

Association Between Hospital-Reported Leapfrog Safe Practices Scores and Inpatient Mortality

Leslie P. Kernisan, MD

Sei J. Lee, MD, MAS

W. John Boscardin, PhD

C. Seth Landefeld, MD

R. Adams Dudley, MD, MBA

THE LEAPFROG GROUP IS A WELL-known nonprofit business coalition that provides information regarding hospital safety and quality to its members (large companies that purchase health care) and to consumers.¹⁻³ Its primary method of evaluating hospitals is via voluntary participation in the Leapfrog Hospital Survey. Initially, these annual surveys assessed hospitals' adoption of 3 initiatives that the organization believed would improve patient safety: computerized physician order entry, staffing of intensive care units by trained intensivist physicians, and evidence-based referrals for high-mortality surgeries. In 2004 a fourth initiative, the Safe Practices Survey, was added to the hospital survey to allow hospitals to report efforts toward implementing the National Quality Forum's Safe Practices for Better Healthcare.⁴

Approximately 1100 urban hospitals have completed the Safe Practices Survey in recent years, and these results are reported to the public on the Internet (available at <http://www.leapfroggroup.org>). Nevertheless, it remains unclear how well quality as assessed by the Safe Practices Survey (consisting of hospitals' self-report of

Context The Leapfrog Hospital Survey allows hospitals to self-report the steps they have taken toward implementing the Safe Practices for Better Healthcare endorsed by the National Quality Forum. The Leapfrog Group currently ranks hospital performance on the safe practices initiative by quartiles and presents this information to the public on its Web site. It is unknown how well a hospital's resulting Safe Practices Score (SPS) correlates with outcomes such as inpatient mortality.

Objective To determine the relationship between hospitals' SPSs and risk-adjusted inpatient mortality rates.

Design, Setting, and Participants Observational analysis of discharge data for all urban US hospitals completing the 2006 safe practices initiative and identifiable in the Nationwide Inpatient Sample. Leapfrog provided an SPS for each hospital as well as 3 alternative scores based on shorter versions of the original survey. Hierarchical logistic regression was used to determine the relationship between quartiles of SPS and risk-adjusted inpatient mortality, after adjusting for hospital discharge volume and teaching status. Subgroup analyses were performed using data from patients older than 65 years and patients with 5% or greater expected mortality risk.

Main Outcome Measures Inpatient risk-adjusted mortality by quartiles of survey score.

Results Of 1075 hospitals completing the 2006 Safe Practices Survey, 155 (14%) were identifiable in the National Inpatient Sample (1 772 064 discharges). Raw observed mortality in the primary sample was 2.09%. Fully adjusted mortality rates by quartile of SPS, from lowest to highest, were 1.97% (95% confidence interval [CI], 1.78%-2.18%), 2.04% (95% CI, 1.84%-2.25%), 1.96% (95% CI, 1.77%-2.16%), and 2.00% (95% CI, 1.80%-2.22%) ($P = .99$ for linear trend). Results were similar in the subgroup analyses. None of the 3 alternative survey scores was associated with risk-adjusted inpatient mortality, although P values for linear trends were lower (.80, .20, and .11).

Conclusion In this sample of hospitals that completed the 2006 Safe Practices Survey, survey scores were not significantly associated with risk-adjusted inpatient mortality.

JAMA. 2009;301(13):1341-1348

www.jama.com

structural and process measures) correlates with outcomes of interest to patients and policy makers, such as mortality. To date, each of the first 3 initiatives has been the subject of mul-

iple peer-reviewed studies,⁵⁻¹¹ and a recent study examining the first 3 initiatives collectively did find some positive associations between survey performance and reduced mortality.¹²

Author Affiliations: Divisions of Geriatrics (Drs Kernisan, Lee, Boscardin, and Landefeld) and Biostatistics (Dr Boscardin) and Philip R. Lee Institute for Health Policy Studies (Dr Dudley), University of California, San Francisco; National VA Quality Scholars Program (Drs Kernisan, Lee, and Landefeld) and Health Services Research Enhancement Award Program (Drs Lee, Boscardin, and Landefeld), San Francisco VA Medical

Center, San Francisco, California; and Center for Advanced Study in the Behavioral Sciences (Dr Landefeld), Stanford University, Palo Alto, California.

Corresponding Author: R. Adams Dudley, MD, MBA, Philip R. Lee Institute for Health Policy Studies, University of California, San Francisco, 3333 California St, Ste 265, San Francisco, CA 94118 (adams.dudley@ucsf.edu).

See also p 1375.

However, little has been published regarding the Safe Practices Survey, which is the most time-consuming component to complete. In particular, to our knowledge it is not yet confirmed that higher scores on the survey correlate with actual outcomes. This issue is pertinent, because survey scores reported on the Internet are ranked by quartiles, which likely suggests to consumers that hospitals in the highest quartile provide safer care than those in lower quartiles.¹³

In this article we present an analysis examining the relationship between scores achieved by urban hospitals on the 2006 Safe Practices Survey and risk-adjusted inpatient mortality. To address questions of generalizability and response bias, we also present comparisons between hospitals that participated in the survey and those that did not.

METHODS

Data and Sample

Mortality data were obtained from the most recent version (2005, available in 2007) of the Nationwide Inpatient Sample (NIS).¹⁴ This database includes inpatient discharge data collected via federal-state partnerships as part of the Agency for Healthcare Research and Quality's Healthcare Cost and Utilization Project. The 2005 NIS contains administrative data on all discharges from 1054 hospitals located in 37 states, approximating a 20% stratified sample of US hospitals (7 995 048 discharges total).¹⁵ Of these 1054 hospitals, 633 are classified as urban. However, only 24 states allow the release of hospital-identifying information, leaving 400 identifiable urban hospitals in the 2005 NIS.

