ORIGINAL ARTICLE

Quality and Performance Indicators in an Academic Department of Head and Neck Surgery

Randal S. Weber, MD; Carol M. Lewis, MD, MPH; Scott D. Eastman, MBA; Ehab Y. Hanna, MD; Olubumi Akiwumi, MS; Amy C. Hessel, MD; Stephen Y. Lai, MD, PhD; Leslie Kian, MBA; Michael E. Kupferman, MD; Dianna B. Roberts, PhD

Objective: To create a method for assessing physician performance and care outcomes that are adjusted for procedure acuity and patient comorbidity.

Design: Between 2004 and 2008 surgical procedures performed by 10 surgeons were stratified into high-acuity procedures (HAPs) and low-acuity procedures (LAPs). Risk adjustment was made for comorbid conditions examined singly or in groups of 2 or more.

Setting: A tertiary care medical center.

Patients: A total of 2618 surgical patients.

Main Outcome Measures: Performance measures included length of stay; return to operating room within 7 days of surgery; and the occurrence of mortality, hospital readmission, transfusion, and wound infection within 30 days of surgery.

Results: The transfusion rate was 2.7% and 40.6% for LAPs and HAPs, respectively. Wound infection rates were 1.4% for LAPs vs 14.1% for HAPs, while 30-day mortality rate was 0.3% and 1.6% for LAPs and HAPs, respectively. The mean (SD) hospital stay for LAPs was 2.1 (3.6) vs 10.5 (7.0) days for HAPs. Negative performance factors were significantly higher for patients who underwent HAPs and had comorbid conditions. Differences among surgeons significantly affect the incidence of negative performance indicators. Factors affecting performance measures were procedure acuity, the surgeon, and comorbidity, in order of decreasing significance. Surgeons were ranked low, middle, and high based on negative performance indicators.

Conclusions: Performance measures following oncologic procedures were significantly affected by comorbid conditions and by procedure acuity. Although the latter most strongly affects quality and performance indicators, both should weigh heavily in physician comparisons. The incidence of negative performance indicators was also influenced by the individual surgeon. These data may serve as a tool to evaluate and improve physician performance and outcomes and to develop risk-adjusted benchmarks. Ultimately, reimbursement may be tied to quantifiable measures of physician and institutional performance.

Arch Otolaryngol Head Neck Surg. 2010;136(12):1212-1218

Author Affiliations: Department of Head and Neck Surgery (Drs Weber, Lewis, Hanna, Hessel, Lai, Kupferman, and Roberts), Office of Performance Improvement (Mr Eastman and Mss Akiwumi and Kian), University of Texas MD Anderson Cancer Center, Houston.

IN BOTH THE PUBLIC AND PRIVATE sector, health care reform is a national priority. Health care costs continue to increase at a rate greater than the consumer price index, and national health care expenditures now account for 7.5% of the gross domestic product (GDP). Between 2009 and 2018, average annual health spending growth (6.1%) is projected to outpace average annual growth in the overall economy (4.4 %), and by 2019, national health spending is expected to reach $4.5 trillion and comprise 19.3% of GDP.[1] By 2050, it is projected that at the current rate of increase, national health care expenditures could exceed the GDP. Obviously, that scenario is untenable and would threaten our economic viability. The strategies for containing the rising costs of health care include reimbursements tied to quality outcomes (as opposed to patient or procedure volume), comparative effectiveness, and directing patients with complex or rare diseases to high-volume regional centers. The emphasis is shifting from episodic care and quantity of care to disease management and quality of care. Waste in our health care delivery system revolves around care administered without proven benefit, errors of omission and commission, and inefficient use or overuse of resources.[2] Improving the quality and efficiency of care and minimizing medical errors have the potential to substantially reduce health care costs.

In the treatment of head and neck cancer, surgery remains an important technique for disease management. Outcome data following surgical intervention for patients with head and neck cancer will be increasingly important as benchmarks of surgical care quality and will be used for comparing the performance of individual surgeons within

and across institutions. Nevertheless, outcomes of surgery have received little attention from the standpoint of improving performance.

A performance indicator is an agreed-on and quantifiable measurement that reflects the critical success factors of a surgical intervention. Surgical performance indicators may be generic or specialty specific. Examples are length of stay (LOS), blood product utilization, 30-day surgical site infection, mortality, readmission, and return to the operating room for an untoward event such as wound dehiscence, hematoma, and/or infection. The National Surgical Quality Improvement Project (NSQIP), sponsored by the American College of Surgeons, is currently examining many of these outcome measures for general surgical procedures and some subspecialty procedures. To our knowledge, no studies have focused on surgical performance measures for patients undergoing complex head and neck surgical procedures.

To assess our departmental and individual performance we collected and analyzed our surgical outcome data. Knowing that an individual surgeon's practice will vary by patient comorbidity and procedure acuity, we adjusted the analysis to account for these confounders. Our objectives were (1) to develop benchmarks to compare the performance of individual surgeons; (2) to identify best practices that could be shared among the group; and (3) to develop a monitoring tool to identify trends requiring intervention to improve quality of care.

