the Trauma Mortality Prediction Model (TMPM)—one based on AIS injury codes and the other based on ICD-9-CM codes—whose statistical performance was superior to existing standard injury models, Injury Severity Score (ISS) and ICISS (an International Classification of Diseases, Ninth Revision-based injury severity score). In 2008, the ACS National Trauma Data Standard was revised to mandate the use of ICD-9-CM codes to characterize injury severity and made AIS injury codes optional. In 2008, we sent hospitals report cards based on either TMPM–ICD-9 or TMPM-AIS (using 2006 data), depending on whether they coded injury data using either ICD-9-CM or AIS codes. Starting in 2009 and then in 2010, participating hospitals received annual report cards based on TMPM–ICD-9 because of the mandated change in coding practices. Because TMPM–ICD-9 and TMPM-AIS are based on 2 completely different sets of injury codes—ICD-9-CM and AIS codes—we limited our analysis to those hospitals that coded injuries using ICD-9-CM codes and received report cards based on TMPM–ICD-9 in 2008-2010 to avoid confounding the intervention effect (nonpublic reporting starting in 2008) with the changeover in injury coding (starting in 2007). Each hospital was provided with benchmarking information on their risk-adjusted in-hospital mortality for their entire patient cohort, as well as separate reports stratified by mechanism of trauma (blunt, gunshot wound, motor vehicle crash, pedestrian, and high risk). A sample report card is shown in the eAppendix (Supplement).

After excluding patients with burns, unspecified injuries, nontraumatic injuries, or missing mechanisms of injury, as well as patients who were dead on admission or transferred out to another hospital, our study sample included 330 700 patients in 44 hospitals (Figure 1). We excluded patients with missing demographic data, invalid ICD-9-CM codes, missing empirical injury severities (MARC values), and missing ICD-9-CM codes. The final analytic sample consisted of 326 206 patients admitted to 44 hospitals (Figure 1).

Statistical Analysis

The aim of this analysis was to determine whether the introduction of nonpublic reporting was associated with lower in-hospital mortality after controlling for patient and hospital factors, as well as controlling for possible temporal trends toward lower mortality rates during the study. This analysis was performed using regression discontinuity modeling, which is an econometric technique that identifies the effect of an intervention (introduction of nonpublic reporting) in a pre-study and poststudy by estimating the intercept shift, controlling for patient and hospital factors and for preexisting temporal trends.

To perform regression discontinuity analysis, we estimated a patient-level logistic regression model to examine the association between in-hospital mortality and the initiation of nonpublic reporting. A dummy variable was used to indicate whether a patient was admitted before or after nonpublic reporting was initiated. We controlled for secular trends by including year of admission as a categorical variable, omitting data from 2008 (when report cards were initiated). We also included an interaction term between report card and year to examine whether the slope of the time trend changed after initiation of reporting. We controlled for patient risk factors using an enhanced version of TMPM–ICD-9: patient age, sex, injury severity of the 5 most severe injuries, transfer status, mechanism of injury (ie, blunt, gunshot wound, motor vehicle crash, stab injury, pedestrian, and low fall), motor component of the Glasgow Coma Scale, and systolic blood pressure. We also controlled for hospital-fixed effects by including a separate indicator variable for each hospital. By including hospital-fixed effects, we were able to identify whether nonpublic reporting led to a mortality reduction within hospitals by allowing each hospital to have a unique intercept.
hospital to act as its own control. Missing values of the motor component of the Glasgow Coma Scale and systolic blood pressure were imputed using the Stata implementation of the multiple imputation by chained equations method described by van Buuren et al.\(^2^2\) Fractional polynomial analysis was used to determine the optimal specification of age.\(^2^3\)

Several sensitivity analyses were performed to examine the robustness of the analysis of the impact of nonpublic reporting. We performed stratified analyses in which we limited the patient cohort to patients with either blunt or penetrating trauma and to either low-risk (predicted probability of death <5%) or high-risk (predicted probability of death ≥5%) patients. We also performed stratified analyses in which we only included low-performance, average-performance, or high-performance hospitals. Hospital performance was estimated in 2006 data using hierarchical logistic regression, based on TMPM–ICD-9, in which hospitals were specified as a random effect. The empirical Bayes estimate of the hospital effect was exponentiated to yield an adjusted odds ratio (AOR).\(^2^4\) Hospitals whose AORs were significantly lower than 1 were classified as high-performance outliers, whereas hospitals with AORs significantly greater than 1 were classified as low-performance outliers.