The NIS data include classification and severity measures for each patient record, including an All Patient Refined Diagnosis Related Group (APR-DRG) and an APR-DRG mortality-risk subclass.¹⁶ The APR-DRG system is a well-known proprietary classification system developed by 3M Health Information Systems (Salt Lake City, Utah) that uses diagnosis codes, procedure codes, and other administra-

tive data to classify patients into base disease categories as well as assign them to 1 of 4 levels of mortality risk (minor, moderate, major, and extreme) within each base disease category.^{17,18}

Leapfrog provided survey data on 1075 urban hospitals that had participated in the 2006 Safe Practices Survey. Of these, 679 were located within the 24 states providing hospital-identified data to the NIS. Because the survey is completed by hospitals in the spring, the 2006 survey captures safety practices in place during the 2005 calendar year.

We discussed our analysis with the survey's authors before starting. They were interested in streamlining their survey and asked us to consider an "Action Safe Practices Score" (ASPS), obtained through a different scoring methodology. This "action-focused" methodology assigned points for a given safe practice solely on the basis of a hospital's answer to the most "actionable" item in the action portion of each Safe Practices Survey section. The original scoring methodology gave credit for establishing systems of awareness, accountability, ability (capacity-building investments), and action. (The online survey is available at <http://leapfrog.medstat.com>.)

In 2008 the survey was reduced from 27 to 13 safe practices: (1) creating a safety culture, (2) ensuring an adequate nursing workforce, (3) ensuring that a pharmacist is active in medication use, (4) not providing patient care summaries from memory, (5) providing patient care information and orders to all clinicians, (6) requiring patient readback of informed consent, (7) documenting resuscitation or end-of-life directives, (8) preventing mislabeled radiographs, (9) providing risk assessment and prevention for deep vein thrombosis/venous thromboembolism, (10) providing anticoagulation services, (11) preventing aspiration, (12) preventing central venous line sepsis, and (13) requiring hand washing. (A list of all National Quality Forum safe practices is available at <http://www.ahrq.gov/qual/30safe.htm>.) We accordingly repeated our analyses using

"SPS-13" score and "ASPS-13" scores based on the 13 retained practices.

The sample for the primary analysis evaluating the association between Safe Practices Score (SPS) and mortality consisted of discharges from those urban US hospitals that had completed the Safe Practices Survey and that were identifiable in the 2005 NIS (155 hospitals). We also planned subgroup analyses for 2 populations in which we hypothesized that inpatient mortality might be more sensitive to adherence to safe practices: patients older than 65 years and patients with greater than 5% expected mortality.

Discharges excluded from the analysis included patients younger than 18 years, oncology patients, recipients of solid organ transplants, and patients transferred to or from another acute-care facility. Application of the exclusion criteria to discharges from urban hospitals in the 24 hospital-identifying NIS states reduced the number of eligible discharges from 4 873 959 to 3 672 146. The exclusions were standard per the recommendations of 3M Health Information Systems for use of the APR-DRG system in risk adjustment.¹⁸

Statistical Analysis

The primary predictor of interest was each hospital's 2006 SPS. The maximum total SPS possible was 1000. An additional predictor of interest was each hospital's 2006 ASPS. We also subsequently tested SPSs and ASPSs based on the 13 safe practices retained in the 2008 survey (SPS-13 and ASPS-13, respectively).

To examine the relationship between survey scores and inpatient mortality, we built hierarchical logistic regression models, known more broadly as generalized linear mixed models,¹⁹ with each discharged patient as the unit of analysis and a random intercept for each hospital to capture the correlation of patients within a hospital. Because the distribution of survey scores was heavily skewed, we categorized the scores into quartiles prior to using them as a predictor in our models. Categorizing the score into quartiles also

seemed appropriate because the Leapfrog Group rates hospitals on its Web site based on which quartile of survey score they fall into.

To adjust for mortality risk, we used the population of discharges from urban hospitals in the 24 hospital-identifying NIS states (3 672 146 discharges) to calculate the observed mortality rate for each combination of APR-DRG and APR-DRG mortality risk score. We then assigned to each patient the mortality risk associated with his or her APR-DRG and APR-DRG mortality risk category and used this expected mortality risk as an adjuster in our logistic regression models.

Additional hospital characteristics included in the model were volume of discharges and whether the hospital was a teaching hospital. These were treated as control variables to account for the possibility that a hospital's SPS may indicate better quality of care through mechanisms not mediated by hospital size and teaching status. Rural hospitals were excluded from our analysis because the Leapfrog Group does not target these hospitals and because the collinearity of rural location and hospital discharge volume would have rendered our models unstable. (Rural hospitals participating in the survey collectively accounted for only 71 169 discharges in the NIS.)

In our models, survey quartile categorizations were entered as 3 dummy variables, with the highest quartile serving as the reference group to produce odds ratios for mortality vs the highest SPS quartile as well as adjusted mortality rates. A test for linear trend of the log-odds of mortality was conducted using a likelihood ratio test with the linear orthogonal polynomial contrast.

A postanalysis power calculation was performed to assess the difference in mortality that our data would permit us to detect. Data were available on 1 772 064 admissions clustered in 155 hospitals of interest. Using a conservative estimate of 0.025 for intrahospital correlation of risk-adjusted mortality (others have estimated this correlation to be below 0.01²⁰) would imply

an effective sample size of 6000. Assuming an overall mortality rate of 2%, we calculated that we had 80% power to detect a 1 percentage point linear increase in mortality from the first to the fourth quartiles of the SPS, assuming 2-tailed $\alpha = .05$. We also calculated that to have 80% power to detect a mortality increase from 1.9% (quartile 1) to 2.1% (quartile 4) would require data from 500 hospitals.

Because the human participants data used for this study were completely deidentified and not collected for the purpose of this research, this study was considered exempt from institutional review board approval, per the policies of the University of California San Francisco's Committee on Human Research. All analyses were performed using StataMP version 10.0 (StataCorp, College Station, Texas), except for the power calculations, which were performed using nQuery Advisor version 6.02 (Statistical Solutions, Saugus, Massachusetts). A 2-sided significance level of .05 was used for all hypothesis tests.

