

# Validating Risk-Adjusted Surgical Outcomes: Site Visit Assessment of Process and Structure

Jennifer Daley, MD,\* Maureen G. Forbes, RN, MPH,<sup>†</sup> Gary J. Young, JD, PhD,<sup>‡</sup>  
Martin P. Charns, DBA,<sup>§</sup> James O. Gibbs, PhD,<sup>||</sup> Kwan Hur, MS,<sup>||</sup>  
William Henderson, PhD,<sup>||</sup> and Shukri F. Khuri, MD, FACS,<sup>†</sup>  
for the participants in the National VA Surgical Risk Study

**Background:** Risk-adjusted mortality and morbidity rates are often used as measures of the quality of surgical care. This study was conducted to determine the validity of risk-adjusted surgical morbidity and mortality rates as measures of quality of care by assessing the process and structure of care in surgical services with higher-than-expected and lower-than-expected risk-adjusted 30-day mortality and morbidity rates.

**Study Design:** A structural survey of 44 Veterans Affairs Medical Center surgical services and site visits to 20 surgical services with higher-than-expected and lower-than-expected risk-adjusted outcomes were conducted. Main outcome measures included assessment of technology and equipment, technical competence of staff, leadership, relationship with other services, monitoring of quality of care, coordination of work, relationship with affiliated institutions, and overall quality of care.

**Results:** Surgical services with lower-than-expected risk-adjusted surgical morbidity and mortality rates had significantly more equipment available in surgical intensive care units than did services with higher-than-expected outcomes (4.3 versus 2.9,  $p < 0.05$ ). Site-visitor ratings of overall quality of care were significantly higher for surgical services with lower-than-expected morbidity and mortality rates (6.1 versus

4.5 for high outliers,  $p < 0.05$ ); technology and equipment were rated significantly better among low-outlier services (7.1 versus 4.8 for high outliers,  $p < 0.001$ ). Masked site-visit teams correctly predicted the outlier status (high versus low) of 17 of the 20 surgical services visited ( $p < 0.001$ ).

**Conclusions:** Significant differences in several dimensions of process and structure of the delivery of surgical care are associated with differences in risk-adjusted surgical morbidity and mortality rates among 44 Veterans Affairs Medical Centers. (J Am Coll Surg 1997;185:341-351. © 1997 by the American College of Surgeons)

During the last decade, interest has grown in comparing the risk-adjusted outcomes of health care providers as a strategy for assessing quality of care. Proprietary and publicly available risk-adjustment systems have been developed for use at the federal, state, regional, and local levels (1). The validity of risk-adjusted outcomes continues to be controversial. For example, clinicians challenge the face validity of the risk factors included in risk-adjustment models; methodologists and clinicians argue over the content and construct validity of both risk factors and outcomes; and the predictive validity of risk-adjustment models, judged by statistical tests of calibration and discrimination, may vary from poor to moderately good (2-4). The primary assumption in using risk-adjusted outcomes as a measure of quality of care is that they reflect true differences among providers in the processes and structure of care. That is, the providers with the best outcomes will be those with the most effective structures and processes for delivering care. Donabedian (5) noted, however, that the observation of outcomes is an indirect means of assessing the process of care.

This article reports the results of an effort to assess the validity of risk-adjusted surgical out-

Received April 15, 1997; Accepted May 27, 1997.

From the \*Brockton/West Roxbury VA Medical Center, West Roxbury, MA, Harvard Medical School, Boston, MA, Division of General Medicine and Primary Care, Beth Israel Deaconess Medical Center, Boston, MA; <sup>†</sup>Brockton/West Roxbury VA Medical Center, West Roxbury, MA, Harvard Medical School, Boston, MA, Brigham and Women's Hospital, Boston, MA; <sup>‡</sup>Management Decision and Research Center, Department of Veterans Affairs, Boston University, School of Public Health, Boston, MA; <sup>§</sup>Management Decision and Research Center, Department of Veterans Affairs; and <sup>||</sup>Hines VA Center for Cooperative Studies in Health Services, Hines, IL.

The National Veterans Affairs Surgical Risk Study was funded by the Office of Quality Management and the Health Services Research and Development Service of the Department of Veterans Affairs. Dr. Daley is a Senior Research Associate in the Career Development Award Program of the Veterans Affairs Health Services Research and Development Service.

Correspondence address: Jennifer Daley, MD, Veterans Affairs Medical Center, 1400 VFW Parkway, West Roxbury, MA 02132.

comes. These risk-adjusted surgical outcomes were developed as part of the National Veterans Affairs Surgical Risk Study (NVASRS). Begun in 1991, the NVASRS established a system for collecting, analyzing, and reporting patient risk factors and 30-day postoperative outcomes on all major noncardiac surgery performed in the 123 Veterans Affairs Medical Centers (VAMCs) that perform major surgery. The first phase of the NVASRS led to the development of risk-adjusted surgical outcomes for 44 university-affiliated VAMCs that perform cardiac and major noncardiac surgery.

We adopted two complementary approaches to assess the validity of the NVASRS risk-adjusted surgical outcomes. In one approach, trained surgeon reviewers performed process-of-care evaluations based on patient charts sampled from those 44 VAMCs that were either high or low outliers for risk-adjusted postoperative mortality and morbidity. The results of the chart review study will be reported separately. In a second approach, we conducted site visits to the high- and low-outlier VAMCs to assess differences in the structures and processes of care that would be evident only through direct observation. We developed a site-visit protocol, interview and observation guidelines, and a site-visit scoring and reporting mechanism. In both approaches, the examiners were masked to the outlier status of the hospitals. The results of these site visits and their relations to risk-adjusted postoperative outcomes are the subjects of this report.

## Methods

*Development of the risk-adjusted surgical mortality and postoperative morbidity models.* The Department of Veterans Affairs (DVA) Office of Quality Management and Health Services Research and Development Service initiated the NVASRS in response to federal legislation requiring the DVA to report comparative risk-adjusted outcomes data. A detailed discussion of the development and results of the risk-adjustment models and outcomes has been presented elsewhere (6-8). Here we provide a brief overview of the risk-adjustment models.