Data management and statistical analyses were performed using Stata SE/MP version 11.0 (StataCorp). Robust variance estimators were used because patient observations were clustered by hospital.\(^2^5\) All statistical tests were 2-tailed and \(P\) values less than .05 were considered significant.

### Results

Patient characteristics before (2006-2007) and after (2008-2010) initiation of nonpublic reporting are shown in eTable 1 in the Supplement. Overall, the median age was 40 years, most patients were male (66%), 24% were transferred in from other hospitals, and most patients sustained injuries from either blunt trauma (42%) or motor vehicle crashes (23%). The observed mortality rate was 4.19%. No clinically significant changes in patient case mix were detected over the 5-year study. Hospital characteristics are shown in eTable 2 in the Supplement. Most hospitals were either level I (43%) or level II (43%) trauma centers, nearly half were university hospitals, and nearly all were nonprofit. Most hospitals had between 200 and 400 beds (39%) or greater than 400 beds (54%). All geographic regions in the United States were well represented: Northeast (16%), South (39%), Midwest (16%), and West (27%).

After controlling for patient characteristics, hospital factors, and preexisting time trends, nonpublic reporting did not have a significant impact on in-hospital mortality (AOR, 0.89; 95% CI, 0.68-1.16; \(P = .39\) (Table 1 and Figure 2A). When we stratified patients by mechanism of trauma, we found similar findings: blunt trauma (AOR, 0.91; 95% CI, 0.69-1.20; \(P = .51\)) and penetrating trauma (AOR, 0.75; 95% CI, 0.44-1.28; \(P = .29\)) (Table 1 and Figure 2B). We also did not find a significant association between nonpublic reporting and in-hospital mor-

<table>
<thead>
<tr>
<th>Variable</th>
<th>All Patients</th>
<th>Blunt</th>
<th>Penetrating</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adjusted Odds Ratio (95% CI)</td>
<td>(P) Value</td>
<td>Adjusted Odds Ratio (95% CI)</td>
</tr>
<tr>
<td>Change in mortality adjusted for patient characteristics</td>
<td>0.96 (0.88-1.05)</td>
<td>.43</td>
<td>0.95 (0.88-1.03)</td>
</tr>
<tr>
<td>Change in mortality adjusted for patient characteristics and hospital effects</td>
<td>0.95 (0.88-1.04)</td>
<td>.26</td>
<td>0.94 (0.87-1.02)</td>
</tr>
<tr>
<td>Change in mortality adjusted for patient characteristics, hospital effects, and time trends</td>
<td>0.89 (0.68-1.16)</td>
<td>.39</td>
<td>0.91 (0.69-1.20)</td>
</tr>
</tbody>
</table>

---

ICD-9 indicates International Classification of Diseases, Ninth Revision; MARC, empirical injury severities.
Similar to all the other analyses, we found that nonpublic reporting was not associated with improved outcomes in patients treated in low- (AOR, 1.17; 95% CI, 0.65-1.23; \( P = .61 \)), average- (AOR, 0.89; 95% CI, 0.65-1.23; \( P = .49 \)), or high-performance (AOR, 0.74; 95% CI, 0.34-1.57; \( P = .43 \)) hospitals (Table 3 and Figure 2E).