RESULTS

Of the 155 urban hospitals in the NIS that had completed the Safe Practices Survey, 34% were teaching hospitals and 66% were nonteaching hospitals (TABLE 1). Within the NIS, hospitals participating in the survey tended to have higher volumes of discharges than those that had not participated. Survey scores for participating hospitals identifiable in the NIS were similar to those for participating hospitals not identifiable (Table 1). The distribution of survey scores was heavily skewed, with most hospitals scoring above 770 (of a possible 1000), regardless of the scoring methodology used.

In 2005 there were 1 772 064 discharges from the 155 hospitals of interest, of which 37 033 resulted in an inpatient death (2.09%). Quartiles of SPS were not a significant predictor of mortality, regardless of whether we adjusted for expected mortality rate or added volume of discharges and teaching status to our models (TABLE 2). From lowest to highest quartile of SPS, inpa-

tient mortality rates adjusted for patient and hospital characteristics were 1.97% (95% confidence interval [CI], 1.78%-2.18%), 2.04% (95% CI, 1.84%-2.25%), 1.96% (95% CI, 1.77%-2.16%), and 2.00% (95% CI, 1.80%-2.22%) ($P = .99$ for trend across all quartiles). Similarly, quartiles of performance using the SPS-13 (which assesses hospitals regarding the 13 safe practices retained in the 2008 survey) did not predict inpatient mortality: corresponding rates from lowest to highest quartile were 1.99% (95% CI, 1.80%-2.21%), 2.00% (95% CI, 1.81%-2.21%), 2.02% (95% CI, 1.82%-2.24%), and 1.95% (95% CI, 1.75%-2.16%) ($P = .80$ for trend across all quartiles).

Quartiles of survey performance using the action-based scoring methodology also did not significantly predict inpatient mortality. Fully adjusted inpatient mortality rates from lowest to highest quartile of ASPs were 2.08% (95% CI, 1.87%-2.30%), 2.02% (95% CI, 1.83%-2.23%), 2.02% (95% CI, 1.83%-2.24%), and 1.86% (95% CI, 1.68%-2.06%) ($P = .20$ for trend across all quartiles). Using ASPs-13, corresponding rates from lowest to highest quartile were 2.12% (95% CI, 1.91%-2.35%), 2.01% (95% CI, 1.82%-2.22%), 2.01% (95% CI, 1.82%-2.22%), and 1.85% (95% CI, 1.66%-2.05%) ($P = .11$ for trend across all quartiles).

To determine if survey scores might be more predictive of mortality in patients who might be at higher risk of death, we repeated our analyses with 2 subgroups of patients, those 65 years and older ($n = 721\,497$) and those with 5% or greater expected mortality risk ($n = 163\,138$). Overall in-hospital mortality rates in these populations were 3.9% and 18.3%, respectively. Quartiles of survey score were not predictive in these subgroups (TABLE 3 and TABLE 4).

We also examined the data to see if hospitals that had chosen to participate in the Safe Practices Survey had fewer inpatient deaths than other hospitals in the NIS (within the 24 states providing hospital-identifying infor-

mation). Fully adjusted mortality among hospitals participating in the survey (1 772 064 discharges within 155 hospitals) was 1.96% (95% CI, 1.83%-2.09%), compared with 2.06% (95% CI, 1.94%-2.19%) among hospitals not participating (1 892 725 discharges within 243 hospitals) ($P = .45$).

COMMENT

On the Internet, hospital performance on the Safe Practices Survey is ranked by quartiles, which may suggest to consumers that hospitals in higher quartiles are safer than hospitals in lower quartiles. In this first study of the relationship between survey scores and hospital outcomes, we studied a national sample of hospitals and found no relationship between quartiles of score and in-hospital mortality, regardless of whether we adjusted for expected mortality risk and certain hospital characteristics (Table 2).

A hospital's performance on each safe practice is evaluated in the survey via several questions assessing institutional systems to promote awareness, accountability, ability (capacity-building investments), and action. Part

of the rationale for designing the survey this way originally may have been to provide "training wheels" and give hospitals credit for creating systems that could eventually support full implementation of a given safe practice. However, awarding survey points for hospital administrative structures raises the possibility that the survey is capturing excessive noise, which may be overwhelming an important signal.

It seemed plausible that inpatient mortality could relate to whether actions were being taken to implement the safe practice. For this reason the survey's creators proposed the alternative "action-based" scoring method, which assigned all points for a safe practice based on whether the hospital indicated that it had implemented the key action for the practice. In our analysis, using this "action-based" scoring method slightly improved the ability of the survey ranking to predict in-hospital mortality, although the association was not statistically significant. Focusing on actions in the future may improve the survey's ability to discriminate between high-quality and low-quality hospitals.

The Safe Practices Survey has recently been shortened from 27 to 13 safe practices, largely in response to feedback regarding the considerable time required to complete the survey. Our findings indicate that scores based on the 13 retained practices are unlikely to be significantly associated with inpatient mortality, even if scoring is limited to actions taken, as was done in the ASPS-13. The findings do not rule out a modest association between the ASPS-13 and hospital mortality. As presented in our results (Table 2), the difference in risk-adjusted mortality between the best and worst quartiles determined by ASPS-13 was 0.27% ($P = .11$ for trend). This difference in absolute mortality risk corresponds to a number needed to treat of approximately 370, which some would find clinically significant. Nonetheless, the likelihood that the observed difference arose by chance is increased by the facts that multiple statistical tests were performed and that associations were not observed in the high-risk groups in which they were most strongly hypothesized.

