Quality Indicators in Head and Neck Operations

A Comparison With Published Benchmarks

Thomas D. Shellenberger, DMD, MD; Rafael Madero-Visbal, MD; Randal S. Weber, MD

Objective: To investigate the reproducibility of quality indicators in the care of patients undergoing operations for head and neck cancer.

Design: A review of specialty-specific surgical quality indicators in a cohort undergoing procedures for definitive treatment of head and neck cancer, stratified by high and low acuity of the surgical procedures and compared with established benchmarks.

Setting: A large tertiary care institution and an associated multidisciplinary cancer center.

Patients: Fifty randomly selected patients with evaluable data who were diagnosed as having head and neck cancer that was definitively treated using any of the 3 modalities (surgical procedures, chemotherapy, and/or radiotherapy) during a 15-month period at our center. Twenty-one patients who underwent operations form the basis of this report.

Main Outcome Measures: Procedures were stratified by acuity on the basis of the extent of the operation.

Data were centered on quality indicators designed to reflect length of stay, readmission within 30 days postoperatively, return to the operating room within 7 days of surgery, use of blood products, 30-day mortality, adequacy of reports on surgical pathologic findings, and surgical site infection.

Results: Diagnoses in the cohort included carcinoma of the oral cavity in 19 patients (39%), oropharynx in 14 (29%), larynx in 13 (27%), and hypopharynx in 3 (6%). High- and low-acuity surgical procedures were performed in 12 and 7 patients, respectively. No statistically significant differences in the measures for quality indicators were found between the cohort and the calculated benchmarks.

Conclusion: Our findings demonstrate the applicability of quality indicators to the care of patients with head and neck cancer treated by surgical intervention stratified by acuity and compared with established benchmarks.


IMS TO DEFINE QUALITY standards in head and neck cancer and surgical treatment bear the promise of improving both the opportunity for optimal care and the value of therapy for patients with cancer. By providing effective guidelines for management of the therapy, quality standards may improve efficiency in the delivery of care and its safety. Patients undergoing care consistent with evidence-based practice and specified by explicit quality measures have shown 1,2 more favorable outcomes after initial treatment compared with patients undergoing care with variable compliance with quality measures. Furthermore, a practice-based system of quality assessment has been demonstrated 3 to be rapid and effective in measuring care quality while allowing comparison among clinicians and practices over time to promote excellence in cancer care.

Objective assessment of quality by the head and neck surgeon managing care for patients with head and neck cancer offers a basis for raising the standard of care. 4 Steps in developing quality-of-care measures have been outlined by the Quality of Care Committee of the American Head and Neck Society 5 to include proposing means by which practitioners can evaluate their treatment practices. Formulating evidence-based quality-of-care measures, validating their applicability, and promoting compliance with such standards can ensure the highest quality of care for patients with head and neck cancer.

Surgical quality indicators provide criteria by which the quality of care can be measured and referenced to defined benchmarks, thereby promoting uniformity in management and identifying trends that lead to process improvement. Quality indicators have been proposed by the Agency...
for Healthcare Research and Quality and widely applied to evaluate the quality of clinical services with a focus on health care outcomes. Specialty-specific quality indicators have been proposed for the treatment of colorectal cancer, trauma care, and breast cancer. Moreover, surgical quality indicators have been established for head and neck surgical oncology to assess care outcomes adjusted for procedure acuity and patient comorbidity. In addition, specific measurable variables have been identified to provide quality data for assessing the extent to which cancer care complies with accepted treatment guidelines and is delivered for patients with oral tongue cancer.

Variability among metrics in quality assessment can limit generalizability and applicability across providers and their institutions. Variability in patient care volume, the disparity of support resources available to clinicians, and the vast milieu of factors influencing the outcome of treatment pose further threats to the validity of quality assessment. Moreover, the heterogeneity of disease and patient factors combined with a paucity of high-level prospective data present a challenge for the hope of uniform treatment strategies for patients with head and neck cancer. Therefore, defining quality indicators for application across surgeons and their multidisciplinary practices and comparing measures with accepted benchmarks may provide an opportunity for improved quality of care for patients undergoing initial treatment for head and neck cancer.

We sought to investigate the reproducibility of surgical quality indicators across institutions in the management of care for patients with head and neck cancer undergoing surgical intervention. To accomplish this, we reviewed specialty-specific surgical quality indicators in a cohort of patients undergoing definitive operations for head and neck cancer at MD Anderson Cancer Center Orlando, Orlando, Florida, in reference to benchmarks established by the Department of Head and Neck Surgery of The University of Texas MD Anderson Cancer Center in Houston, Texas. The results of our study support the notion that the quality of surgical care for patients with head and neck cancer can be objectively measured in a reproducible manner. Moreover, we propose that the assessment of surgical quality indicators can guide improvement in the quality of care rendered for this population with the aim of decreasing complications of treatment.