The surgical services of 44 VAMCs participated in the first phase of the study (October 1, 1991 to December 31, 1993). Each participating center is affiliated with a medical school and is involved in graduate medical education in surgery, surgical subspecialties, and other medical specialties. Surgical cases were eligible for inclusion in the study if the patient received general, spinal, or epidural anesthesia. Surgical procedures with very low ob-

served mortality or postoperative morbidity rates were excluded (eg, vascular shunt revision, diagnostic endoscopy, simple incision and drainage of skin abscesses). Operations were collected in eight surgical subspecialties: general surgery, thoracic surgery, peripheral vascular surgery, neurosurgery, orthopaedics, urology, otolaryngology, and plastic surgery. Cardiac surgery information was collected but was analyzed separately as part of the Continuous Quality Improvement in Cardiac Surgery program based at the Denver VAMC (9-13).

Clinical patient risk factors, intraoperative information, and 30-day postoperative outcomes were collected prospectively on 87,078 surgical procedures over a 27-month period. Postoperative outcomes included both mortality and 21 predefined morbidities occurring in the 30 days after the index operation. Preoperative clinical risk factors (total factors = 67), preoperative laboratory results (total results = 17), and intraoperative variables such as the Common Procedural Terminology (CPT)-4 code (there were 16 total variables) of the principal operative procedure were noted, as were other associated operative procedures and total operative time.

We used the preoperative variables to develop risk-adjustment models for mortality and postoperative morbidity. Logistic regression was used for surgical mortality which was measured as whether the patient died within 30 days of the index surgical procedure. Logistic regression was also used to model postoperative morbidity, defined as the presence of one or more of the 21 predefined morbidities occurring in the 30 days postoperatively. To account for the underlying risk of the surgical procedure independent of patient risk, we used the ratings of the complexity of each of the CPT-4 codes in the data set by panels of Veterans Administration surgical subspecialists. The panelists rated the complexity of each CPT-4 code on a scale from 1 to 5. The surgical panel ratings were highly correlated with the Resource Based Relative Value Scale weighting scales ( $r = 0.76$  for all operations;  $p = 0.0001$ ) (14, 15).

The mortality risk-adjustment model for all operations had a total of 34 patient risk factors and had a c-index of 0.89. Leading predictor variables for all operations included preoperative serum albumin level, American Society of Anesthesia class, emergency scheduling of the operation, the presence of disseminated cancer, age, a blood urea nitrogen  $> 40$  mg/dL, the do-not-resuscitate status of the patient before operation, platelet count  $\leq 150,000$ /mL, weight loss  $> 10\%$  in the 6

months before operation, and an elevated serum glutamic oxaloacetic transaminase level (7).

The postoperative morbidity model for all operations had 18 patient risk factors and a c-index of 0.78. The leading predictor variables for postoperative morbidity were similar, but not identical, to those for the mortality models and included preoperative serum albumin, American Society of Anesthesia class, the complexity score of the principal operation, emergency scheduling of the operation, the patient's functional status before the operation, use of ventilation before operation, wound infection before operation, a history of chronic obstructive pulmonary disease, and a hematocrit  $\leq 38\%$  (8).

The logistic regression models were used to estimate the probability of death or postoperative morbidity for each patient. The expected mortality rate for each VAMC surgical service was calculated by summing the probabilities of death or morbidity for each of its cases. The observed VAMC mortality rate was then compared with the expected VAMC mortality rate in the form of an observed-to-expected ratio (O/E ratio). A similar O/E ratio was calculated for postoperative morbidity. Statistical outliers from the average O/E ratio of 1.0 were calculated using an exact method to compute the confidence interval for a binomial proportion. Confidence intervals that did not include 1.0 indicated high or low outlier hospitals. The O/E ratios were kept in strict confidence for the duration of the study. Of the 44 VAMCs, 13 were statistical outliers for mortality at the 90% confidence interval (6 high and 7 low), and 20 were statistical outliers for postoperative morbidity at a 99% confidence interval (8 high and 12 low).

#### *Site-visit methodology.*

#### **Site-visit rationale and conceptual framework.**

Site visits were used to test the validity of the DVA's risk-adjusted surgical outcomes as a measure of quality of care. We conducted this test by examining the relationship between a VAMC's risk-adjusted surgical outcomes and a site-visit team's ratings of its structures and processes of care. We hypothesized that VAMCs with higher ratings on structures and processes would have lower-than-expected risk-adjusted adverse outcomes; conversely, VAMCs with low ratings of processes and structures would have higher-than-expected risk-adjusted adverse outcomes.

The rationale for the site visits comes from a growing body of theoretic and empiric research indicating that the structural characteristics and organizational processes of health care institutions influence patient outcomes (16–20). Site visits are

a traditional and widely practiced method for assessing structure and processes of care. Most accrediting organizations in the health care industry use site visits as an integral part of their assessment strategy (21–23).

Our site-visit methodology emphasized the expert judgment of site visitors in assessing the structures and processes of surgical services. To ensure consistency across sites, we developed a site-visit protocol and a scoring system in relation to seven conceptual dimensions of surgical-service structure and process: technology and equipment, technical competence of clinical staff (surgeons, anesthesiologists, nurses, and house staff), the interface with other clinical and support services in the hospital, relationships with the affiliated medical school and university hospitals, monitoring of quality of care, communication and coordination of work, and surgical service and hospital clinical leadership. We selected these dimensions from the available literature addressing quality of care in health care institutions. Table 1 provides brief definitions of each of these dimensions.

#### **Selection of surgical services for site visits.**

Twenty of the 44 participating VAMCs were selected for site visits. We selected the five VAMC surgical services with the highest O/E ratios for mortality (range, 1.28–1.53), the five with the lowest mortality O/E ratios (range, 0.49–0.76), the five with the highest postoperative morbidity O/E ratios (range, 1.20–1.46), and the five with the lowest postoperative morbidity O/E ratios (range, 0.49–0.82) for all operations combined. All selected VAMCs were statistical outliers for mortality or postoperative morbidity. One VAMC was a low outlier for mortality and postoperative morbidity. This VAMC was selected for its mortality outlier status, and a sixth hospital was chosen as a low postoperative morbidity outlier.