Given the voluntary nature of this self-reported survey, it is also plausible that a lack of correlation with mortality might

Table 1. Urban Hospital Characteristics in 24 States That Identify Individual Hospitals in the Nationwide Inpatient Sample (NIS)^a

	Hospitals Identifiable in the NIS			Non-NIS Hospitals Participating in Safe Practices Survey (n = 524)
	Overall (n = 400) ^b	Not Participating in Safe Practices Survey (n = 243)	Participating in Safe Practices Survey (n = 155)	
Type, No. (%)				
Teaching	111 (28)	59 (24)	52 (34)	193 (37) ^c
Nonteaching	289 (73)	184 (76)	103 (66)	331 (63) ^c
Region, No. (%)				
Northeast	108 (27)	72 (30)	36 (23)	83 (16) ^c
Midwest	84 (21)	53 (22)	31 (20)	116 (22) ^c
South	85 (21)	63 (26)	22 (14)	65 (12) ^c
West	123 (31)	55 (23)	66 (43)	260 (50) ^c
Total discharges, median (IQR)	9315 (4309-17 592)	7215 (2383-14 452)	12 987 (7344-20 683)	
Total scores, median (IQR) ^d				
SPS			905 (848-963)	927 (847-970)
ASPS			919 (808-963)	921 (820-961)
SPS-13			896 (834-962)	924 (829-972)
ASPS-13			911 (777-976)	917 (800-966)

Abbreviations: ASPS, Action Safe Practices Score; ASPS-13: Action Safe Practices Score based on 13 safe practices; IQR, interquartile range; SPS, Safe Practices Score; SPS-13, Safe Practices Score based on 13 safe practices.

^aStates providing hospital identifying information in the NIS: Arizona, California, Colorado, Connecticut, Florida, Iowa, Illinois, Kentucky, Massachusetts, Maryland, Minnesota, Missouri, North Carolina, New Hampshire, New Jersey, Nevada, New York, Oregon, Rhode Island, Utah, Vermont, Washington, Wisconsin, and West Virginia.

^bIncludes 2 hospitals that participated in initiatives 1 through 3 but not initiative 4.

^cData obtained from 2006 American Hospital Association Annual Survey.

^dOf a possible 1000.

be due to confounding associated with a “healthy volunteer” effect. Hospitals that have already engaged in improving quality of care may be more likely to want to participate in the Safe Practices Survey. However, in our study we found that participation in the survey was not predictive of lower risk-adjusted mortality. This would suggest that our negative findings are unlikely to be owing to safer hospitals being more likely to participate in the survey.

To our knowledge, this is the first peer-reviewed analysis that has sought to assess whether better performance on the Safe Practices Survey correlates with outcomes indicative of improved

patient safety. Given that 2 recently published analyses^{21,22} in the business/medical literature used the Safe Practices Survey as the metric of hospital quality, it seems clear that additional efforts to explore the validity and value of this well-publicized^{1,23,24} quality measure are needed. As the patient-safety movement increases in importance, hospitals face increasingly complex choices regarding improvement and reporting to the public. Likewise, consumers are faced with multiple sources of information on hospital quality and are encouraged to choose a facility based on this information. Despite a lack of evidence demonstrating the

validity of the Safe Practices Survey, the survey is well known and influential. For this reason, validating the survey rankings as a measure and as a source of accurate information for consumers and researchers is important.

Our findings suggest, however, that the survey as currently designed does not discriminate between hospitals with higher and lower inpatient mortality. Some will question our choice of overall inpatient mortality as the outcome of interest. We acknowledge that very valid concerns about use of mortality rates as a measure of hospital quality of care have been raised.²⁵⁻²⁷ Despite this, risk-adjusted mortality rates remain among the most

Table 2. Inpatient Mortality Rates, by Quartiles of Performance on Safe Practices Survey

Rate	SPS Quartile ^a				P Value for Trend	ASPS Quartile ^a				P Value for Trend
	1	2	3	4		1	2	3	4	
Score	123-848	848-905	905-963	963-1000		274-808	808-919	919-963	963-1000	
Raw mortality										
Rate, %	1.86	2.20	1.83	1.88		1.96	2.04	1.96	1.80	
(95% CI)	(1.65-2.10)	(1.96-2.47)	(1.62-2.05)	(1.67-2.12)		(1.74-2.21)	(1.81-2.29)	(1.74-2.21)	(1.59-2.03)	
OR	0.99	1.17	0.97	1 [Reference]	.55	1.09	1.14	1.10	1 [Reference]	.28
(95% CI)	(0.83-1.18)	(0.99-1.39)	(0.82-1.15)			(0.92-1.30)	(0.96-1.35)	(0.92-1.30)		
Risk-adjusted mortality										
Rate, %	2.02	2.12	2.00	2.06		2.13	2.09	2.08	1.90	
(95% CI) ^b	(1.83-2.24)	(1.92-2.34)	(1.80-2.21)	(1.86-2.29)		(1.93-2.36)	(1.89-2.31)	(1.88-2.30)	(1.72-2.11)	
OR	0.98	1.03	0.97	1 [Reference]	.98	1.12	1.10	1.10	1 [Reference]	.14
(95% CI)	(0.85-1.14)	(0.89-1.19)	(0.84-1.12)			(0.97-1.30)	(0.95-1.27)	(0.95-1.27)		
Risk- and hospital-adjusted mortality										
Rate, %	1.97	2.04	1.96	2.00		2.08	2.02	2.02	1.86	
(95% CI) ^c	(1.78-2.18)	(1.84-2.25)	(1.77-2.16)	(1.80-2.22)		(1.87-2.30)	(1.83-2.23)	(1.83-2.24)	(1.68-2.06)	
OR	0.98	1.02	0.98	1 [Reference]	.99	1.11	1.08	1.09	1 [Reference]	.20
(95% CI)	(0.85-1.14)	(0.89-1.18)	(0.85-1.13)			(0.96-1.27)	(0.94-1.25)	(0.94-1.25)		
	SPS-13 Quartile ^a					ASPS-13 Quartile ^a				
	1	2	3	4		1	2	3	4	
Score	97-834	834-896	896-962	962-1000		132-777	777-911	911-976	976-1000	
Raw mortality										
Rate, %	1.90	2.15	1.82	1.89		1.99	1.98	2.01	1.77	
(95% CI)	(1.69-2.14)	(1.91-2.42)	(1.62-2.05)	(1.68-2.13)		(1.76-2.25)	(1.76-2.22)	(1.78-2.26)	(1.57-2.01)	
OR	1.00	1.14	0.96	1 [Reference]	.50	1.13	1.12	1.13	1 [Reference]	.24
(95% CI)	(0.85-1.19)	(0.96-1.35)	(0.81-1.14)			(0.94-1.34)	(0.94-1.33)	(0.95-1.35)		
Risk-adjusted mortality										
Rate, %	2.05	2.08	2.07	2.00		2.17	2.08	2.06	1.89	
(95% CI) ^b	(1.85-2.27)	(1.88-2.30)	(1.87-2.29)	(1.81-2.22)		(1.96-2.40)	(1.89-2.30)	(1.87-2.28)	(1.70-2.09)	
OR	1.03	1.04	1.03	1 [Reference]	.72	1.15	1.11	1.10	1 [Reference]	.07
(95% CI)	(0.89-1.19)	(0.90-1.20)	(0.89-1.20)			(1.00-1.34)	(0.96-1.28)	(0.95-1.27)		
Risk- and hospital-adjusted mortality										
Rate, %	1.99	2.00	2.02	1.95		2.12	2.01	2.01	1.85	
(95% CI) ^c	(1.80-2.21)	(1.81-2.21)	(1.82-2.24)	(1.75-2.16)		(1.91-2.35)	(1.82-2.22)	(1.82-2.22)	(1.66-2.05)	
OR	1.02	1.02	1.03	1 [Reference]	.80	1.13	1.08	1.09	1 [Reference]	.11
(95% CI)	(0.89-1.18)	(0.88-1.18)	(0.89-1.19)			(0.98-1.31)	(0.94-1.25)	(0.94-1.25)		