METHODS

A review of the oncology care practices of MD Anderson Cancer Center Orlando was performed for patients diagnosed as having head and neck cancer from June 1, 2008, through September 30, 2009. The review was facilitated by the MD Anderson Physicians Network, a quality management and best practice organization that delivers cancer management services by assessing quality of care across the MD Anderson network. Patients were cared for by physicians of the MD Anderson Cancer Center Orlando. Data abstractors were from MD Anderson Cancer Center Orlando, with training and support provided by the MD Anderson Physicians Network. During the record review, MD Anderson Cancer Center Orlando auditors reviewed electronic and paper records for the hospital and physician’s office practice locations. The review was performed in compliance with the policies and procedures of the Oncology Institutional Review Board of MD Anderson Cancer Center Orlando.

Each case was reviewed and all the reporting was approved by the MD Anderson Physicians Network medical director. If cases were not concordant for treatment, MD Anderson Cancer Center faculty provided subspecialty expert review of the individualized treatment decisions.

PARTICIPANTS

Quality indicator assessment required collaboration and cooperation between multiple departments and divisions at MD Anderson Physicians Network, The University of Texas MD Anderson Cancer Center, and MD Anderson Cancer Center Orlando. Primary participants included the Department of Biostatistics and the Computer Applications and Support Department from The University of Texas MD Anderson Cancer Center; the medical director, Quality Management, Healthcare Services, and Account Management from MD Anderson Physicians Network; and Health Information Management, Tumor Registry, and Cancer Services from MD Anderson Cancer Center Orlando. Quality indicators were assessed in an eligible, randomized sample of patients.

MEASUREMENT

Electronic audit and data entry tools that were developed and designed to eliminate bias were used consistently throughout the assessment. Physician members of the head and neck multidisciplinary team were consulted for subspecialty expert review when appropriate.

RANDOMIZATION

Accession number files from the MD Anderson Cancer Center Orlando Tumor Registry were submitted to MD Anderson Physicians Network Quality Management and forwarded to the Computer Applications and Support Department for randomization using commercial software (SPSS; SPSS Inc, Chicago, Illinois). Once the randomized files of accession numbers were created, the files were returned to MD Anderson Cancer Center Orlando for record retrieval. The Tumor Registry provided an electronic file of routine data elements to upload to the database. From all cases diagnosed during the study, accession numbers were randomly selected for evaluation and screened for inclusion. Training, technical support, and clinical support were provided by MD Anderson Physicians Network. Data abstractors from MD Anderson Cancer Center Orlando entered data via the Internet to the database of MD Anderson Physicians Network. Record review and data entry were conducted by abstractors from MD Anderson Cancer Center Orlando; the abstractors also identified and provided source documents. All data were verified by review of source documents by MD Anderson Physicians Network. The medical director reviewed each case.

PATIENT SELECTION

Evaluable cases included head and neck cancer definitively treated by any of the 3 modalities: surgical procedures, chemotherapy, and/or radiotherapy. Fifteen months of data were reviewed to include patients with head and neck cancer diagnosed between June 1, 2008, and September 30, 2009, as identified by the Tumor Registry. Once 50 cases that appeared to
be evaluable were identified during the data collection, no further cases were selected.

EXCLUSION

Cases were excluded if there was widespread metastatic disease, recurrent disease with metastasis, a dual synchronous primary tumor, complex histologic findings or complex comorbidities or if comparison with Clinical Care Guidelines of MD Anderson Cancer Center would be inappropriate. Also excluded were transient cases in which information was noncontributory or evaluation was for a second opinion.

QUALITY INDICATORS ASSESSMENT END POINTS

Data centered on established quality indicators that reflect measures endorsed by both MD Anderson Cancer Center and current professional and quality organizations, including the Advisory Board of the American Society of Clinical Oncology, the National Comprehensive Cancer Center Network, the Commission on Cancer of the American College of Surgeons, and National Quality Forum. Surgical procedures were categorized by acuity (low vs high). Benchmark values for the quality indicators were established by the Department of Head and Neck Surgery of MD Anderson Cancer Center.10,14

Procedures were categorized by acuity to allow meaningful comparison with benchmarks established by the Department of Head and Neck Surgery. Low-acuity procedures comprised partial endoscopic laryngectomy, laryngoscopy, lymphadenectomy, glossectomy without reconstruction, parotidectomy, and thyroidectomy. High-acuity procedures included glossectomy with reconstruction, partial open laryngectomy, total laryngectomy, total mandibulectomy, and pharyngolaryngectomy. Other procedures included those not falling into the low- or high-acuity category.