**Site-visit teams.** Each site-visit team consisted of three people: a chief of surgery from a VAMC among the 44 VAMCs, a VAMC surgical nurse intensivist with clinical and management experience in postoperative intensive care, and a member of the study team. Each site-visit team used the same interview protocols and guides at each visit. All of the site visitors attended an orientation meeting with study investigators, at which time they were briefed about the goals of the study, the interview and site-visit observation protocols, and the scoring instruments. Ten surgeons, five nurses, and four study investigators participated in the site visits. Both the site visitors and the 20 participating VAMCs were aware that outlier hospitals were being visited but were masked to the

Table 1. Dimensions of Quality of Care and Definitions

Dimension	Definition
Technology and equipment	This component refers to the availability and quality of surgical technology and equipment in the operating rooms, anesthesia preparation areas, surgical intensive care units, surgical floors, and ambulatory areas. Consideration should be given to the factors underlying effective or ineffective resource allocation and resource management.
Technical competence of staff	This component refers to the technical skills of surgical, nursing, and anesthesia staff in the provision of care. Consideration is to be given to whether surgical, nursing, and anesthesia staff and house staff have sufficient training, experience, and supervision to carry out their responsibilities on the surgical service.
Interface with other services	This component refers to the surgical service's relationship with other services in the hospital (eg, medical service, laboratory, radiology, medical administration, information resource management) on which it depends for resources and technical support. Consideration is given to factors that promote or impede effective relationships and the implications for patient care.
Relationships with affiliated institutions	This component refers to the relationship between the VA surgical service and the department of surgery and other surgical subspecialties of the academic affiliated institution. Consideration is given to the degree to which the VA surgical service is integrated into the affiliated institution through committee participation and shared physical resources and staff, and the overall mutual cooperation and support that exists between the institutions.
Monitoring of quality of care	This component refers to the processes and activities designed to improve quality of care. Consideration is given to how quality-of-care problems are identified, evaluated, and corrected.
Coordination of work	This component refers to the activities, processes, and systems that coordinate work of the surgical staff. Consideration is given to the use of several coordination mechanisms: staff training, policies and regulations, staff meetings, patient care conferences, and supervision.
Leadership	This component refers to the quality of clinical and management leadership on the surgical service. Consideration is given to whether the leadership of the surgical service (ie, the chief of surgery, chief of anesthesia, and nurse director) is aware of and sensitive to the ideas and concerns of surgical staff and whether the leadership maintains high quality standards for staff.

VA, Veterans Administration.

outlier status of the surgical services at the time of the site visit.

**Site-visit protocol.** The last case was entered into the NVAERS on December 31, 1993, and data collection was completed on February 28, 1994. Data editing, model development, and analysis of interhospital variation were completed in October 1994. Sites were notified of their selection for site visits in November and December 1994. Site visits were conducted between February 1, 1995, and April 30, 1995. Each site visit was 2 days long. Before the site visit, each team reviewed previously collected descriptive data on the VAMC surgical service. The descriptive data came from a survey (Background and Structure Questionnaire) of the chiefs of surgery at all 44 VAMCs that participated in the first phase of the NVAERS. This questionnaire asked about structural characteristics of the surgical services, including the number of surgical beds, surgical volume, academic affiliations and residency programs, staffing of the surgical service (surgeons, anesthesiologists, and nurses), technology in the surgical intensive care unit (SICU) and operating rooms (ORs), and quality-of-care monitoring program.

The site-visit team conducted both individual and group interviews in accordance with a structured interview protocol designed for each individual or group interview. Questions in the interview protocol corresponded to one of the seven

quality-of-care dimensions presented in Table 1. For each VAMC, individual interviews were conducted with each chief of surgery, chief of staff, chief of anesthesiology, chief of medicine, senior nurse responsible for the surgical service, chief surgical resident, and the person responsible for quality assurance on the surgical service. Group interviews were conducted with staff surgeons (vascular and general surgeons, orthopaedists, and urologists); nurse managers responsible for the ORs and the SICUs; nurse managers responsible for each floor caring for surgical patients; and a group of surgeons, anesthesiologists, nurses, and clinical support staff (eg, a ward clerk or secretary) from a surgical floor. The site-visit team also attended "walk rounds" in the SICU and had a tour of the entire surgical service (ie, ORs, recovery room, SICU, surgical floors, and ambulatory surgery areas). The nurse site visitor also spent 1-2 hours observing the physical layout and organization of care in the SICU. In addition, the site-visit team held an exit meeting with the chief of surgery. The site-visit team used the exit meeting to debrief the chief in general about the assessment and to clarify outstanding questions or issues that arose during the 2-day visit.

**Rating and assessment of processes and structures by site-visit team.** The site-visit teams conducted their assessment of the 20 VAMCs in two steps. First, immediately after the completion of a

Table 2. Selected Surgical Service Structural Characteristics for All Hospitals and for High- and Low-Outlier Hospitals\*

Structural characteristic	All hospitals (n = 44)	High-outlier hospitals (n = 10)	Low-outlier hospitals (n = 10)	p value <sup>†</sup>
Surgical volume				
Mean number of operations annually	2,804 ± 952 (1,132–5,000)	2,681 ± 1,060	2,494 ± 649	0.64
Surgical residency program				
Mean residents/month on all services	24.2 ± 7.4 (10–50)	22.8 ± 4.6	22.8 ± 4.6	1
Surgeon staffing				
Mean number full-time surgeons	3.57 ± 3.9 (0–18)	2.6 ± 2.1	4.1 ± 3.2	0.24
Mean number part-time surgeons	26.4 ± 11.1 (1–51)	22.6 ± 13.4	28.2 ± 15.5	0.4
Use of allied health advanced practitioners				
Programs that use such practitioners (%)	70	60	80	0.2
Availability of technology				
Mean SICU technology score (possible range 0–8)	4.5 ± 1.8	2.9 ± 1.3	4.3 ± 1.3	0.02
Mean OR technology score (possible range 0–10)	7.5 ± 2.1	6.9 ± 2.6	7.2 ± 2.5	0.79

\*Data are mean ± standard deviation (range), except as noted.

<sup>†</sup>Statistical comparisons are between high- and low-outlier hospitals.  
SICU, surgical intensive care unit; OR, operating room.

site visit, each site-visit team member independently rated the VAMC surgical service from 1 (poor) to 9 (excellent) for each of the seven dimensions and for overall quality of care. In addition, each site-visit team member was asked to predict the outlier status of the VAMC (ie, higher-than-expected or lower-than-expected risk-adjusted mortality and/or postoperative morbidity). Second, after the members of a site-visit team independently rated the VAMC, they met and discussed their ratings and determined a team rating for each dimension and for overall quality of care. They also made a team prediction of the outlier status of the surgical service. In addition, each site-visit team member submitted a written justification for his or her ratings as well as a summary of the strengths and areas for improvement for each VAMC. The team members were also asked to describe innovative programs, policies, or procedures they had identified during the site visits and obstacles to providing high-quality care.