METHODS

Between 2004 and 2008, outcome data were collected for surgical procedures performed by the faculty in the Department of Head and Neck Surgery at the University of Texas MD Anderson Cancer Center. The data set included patient demographics, diagnoses, procedures (including whether reconstruction was performed), and patient comorbidity. Outcome measures were LOS for selected head and neck surgical procedures, blood product utilization, return to the operating room within 7 days of surgery, and the following variables within 30 days of surgery: mortality, hospital readmission, and surgical site infections. Data sources included the institutional enterprise information warehouse (EIW), which contains the surgical indexing database; patient account tables; and the tumor registry. Information on the surgery date, case number, *Current Procedural Terminology* codes, and surgeon were obtained from the surgical indexing database. The inpatient encounter, admission, and discharge dates and *International Classification of Diseases, Ninth Revision,* diagnosis codes were obtained from the patient account tables. Vital status and date of death were obtained from the tumor registry. Other databases within the institution, such as those maintained by the Department of Anesthesiology, were accessed when appropriate.

To risk adjust for comorbid conditions and procedure acuity, we collected the diagnoses recorded in the patient's medical record. These included diabetes, cardiovascular disease, history of congestive heart failure, chronic obstructive pulmonary disease, liver disease, and renal disease. Comorbid conditions were examined singly and in groups of 2 or more and were used to adjust for risk.

The surgical procedures selected were stratified into high and low acuity. High-acuity procedures (HAPs) were those requiring mandibulectomy, pharyngolaryngectomy, and major glossectomy. All of the HAPs required either pedicled flap or microvascular free-tissue transfer reconstruction. Since most

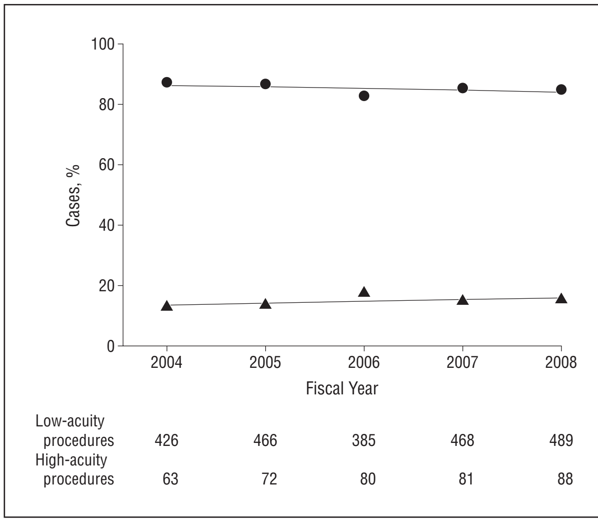

Figure 1. High- and low-acuity procedures performed at the study institution during fiscal years 2004 through 2008.

reconstructions are performed by a plastic surgeon, flap failures were not considered negative indicators. The low-acuity procedures (LAPs) included laryngoscopy, neck dissection and/or glossectomy, parotidectomy, and thyroidectomy with or without neck dissection. **Figure 1** illustrates the number of HAPs and LAPs performed over the 5-year period.

Performance indicators for LAPs and HAPs were the same, with 2 exceptions: For low-acuity procedures, the expected LOS was 2 or more days, and transfusion involved 0 to 1 U of packed red blood cells, whereas for the HAPs, the expected LOS was 11 days, and transfusions involved 0 to 2 U or more than 2 U of packed red blood cells. When LOS or transfusions exceeded the defined cutoff points, they were considered negative performance indicators, as were return to the operating room, readmission, infections, and mortality.

It may appear liberal to use a transfusion cutoff of 0 to 1 U of packed red blood cells for low-acuity procedures, which include direct laryngoscopy, parotidectomy, and thyroidectomy. However, this cutoff was determined by assessing the 75th percentile of the incidence of blood transfusion: the 75th percentile is the standard cutoff for evaluating adverse events. In addition, although infections and blood transfusions are extremely rare with direct laryngoscopy, this procedure was not excluded from analysis of low-acuity procedures because this cohort of patients is important in regard to associated comorbidities. Many patients undergoing very-low-acuity procedures have comorbidities, making it incumbent on the surgeon to ensure that medical preparation is appropriate and adequate.

Descriptive statistics for scaled values and frequencies of study patients within the categories for each of the parameters of interest were enumerated with the assistance of commercial statistical software. Cutoff points for defining negative indicators for scaled parameters for LAPs and HAPs were set at the 75th percentile. Correlations between variables were assessed by Pearson χ^2 or, where there were fewer than 10 subjects in any cell of a 2×2 grid, by the 2-tailed Fisher exact test. These statistical tests were performed with the assistance of the Statistica (StatSoft Inc, Tulsa, Oklahoma) and (SPSS for Windows, version 17; SPSS Inc, Chicago, Illinois) statistical software applications.