Abbreviations: ASPS, Action Safe Practices Score; ASPS-13, Action Safe Practices Score based on 13 safe practices; CI, confidence interval; OR, odds ratio; SPS, Safe Practices Score; SPS-13, Safe Practices Score based on 13 safe practices.

^a Quartile cutoffs rounded to nearest whole integers; no actual overlap exists between quartiles.

^b Adjusted for each patient’s expected risk of mortality (based on All Patient Refined Diagnosis Related Group classification).

^c Adjusted for each patient’s expected risk of mortality, for hospital discharge volume, and for hospital teaching status.

commonly reported outcomes in both the published literature and in public reports of hospital quality. Furthermore, the Institute of Medicine has cited prevention of inpatient deaths as an important reason to focus on patient safety.²⁸ Until consumer-oriented hospital quality reports become explicit as to what benefits consumers can expect from a “safer” hospital, consumers will almost certainly assume a safer hospital is one in which a patient is less likely to die.²⁹ Thus, our analyses are consistent with the most likely consumer interpretation of the data presented on the Internet.

Our study results suggest several points regarding the Safe Practices Sur-

vey that would be valuable for the greater patient safety community to consider. An important issue is whether the SPS is measuring what needs to be measured. Many of the safe practices are processes to improve care, yet in its current form the survey is measuring the “processes around the process.” This often gives hospitals credit for what essentially may be good intentions. This also gives points for having structures that may support implementation of a safe practice, rather than awarding credit only when the safe practice is being consistently followed. Such a scoring system likely is vulnerable to inflation of scores. Of note, most hospitals score

quite well on the survey (Table 1). It may be that all hospitals truly are doing well on the safe practices; however, it seems more likely that the survey as currently designed is unable to discriminate between truly high and low adherence to the safe practices.

It also is possible that too much is being measured. Steps have already been taken to address this by reducing the survey to 13 practices, but our results suggest that this alone is unlikely to improve the survey’s ability to correlate with inpatient mortality.

It is unknown how well a hospital executive’s report of actions being taken to support and implement safety prac-

Table 3. Inpatient Mortality Rates in Patients 65 Years or Older, by Quartiles of Performance on Safe Practices Survey

Rate	SPS Quartile ^a				P Value for Trend	ASPS Quartile ^a				P Value for Trend
	1	2	3	4		1	2	3	4	
Score	123-848	848-905	905-963	963-1000		274-808	808-919	919-963	963-1000	
Raw mortality										
Rate, %	3.70	3.93	3.56	3.63		3.89	3.87	3.69	3.39	
(95% CI)	(3.35-4.09)	(3.57-4.33)	(3.22-3.93)	(3.28-4.02)		(3.53-4.30)	(3.51-4.26)	(3.34-4.06)	(3.07-3.76)	
OR	1.02	1.09	0.98	1 [Reference]	.48	1.15	1.14	1.09	1 [Reference]	.04
(95% CI)	(0.88-1.18)	(0.94-1.26)	(0.84-1.14)			(1.00-1.34)	(0.99-1.32)	(0.94-1.26)		
Risk-adjusted mortality										
Rate, %	4.01	4.23	3.97	4.16		4.23	4.21	4.12	3.80	
(95% CI) ^b	(3.61-4.45)	(3.83-4.68)	(3.57-4.40)	(3.73-4.62)		(3.81-4.69)	(3.80-4.66)	(3.72-4.56)	(3.42-4.23)	
OR	0.96	1.02	0.95	1 [Reference]	.88	1.12	1.11	1.09	1 [Reference]	.15
(95% CI)	(0.83-1.13)	(0.88-1.19)	(0.82-1.11)			(0.96-1.30)	(0.96-1.30)	(0.93-1.27)		
Risk- and hospital-adjusted mortality										
Rate, %	3.90	4.08	3.89	4.03		4.13	4.07	4.00	3.74	
(95% CI) ^c	(3.51-4.33)	(3.68-4.51)	(3.51-4.31)	(3.62-4.50)		(3.72-4.59)	(3.67-4.50)	(3.61-4.43)	(3.36-4.16)	
OR	0.96	1.01	0.96	1 [Reference]	.82	1.10	1.08	1.07	1 [Reference]	.24
(95% CI)	(0.83-1.12)	(0.87-1.18)	(0.83-1.12)			(0.94-1.27)	(0.93-1.26)	(0.92-1.24)		
	SPS-13 Quartile ^a					ASPS-13 Quartile ^a				
	1	2	3	4		1	2	3	4	
Score	97-834	834-896	896-962	962-1000		132-777	777-911	911-976	976-1000	
Raw mortality										
Rate, %	3.78	3.83	3.65	3.57		3.95	3.81	3.75	3.33	
(95% CI)	(3.42-4.17)	(3.47-4.22)	(3.30-4.04)	(3.23-3.95)		(3.57-4.36)	(3.46-4.18)	(3.41-4.13)	(3.01-3.69)	
OR	1.06	1.07	1.02	1 [Reference]	.35	1.19	1.15	1.13	1 [Reference]	.02
(95% CI)	(0.91-1.23)	(0.93-1.24)	(0.88-1.19)			(1.03-1.38)	(0.99-1.33)	(0.98-1.31)		
Risk-adjusted mortality										
Rate, %	4.06	4.14	4.16	4.00		4.32	4.19	4.10	3.76	
(95% CI) ^b	(3.66-4.51)	(3.74-4.59)	(3.75-4.61)	(3.60-4.45)		(3.89-4.79)	(3.79-4.63)	(3.70-4.53)	(3.38-4.19)	
OR	1.02	1.04	1.04	1 [Reference]	.87	1.15	1.12	1.09	1 [Reference]	.07
(95% CI)	(0.87-1.18)	(0.89-1.21)	(0.89-1.21)			(0.99-1.35)	(0.96-1.30)	(0.94-1.27)		
Risk- and hospital-adjusted mortality										
Rate, %	3.94	3.98	4.07	3.91		4.21	4.04	4.00	3.70	
(95% CI) ^c	(3.55-4.38)	(3.59-4.41)	(3.66-4.51)	(3.51-4.35)		(3.79-4.68)	(3.65-4.47)	(3.62-4.42)	(3.32-4.12)	
OR	1.01	1.01	1.04	1 [Reference]	.99	1.13	1.09	1.08	1 [Reference]	.13
(95% CI)	(0.87-1.17)	(0.87-1.18)	(0.89-1.20)			(0.97-1.32)	(0.94-1.27)	(0.93-1.25)		