Data on comorbid conditions were collected according to diagnosis from the patient’s medical record. These included diabetes mellitus, cardiovascular disease, history of congestive heart failure, chronic obstructive pulmonary disease, liver disease, and renal disease.

Benchmarks for the surgical quality indicators adjusted for procedure acuity and patient comorbidities were developed in a review of procedures performed by 10 surgeons in 2618 patients during a 5-year period in the Department of Head and Neck Surgery at The University of Texas MD Anderson Cancer Center.10 For length of stay and use of blood products, exceeding the defined cutoff points for low- and high-acuity procedures was considered a negative performance indicator. The standard cutoff for evaluating adverse events was applied by assessing the 75th percentile for the scaled variables, i.e., the number of days of hospitalization and the incidence of blood transfusion. In contrast, for return to the operating room, surgical site infection, and mortality, any event independent of the acuity of surgery was considered a negative performance indicator.

LENGTH OF STAY

The duration of hospitalization was quantified as the number of days, in full or in part, from the calendar date of admission through the date of discharge and was stratified between low- and high-acuity procedures. The benchmarks established for length of stay by acuity of procedure are that at least 75% of patients stay 3 days or less after undergoing a low-acuity procedure and 12 days or less after undergoing a high-acuity procedure.

READMISSION WITHIN 30 DAYS OF THE OPERATION

Patients were assessed for the presence or absence of readmission to the hospital and the number of days from the date of the operation. Readmission was considered present in any patient who was admitted to an inpatient unit within the center for any cause within 30 calendar days of the date of the procedure. The benchmark for this variable is less than 5% of patients after a low-acuity procedure and less than 13% of patients after a high-acuity procedure.

RETURN TO THE OPERATING ROOM WITHIN 7 DAYS OF THE OPERATION

Data were obtained on patients who returned to the operating room for any reason within 7 calendar days of the date of the procedure. The benchmark for return to the operating room within 7 days of surgery is less than 2% of patients after a low-acuity procedure and less than 10% of patients after a high-acuity procedure.

30-DAY MORTALITY

Mortality after surgery is an important indicator for surgical outcomes. Perioperative mortality was defined as death within 30 days of the procedure. The benchmark for 30-day mortality is less than 0.3% of patients after a low-acuity procedure and less than 2% of patients after a high-acuity procedure.

USE OF BLOOD PRODUCTS

The need for transfusion was assessed for the surgical cohort, and the number of units transfused for each patient was recorded. Use of blood products was defined as the administration of packed red blood cells either intraoperatively or during the remainder of the postoperative period of hospitalization. The benchmark for use of blood products is that 75% of patients undergoing a low-acuity procedure receive less than 1 U of blood and 75% of those undergoing a high-acuity procedure receive less than 3 U.

ADEQUACY OF PATHOLOGY REPORTS

The surgical pathologic report for each procedure was assessed for completeness as determined by the reporting standards of the College of American Pathologists (http://www.cap.org/apps/cap.portal). The standard data elements for invasive carcinomas of the upper aerodigestive tract are site specific for primary tumors of the oral cavity, pharynx, larynx, paranasal sinuses, and salivary glands. The elements include tumor size, histologic type and grade, depth of invasion, extracapsular spread, number of nodes removed, number of positive nodes, size of the largest node, margin status, venous/lymphatic invasion, perineural invasion, and additional pathologic findings, as well as pathologic staging. Reports were considered inadequate if missing any of the data elements. The missing element or reasons for omission were recorded. Deviation from standard reporting elements was captured for all patients undergoing a low- or high-acuity procedure.

SURGICAL SITE INFECTIONS

Patients were assessed for the presence or absence of surgical site infection 30 days from the date of the operation. The surgical site was considered free of infection in the absence of documented
clinical signs or symptoms of infection within 30 calendar days from the date of the operation. The benchmark for this factor is 98% of patients undergoing a low- or high-acuity procedure.

STATISTICAL ANALYSIS

Counts and percentages of patients with measures of the end points were determined for the quality indicators assessed. Patients were stratified according to the primary surgical procedure into low- and high-acuity groups. Frequencies of patients within groups and among the combined groups for each quality indicator were enumerated, and descriptive statistics on scaled data (ie, length of stay and use of blood products) were calculated. Benchmarks were calculated for each quality indicator according to the number of patients within the low- and high-acuity groups and among the combined groups when appropriate. Frequencies for benchmark values in groups and among the combined groups were compared with the calculated benchmarks by the 1-tailed Fisher exact test. \( P < .05 \) was considered statistically significant. The relative risk (RR) and 95% CI were calculated for each indicator. Statistical analysis was performed with commercial software (Number Cruncher Statistical Systems; Kaysville, Utah, http://www.ncss.com).