The graphs show the trends in the adjusted mortality rates for all patients (A), patients with blunt or penetrating trauma (B), low-risk patients (predicted probability of death <5%) (C), high-risk patients (predicted probability of death \( \geq 5 \% \)) (D), and hospitals stratified by performance strata, adjusting for patient risk factors (E). Nonpublic reporting was initiated in 2008.
and complication rates by 42%. Under the leadership of the ACS, the NSQIP was expanded to include hospitals outside of the VA. The SMARTT (Survival Measurement and Reporting Trial for Trauma) reporting initiative did not find that the program was associated with decreases in the mortality rates for patients hospitalized with acute myocardial infarction, heart failure, and pneumonia since 2005. Controlling for preexisting time trends in mortality, there is no evidence that this public reporting initiative resulted in a mortality reduction for acute myocardial infarction and pneumonia, and only a modest 3% relative risk reduction for heart failure.

The question arises as to why we found no association between the SMARTT (Survival Measurement and Reporting Trial for Trauma) reporting initiative and mortality. It is possible that performance benchmarking is not as effective as once believed. Although the early results from the VA NSQIP were very impressive, the effect of the ACS NSQIP on mortality outcomes were less dramatic and did not account for the fact that the benefit observed with nonpublic reporting may have been owing instead to preexisting temporal trends and unrelated to performance feedback. Furthermore, the negative findings of the more methodologically rigorous and larger Hospital Compare study in nonsurgical patients stand in sharp contrast to the earlier NSQIP studies.

There are historical precedents to suggest that nonpublic reporting could be expected to lead to improved outcomes. The VA National Surgical Quality Improvement Program (NSQIP) was established by congressional mandate in response to concerns about the quality of surgical care in VA hospitals. The VA prospectively collected clinical data on patient risk and outcomes and provided VA hospitals with information on how to improve patient outcomes. In the VA, hospital-sharable site visits, best practices from high-performance hospitals are disseminated to all hospitals in the VA and ACS NSQIP.

### Table 2. In-Hospital Mortality Before and After Implementation of Nonpublic Reporting

<table>
<thead>
<tr>
<th>Variable</th>
<th>All Patients</th>
<th>Mortality Risk</th>
<th>Low</th>
<th>Average</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in mortality adjusted for patient characteristics</td>
<td>Adjusted Odds Ratio (95% CI)</td>
<td>P Value</td>
<td>Adjusted Odds Ratio (95% CI)</td>
<td>P Value</td>
<td>Adjusted Odds Ratio (95% CI)</td>
</tr>
<tr>
<td>Change in mortality adjusted for patient characteristics and hospital effects</td>
<td>0.95 (0.88-1.04)</td>
<td>.26</td>
<td>0.99 (0.89-1.10)</td>
<td>.82</td>
<td>0.93 (0.86-1.01)</td>
</tr>
<tr>
<td>Change in mortality adjusted for patient characteristics, hospital effects, and time trends</td>
<td>0.89 (0.68-1.16)</td>
<td>.39</td>
<td>0.84 (0.57-1.25)</td>
<td>.40</td>
<td>0.88 (0.67-1.17)</td>
</tr>
</tbody>
</table>

### Table 3. In-Hospital Mortality Before and After Implementation of Nonpublic Reporting

<table>
<thead>
<tr>
<th>Variable</th>
<th>Performance Hospitals</th>
<th>Low</th>
<th>Average</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in mortality adjusted for patient characteristics</td>
<td>Adjusted Odds Ratio (95% CI)</td>
<td>P Value</td>
<td>Adjusted Odds Ratio (95% CI)</td>
<td>P Value</td>
</tr>
<tr>
<td>Change in mortality adjusted for patient characteristics and hospital effects</td>
<td>0.94 (0.74-1.18)</td>
<td>.58</td>
<td>0.93 (0.84-1.03)</td>
<td>.18</td>
</tr>
<tr>
<td>Change in mortality adjusted for patient characteristics, hospital effects, and time trends</td>
<td>1.17 (0.65-2.10)</td>
<td>.61</td>
<td>0.89 (0.65-1.23)</td>
<td>.49</td>
</tr>
</tbody>
</table>

### Discussion

We did not find significant evidence that providing hospitals with nonpublic reports of their risk-adjusted trauma mortality rates was associated with improvement in trauma mortality, even after controlling for patient case mix and preexisting time trends. We found similar findings when we limited our analysis to patients with either blunt trauma, penetrating trauma, or either a low risk for death or high risk for death. We also did not find evidence that nonpublic reporting had a differential effect depending on whether a hospital was a low-, average-, or high-performance hospital.