Abbreviations: ASPS, Action Safe Practices Score; ASPS-13, Action Safe Practices Score based on 13 safe practices; CI, confidence interval; OR, odds ratio; SPS, Safe Practices Score; SPS-13, Safe Practices Score based on 13 safe practices.

^a Quartile cutoffs rounded to nearest whole integers; no actual overlap exists between quartiles.

^b Adjusted for each patient’s expected risk of mortality (based on All Patient Refined Diagnosis Related Group classification).

^c Adjusted for each patient’s expected risk of mortality, for hospital discharge volume, and for hospital teaching status.

tics correlates with actual activity within that hospital. It may not be reasonable to assume that hospitals are doing what their executives say they are. Our study results call into question the use of a lengthy unaudited survey as a tool for measuring adherence to safe practices. Further research to examine how well self-reported activities regarding hospital safety practices correlate with actual activities within the hospital will be helpful in determining the value of self-reported safety data.

Our study has important limitations. The most significant is that our main analysis had enough power to detect only 1% or greater differences in

mortality. Although at a policy and epidemiology level the observed 0.2% difference in mortality rate between the first and fourth quartiles of performance using ASPS scoring is potentially important, we calculated that we would have needed on the order of 11 million admissions grouped by 500 hospitals to conclude statistical significance for a difference of this magnitude. Such an analysis was not feasible given the data available from the largest hospital data set in the nation. Furthermore, even if a statistically significant association were to be found with a larger sample, our results indicate that the overall magnitude of the relation-

ship between survey score and mortality would almost certainly remain quite small and would be of unclear usefulness to individual consumers.

A second limitation is that we did not study other outcomes that might be responsive to adoption of the safe practices, such as complications. It is possible that high performance on the survey does correlate with decreased complication rates or with other outcomes of interest to purchasers and policy makers. However, complication rates are difficult to accurately measure using administrative data. Finally, our primary analysis examines 155 of the 1075 urban hospitals that

Table 4. Inpatient Mortality Rates in Patients With >5% Mortality Risk, by Quartiles of Performance on Safe Practices Survey

Rate	SPS Quartile ^a				P Value for Trend	ASPS Quartile ^a				P Value for Trend
	1	2	3	4		1	2	3	4	
Score	123-848	848-905	905-963	963-1000		274-808	808-919	919-963	963-1000	
Raw mortality										
Rate, %	17.59	18.58	17.36	18.17		18.33	18.78	17.52	17.08	
(95% CI)	(16.35-18.90)	(17.34-19.90)	(16.15-18.63)	(16.89-19.52)		(17.07-19.67)	(17.52-20.11)	(16.31-18.80)	(15.89-18.35)	
OR	0.96	1.03	0.95	1 [Reference]	.89	1.09	1.12	1.03	1 [Reference]	.08
(95% CI)	(0.85-1.09)	(0.91-1.16)	(0.84-1.07)			(0.96-1.23)	(0.99-1.27)	(0.91-1.17)		
Risk-adjusted mortality										
Rate, %	18.03	18.88	18.00	19.09		18.62	19.00	18.65	17.71	
(95% CI) ^b	(16.62-19.53)	(17.47-20.38)	(16.61-19.48)	(17.60-20.67)		(17.18-20.16)	(17.56-20.52)	(17.21-20.17)	(16.32-19.19)	
OR	0.93	0.99	0.93	1 [Reference]	.51	1.06	1.09	1.07	1 [Reference]	.35
(95% CI)	(0.81-1.07)	(0.86-1.13)	(0.81-1.07)			(0.93-1.22)	(0.95-1.25)	(0.93-1.22)		
Risk- and hospital-adjusted mortality										
Rate, %	17.63	18.33	17.71	18.52		18.25	18.55	18.18	17.37	
(95% CI) ^c	(16.24-19.11)	(16.94-19.81)	(16.35-19.17)	(17.03-20.12)		(16.82-19.77)	(17.15-20.05)	(16.75-19.70)	(15.98-18.85)	
OR	0.93	0.98	0.94	1 [Reference]	.42	1.05	1.06	1.05	1 [Reference]	.50
(95% CI)	(0.81-1.07)	(0.85-1.12)	(0.82-1.07)			(0.91-1.20)	(0.93-1.22)	(0.92-1.21)		
	SPS-13 Quartile ^a					ASPS-13 Quartile ^a				
	1	2	3	4		1	2	3	4	
Score	97-834	834-896	896-962	962-1000		132-777	777-911	911-976	976-1000	
Raw mortality										
Rate, %	17.77	18.33	17.86	17.74		18.54	18.36	17.78	17.03	
(95% CI)	(16.52-19.09)	(17.07-19.65)	(16.62-19.17)	(16.50-19.05)		(17.24-19.92)	(17.13-19.65)	(16.57-19.06)	(15.81-18.32)	
OR	1.00	1.04	1.01	1 [Reference]	.85	1.11	1.10	1.05	1 [Reference]	.08
(95% CI)	(0.89-1.13)	(0.92-1.18)	(0.89-1.14)			(0.98-1.26)	(0.97-1.24)	(0.93-1.19)		
Risk-adjusted mortality										
Rate, %	18.15	18.70	18.54	18.59		18.91	18.75	18.68	17.62	
(95% CI) ^b	(16.73-19.67)	(17.27-20.22)	(17.11-20.06)	(17.15-20.12)		(17.43-20.49)	(17.35-20.24)	(17.27-20.18)	(16.21-19.11)	
OR	0.97	1.01	1.00	1 [Reference]	.74	1.09	1.08	1.07	1 [Reference]	.25
(95% CI)	(0.85-1.12)	(0.88-1.16)	(0.87-1.14)			(0.95-1.26)	(0.94-1.24)	(0.94-1.23)		
Risk- and hospital-adjusted mortality										
Rate, %	17.73	18.18	18.16	18.13		18.53	18.27	18.26	17.26	
(95% CI) ^c	(16.34-19.23)	(16.78-19.67)	(16.74-19.67)	(16.70-19.66)		(17.07-20.09)	(16.89-19.74)	(16.86-19.74)	(15.86-18.75)	
OR	0.96	0.98	0.99	1 [Reference]	.58	1.07	1.06	1.06	1 [Reference]	.37
(95% CI)	(0.84-1.10)	(0.85-1.13)	(0.87-1.14)			(0.93-1.23)	(0.92-1.21)	(0.93-1.22)		