The influence of surgical acuity on the outcome measures was first assessed to determine its impact on performance and quality outcomes. Then, the impact of comorbid conditions on the outcome measures listed for both HAPs and LAPs was evaluated. Finally, the influence of the individual surgeons was added to the model to estimate its impact on performance outcome.

Table 1. Incidence of Negative Surgery Performance Indicators in Low- and High-Acuity Procedures[a]

Negative Indicator	Procedure Type: Low Acuity (n=2234)	Procedure Type: High Acuity (n=384)	*P* Value[b]
Stay of ≥6 d	219 (9.8)	371 (96.6)	<.001
Red blood cell usage	61 (2.7)	156 (40.6)	<.001
Return to operating room ≤7 d	54 (2.4)	35 (9.1)	<.001
Site infected in ≤30 d	26 (1.2)	54 (14.1)	<.001
Readmission in ≤30 d	83/1309 (6.3)	54/379 (14.3)	<.001
Died in ≤30 d	7 (0.3)	6 (1.6)	.001
Any negative indicator	873 (39.1)	234 (60.9)	<.001
≥2 Negative indicators	151 (6.8)	102 (40.3)	<.001

[a] Unless otherwise noted, data are number (percentage) of patients.
[b] *P* values from the Pearson χ^2 test

RESULTS

Over the 5-year study period, 2618 patients were included in the analysis. A total of 2234 patients underwent LAPs, while the remainder underwent high-acuity resections with reconstruction (Figure 1). Low-acuity procedures included direct laryngoscopy and neck dissection, which accounted for 19.7% and 39.1% of all LAPs, respectively.

Table 1 lists the incidence of negative performance indicators according to procedure acuity. For low-acuity procedures, the transfusion rate was 2.7% compared with 40.6% for patients with high-acuity operations. The overall wound infection rate for low- vs high-acuity surgery was 1.2% vs 14.1%. Overall, the 30-day mortality rate was low at 0.3% and 1.6% for LAPs and HAPs, respectively. The mean (SD) hospital stay for LAPs was 2.1 (3.6) days vs 10.5 (7.0) days for HAPs. The incidence of all the negative performance factors tested was statistically significantly higher in patients who underwent HAPs. Because of the differences in LOS and blood usage between LAPs and HAPs, the values for negative performance for each type of procedure were adjusted to the 75th percentile for use in further analysis. For LOS, the values were 2 or more days and 11 or more days for LAPs and HAPs, respectively. For blood usage, the values were 1 U or more and 2 U or more for LAPs and HAPs, respectively.

Possible associations of preexisting comorbid conditions with negative surgical performance measures were analyzed (**Table 2**). There was a significantly higher incidence of at least some negative indicators among patients with each of the comorbid conditions examined, but the effects of diabetes and liver disease appeared to affect fewer indicators than the others. This may have been influenced by their relatively low incidences. The presence of any comorbid condition was significantly correlated with higher levels of each of the negative indicators, having at least 1 negative indicator, and having 2 or more negative indicators.

Furthermore, patients who underwent HAPs also had higher rates of comorbid conditions. **Table 3** demonstrates the relationship between incidences of patients' comorbid conditions stratified by procedure acuity. With the exception of diabetes and liver disease there was a significantly higher incidence of comorbid conditions in the patients who underwent HAPs. In order to examine the association of comorbid conditions with negative performance indicators, it was necessary to stratify by the acuity of the surgical procedure.

Table 4 summarizes the significant correlations between having 1 or more comorbid condition and 1 or more or 2 or more negative indicators for both the low-acuity and high-acuity surgery patients. There were also significant correlations between having 2 comorbid conditions and 1 or more negative indicators among the LAP patients and having 1 or more and 2 or more negative indicators among the HAP patients, The correlation with negative performance indicators that was encountered among patients with 2 or more comorbid conditions appeared more pronounced for patients who underwent HAPs.

To compare physician performance as measured by negative indicators, we examined the percentages of HAPs vs LAPs performed by the individual surgeons and the incidences of patients with comorbid conditions on whom they performed those procedures. The activity of the 10 surgeons included in this study with regard to numbers of total cases, stratified by high and low acuity, and the percentage of cases that were HAPs were enumerated for each individual surgeon. The 25th and 75th percentile values were calculated to divide the surgeons into 3 groups based on the percentages of their patients who underwent HAPs. The upper 25th percentile group had significantly higher proportions of high-acuity cases than either the lower 75th percentile group or the middle group (**Table 5**). The lower 75th percentile group had a significantly lower proportion of high-acuity cases than the middle group as well.

Similarly, to determine whether there was variability among the surgeons in the proportions of their patients who had coexisting comorbid conditions, we noted the proportion of cases with at least 1 comorbid condition prior to surgery for each of the surgeons and calculated the upper 25th and lower 75th percentiles. There was significant variation between the groups, notably between the upper 25th and lower 75th percentile groups, and between the middle group and the lower 75th percentile group (**Table 6**).