*Analysis of site-visit data.* We analyzed the data from the site visits in three ways. First, we compared the site-visit team ratings of the 10 VAMCs with lower-than-expected outcomes with those of the 10 VAMCs with higher-than-expected outcomes. Differences between the team ratings on each of the seven dimensions of quality of care and overall quality of care for high- and low-outlier hospitals were tested for statistical significance using unpaired Student's *t*-tests. Second, we compared the team prediction of the outlier status for each of the 20 surgical services with the outlier status based on the risk-adjusted outcomes. Percent exact agreement and kappa statistics were

calculated. Third, we analyzed the written site-visit reports to assess qualitative differences between the high- and low-outlier VAMCs regarding the seven quality-of-care dimensions used in the site visits.

## Results

*Descriptive characteristics of the surgical services.* Table 2 presents selected structural characteristics for the 10 high- and the 10 low-outlier surgical services and for all 44 VAMCs in the NVASRS. Although a considerable range existed in the number of surgical procedures performed annually among the 44 VAMCs (1,132–5,000 operations per year), the mean number of operations for the high and low outliers was similar. In addition, there were no statistically significant differences between the high and low outliers with respect to the mean number of surgical residents rotating on each surgical service monthly, the mean number of full- and part-time surgeons, and the use of allied health advanced practitioners (eg, nurse practitioners and physician assistants).

Table 2 also presents information comparing the high and low outliers regarding the availability of technology in the SICUs and ORs. Each of the 44 VAMCs received a technology-availability score based on how many of 8 possible items were available in the SICU<sup>1</sup> and how many of 10 possible

<sup>1</sup> Two-channel electrocardiograms, three pressure channel modes for pulmonary artery wave measurement, gastric pH meters, continuous mixed venous oxygen saturation measurement equipment, end-tidal carbon dioxide measurement, automated data-management systems (eg, Careview, Hewlett Packard, CA), in-unit blood gas and blood electrolyte measurement.

Table 3. Site-Visitor Team Ratings of Dimensions of Quality of Care by High- and Low-Outlier Hospitals\*

Dimension	Low-outlier hospitals (n = 10)	High-outlier hospitals (n = 10)	p Value
Technology and equipment	7.1 ± 1.1	4.8 ± 1.5	0.001
Technical competence	6.8 ± 1.0	6.0 ± 0.9	0.08
Interface with other services	5.4 ± 1.8	4.5 ± 1.2	0.2
Relationship with affiliated institutions	6.3 ± 1.1	5.1 ± 1.7	0.08
Monitoring of quality of care	5.4 ± 1.4	4.4 ± 1.6	0.1
Coordination of work	5.5 ± 1.8	4.6 ± 1.2	0.2
Leadership	6.0 ± 2.2	5.0 ± 1.2	0.2
Overall quality of care	6.1 ± 1.9	4.6 ± 1.0	0.04

\*Ratings of dimensions of quality of care ranged from 1 = poor to 9 = excellent. Data are mean ± standard deviation.

items were available in the OR.<sup>2</sup> Although the high- and low-outlier surgical services appeared to be similar in the availability of OR equipment, the low outliers did have significantly more equipment in the SICU than did the high outliers (4.3 versus 2.9,  $p < 0.05$ ).

*Site-visitor ratings for the dimensions of quality of care.* The comparison of the team ratings for the high- and low-outlier surgical services produced results that were largely as expected (Table 3). The team rating of overall quality of care was 6.1 for the low-outlier surgical services, compared with 4.6 for the high outliers ( $p < 0.05$ ). Low outliers also received significantly higher summary team ratings than the high outliers for the technology and equipment dimension (7.1 versus 4.8,  $p < 0.001$ ). The team ratings on the other six dimensions (technical competence, interface with other services, relationship with affiliated institutions, monitoring of quality of care, coordination of work, and leadership) were consistently higher for low-outlier hospitals, although the differences reached only borderline statistical significance.

*Site-visitor prediction of outlier status of surgical services.* The site visitors correctly identified the outlier status of 17 of the 20 VAMCs (Table 4). Percent exact agreement was 85% ( $p < 0.001$ ); the kappa statistic was 0.7. They correctly identified nine of the mortality-outlier VAMCs and eight of the postoperative morbidity outliers. They incorrectly identified one low-mortality outlier as a high outlier. For the two other VAMCs, one site-visit team incorrectly identified a high-morbidity outlier as a low outlier, and another team incorrectly identified a low-morbidity outlier as a high outlier. The site visitors were not able to identify whether

the VAMCs were outliers for mortality as opposed to postoperative morbidity.

*Qualitative content analysis of the site-visit reports.* The qualitative content analysis of the site visitors' written reports provided detailed information about differences between the high- and low-outlier surgical services for each of the seven quality-of-care dimensions. Characteristics of low-outlier hospitals were noted if they were observed in six or more of the site-visit reports and were noted in three or fewer of the high-outlier hospitals. Concerning technology and equipment, the site visitors noted that the physical layout of the low-outlier hospitals (eg, space between beds in the SICU, square footage in the OR), particularly in the SICU, tended to be more spacious and less constricted than in the high outliers. In the area of technical competence, the site visitors commented on the higher use of per diem nursing staff in the SICU and a lower skill level among nursing staff in high outliers, in part because of the loss of clinical nurse specialists from layoffs and cutbacks. With respect to the relationship between the VAMC surgical service and the affiliated university surgery program, the site visitors identified two distinguishing characteristics of low outliers: a high level of integration of faculty (eg, salary structures, academic conferences, faculty sharing of resources) and short distances between the university hospital and the VAMC.

Table 4. Site-Visitor Team Judgment of Outlier Status by Risk-Adjusted Outcomes Model Outlier Status

Risk-adjusted outcomes model	Site-visit team judgment		Total
	High	Low	
High	9	1	10
Low	2	8	10
Total	11	9	20

Percent exact agreement = 85% ( $p < 0.001$ ); kappa = 0.7.

<sup>2</sup> Two-channel electrocardiograms, four channel modes for pulmonary artery waveforms, cell saver, intraoperative ultrasonography, transesophageal echography, left ventricular assist devices, capnometers, oximetric catheters, pulse oximetry, and DragerNarkomed 4 anesthesia equipment.