RESULTS

One-hundred one cases of patients with head and neck cancer evaluated during the study were randomly selected from the Tumor Registry and screened for inclusion. Fifty-one patients were eliminated from consideration because they were not selected for the study (n=14), were treated outside the established time frame (n=7), underwent prior treatment (n=7), underwent treatment elsewhere (n=7), received no treatment (n=4), had histologic findings other than squamous cell carcinoma (n=3), were found with exclusion criteria after review of abstracted data (n=3), had dual synchronous primary tumors (n=2), had distant metastasis (n=1), or had a chart with insufficient documentation (n=1). After inclusion into the study, 1 patient was eliminated because deterioration of the medical condition had precluded the continuation of definitive treatment. The final cohort included 49 patients.

Demographic and clinical characteristics of patients forming the cohort are reported in Table 1. The mean age was 59.4 years, and 42 (86%) of the patients were male. The T and N stages of cancer were evenly distributed throughout the cohort. Fifteen patients (31%) had never smoked. The number and percentage of patients categorized by primary site and subsite of head and neck cancer are reported in Table 2. Oral cavity carcinoma was present in 19 patients (39%), oropharyngeal carcinoma in 14 patients (29%), laryngeal carcinoma in 13 patients (27%), and hypopharyngeal carcinoma in 3 patients (6%).

The incidence of comorbid conditions is reported in Table 3. Cardiovascular disease was the most prevalent comorbidity (24 patients [49%]). Thirty-five patients (71%) had at least 1 comorbid disease and 37% of the patients (n=18) had at least 2 comorbidities.

Twenty-one of the 49 patients received surgical intervention and form the basis of this report. Twelve of these patients underwent a high-acuity procedure, including glossectomy with reconstruction (n=5), total laryngectomy (n=1), mandibulectomy (n=3), mandibulectomy with re-
construction (n = 1), pharyngolaryngectomy (n = 1), and pharyngolaryngectomy with reconstruction (n = 1). Seven patients underwent a low-acuity procedure, including glossectomy without reconstruction (n = 4), modified radical neck dissection (n = 1), and selective neck dissection (n = 2). Two patients underwent a procedure not categorized as high or low acuity, including maxillectomy (n = 1) and other type of resection (n = 1). Twenty-eight patients underwent definitive treatment with chemotherapy and/or radiotherapy without an operation.

QUALITY INDICATOR ASSESSMENT

The number and percentage of patients with measures of quality indicators and their associated benchmarks are reported in Table 4.

Length of Stay

Three of 7 patients (43%) who underwent a low-acuity procedure stayed 3 days or less, which was below the calculated benchmark of 6 patients or 75% of the cohort (P = .09). The RR for length of stay more than 3 days was 2.01 (95% CI, 0.63-4.92). Seven of 12 patients (58%) who underwent a high-acuity procedure stayed 12 days or less, which was below the calculated benchmark of 9 patients or 75% of the cohort (P = .33). The RR for length of stay more than 12 days was 1.29 (95% CI, 0.66-2.62). Thus, the length of stay fell below benchmarks for both low- and high-acuity procedures, although this difference was not statistically significant.

Readmission Within 30 Days After the Operation

One of 7 patients (14%) who underwent a low-acuity procedure were readmitted within 30 days after the operation, which was above the calculated benchmark of no patients or less than 5% of the cohort (P = .33). The RR for readmission for patients who underwent a low-acuity procedure was 2.82 (95% CI, 0.01-23.73). Two of 12 patients (17%) who underwent a high-acuity procedure were readmitted within 30 days after the operation, which was above the calculated benchmark of 1 patient or less than 13% of the cohort (P = .78). The RR for that group was 1.31 (95% CI, 0.07-8.24).

Table 3. Comorbidities

<table>
<thead>
<tr>
<th>Comorbidity</th>
<th>Patients, No. (%) (N = 49)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diabetes mellitus</td>
<td>6 (12)</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>24 (49)</td>
</tr>
<tr>
<td>Previous CHF</td>
<td>2 (4)</td>
</tr>
<tr>
<td>COPD</td>
<td>1 (2)</td>
</tr>
<tr>
<td>Liver disease</td>
<td>4 (8)</td>
</tr>
<tr>
<td>None</td>
<td>14 (29)</td>
</tr>
<tr>
<td>Comorbid conditions</td>
<td></td>
</tr>
<tr>
<td>≥1</td>
<td>35 (71)</td>
</tr>
<tr>
<td>≥2</td>
<td>18 (37)</td>
</tr>
</tbody>
</table>

Abbreviations: CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease.