To assess whether there was variability among the surgeons in the proportions of their cases with negative performance indicators, we enumerated the percentage of cases that had 2 or more negative indicators for each surgeon. The lowest 25th (fewest negative indicators) and highest 25th percentiles were then calculated. The lowest 25th percentile group had statistically significantly lower proportions of patients who had 2 or more negative performance indicators than either the highest 25th percentile group or the middle group (**Table 7**). The highest 25th percentile group had a significantly lower proportion of cases with negative indicators than the middle group as well.

The incidences of negative performance indicators were correlated with the acuity of the surgical procedures performed, the patient's comorbidity, and the performance of the surgeon (**Figure 2**). Procedure acuity had the greatest impact, followed by the individual surgeon, and then by comorbidity. There is considerable overlap in the confidence limits of the risk estimates between procedure acuity and physician as they relate to the presence or absence of 2 or more negative indicators and, to a lesser degree, between physician and comorbid conditions.

Table 2. Associations Between Comorbidities and Surgery Quality Indicators[a]

Comorbid Condition/Indicator	Cardiovascular Disease	COPD	Diabetes	Liver Disease	Prior Heart Failure	Renal Disease	Any Comorbidity	≥2 Comorbidities
Stay of ≥6 d	<.001	<.001	.03	.50	<.001	<.001	<.001	<.001
Red blood cell usage	<.001	<.001	.006	.22	<.001	.02	<.001	<.001
Return to operating room ≤7 d	.04	.21	.36	.12	.50	.57	.004	.11
Site infected in ≤30 d	.006	.54	.97	.004	.08	.88	.004	.11
Readmission in ≤30 d	.11	.005	.49	.02	.34	<.001	.01	.01
Died in ≤30 d	.31	<.001	.25	.63	<.001	<.001	.004	<.001
Any negative indicator	<.001	<.001	.20	.77	<.001	.007	<.001	<.001
≥2 Negative indicators	<.001	<.001	.04	.06	<.001	.006	<.001	<.001

Abbreviation: COPD, chronic obstructive pulmonary disease.
[a] All data are *P* values from the Pearson χ^2 test.

Where each of the surgeons included in the study fell in the 2-dimensional matrices showing percentages of patients with at least 1 comorbid condition and percentages of patients with 2 or more negative performance indicators is shown in **Figure 3**. The position of each surgeon in the 3-dimensional matrix showing percentages of patients who underwent HAPs, had at least 1 comorbid condition, and had 2 or more negative performance indicators is shown in **Figure 4**.

COMMENT

With ongoing health care reform in the United States, emphasis will be placed on access to care, controlling costs, and tying reimbursement to the quality of care and outcomes. High-performing practitioners and institutions may benefit in this environment through regionalization of cancer care to improve outcomes. Regionalization is defined as the delivery of care at a limited number of selected provider sites. In that environment, patients and payers will seek physicians and institutional health care providers who have demonstrated superior performance as measured by more efficient resource utilization and superior outcomes, including better local-regional tumor control and survival. For major oncologic procedures, such as esophagectomy, pancreatectomy, and colectomy, high-volume centers have demonstrated superior outcomes with respect to LOS and perioperative mortality.[3,4] Long et al[5] demonstrated lower perioperative mortality and LOS for patients undergoing craniotomy in high-volume vs low-volume institutions. In their study, the authors noted that costs are higher when care is provided in high-volume centers, which is offset by improved outcomes. Striking benefits of regionalization were found by Gordon et al,[6] who examined outcomes following pancreaticoduodenectomies performed across the region and compared these with outcomes at 1 high-volume tertiary academic center that performed over 50% of the procedures in the region. Mortality rate, mean LOS, and intensive care unit LOS were lower in the regional center than in the nonregional centers. Total charges were all statistically significantly lower in the regional high-volume center than in the other community hospitals. In general, improvement in outcomes following complex surgical procedures can be expected in high-volume centers.

Table 3. Incidence of Comorbid Conditions in Patients Undergoing Low- or High-Acuity Procedures[a]

	Procedure Type		
Comorbid Condition	Low Acuity (n=2234)	High Acuity (n=384)	*P* Value[b]
Diabetes	270 (12.1)	57 (14.8)	.13
Cardiovascular disease	1020 (45.7)	223 (58.1)	<.001
Previous heart failure	75 (3.4)	34 (8.9)	<.001
Chronic obstructive pulmonary disease	100 (4.5)	65 (16.9)	<.001
Liver disease	35 (1.6)	10 (2.6)	.15
Renal disease	40 (1.8)	13 (3.4)	.04
Any comorbid condition	1113 (49.8)	259 (67.5)	<.001
≥2 Comorbid conditions	341 (15.3)	110 (28.7)	<.001

[a] Unless otherwise noted, data are given as number (percentage) of patients.
[b] *P* values from the Pearson χ^2 test

As a by-product of regionalization, greater resource utilization and cost of care may result from treating patients with higher disease acuity and comorbid conditions at centers with high volumes and demonstrably higher quality of care. Institutions will place greater emphasis on efficiency and performance to reduce financial risk and to remain viable. Reimbursement may be tied to risk-adjusted outcomes, thereby creating incentives to increase efficiency and prospectively manage comorbidity. Treating physicians will likely be profiled and ranked according to their resource utilization and performance. It will be incumbent on physicians to provide care efficiently and to prospectively manage comorbid conditions that will have a direct effect on performance measures.[7] Providing physicians with risk-adjusted data on individual performance and peer comparisons will likely impact physician practices.