In the dimension concerned with the interface between the surgical service and other clinical departments and support functions in the VAMC (eg, internal medicine, cardiology, radiology, clinical laboratory, pharmacy, social service, blood bank), the site visitors noted that the low outliers had strong relationships with internal medicine and cardiology; these relationships appeared to contribute to the timeliness of cardiology consultations. Among the low outliers, they also noted the high quality of personal and professional relationships between the chiefs of surgery and internal medicine; the two chiefs met regularly to discuss issues of mutual concern and interest and showed mutual respect and trust. In terms of monitoring of quality of care, although each VAMC had a quality-assurance program in place, the site visitors reported that the low outliers were far more effective than the high outliers in collecting and monitoring performance indicators and were more likely to use the information to develop quality-improvement initiatives within the service on a regular basis. Moreover, the site visitors observed that the personnel in the department responsible for quality monitoring and improvement had a close working relationship with the chief of surgery. This was characterized by the proximity of their offices, frequently scheduled meetings, and high informal access to the chief for communication and problem solving.

In the area of communication and coordination of work, the site visitors noted several differences between the high- and low-outlier VAMCs. First, more interaction between surgical and nursing staff was observed at both the administrative and patient care levels in the low-outlier hospitals. For example, at all the low-outlier surgical services, the chief of surgery met regularly with the nursing staff responsible for management of the OR, SICU, and the surgical floors. In addition, the surgeons and nurses were more likely to round together on the patient floors at the low outliers than at the high outliers. Second, the relationship between anesthesia and surgery staff was substantially better at the low-outlier surgical services. In particular, the chiefs of surgery and of anesthesia were more likely to meet regularly and work together with nursing leadership to address issues about OR scheduling. Third, the low-outlier surgical services were more progressive in the development and implementation of clinical pathways and practice guidelines. Fourth, clinical supervision of the surgical residents by attending surgeons was better at the low-outlier hospitals, particularly in the postoperative period.

In terms of leadership, the site-visitor comments suggested that the chief of surgery was more involved in the day-to-day management of the surgical service at low outliers than at high outliers. For example, at low-outlier hospitals, the chief of surgery held regular staff meetings that included surgical attending staff, house staff, anesthesia personnel, and representatives of nursing from all the surgical units (ambulatory, OR, SICU, and surgical floors).

### Discussion

This study confirmed our hypothesis of an association between the risk-adjusted adverse outcomes of surgical mortality and postoperative morbidity in 20 VAMC surgical services and the implicit judgment of peer health professionals about the quality of care (processes and structures) of those surgical services. Site-visit teams consisting of a chief of surgery, a nurse with expertise in the management and delivery of surgical care, and a study investigator were able to identify those surgical services with better-than-expected risk-adjusted outcomes (ie, low outliers) and those with worse-than-expected risk-adjusted outcomes (ie, high outliers) with 85% exact agreement ( $\kappa = 0.7$ ). These site-visit teams also rated each of the 20 VAMCs on seven dimensions of quality of care. Although detection of statistically significant differences between the high and low outliers on these ratings was limited by the small number of sites visited (20), the ratings were consistently higher for the low outliers for each of the seven dimensions and reached statistical significance at the 0.05 level for overall quality of care and for technology and equipment. Additional qualitative analysis of site-visitor reports on the 20 VAMCs revealed a number of differences in the structures and processes of surgical care between the high and low outliers.

Previous investigations of the association between risk-adjusted adverse outcomes and processes and structures of care have yielded conflicting evidence. Research using large administrative data bases has focused on the relations between surgical mortality and surgical volume (24–26) and other structural features of hospitals and surgical services, such as the organization of the medical staff and the ownership of hospitals (27–29). Reviews of the process of care using both explicit and implicit medical record audits have yielded conflicting results. In a study of medical patients with stroke, pneumonia, and acute myocardial infarction from hospitals with higher- and lower-

than-expected mortality rates, Dubois and associates (30) found no differences in quality of care on explicit chart review, but using implicit chart review, they found higher rates of preventable deaths among pneumonia patients (but not for those with acute myocardial infarction and stroke) in the high-outlier facilities. Kahn and associates (31) and Rubenstein and colleagues (32) found significantly higher risk-adjusted mortality for patients having poor quality of care, as judged by chart review among Medicare beneficiaries, as part of their assessment of the impact of the implementation of prospective payment. This finding was apparent for the four medical conditions assessed, but not for the surgical condition studied (hip fracture).

The associations between risk-adjusted mortality rates and processes and structures of care in surgery have been explored almost exclusively in cardiac surgery. In a study of cardiac surgery cases in New York State, Hannan and colleagues (33) identified high- and low-outlier hospitals among the hospitals performing cardiac surgery. Chart reviews of patients undergoing coronary artery bypass grafting in the seven outlier hospitals found quality-of-care problems significantly more often among the high-outlier hospitals (33). Thomas and coworkers (34) compared the validity of judgments of quality of care by peer-review organization reviewers with risk-adjusted mortality rates derived from a discharge abstract data base and confirmed a significant association between the two for angina and cardiac surgery but only a weak association for acute myocardial infarction and no association for septicemia.

This study is consistent with a small number of previous efforts to validate risk-adjusted outcomes using a site-visit methodology. Knaus and associates (19) sought to validate risk-adjusted mortality rates for a sample of intensive care units by examining the unit structures and processes. They found that intensive care units with lower mortality rates were more likely to have a full-time medical director, tended to be located in a hospital with a comprehensive nurse continuing education program, and typically had a high level of coordination among clinical staff (19). Another study was conducted among five hospitals with cardiac surgery programs in northern New England with a 2.5-fold variation in risk-adjusted mortality rates. Site visits to these hospitals revealed considerable variability in several aspects of the process of care, including technical process of care, process ownership and leadership, communication among and between participants, decision making, recording

and use of data, and underlying factors such as training levels, staff fatigue, and environmental characteristics (35). These investigators have reported a significant reduction in cardiac surgical mortality associated with quality-improvement efforts, based partially on observations made during these site visits (36).

The current study joins a small number of studies that have included adverse surgical outcomes in addition to postoperative mortality (37-41). Major postoperative morbidities such as those recorded in the NVASRS and their relationship to poor process and structure of care are the basis of many clinical quality-assurance, quality-improvement, and infection-control programs in hospitals and of clinical exercises such as surgical morbidity and mortality conferences ("M&Ms"). Yet relatively little is known about the relationship between the structure and process of care and their impact on surgical morbidity other than for surgical wound infections (42). Although the site-visit teams in our investigation did not attempt to differentiate among the observed structures and processes and their relative influences on mortality versus postoperative morbidity, future research may be able to disentangle such causal relations.