Return to the Operating Room Within 7 Days After the Procedure

No patients who underwent a low-acuity procedure returned to the operating room within 7 days after the procedure, meeting the calculated benchmark of no patients or less than 2% of the cohort (P = .64). The RR for patients who underwent a low-acuity procedure was 1.03 (95% CI, 0.41-2.42). Two patients (17%) of those who underwent a high-acuity procedure returned to the operating room within 7 days of after the procedure, which was above the calculated benchmark of 1 patient or less than 10% of the cohort (P = .28). The RR for this group was 1.73 (95% CI, 0.03-6.72).

30-Day Mortality

No patients died within 30 days after the operation, meeting the calculated benchmark of no patients or less than 0.3% of the cohort after a low-acuity procedure (RR, 0.99; 95% CI, 0.99-1.04; P = .50) and no patients or less than 2% of the cohort after a high-acuity procedure (RR, 0.98; 95% CI, 0.98-1.01; P = .50).

Use of Blood Products

None of the 7 patients who underwent a low-acuity procedure received blood products, which was above the calculated benchmark of 6 patients or 75% of the cohort (RR, 0.75; 95% CI, 0.78-1.13; P = .68). Ten patients (83%) who underwent a high-acuity procedure received less than 3 U of blood products, which was above the calculated benchmark of 9 patients or 75% of the cohort (RR, 0.92; 95% CI, 0.64-1.42; P = .84).

Adequacy of Reports on Pathologic Findings

Seventeen patients (89%) of the combined cohorts had reports on pathologic findings deemed adequate according to the standards of the College of American Pathology, which was below the calculated benchmark of 19 patients or 100% of the cohort (P = .50 for low and high acuity). The missing element for which the report failed to meet the College of American Pathologists standards was pathologic staging in 2 patients. The RR of an inadequate report was 1.17 (95% CI, 0.88-1.13) for low-acuity procedures and 1.09 (0.93-1.13) for high-acuity procedures.

Free of Surgical Site Infection

All 7 patients who underwent a low-acuity procedure were free of surgical site infection at 30 days, meeting the calculated benchmark of 7 patients or 98% of the cohort (P = .50). Ten of the patients (83%) who underwent a high-acuity procedure were free of surgical site infection at 30 days, which was below the benchmark of 12 patients or 98% of the cohort (P = .24). The RR for surgical site infection at 30 days was 0.98 (95% CI, 0.98-1.01) for low-acuity procedures and 1.13 (95% CI, 0.90-1.18) for high-acuity procedures.
The findings of our study demonstrate the applicability of surgical quality indicators to the care of patients treated for head and neck cancer with operations stratified by acuity and compared with established benchmarks. In patients undergoing operations at MD Anderson Cancer Center Orlando, procedures were stratified uniformly by acuity and assessed for the quality indicators established by The University of Texas MD Anderson Cancer Center. The results were quantified and compared with calculated benchmarks to identify differences or trends. Although no statistically significant differences in the measures for quality indicators were found, the point estimates for RRs reflected relatively wide CIs owing to the small sample sizes. Nonetheless, clinically relevant differences were discovered for length of stay and for readmission within 30 days after the surgical procedure that suggest targets for process improvement at our center.

An extended length of stay and readmission within 30 days after the procedure imply complications in wound healing or delay in return of function. However, the length of stay and readmission within 30 days of the operation after a high-acuity procedure are far more likely than after a low-acuity procedure to be influenced by factors not directly related to the procedure, such as medical comorbidities, family support system, and resources available for care after discharge. Therefore, length of stay may be increased after a high-acuity procedure, which presents greater challenges to medical compensation of comorbid conditions, the education of family members by nursing staff on support measures in the home, and the qualifications for admission to a rehabilitation or skilled nursing facility. Identifying the cause of extended length of stay and readmission was thus critical to efforts for more efficient use of inpatient resources for our patients.

Thirty-day mortality, even more than length of stay and readmission, implies a major complication of the operation that resulted in unanticipated death. Nonetheless, patient factors and tumor factors that may not be reflected by the acuity of the procedure can be associated with the rare event of death. Although 30-day mortality was not encountered in our study, the risk of death after a high- or low-acuity procedure was compared with established benchmarks that were extrapolated from large numbers of patients from external data sources.