Abbreviations: ASPS, Action Safe Practices Score; ASPS-13, Action Safe Practices Score based on 13 safe practices; CI, confidence interval; OR, odds ratio; SPS, Safe Practices Score; SPS-13, Safe Practices Score based on 13 safe practices.

^a Quartile cutoffs rounded to nearest whole integers; no actual overlap exists between quartiles.

^b Adjusted for each patient's expected risk of mortality (based on All Patient Refined Diagnosis Related Group classification).

^c Adjusted for each patient's expected risk of mortality, for hospital discharge volume, and for hospital teaching status.

participated in the Safe Practices Survey, raising the question of whether our findings can be generalized to the many survey participants not in the NIS. However, the NIS is designed to approximate a stratified random sample of US hospitals, and individual hospitals cannot choose whether to be included, so there is no reason to believe that survey participants included in the NIS differed in any systematic fashion from those not included. Furthermore, we found that survey participants had similar survey scores, regardless of whether they were in the NIS. Hence, although we cannot exclude the possibility that our findings are not generalizable to other hospitals participating in the survey, we have not identified any specific reason to question the generalizability of our findings.

In summary, although a recent study has found that hospitals performing well on the first 3 initiatives of the Leapfrog Hospital Survey do have lower risk-adjusted mortality,¹² our analysis was unable to find a correlation between better performance on the fourth initiative (the Safe Practices Survey) and lower risk-adjusted mortality. It is possible that inviting hospitals to self-report on their patient safety practices and then assigning them to quartiles of score is not an effective way to assess hospital quality and safety. Our findings should not be interpreted, however, as indicating that the safe practices are not important or that they cannot be measured in an informative and valid way. Rather, future work should seek to establish valid methods for assessing adherence to the safe practices. Further research is needed to determine how performance on the Safe Practices Survey or other instruments designed to measure safe practices performance may correlate with other outcomes of interest to patients and policy makers.

Author Contributions: Dr Kernisan 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: Kernisan, Dudley.

Acquisition of data: Kernisan, Dudley.

Analysis and interpretation of data: Kernisan, Lee, Boscardin, Landefeld, Dudley.

Drafting of the manuscript: Kernisan, Dudley.

Critical revision of the manuscript for important intellectual content: Kernisan, Lee, Boscardin, Landefeld, Dudley.

Statistical analysis: Kernisan, Boscardin, Dudley.

Obtained funding: Dudley.

Administrative, technical, or material support: Kernisan, Dudley.

Study supervision: Lee, Landefeld, Dudley.

Financial Disclosures: Dr Dudley reported serving as an occasional consultant for the Leapfrog Group (but that he is not paid for this work) and serving on the Leapfrog Group's Hospital Rewards Program Steering Committee. Work on this article was not financially supported by the Leapfrog Group. Drs Kernisan, Lee, Boscardin, and Landefeld reported no financial or other conflicts of interest.

Funding/Support: This study was supported by the UCSF Division of Geriatrics, the UCSF Institute for Health Policy Studies, and the San Francisco VA Medical Center. Dr Kernisan was supported by T32 research fellowship grant T32 AG000212-16 from the National Institute on Aging, National Institutes of Health, followed by a VA Quality Scholars fellowship. Dr Dudley's work was supported by a Robert Wood Johnson Foundation Investigator Award in Health Policy. This study was also supported by the Leapfrog Group, which provided its data at no cost.

Role of the Sponsor: The Leapfrog Group was involved in the design of the study and provided data but had no role in the analysis or interpretation of the data; the group had no role in the preparation of the manuscript and was not asked to approve the final manuscript but did review the manuscript and provided minor corrections of factual statements about its survey program. The Robert Wood Johnson Foundation had no role in the design of the study, the analysis of data, or the preparation of the manuscript.

Additional Contributions: We thank the Leapfrog Group for sharing its data with us.