Our study has several limitations. First, we were limited by resource constraints to visiting fewer than half of the surgical services participating in the NVASRS. This small sample size limits the statistical power of the study. Second, to maximize the variation in the risk-adjusted outcomes, we did not visit surgical services with O/E ratios of 1.0, in the middle of the distribution of outcomes. As a result, the site visitors may have had an easier task determining the outlier status of the surgical services they visited; the results of our study, therefore, cannot be used to confirm the validity of site visits as a measure of quality of care. Third, the site visits were conducted 1 year after the last patients were entered into the study. Site visitors were careful to assess differences in process and structure between the time of data collection and the site visits and reported few differences.

Postoperative mortality and morbidity rates are only part of the spectrum of clinically relevant outcomes for patients undergoing operations. We have not included in our assessment of these surgical services other measures of outcomes, such as changes in health status for patients (eg, disease-specific or generic measures of health status); patient perceptions of the nontechnical aspects of their care, or the efficient use of resources in providing care (eg, utilization of hospital resources and costs). The DVA is developing such a compre-

hensive system of performance measurement, which is being incorporated into the comparative assessment of all 123 VAMCs performing surgery. Further, the generalizability of this study may be limited. The patient population in the DVA is predominantly middle-aged and elderly men. The attending surgeons and surgical residents who operate in these 44 VAMCs also operate in academic university medical centers.

Given the results of this site-visit study, we are now more confident that differences exist in how surgical care is provided and structured between surgical services with lower-than-expected and higher-than-expected postoperative adverse events. The intent of the NVAERS was to improve the quality of surgical care provided to all veterans undergoing operations, rather than to identify a group of surgical services with higher-than-expected surgical morbidity and mortality rates as “bad apples” (43). How can we use the information from this study to improve the practice of surgery? The site-visit teams were often met with skepticism and defensiveness by the clinical staff of the surgical services. Many surgeons, nurses, and anesthesiologists with whom we met were skeptical that risk-adjusted surgical mortality and postoperative morbidity rates are indicative of the quality of the care that they provide. Given the results of the site visits, we can provide evidence that the risk-adjusted outcomes are associated with differences in those structures and processes of surgical services that influence quality of care. In the spirit of quality improvement, we hope that the “best practices” of those surgical services with better-than-expected outcomes will serve as the source of innovation and possible process redesign for other surgical services in and outside of the DVA.

### Acknowledgments

The authors would like to thank the site visitors who gave generously of their time and expertise to visit the surgical services of the 20 VAMCs: Drs. John Stremple, Alex Solowey, Robert Nesbit, Jr, John Gleysteen, Nathan Couch, Walter Merrill, Theodore Pappas, Willard Johnson, Raymond Reed, and Edward Passaro, Jr; and Anne George, RN, Alice Perry, RN, Geraldine Feaster, John Tatarakis, RN, and Mary Beth Harrington, RN. The authors also wish to thank the surgeons, anesthesiologists, nurses, and administrative staff at each of the VAMCs who shared their experience and insights with the site-visit teams, and, in particular, the chiefs of surgery who permitted us to visit their services.

## APPENDIX

### Study Participants

Chairperson's Office: Shukri F. Khuri, MD (chairperson); Jennifer Daley, MD (co-chairperson); Maureen Forbes, MS, RN, MPH (health services researcher); Lynn-Marie Herlihy, BA (program assistant); Candace Savage (national administrative coordinator); and Jeannette Spencer, MS, RN (national clinical coordinator). *Hines Center for Cooperative Studies in Health Services*: John Demakis, MD, and William Henderson, PhD (codirectors); James Gibbs, PhD (health services researcher); Kwan Hur, MS (biostatistician); Bharat Thakkar (statistical programmer); and Sharon Urbanski (statistical assistant). *Traveling nurse coordinators*: Nancy Deegan, MSN, RN, San Antonio, TX; Jeannette Spencer, MS, RN, Brockton/West Roxbury, MA; and Debra Wilcox, MSN, RN, Denver, CO. *VA central office*: Galen Barbour, MD (Associate Chief Medical Director for Quality Management); Scott Beck, ME, CPQA (Special Assistant to AsCMD for Quality Management); Daniel Deykin, MD (Chief, Cooperative Studies Program and Director, Health Services R&D); and Gerald McDonald, MD (Deputy Chief of Surgery). *Executive Committee*: Shukri F. Khuri, MD (chairperson), Brockton/West Roxbury, MA; Jennifer Daley, MD (co-chairperson), Brockton/West Roxbury, MA; J. Bradley Aust, MD, San Antonio, TX; Scott Beck, ME, CPQA, Washington, DC; John Demakis, MD, Hines, IL; Peter J. Fabri, MD, Tampa, FL; James Gibbs, PhD, Hines, IL; Frederick Grover, MD, Denver, CO; Karl Hammermeister, MD, Denver, CO; William Henderson, PhD, Hines, IL; George L. Irvin III, MD, Miami, FL; Gerald McDonald, MD, Washington, DC; Edward Passaro, Jr, MD, West Los Angeles, CA; Jeannette Spencer, MS, RN, Brockton/West Roxbury, MA; and John Stremple, MD, Pittsburgh, PA. *Expert Advisory Committee*: Barbara McNeil, MD, PhD (chairperson), Boston, MA; J. Bradley Aust, MD, San Antonio, TX; Paul Ebert, MD, Chicago, IL; Frank Harrell, PhD, Durham, NC; Lisa Iezzoni, MD, Boston, MA; John Mannick, MD, Boston, MA; L. Richard Smith, PhD, Durham, NC; and J. William Thomas, PhD, Ann Arbor, MI. *Birmingham Information Systems Center*: Alan Monosky (computer specialist); Michael Montali (senior systems analyst); and Steve Musgrove (computer specialist). *St. Louis Continuing Education Center*: Donna Schoonover, RNCS, MSN (Program Director). *Durham Regional Medical Education Center*: Linda Exner, MS, RN (Chief, Special Projects).