Table 4. Quality Indicators of Cohort and Associated Benchmarks

<table>
<thead>
<tr>
<th>Quality Indicator</th>
<th>Total</th>
<th>Cohort (n = 7)</th>
<th>Benchmark for 7 Patients</th>
<th>P Valueb</th>
<th>RR (95% CI)b</th>
<th>Total</th>
<th>Cohort (n = 12)</th>
<th>Benchmark for 12 Patients</th>
<th>P Valueb</th>
<th>RR (95% CI)b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of stay, d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-3</td>
<td>4</td>
<td>3 (43)</td>
<td>6 (75)</td>
<td>.09</td>
<td>2.01 (0.63-4.92)</td>
<td>1</td>
<td>1 (8)</td>
<td>9 (75)</td>
<td>.64</td>
<td>1.03 (0.41-2.42)</td>
</tr>
<tr>
<td>4-5</td>
<td>2</td>
<td>1 (14)</td>
<td>1 (8)</td>
<td></td>
<td></td>
<td>1</td>
<td>1 (8)</td>
<td>1 (10)</td>
<td>.33</td>
<td>1.31 (0.07-8.24)</td>
</tr>
<tr>
<td>6-8</td>
<td>6</td>
<td>2 (29)</td>
<td>4 (33)</td>
<td></td>
<td></td>
<td>4</td>
<td>3 (25)</td>
<td>3 (25)</td>
<td>.78</td>
<td>1.31 (0.07-8.24)</td>
</tr>
<tr>
<td>9-12</td>
<td>1</td>
<td>0</td>
<td>0 (0)</td>
<td>.33</td>
<td>1.29 (0.66-2.62)</td>
<td>0</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;12</td>
<td>6</td>
<td>1 (14)</td>
<td>5 (42)</td>
<td></td>
<td></td>
<td>5</td>
<td>4 (33)</td>
<td>4 (33)</td>
<td>.33</td>
<td>1.29 (0.66-2.62)</td>
</tr>
<tr>
<td>Readmission ≤30 d after procedure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>3</td>
<td>1 (14)</td>
<td>0 (0)</td>
<td>.33</td>
<td>2.82 (0.01-23.73)</td>
<td>2</td>
<td>1 (17)</td>
<td>1 (&lt;13)</td>
<td>.78</td>
<td>1.31 (0.07-8.24)</td>
</tr>
<tr>
<td>No</td>
<td>16</td>
<td>6 (86)</td>
<td>10 (83)</td>
<td></td>
<td></td>
<td>10</td>
<td>8 (67)</td>
<td>10 (83)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Return to OR ≤7 d after surgery</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>2</td>
<td>0</td>
<td>0 (0)</td>
<td>.64</td>
<td>1.03 (0.41-2.42)</td>
<td>2</td>
<td>1 (17)</td>
<td>1 (&lt;10)</td>
<td>.28</td>
<td>1.73 (0.03-6.72)</td>
</tr>
<tr>
<td>No</td>
<td>17</td>
<td>7 (100)</td>
<td>12 (100)</td>
<td></td>
<td></td>
<td>10</td>
<td>8 (67)</td>
<td>10 (83)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30-d Mortality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>0</td>
<td>0</td>
<td>0 (0)</td>
<td>.50</td>
<td>0.99 (0.99-1.04)</td>
<td>0</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>.50</td>
<td>0.98 (0.98-1.01)</td>
</tr>
<tr>
<td>No</td>
<td>19</td>
<td>7 (100)</td>
<td>12 (100)</td>
<td></td>
<td></td>
<td>10</td>
<td>8 (67)</td>
<td>10 (83)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blood product use, U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>17</td>
<td>7 (100)</td>
<td>10 (83)</td>
<td>.68</td>
<td>0.75 (0.78-1.13)</td>
<td>9</td>
<td>7 (58)</td>
<td>9 (75)</td>
<td>.84</td>
<td>0.92 (0.64-1.42)</td>
</tr>
<tr>
<td>1-3</td>
<td>1</td>
<td>0</td>
<td>1 (8)</td>
<td>.50</td>
<td>0.98 (0.98-1.01)</td>
<td>1</td>
<td>0 (0)</td>
<td>1 (8)</td>
<td>.50</td>
<td>0.98 (0.98-1.01)</td>
</tr>
<tr>
<td>&gt;3</td>
<td>1</td>
<td>0</td>
<td>1 (8)</td>
<td>.50</td>
<td>0.98 (0.98-1.01)</td>
<td>1</td>
<td>0 (0)</td>
<td>1 (8)</td>
<td>.50</td>
<td>0.98 (0.98-1.01)</td>
</tr>
<tr>
<td>Adequacy of pathology reports</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>17</td>
<td>6 (86)</td>
<td>10 (83)</td>
<td>.50</td>
<td>1.17 (0.88-1.13)</td>
<td>11</td>
<td>9 (75)</td>
<td>9 (75)</td>
<td>.78</td>
<td>1.31 (0.07-8.24)</td>
</tr>
<tr>
<td>No</td>
<td>2</td>
<td>1 (14)</td>
<td>1 (8)</td>
<td>.50</td>
<td>1.09 (0.93-1.13)</td>
<td>1</td>
<td>1 (8)</td>
<td>1 (8)</td>
<td>.50</td>
<td>1.09 (0.93-1.13)</td>
</tr>
<tr>
<td>Surgical site free of infection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>18</td>
<td>7 (100)</td>
<td>12 (83)</td>
<td>.50</td>
<td>0.98 (0.98-1.01)</td>
<td>10</td>
<td>8 (67)</td>
<td>12 (98)</td>
<td>.24</td>
<td>1.13 (0.90-1.18)</td>
</tr>
<tr>
<td>No</td>
<td>1</td>
<td>0</td>
<td>2 (17)</td>
<td>.50</td>
<td>1.09 (0.93-1.13)</td>
<td>2</td>
<td>1 (14)</td>
<td>1 (14)</td>
<td>.50</td>
<td>1.09 (0.93-1.13)</td>
</tr>
</tbody>
</table>