The model we have developed may be used as a tool for changing practice behavior, identifying best practices, and recognizing adverse trends that require an in-depth cause analysis. For example, scrutinizing management practices that adversely affect outcomes could require modification of technique or antibiotic wound prophylaxis, depending on the outcome measure in need of improvement. If an upward trend in perioperative

Table 4. Association of Comorbid Conditions and Negative Quality Indicators in Patients Undergoing Low- or High-Acuity Procedures

		≥1 Negative Indicator (n=552)		≥2 Negative Indicators (n=238)	
Procedure Type	**Total**	**Patients, No. (%)**	***P* Value**[a]	**Patients, No. (%)**	***P* Value**[a]
Low acuity	2234	873 (39.1)		151 (6.8)	
Any comorbid condition	1113	478 (43.0)	<.001	95 (8.5)	<.001
≥2 comorbid conditions	341	154 (45.2)	.01	28 (8.2)	.25
High acuity	384	234 (60.9)		158 (53.4)	
Any comorbid condition	259	174 (67.2)	<.001	102 (26.6)	.006
≥2 comorbid conditions	110	86 (78.2)	<.001	45 (40.9)	<.001

[a] *P* values from Pearson χ^2 test.

Table 5. Low- and High-Acuity Procedures Among Surgeons

High-Acuity Group	Surgeons, No.	Low-Acuity Cases, No.	High-Acuity Cases, No. (%)
Upper 25%[a,b]	3	538	138 (20.4)
Middle 50%[c]	4	897	165 (15.5)
Lower 25%	3	799	81 (9.2)
Total	**10**	**2234**	**384** (14.7)

[a] Upper 25% vs middle 50%, $P < .001$.
[b] Upper 25% vs lower 25%, $P < .001$.
[c] Middle 50% vs lower 25%, $P = .009$.

Table 6. Proportions of Cases With Comorbid Conditions Among Surgeons

≥1 Comorbidity Group	Surgeons, No.	Patients Without Comorbid Condition(s), No.	Patients With Comorbid Condition(s), No. (%)
Upper 25%[a,b]	3	244	346 (58.6)
Middle 50%[c]	4	415	521 (55.7)
Lower 25%	3	587	505 (36.8)
Total	10	1246	1372 (52.4)

[a] Upper 25% vs middle 50%, $P = .25$.
[b] Upper 25% vs lower 25%, $P < .001$.
[c] Middle 50% vs lower 25%, $P < .001$.

Table 7. Proportions of Cases With 2 or More Negative Performance Indicators Among Surgeons

≥2 NI Group	Surgeons, No.	Patients Without ≥2 NI	Patients With ≥2 NI, No. (%)
Fewest 25%[a,b]	3	834	46 (5.23)
Middle 50%[c]	3	839	85 (9.12)
Highest 25%	4	692	122 (14.99)
Total	**10**	**2365**	**253** (9.66)

Abbreviation: NI, negative indicators.
[a] Fewest 25% vs middle 50%, $P = .001$.
[b] Fewest 25% vs highest 25%, $P < .001$.
[c] Middle 50% vs highest 25%, $P < .001$.

wound infections was noted among 1 or more physicians in the cohort, an in-depth analysis of the cause might identify etiologic factors requiring remediation.[8]

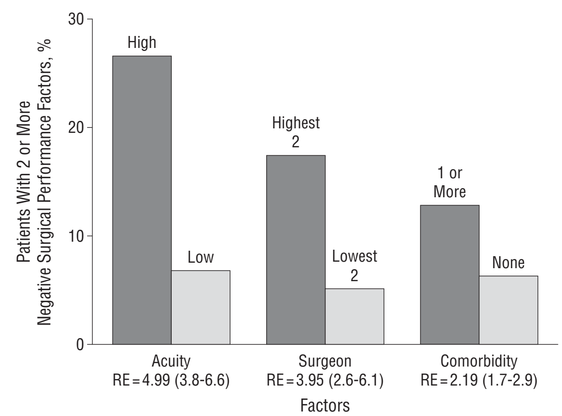

Figure 2. Factors affecting incidence of 2 or more negative surgical performance indicators. RE indicates risk estimate. $P < .001$ for all comparisons.