### Participating VA Medical Centers, Chiefs of Surgery, and Surgical Clinical Nurse Reviewers

Albuquerque, NM: Walter Dickinson, MD; Donald Morris, MD; Tony Lantzer, RN. Ann Arbor, MI: Daniel B. Hinshaw, MD; Linda Brooks, RN. Asheville, NC: Stewart M. Scott, MD; Valerie Allen, RN; Cindy McCutcheon, RN; Kitty Williams, RN. Atlanta, GA: Maurice J. Jurkiewicz, MD; Rita Walker, RN; Susan Zarter, RN. Augusta, GA: Robert R. Nesbit, Jr, MD; Connie Miller, RN. Birmingham, AL: John J. Gleysteen, MD; Linda Helm-Little, RN. Brockton/West Roxbury, MA: Shukri F. Khuri, MD; Jeannette Spencer, MS, RN. Buffalo, NY: William Lynn Weaver, MD; Leon Levinsky, MD; Natalie McIntosh, MS, RN; Ruth Strebe, MA, RN; Delores Young, MS, RN. Charleston, SC: John Allison, MD; Stephen Johnston, MSN, RN. Chicago, IL: Donald K. Wood, MD; Paz Encomienda, RN. Cleveland, OH: Jerrel Benson, MD; Linda Graham, MD; Debra Graham, MD (PI); Douglas Nemecek, MD (former PI); Mary Ann E. Bobulsky, BSN, RN. Columbia, MO: Jerry Templer, MD; Loren Humphrey, MD; Barbara von Thun, RN; Paula Baker, RN. Dallas, TX: Stuart Myers, MD; Jan Dodd-Jolly, RN; Bernice Willis, RN. Denver, CO: Frederick Grover, MD; Marilyn Villaverde, BSN, RN. Durham, NC: Theodore Pappas, MD; Walter G. Wolfe, MD; Joseph A. Moylan, MD; Jean Hanchey, MSN, RN; Beth Radecke, RN. Gainesville, FL: Timothy Flynn, MD; Linda Carter, RN. Hines, IL: George Mason, MD; Pamela Scott. Houston, TX: James W. Jones, MD; Dorothy Forte, BSN, RN. Indianapolis, IN: Robert Lempke, MD; Priscilla Mullins,

CPHQ, RN; Kathleen Galloway, BSN, RN, CPQA. Lexington, KY: Andris Kazmers, MD; Rose Mary Collins, RN. Little Rock, AR: Raymond C. Read, MD; Richard D. Bloesch, RN. Madison, WI: James R. Starling, MD; Lori Hagen, RN. Memphis, TN: Eugene C. Mangiante, MD; Anita Garrison, BSN, RN, CCRN. Miami, FL: George L. Irvin III, MD; Patricia Lee, RN. Milwaukee, WI: William Schulte, MD; Christine Tyler, RN. Minneapolis, MN: Don McQuarrie, MD; Edward W. Humphrey, MD; Ellie Carson, RN; Carolyn Robinson, RN. Nashville, TN: Walter Merrill, MD; Achilles Demetriou, MD; Jeanette Pujol, RN. New Orleans, LA: Paul Hastings, MD; Adrienne Thymes-Leflore, RN; Kathleen Wright, RN. New York, NY: Alex C. Solowey, MD; Jacqueline Parker, RN. Oklahoma City, OK: Donald R. Carter, MD; Rouchelle Osborn, MS, RN; Laura Hall, RN, MPH. Palo Alto, CA: Shoichi Kohatsu, MD; Anne Grillot, MSN, RN. Pittsburgh, PA: John F. Stremple, MD; Anita Steinmiller, MSN, RN, CPQA. Portland, OR: Clifford Deveney, MD; Michael Montana, MPA, RN. Richmond, VA: Hunter McGuire, Jr, MD; Susan Clark, MSN, RN, CCRN. Salt Lake City, UT: Mark W. Wolcott, MD; Ida Marie Bickle, MS, RN. San Antonio, TX: Waid Rogers, MD; Linda Porazzi, RN, MSN, CNS; Patricia Grota, RN. San Diego, CA: Bruce Stabile, MD; Gail Maxwell, RN. San Francisco, CA: Lawrence W. Way, MD; Gloria Mason, RNP, BSN; Rita J. Sears, RN, MS. Seattle, WA: Ted Kohler, MD; H. M. Radke, MD; Ann Neil, RN; Nancy Ambur, RN. Tampa, FL: Peter J. Fabri, MD; Katherine Bowns, RN. Tucson, AZ: Charles Zukoski, MD; Christopher Brown, RN; Patricia Hogan, RN. Washington, DC: John Harmon, MD; Deborah Fleming, BSN, RN; Mary Ann King, BSN, RN. West Haven, CT: Barbara Kinder, MD; Kathleen Maher-Clearly, BSN, RN; Carol Sapienza, MSN, RN; ReBecca Salas, RN. West Los Angeles, CA: Edward Passaro, Jr, MD; Marilyn DeGroot, RN; Sheila Cooley, MS, RN.