Abbreviations: OR, operating room; RR, relative risk.

a All P values determined with the Fisher exact test.
b Relative risks of negative performance in cohort compared with calculated benchmark values.
retrospective studies involving similar surgical procedures. High-volume and low-volume hospitals differ in many aspects of perioperative care, and mechanisms underlying the association of volume to outcome have not been identified. Further efforts to reduce the effects of confounding are most likely to result from uniform stratification of patients with head and neck cancer by comorbidity, as advanced by Hall et al. All patients in our study with more than 1 comorbidity underwent preoperative medical evaluation for optimization of medical conditions and risk stratification, with continuity of medical care in the perioperative period and after hospital discharge.

Use of blood products reflects inherent added morbidity of the primary surgical procedure. A requirement for blood transfusion during or after a low-acuity procedure would seem difficult to justify in most cases; however, blood transfusion in association with a high-acuity procedure may reflect variables aside from the quality of surgery, such as preoperative blood cell count, prior treatment with chemotherapy, threshold for transfusion on medical grounds, and blood coagulation status. Nonetheless, we found use of blood products for high-acuity procedures to be directly related to the duration of surgery (data not shown), a well-described risk factor for perioperative morbidity and extended length of stay.

The adequacy of reports on pathologic findings is a surgical quality indicator that more accurately indicates the role for adjuvant treatment with radiation or chemotherapy than the quality of the operation. However, specific elements of the data set carry implications regarding the quality of surgical intervention and might clarify the need for an additional operation after the primary procedure with the inherent added morbidity, such as a bone margin found to include cancer cells only on final pathology analysis. Although it is not part of the standards of reporting for the College of American Pathologists, intraoperative frozen section perhaps represents the most indicative of factors reflecting the completeness of resection and the efficacy of the surgical procedure. Moreover, accurate pathologic interpretation of intraoperative frozen sections poses a far greater challenge than the completion of a template for pathologic analyses of formalin-fixed specimens. To that end, we have incorporated additional quality indicators for intraoperative decision making into the data set in ongoing evaluation of surgical quality. In addition, following a pathology template standardized for disease and site ensures appropriate documentation of negative as well as positive findings and promotes uniform adjuvant treatment decision making.

Accurate assessment of quality indicators is threatened by several inherent limitations, including selection and recall bias, the retrospective nature of data review, and the completeness of documentation in the medical record. Although many of these obstacles were overcome in our study, great effort was expended in resources for the manpower of data abstraction and the expertise of technical support. Such efforts will unlikely prove efficient for incorporation into ongoing programs of quality assessment. Nonetheless, our current efforts offer promise by way of a structured clinical document as part of the electronic medical record that ensures uniformity in data input at the point of care and allows query of quality indicators in real time. We also have modified data entry templates in Otobase (American Head and Neck Society) to permit prospective entry of the quality indicator data for patients undergoing operations, as well as radiotherapy and chemotherapy.