Within head and neck surgical oncology practices, a clinical tool, such as a care pathway, can be incorporated into perioperative management of patients to improve efficiency and quality of care. Care pathways are algorithms designed to achieve realistic goals for clinical progress through coordination of a multidisciplinary management team that simultaneously seeks to maximize efficient use of hospital resources and maintain accepted standards of care.[9] Pathways incorporate the highest level of evidence available or best practices for a particular intervention in the perioperative management of patients. The care pathway paradigm can help to standardize decision making and introduce uniformity in resource utilization. Analogous to the high-acuity patients in this study, Husbands et al[9] demonstrated that care pathway implementation reduced LOS for patients requiring complex head and neck reconstruction from a mean of 13 days to 8 days. Gendron et al[10] found a significantly decreased incidence of postoperative nosocomial pneumonia following major head and neck cancer among patients who were managed on a clinical pathway vs those who received routine postoperative management. Decreased pulmonary complications seen in patients on the clinical pathway was believed to result from the mandatory daily interventions that were part of the postoperative physician order sets. These automatically began aggressive pulmonary toilet, ambulation, and respiratory therapy intervention at prescribed times during the postoperative period. Pathways are commonly used in current

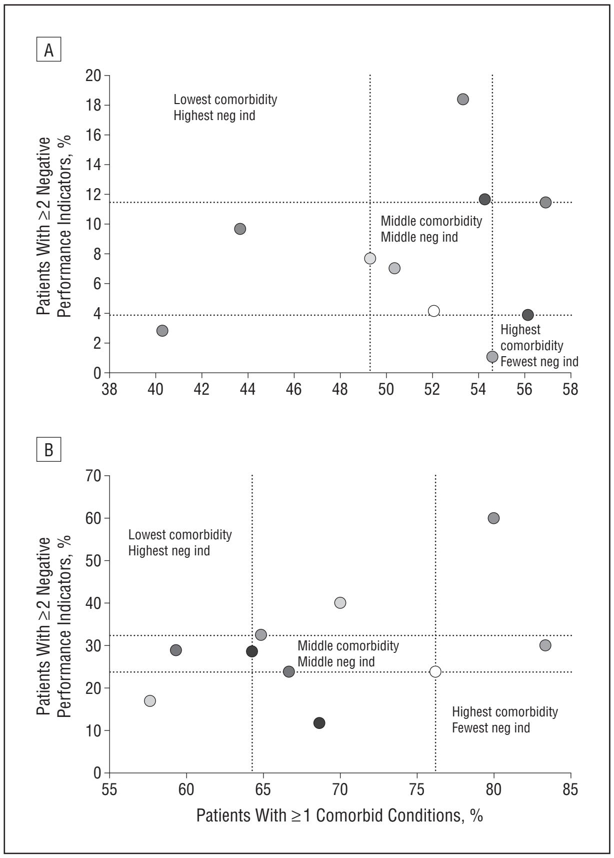

Figure 3. Comorbidity and negative indicators (neg ind) in low-acuity (A) and high-acuity (B) procedures by surgeon. Where each of the surgeons included in the study fell in the 2-dimensional matrices showing percentages of patients with at least 1 comorbid condition and percentages of patients who had 2 or more negative performance indicators is shown. The dashed lines indicate the boundaries of the highest and lowest 25th percentiles. In both panels, the lower right corner would have the surgeons whose patients had the lowest incidence of 2 or more negative indicators and the highest incidence of comorbidity, while the upper left corner would contain the surgeons whose patients had both the highest incidence of negative indicators and the lowest proportion with comorbid conditions.

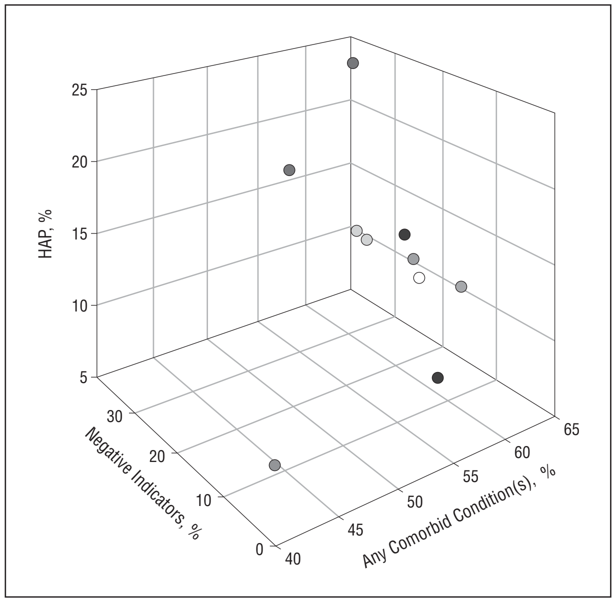

Figure 4. High-acuity procedures (HAPs), comorbidity, and negative performance indicators by surgeon. The position of each surgeon in the 3-dimensional matrix shows percentages of patients who underwent HAPs, had at least 1 comorbid condition, and had 2 or more negative performance indicators. While Figure 3 is specific for either HAPs or low-acuity procedures, this figure highlights the percentage of HAPs in a given surgeon's case mix. Most of the surgeons are grouped in the middle of the diagram. The surgeon represented by the dot in the lower left corner, however, has the lowest percentage of HAPs, fewest patients with comorbid conditions, and fewest cases with 2 or more negative performance indicators. On the other hand, the surgeon represented by the dot at the top of the diagram had the highest percentage of HAPs, a high percentage of patients with comorbid conditions, and a high percentage of cases with 2 or more negative performance indicators.

surgical practices to ensure consistency of care and interventions designed to achieve milestones that permit hospital discharge on a targeted date.