There are historical precedents to suggest that nonpublic reporting could be expected to lead to improved outcomes. The VA National Surgical Quality Improvement Program (NSQIP) was established by congressional mandate in response to concerns about the quality of surgical care in VA hospitals. The VA prospectively collected clinical data on patient risk and outcomes and provided VA hospitals with information on how to improve patient outcomes. In the VA, hospital-sharable site visits, best practices from high-performance hospitals are disseminated to all hospitals in the VA and ACS NSQIP.

mortality for all Medicare patients with acute myocardial infarction, heart failure, and pneumonia since 2005. Controlling for preexisting time trends in mortality, there is no evidence that this public reporting initiative resulted in a mortality reduction for acute myocardial infarction and pneumonia, and only a modest 3% relative risk reduction for heart failure.

The question arises as to why we found no association between the SMARTT (Survival Measurement and Reporting Trial for Trauma) reporting initiative and mortality. It is possible that performance benchmarking is not as effective as once believed. Although the early results from the VA NSQIP were very impressive, the effect of the ACS NSQIP on mortality outcomes were less dramatic and did not account for the fact that the benefit observed with nonpublic reporting may have been owing instead to preexisting temporal trends and unrelated to performance feedback. Furthermore, the negative findings of the more methodologically rigorous and larger Hospital Compare study in nonsurgical patients stand in sharp contrast to the earlier NSQIP studies.

It is also possible that benchmarking alone is not sufficient and that reporting initiatives need to be tied to financial incentives. However, evaluations of the Premier and Centers for Medicare and Medicaid Services Hospital Quality Incentive Demonstration program did not find that the program was associated with decreases in the mortality rates for patients hospitalized with acute myocardial infarction, heart failure, or pneumonia. Finally, it is also possible that providing hospitals with benchmarking information, with or without financial incentives, is not enough to improve outcomes without also providing hospitals with information on how to improve patient outcomes. In the VA, hospitals have the option of inviting the NSQIP to conduct consultative structured site visits, and best practices from high-performance hospitals are disseminated to all hospitals in the VA and ACS NSQIP.
ever, the evidence that improved adherence to best practices is associated with better outcomes is relatively modest in both surgical and nonsurgical patients.31,32

Our study has several potential limitations. First, we did not know to what extent participating hospitals actually used the benchmarking information to guide their quality improvement efforts. Second, our hospital study cohort was limited to a subset of hospitals in the NTDB with a low incidence of missing data and to hospitals that used ICD-9-CM codes throughout the study period; therefore, it is not necessarily representative of all hospitals caring for trauma patients. Although we found no evidence of important and sustained effects, a larger hospital sample would have provided greater statistical precision. Third, we used ICD-9-CM diagnostic codes as the basis for injury coding, as opposed to clinical AIS injury codes, because of the change in the national standard for trauma coding that occurred during the study. However, the ICD-9-based trauma mortality prediction model used to produce the quality reports in our analyses has been previously validated and shown to have excellent statistical properties.18

Fourth, because the NTDB does not have reliable comorbidity data, we did not include comorbidities in our benchmarking reports, nor did we include them in our analyses. However, we have previously shown, using all-payer administrative data, that omitting comorbidity data from risk adjustment does not have a substantial impact on hospital quality measurement in trauma.19 Finally, our analysis was limited to in-hospital mortality and did not examine complications, functional outcomes, readmissions, or long-term outcomes because of limitations of the NTDB.