REFERENCES

1. Scanlon DP, Christianson JB, Ford EW. Hospital responses to the Leapfrog Group in local markets. *Med Care Res Rev.* 2008;65(2):207-231.
2. Callcut RA, Breslin TM. Shaping the future of surgery: the role of private regulation in determining quality standards. *Ann Surg.* 2006;243(3):304-312.
3. Manthous CA. Leapfrog and critical care: evidence- and reality-based intensive care for the 21st century. *Am J Med.* 2004;116(3):188-193.
4. Factsheet: the Leapfrog safe practices score leap. Leapfrog Group Web site. http://www.leapfroggroup.org/media/file/Leapfrog-Safe_Practices_Score_Fact_Sheet.pdf. Accessed August 6, 2007.
5. Brooke BS, Perler BA, Dominici F, Makary MA, Pronovost PJ. Reduction of in-hospital mortality among California hospitals meeting Leapfrog evidence-based standards for abdominal aortic aneurysm repair. *J Vasc Surg.* 2008;47(6):1155-1156, 1163-1164.
6. Goshima KR, Mills JL Sr, Awari K, Pike SL, Hughes JD. Measure what matters: institutional outcome data are superior to the use of surrogate markers to define "center of excellence" for abdominal aortic aneurysm repair. *Ann Vasc Surg.* 2008;22(3):328-334.
7. Rodgers M, Jobe BA, O'Rourke RW, Sheppard B, Diggs B, Hunter JG. Case volume as a predictor of inpatient mortality after esophagectomy. *Arch Surg.* 2007;142(9):829-839.
8. Kilbridge PM, Welebob EM, Classen DC. Development of the Leapfrog methodology for evaluating hospital implemented inpatient computerized physician order entry systems. *Qual Saf Health Care.* 2006;15(2):81-84.
9. Welke KF, Barnett MJ, Sarrazin MS, Rosenthal GE. Limitations of hospital volume as a measure of quality of care for coronary artery bypass graft surgery. *Ann Thorac Surg.* 2005;80(6):2114-2119.
10. Birkmeyer JD, Dimick JB. Potential benefits of the new Leapfrog standards: effect of process and outcomes measures. *Surgery.* 2004;135(6):569-575.
11. Young MP, Birkmeyer JD. Potential reduction in mortality rates using an intensivist model to manage intensive care units. *Eff Clin Pract.* 2000;3(6):284-289.
12. Jha AK, Orav EJ, Ridgway AB, Zheng J, Epstein AM. Does the Leapfrog program help identify high-quality hospitals? *Jt Comm J Qual Patient Saf.* 2008;34(6):318-325.
13. Leapfrog Safe Practices Score. Leapfrog Group Web site. http://www.leapfroggroup.org/for_consumers/scoring_info/safe_practices_score. Accessed July 12, 2008.
14. Overview of the Nationwide Inpatient Sample (NIS). Healthcare Cost and Utilization Project (HCUP) Web site. <http://www.hcup-us.ahrq.gov/nisoverview.jsp>. 2005. Accessibility verified February 26, 2009.
15. Introduction to the HCUP Nationwide Inpatient Sample (NIS), 2005. Healthcare Cost and Utilization Project (HCUP) Web site. http://www.hcup-us.ahrq.gov/db/nation/nis/NIS_Introduction_2005.jsp. Accessed August 1, 2007.
16. Overview of disease severity measures disseminated with the Nationwide Inpatient Sample (NIS) and Kids' Inpatient Database (KID): executive summary. Healthcare Cost and Utilization Project (HCUP) Web site. http://www.hcup-us.ahrq.gov/db/nation/nis/severity_overview.jsp. March 10, 2006. Accessed September 10, 2008.
17. Romano PS, Chan BK. Risk-adjusting acute myocardial infarction mortality: are APR-DRGs the right tool? *Health Serv Res.* 2000;34(7):1469-1489.
18. All Patient Refined Diagnosis Related Groups (APR-DRGs) version 20.0: methodology overview. Agency for Healthcare Research and Quality Web site. <http://www.hcup-us.ahrq.gov/db/nation/nis/APR-DRGsV20MethodologyOverviewandBibliography.pdf>. Accessed September 10, 2008.
19. Breslow NE, Clayton DG. Approximate inference in generalized linear mixed models. *J Am Stat Assoc.* 1993;88(421):9-25.
20. Meurer LN, Yang H, Guse CE, Russo C, Brasel KJ, Layde PM. Excess mortality caused by medical injury. *Ann Fam Med.* 2006;4(5):410-416.
21. Huerta TR, Ford EW, Peterson LT, Brigham KH. Testing the hospital value proposition: an empirical analysis of efficiency and quality. *Health Care Manage Rev.* 2008;33(4):341-349.
22. Olson JR, Belohlav JA, Cook LS, Hays JM. Examining quality improvement programs: the case of Minnesota hospitals. *Health Serv Res.* 2008;43(5, pt 2):1787-1806.
23. The Leapfrog Group: HAI prevention efforts still have a long way to go. *Healthcare Benchmarks Qual Improv.* 2007;14(11):121-125.
24. Leapfrog Group releases its first "top hospitals" list. *Healthcare Benchmarks Qual Improv.* 2007;14(1):1-4.
25. Thomas JW, Hofer TP. Accuracy of risk-adjusted mortality rate as a measure of hospital quality of care. *Med Care.* 1999;37(1):83-92.
26. Shojania KG, Forster AJ. Hospital mortality: when failure is not a good measure of success. *CMAJ.* 2008;179(2):153-157.
27. Krumholz HM, Rathore SS, Chen J, Wang Y, Radford MJ. Evaluation of a consumer-oriented internet health care report card: the risk of quality ratings based on mortality data. *JAMA.* 2002;287(10):1277-1287.
28. Kohn LT, Corrigan J, Donaldson MS; Institute of Medicine, (U.S.) Committee on Quality of Health Care in America. *To Err Is Human: Building a Safer Health System.* Washington, DC: National Academies Press; 1999.
29. Leonardi MJ, McGory ML, Ko CY. Publicly available hospital comparison web sites: determination of useful, valid, and appropriate information for comparing surgical quality. *Arch Surg.* 2007;142(9):863-869.