Owing to the variability in case mix and comorbidity particular to each surgeon's practice, it was necessary to adjust for these factors to isolate the impact of an individual surgeon on performance measures. Adjustment is particularly necessary for HAPs, since these patients tend to have higher rates of comorbid conditions. As summarized in Table 3, the influence of specific comorbidities was also addressed. Although cardiovascular disease, previous heart failure, and chronic obstructive pulmonary disease were the only comorbidities found to have significant differences between low- and high-acuity cases, this finding may be confounded by the low incidences of renal and liver diseases. As our data demonstrate, there was a 78% incidence of 2 or more negative performance measures encountered for patients undergoing HAPs who had multiple comorbid conditions. In contrast, 2 or more negative performance measures were less likely to occur as an outcome among patients undergoing low-acuity procedures, regardless of their comorbidity. For the LAP patients with 2 or more comorbid conditions, the likelihood of encountering 2 or more negative performance indicators was 8.2%, or approximately 9 times less than for HAPs. Risk adjustment is therefore critical for comparing surgical outcomes given the heterogeneity of the patients' medical conditions and procedure acuity.

It should be noted that most of our reconstructions were performed by a second surgeon, which means that more than 1 surgeon was involved in the care of most patients undergoing HAPs. Although the head and neck surgeon manages all perioperative care, we excluded flap failure as a negative indicator. Future studies, however, will take into account all team cases and all resultant complications.

Through this methodology, we were able to divide surgeons' performance for HAPs into 3 groups adjusted for comorbidity and to identify those with the most favorable performance profile. These data are useful for allowing each surgeon to assess individual outcomes in relation to those of his or her peers. These data are provided to the faculty surgeons on a semiannual basis, with cumulative data added and analyzed over time. The longer-term analysis is especially useful for determining positive and negative trends over time.

A future planned extension of this work is to perform a concordance analysis of a faculty head and neck surgeon's performance whose practice is based at an affiliated institution. Using the same criteria for procedure acuity and risk adjustment should permit a comparison of outcomes between surgeons practicing in 2 affiliated institutions. Relative to our current data, when we adjusted for comorbidi-

ties and surgical acuity, surgeon volume did not have significance. This is likely confounded by our status as a high-volume institution, and comparison to a lower-volume center would be necessary to appropriately address whether surgeons with higher volumes have better outcomes.

In the future, as care is increasingly more regionalized, reporting intrainstitutional and interinstitutional provider outcomes will likely be mandatory, thereby permitting patients to seek care based on objective outcomes measures.[11-14] Developing and reporting these data sets will require robust prospectively acquired data and consistent methodology for analysis and reporting. Governmental agencies, such as the Agency for Healthcare Research and Quality and the Center for Medicare and Medicaid Services, and nongovernmental organizations, such as the National Quality Forum and the NSQIP, will provide the impetus and development funding for this effort.[15-17]

Within the context of payment reform for health care services, these data will be critical as national benchmarks are established and achieving outcomes that meet or exceed these targets will be directly tied to reimbursement (pay for performance).[18,19] The concept of bundled payments for cancer care that cover a range of services related to a defined cancer diagnosis is also being seriously considered. Bundled payments will be paid for all of the care delivered to a patient with cancer for a particular diagnosis over a defined interval. How these payments will be divided among the involved health care providers and treating institutions is unclear. Analogous to Medicare payments based on the admitting diagnosis and diagnosis-related group, bundled payments will place the onus on hospitals and health care providers to manage their costs of care by increasing efficiency, performance, and quality.

CONCLUSION

In evaluating physician performance comparing the relative incidence of negative surgical quality and performance indicators, other factors that affect outcomes must be included. Two important factors are procedure acuity and the patient's comorbid conditions. Although both factors have significant effects, the acuity of the procedure is the strongest determinant of the incidence of negative quality and performance indicators and should be heavily weighted in physician comparisons. These data may serve as a tool to evaluate physician performance and to positively impact outcomes through identifying best practices and providing data to individual surgeons who, through self-improvement, may positively affect their patients' outcomes. In the future, regionalization of complex care will likely direct patients who need resource-intensive therapy to centers that have objective data demonstrating that the health care providers and the institution meet national benchmarks for performance and quality. Finally, reimbursement may eventually be tied to quantifiable measures of physician and institutional performance.

Accepted for Publication: September 2, 2010.
Correspondence: Randal S. Weber, MD, Department of Head and Neck Surgery, University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd, Unit 1445, Houston, TX 77030 (rsweber@mdanderson.org).