The sample size of the present study was limited as a result of several factors. With the methodologic design, review of the data points of even 50 patients proved to be a laborious task. For instance, the MD Anderson Physicians Network team included the medical director, a head and neck surgical faculty consultant, an information technology team member, 3 registered nurse reviewers, an account manager, and an administrative assistant. The MD Anderson Cancer Center Orlando team included a tumor registrar, a registered nurse data abstractor, and assistance from the medical records and information technology departments. Total time was estimated at 20.5 hours for administration, 7.0 hours for tumor registry access, 17.5 hours for information technology services, and 248.5 hours for data abstraction and review during a 6-week period. Although exhaustive, the approach met the primary goal of the study by eliminating bias to objectively acquire the data in testing the hypothesis that the established quality indicators and associated benchmarks are reproducible across institutions. Undoubtedly, as performance indicators are validated further, the coming challenge will lie in more efficient means of performing the endeavors of quality programs.

Sample sizes adequately powered to detect all but large differences in observed values relative to calculated benchmarks are unlikely to be feasible at single institutions. For example, to detect a rate of readmission that exceeded the benchmark by 50% (20% in the cohort relative to 13% for the benchmark) with 80% power would require 469 patients undergoing a high-acuity procedure. Even in the highest-volume centers, the time required to accrue such data would limit their interpretation. Nonetheless, in our study, the 95% CIs around the point estimates for RR of negative performance were sufficiently narrow to draw meaningful conclusions for most values (with the exception of readmission ≤ 30 days after a low-acuity procedure). Perhaps even more important than testing statistical significance may be defining the extent of deviation from benchmarks that are clinically relevant to identify trends for targeting improvement in the quality of care.

Another shortcoming in quality assessment is the absence of high-level data associating the quality indicators of our study with the outcomes most important to patients, such as satisfaction with care, quality of life, and disease-free survival. Moreover, the global nature of the quality indicators stratified by acuity may not apply as accurately as indicators specific for subsites, diseases, or procedures, as evidenced by the conclusions of our previous findings in patients with oral tongue squamous cell carcinoma undergoing an operation. Finally, separating the quality of surgical care for patients who undergo multimodality therapy for head and neck cancer from the quality of radiotherapy and chemotherapy may prove to be an even greater challenge. The quality of care...
delivered for patients with head and neck cancer is determined not only by our performance standards related to operations; indeed, improving outcomes in patients will depend on superior coordination, information sharing, and teamwork in the multidisciplinary setting. Effective team building will underpin future efforts to shift to disease management and quality of care by team members united by common culture. In any case, determining the most appropriate quality indicators will most likely depend on the results of studies focusing on compliance with treatment guidelines, as well as assessment of quality indicators, and will be based on prospective, multi-institutional approaches.

Despite the difficulties cited in the assessment of quality indicators, the results of our study suggest that the measures can be assessed as data endpoints in the medical record for patients undergoing operations for head and neck cancer and stratified by acuity. Benchmarks can be established for quality indicators and applied to clinical practices within or among institutions to objectively measure quality of care. The results of comparisons with benchmarks can be used to review clinical practices and identify areas for process improvement. Even among treating physicians and between institutions in which practice patterns and the availability of resources differ, quality of care can be objectively assessed to chart further courses. The resultant changes in practice can raise the satisfaction of patients with their treatment, the quality of their lives during and after cancer treatment, and the expectations for surviving the disease.

Submitted for Publication: March 16, 2011; final revision received May 23, 2011; accepted July 26, 2011.

Correspondence: Thomas D. Shellenberger, DMD, MD, Head and Neck Surgical Oncology, MD Anderson Cancer Center Orlando, 1400 S Orange Ave, MP 760, Orlando, FL 32806 (thomas.shellenberger@orlandohealth.com).

Author Contributions: All authors 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: Shellenberger and Weber. Acquisition of data: Shellenberger, Madero-Visbal, and Weber. Analysis and interpretation of data: Shellenberger, Madero-Visbal, and Weber. Drafting of the manuscript: Shellenberger and Madero-Visbal. Critical revision of the manuscript for important intellectual content: Shellenberger, Madero-Visbal, and Weber. Administrative, technical, and material support: Shellenberger. Study supervision: Shellenberger and Weber.

Financial Disclosure: None reported.

Previous Presentation: This study was presented at the American Head and Neck Society 2011 Annual Meeting; April 28, 2011, Chicago, Illinois.

Additional Contributions: We acknowledge Helmuth Goepfert, MD, and Richard Babaian, MD, for their dedication to this work and their commitment to improving the quality of cancer care, as well as the indispensable contributions of Lori Sturdevant, BA, Dawn Sandelin, BA, Karen Cortez, RN, OCN, Anne Smylie, Rhonda Earls, and William Simeone, MHA, from MD Anderson Physicians Network and Connie Popper, RN, MSN, and Gina McNellis, MA, RHIA, of MD Anderson Cancer Center Orlando.

REFERENCES