## References

- Iezzoni LI. Risk adjustment and current health policy debates. In: Iezzoni, LI, ed. *Risk Adjustment for Measuring Health Care Outcomes*. Second edition. Ann Arbor, MI: Health Administration Press; 1997:517-595.
- Blumberg MS. Risk-adjusting health care outcomes: a methodologic review. *Med Care Rev* 1986;43:351-393.
- Daley J. Validity of risk adjustment methods. In: Iezzoni LI, ed. *Risk Adjustment for Measuring Health Care Outcomes*. Second edition. Ann Arbor, MI: Health Administration Press; 1997:331-364.
- Ash AS, and Shwartz M. Evaluating the performance of risk-adjustment methods: dichotomous measures. In: Iezzoni LI, ed. *Risk Adjustment for Measuring Health Care Outcomes*. Second edition. Ann Arbor, MI: Health Administration Press; 1997:427-470.
- Donabedian A. *Explorations in Quality Assessment and Monitoring*. Vol. 1. The Definition of Quality and Approaches to its Assessment. Ann Arbor, MI: Health Administration Press; 1980.
- Khuri SK, Daley J, Henderson WH, et al. The National Veterans Administration Surgical Risk Study: adjustment for the comparative assessment of the quality of surgical care. *J Am Coll Surg* 1995;180:519-531.
- Khuri SK, Daley J, Henderson WH, et al. Comparative assessment of thirty-day mortality following major surgery: results of the National VA Surgical Risk Study. *J Am Coll Surg* 1997;185:315-327.
- Daley J, Khuri SK, Henderson WH, et al. Comparative assessment of thirty-day morbidity following major surgery: results of the National VA Surgical Risk Study. *J Am Coll Surg* 1997;185:328-340.
- Grover FL, Hammermeister KE, and Burchfiel C. Initial report of the Veterans Administration Preoperative Risk Assessment Study for Cardiac Surgery. *Ann Thorac Surg* 1990;50:12-28.
- Hammermeister KE, Burchfiel C, Johnson R, and Grover FL, for the VA surgeons and cardiologists. Identification of patients at greatest risk for developing major complications at cardiac surgery. *Circulation* 1990;82(Suppl 4):380-389.
- Grover FL, Johnson RR, Hammermeister KE, and the VA Cardiac Surgeons. Factors predictive of operative mortality among coronary artery bypass subsets. *Ann Thorac Surg* 1993;56:1296-1307.
- Grover FL, Johnson RR, Shroyer AL, et al. The Veterans Affairs Continuous Improvement in Cardiac Surgery Study. *Ann Thorac Surg* 1994;58:1845-1851.
- Hammermeister KE, Johnson R, Marshall G, and Grover FL. Continuous assessment and improvement in quality of care. A model from the Department of Veterans Affairs. *Ann Surg* 1994;219:281-290.
- Hsiao WC, Couch NT, Causino N, et al. Resource-based relative values for invasive procedures performed by eight surgical specialties. *JAMA* 1988;260:2418-2424.
- Hsiao WC, Braun P, Yntema D, and Becker ER. Estimating physicians' work for resource-based relative value scale. *N Engl J Med* 1988;319:835-841.
- Flood AB, and Scott WR. *Hospital Structure and Performance*. Baltimore: Johns Hopkins, 1987.
- Shortell SM, Zimmerman JE, Rousseau DM, et al. The performance of intensive care units: does good management make a difference? *Med Care* 1994;32:508-525.
- Shortell SM, and Logerfo JP. Hospital medical staff organization and quality of care: results for myocardial infarction and appendectomy. *Med Care* 1981;19:1041-1055.
- Knaus WA, Draper EA, Wagner DP, et al. An evaluation of outcome from intensive care units in major medical centers. *Ann Intern Med* 1986;104:410-418.
- Baggs JG, Ryan SA, Phelps CE, et al. The association between interdisciplinary collaboration and patient outcomes in a medical intensive care unit. *Heart Lung* 1992;21:18-24.
- The Joint Commission on Accreditation of Healthcare Organizations. *1996 Comprehensive Accreditation Manual for Hospitals*. Oakbrook Terrace, IL: JCAHO; 1995.
- Jennings K, and Westfall F. A survey-based benchmarking approach for health care using the Baldrige quality criteria. *Jt Comm J Qual Improv* 1994;20:500-509.
- Hertz HS, Reimann CW, and Bostwick MC. The Malcolm Baldrige National Quality Award concept: could it help stimulate or accelerate health care quality improvement? *Qual Manage Health Care* 1994;2:63-72.
- Luft HS. The relation between surgical volume and mortality: an exploration of causal factors and alternative models. *Med Care* 1980;18:940-959.
- Riley G, and Lubitz J. Outcomes of surgery among the Medicare aged: surgical volume and mortality. *HCFA Review* 1988; 7:37-47.
- Sloan FA, Perrin JM, and Valvona J. In-hospital mortality of surgical patients: is there an empiric basis for standard setting? *Surgery* 1986;99:446-454.
- Kelly JV, and Hellinger FJ. Physician and hospital factors associated with mortality of surgical patients. *Med Care* 1986; 24:785-800.
- Shortell SM, and Hughes EFX. The effects of regulation, competition, and ownership on mortality rates among hospital inpatients. *N Engl J Med* 1988;318:1100-1107.
- Hartz AJ, Krakauer H, Kuhnen, et al. Hospital characteristics and mortality rates. *N Engl J Med* 1989;321:1720-1725.
- Dubois RW, Rogers WH, Moxley JH, et al. Hospital inpatient mortality: is it a predictor of quality? *N Engl J Med* 1987;317: 1674-1680.
- Kahn KL, Rogers WH, Rubenstein LV, et al. Measuring quality of care with explicit process criteria before and after implementation of the DRG-based prospective payment system. *JAMA* 1990;264:1969-1973.
- Rubenstein LV, Kahn KL, Reinish EJ, et al. Changes in quality of care for five diseases using implicit review, 1981 to 1986. *JAMA* 1990;264:1974-1979.
- Hannan EL, Kilburn H, O'Donnell JF, et al. Adult open heart surgery in New York State: an analysis of risk factors and hospital mortality rates. *JAMA* 1990;264:2768-2774.

34. Thomas JW, Holloway JJ, and Guire KE. Validating risk-adjusted mortality as an indicator of quality of care. *Inquiry* 1993;30:6-22.
35. Kasper JF, Plume SK, and O'Connor GT. A methodology for QI in the coronary artery bypass grafting procedure involving comparative process analysis. *QRB Qual Rev Bull* 1992;18:129-133.
36. O'Connor GT, Plume SK, Olmstead EM, et al. A regional intervention to improve hospital mortality associated with coronary artery bypass graft surgery. *JAMA* 1996;275:841-846.
37. Roos LL, Cageorge SM, Roos NP, et al. Centralization, certification, and monitoring. Readmissions and complications after surgery. *Med Care* 1986;24:1044-1066.
38. Desharnais SL, McMahon LF, Wroblewski RT, et al. Measuring outcomes of hospital care using multiple risk-adjusted indexes. *Med Care* 1990;28:1127-1141.
39. Rosen AK, Geraci JM, Ash AS, et al. Postoperative adverse events of common surgical procedures in the Medicare population. *Med Care* 1992;30:753-765.
40. Iezzoni LI, Daley J, Heeren T, et al. Using administrative data to screen hospitals for high complication rates. *Inquiry* 1994;31:40-55.
41. Riley G, Lubitz J, Gornick M, et al. Medicare beneficiaries: adverse outcomes after hospitalization for eight procedures. *Med Care* 1993;31:921-949.
42. Sawyer RG, and Pruett TL. Wound infections. *Surg Clin North Am* 1994;74:519-536.
43. Berwick DN. Continuous improvement as an ideal in health care. *N Engl J Med* 1989;320:53-56.