Author Contributions: Drs Weber, Hanna, and Lai and Mr Eastman had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. *Study concept and design:* Weber, Hanna, Hessel, and Lai. *Acquisition of data:* Hanna, Akiwumi, and Kian. *Analysis and interpretation of data:* Weber, Lewis, Eastman, Kupferman, and Roberts. *Drafting of the manuscript:* Weber and Roberts. *Critical revision of the manuscript for important intellectual content:* Weber, Lewis, Eastman, Hanna, Akiwumi, Hessel, Lai, Kian, Kupferman, and Roberts. *Statistical analysis:* Roberts. *Administrative, technical, and material support:* Weber, Lewis, Eastman, Akiwumi, Kian, and Kupferman. *Study supervision:* Weber, Hanna, Hessel, and Kian.
Financial Disclosure: None reported.
Previous Presentation: This study was presented at the annual meeting of the American Head and Neck Society; April 28, 2010; Las Vegas, Nevada.

REFERENCES

1. Sisko A, Truffer C, Smith S, et al. Health spending projections through 2018: recession effects add uncertainty to the outlook. *Health Aff (Millwood)*. 2009;28(2):w346-w357.
2. US Institute of Medicine; Committee on Quality of Health Care in America. *Crossing the Quality Chasm: A New Health System for the 21st Century.* Washington, DC: National Academy Press; 2001.
3. Stitzenberg KB, Sigurdson ER, Egleston BL, Starkey RB, Meropol NJ. Centralization of cancer surgery: implications for patient access to optimal care. *J Clin Oncol*. 2009;27(28):4671-4678.
4. Wouters MW, Wijnhoven BP, Karim-Kos HE, et al. High-volume versus low-volume for esophageal resections for cancer: the essential role of case-mix adjustments based on clinical data. *Ann Surg Oncol*. 2008;15(1):80-87.
5. Long DM, Gordon T, Bowman H, et al. Outcome and cost of craniotomy performed to treat tumors in regional academic referral centers. *Neurosurgery*. 2003;52(5):1056-1065.
6. Gordon TA, Burleyson GP, Tielsch JM, Cameron JL. The effects of regionalization on cost and outcome for one general high-risk surgical procedure. *Ann Surg*. 1995;221(1):43-49.
7. Reinke T. The push for health care value spurs greater physician scrutiny. *Physician Exec*. 2007;33(4):42-44, 46-47.
8. Chalasani V, Abdelhady M, Stitt L, Izawa JI. Quality assurance and benchmarking for radical cystectomy: monitoring early complications and mortality using cumulative summation charts. *J Urol*. 2009;181(4):1581-1586.
9. Husbands JM, Weber RS, Karpati RL, et al. Clinical care pathways: decreasing resource utilization in head and neck surgical patients. *Otolaryngol Head Neck Surg*. 1999;121(6):755-759.
10. Gendron KM, Lai SY, Weinstein GS, et al. Clinical care pathway for head and neck cancer: a valuable tool for decreasing resource utilization. *Arch Otolaryngol Head Neck Surg*. 2002;128(3):258-262.
11. Staiger DO, Dimick JB, Baser O, Fan Z, Birkmeyer JD. Empirically derived composite measures of surgical performance. *Med Care*. 2009;47(2):226-233.
12. Polk HC Jr. Renewal of surgical quality and safety initiatives: a multispecialty challenge. *Mayo Clin Proc*. 2006;81(3):345-352.
13. Shahian DM, Edwards FH, Ferraris VA, et al; Society of Thoracic Surgeons Quality Measurement Task Force. Quality measurement in adult cardiac surgery: part 1—conceptual framework and measure selection. *Ann Thorac Surg*. 2007;83(4)(suppl):S3-S12.
14. O'Brien SM, Shahian DM, DeLong ER, et al. Quality measurement in adult cardiac surgery: part 2—statistical considerations in composite measure scoring and provider rating. *Ann Thorac Surg*. 2007;83(4)(suppl):S13-S26.
15. Birkmeyer JD, Dimick JB. Potential benefits of the new Leapfrog standards: effect of process and outcomes measures. *Surgery*. 2004;135(6):569-575.
16. Siegel JE, Byron SC, Lawrence WF. Federal sponsorship of cost-effectiveness and related research in health care: 1997-2001. *Value Health*. 2005;8(3):223-236.
17. Bilimoria KY, Bentrem DJ, Stewart AK, et al. Lymph node evaluation as a colon cancer quality measure: a national hospital report card. *J Natl Cancer Inst*. 2008;100(18):1310-1317.
18. Lee PV, Berenson RA, Tooker J. Payment reform: the need to harmonize approaches in Medicare and the private sector. *N Engl J Med*. 2010;362(1):3-5.
19. Porembka MR, Hall BL, Hirbe M, Strasberg SM. Quantitative weighting of postoperative complications based on the accordion severity grading system: demonstration of potential impact using the American college of surgeons national surgical quality improvement program. *J Am Coll Surg*. 2010;210(3):286-298.