Our findings have potentially important implications for the ACS as it continues to expand the ACS Trauma Quality Improvement Program (TQIP). This national trauma benchmarking program grew out of a pilot study of 23 centers in 2008 to include more than 160 trauma centers.33,34 The TQIP is modeled after the ACS NSQIP and assumes that TQIP benchmarking reports will serve as a catalyst for performance improvement and lead to better patient outcomes. Our findings point to the limitations of performance feedback as a mechanism for improving trauma outcomes. However, current efforts to shift financial risk to hospital and physician groups under the Affordable Health Care Act35 may provide very strong incentives for hospitals to decrease costly complications and improve patient outcomes. As we move toward a future health care model where hospitals and physicians are increasingly held accountable for patient and financial outcomes, detailed benchmarking reports from the ACS TQIP may provide critical information that trauma centers can use to identify and more effectively target quality problems.

Conclusions

In summary, our study suggests that nonpublic reporting of hospital risk-adjusted mortality rates does not lead to improved trauma mortality outcomes. It is possible that adding other interventions to performance feedback, such as structured site visits, efforts to identify and disseminate best practices, and meaningful financial incentives, will lead to improved trauma outcomes. The findings of this study may prove useful to the ACS leadership as it moves ahead to further develop and expand its national trauma benchmarking program. It may be possible for the ACS TQIP, using a much larger hospital sample, to reexamine the value of nonpublic reporting, possibly in combination with other quality-improvement strategies.

ARTICLE INFORMATION

Accepted for Publication: April 19, 2013.
Published Online: December 11, 2013.

Author Contributions: Dr Glance 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 of data: Glance.
Analysis and interpretation of data: Glance, Osler, Mukamel, Dick.
Drafting of the manuscript: Glance.
Critical revision of the manuscript for important intellectual content: All authors.
Statistical analysis: Glance, Osler, Mukamel, Dick.
Obtained funding: Glance, Mukamel.
Administrative, technical, or material support: Glance, Meredith.
Study supervision: Glance.
Conflict of Interest Disclosures: None reported.
Funding/Support: This project was supported by grant ROI HS 16737 from the Agency for Healthcare Research and Quality.

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

Disclaimer: The views presented in this manuscript are those of the authors and may not reflect those of the Agency for Healthcare Research and Quality.

REFERENCES

7. Ryan AM, Nallamothu BK, Dimick JB. Medicare’s public reporting initiative on hospital quality had modest or no impact on mortality from three key conditions. Health Aff (Millwood). 2012;31(3):585-592.
8. VanLare JM, Blum JD, Conway PH. Linking performance with payment: implementing the Physician Value-Based Payment Modifier. JAMA. 2012;308(20):2089-2090.
Hospital Report Cards Necessary but Not Sufficient?
Justin B. Dimick, MD, MPH; Samantha K. Hendren, MD, MPH

Performance feedback is increasingly used as a strategy to motivate surgeons to improve. Numerous surgical specialties have launched clinical registries to generate reports of risk-adjusted outcomes including cardiac, bariatric, general, vascular, and trauma surgical procedures. These reports are fed back to hospitals, allowing them to identify problem areas such as higher-than-expected complication rates. In theory, hospitals then implement changes to address these shortcomings. Proponents of performance feedback argue that hospitals and surgeons, once they know where they do not measure up, will make the necessary changes to improve outcomes. However, many believe that this view is too optimistic; that is, that feedback alone is not enough to spur improvement. In this issue of JAMA Surgery, Glance et al provide strong evidence supporting the latter.

Glance et al conducted a rigorous evaluation of the impact of performance reports on risk-adjusted mortality for trauma patients. The study was a natural experiment in the setting of the National Trauma Data Bank (NTDB) maintained by the American College of Surgeons. Using an interrupted time-series design, Glance et al found no improvement in mortality (no change in the slope) after NTDB began providing benchmarking reports to hospitals. The methods in the study are state of the art and leave very little room for criticism. Therefore, we believe the study provides strong evidence that, in this context, performance reports are not enough to motivate improvement.

Although this study is among the first to rigorously evaluate performance feedback in surgery, there is a robust literature in other health care settings. A recent Cochrane meta-analysis demonstrated a positive, but very small, benefit of the effectiveness of audit and feedback alone on physician performance. Because the reports in the NTDB are provided