The Comparative Assessment and Improvement of Quality of Surgical Care in the Department of Veterans Affairs

Shukri F. Khuri, MD; Jennifer Daley, MD; William G. Henderson, PhD

Prompted by the need to assess comparatively the quality of surgical care in 133 Veterans Affairs (VA) hospitals, the Department of Veterans Affairs conducted the National VA Surgical Risk Study between October 1, 1991, and December 31, 1993, in 44 VA medical centers. The study developed and validated models for risk adjustment of 30-day morbidity and 30-day mortality after major surgery in 8 noncardiac surgical specialties. Similar models were developed for cardiac surgery by the VA’s Continuous Improvement in Cardiac Surgery Program. Based on the results of the National VA Surgical Risk Study and the Continuous Improvement in Cardiac Surgery Program, the VA established in 1994 a VA National Surgical Quality Improvement Program (NSQIP), in which all the medical centers performing major surgery participated. An NSQIP nurse at each center oversees the prospective collection of data and their electronic transmission for analysis at 1 of 2 data coordinating centers. Feedback to the providers and managers is aimed at achieving continuous quality improvement. It consists of (1) comparative, site-specific, and outcome-based annual reports; (2) periodic assessment of performance; (3) self-assessment tools; (4) structured site visits; and (5) dissemination of best practices. The NSQIP also provides an infrastructure to enable the VA investigators to query the database and produce scientific presentations and publications. Since the inception of the NSQIP data collection process, the 30-day postoperative mortality after major surgery in the VA has decreased by 27%, and the 30-day morbidity by 45%. The future of the NSQIP lies in expanding it to the private sector and in enhancing its capabilities by incorporating additional measures of outcome, structure, process, and cost.

Arch Surg. 2002;137:20-27

The Veterans Health Administration in the Department of Veterans Affairs (VA) is the largest single health care provider in the United States. It comprises more than 159 medical centers, 376 outpatient clinics, and 165 long-term care facilities. One hundred twenty-eight VA medical centers perform major surgery, of which 42 perform cardiac surgery. Since the mid 1980s, the VA has been expending major efforts in the development and application of systems for the comparative assessment and improvement of the quality of surgical care in the various VA medical centers that perform cardiac and noncardiac surgery. The efforts in cardiac surgery have culminated in the establishment of the Continuous Improvement in Cardiac Surgery Program, while the efforts in noncardiac surgery have resulted in the establishment of the VA National Surgical Quality Improvement Program (NSQIP). The NSQIP, which also collects data for the Continuous Improvement in Cardiac Surgery Program, is the first nationally validated, outcome-based, risk-adjusted, and peer-controlled program for the measurement and enhancement of the quality of care in major surgical specialties. This article describes the evolution, methods, and future plans of the NSQIP.

From the Veterans Affairs Boston Healthcare System, West Roxbury (Drs Khuri and Daley), and Brigham and Women’s Hospital (Dr Khuri) and Institute for Health Policy, Massachusetts General Hospital/Partners Healthcare System (Dr Daley), Harvard Medical School, Boston, Mass; and Hines Veterans Affairs Cooperative Studies Program Coordinating Center, Hines, Ill (Dr Henderson).
BACKGROUND

The Congressional Mandate

The need for a proper system for the comparative assessment of the quality of surgical care in 133 VA hospitals became critical after the US Congress passed a law in December 1986 that mandated that the VA report its surgical outcomes annually compared with the "national average," and that these outcomes should be adjusted for the severity of patients' illnesses. A group of VA surgeons who were asked to consult with the VA Central Office on how to respond to the congressional mandate advised that there had been no known national averages for outcomes of surgical procedures and no known risk adjustment models that could be applicable to the various surgical specialties. They recognized, however, that the VA, by virtue of its centralized authority and its advanced medical informatics infrastructure, was in a unique position to develop national averages and risk adjustment models, and to set up a model system for the comparative assessment of the quality of surgical care among its various institutions. Toward that goal, these surgeons prompted the VA to conduct the National VA Surgical Risk Study (NVASRS) between October 1, 1991, and December 31, 1993, in 44 VA medical centers.3-5 The study was conducted by an executive committee composed of VA surgeons, health services researchers, and biostatisticians. An advisory panel of experts from outside the VA provided the scientific oversight for the design and conduct of the study (Table 1).

The Rationale

The rationale underlying the NVASRS was based on Iezzoni's "algebra of effectiveness" (Figure 1), a conceptual framework in which outcomes of health care are determined by the sum of 3 major factors: patients' risk factors before surgery, the effectiveness (quality) of the patients' care, and random variation.6 If one accounts for the severity of the patients' illnesses by proper risk adjustment and for random events by proper statistical methods, one can then equate outcome with effectiveness of care. Hence, to enable the use of outcome as a measure of quality of surgical care, the NVASRS had to (1) develop a reliable clinical database of patients' relevant preoperative risk factors and postoperative outcomes and (2) develop analytic tools for proper risk adjustment and to account for random events (Figure 1). The NVASRS recognized that surgical care was ideally suited for the use of outcome rather than process measures in the comparative assessment of quality of care, because surgical care revolved primarily around a single event (the operation), which, in most cases, had an expected measurable outcome.7

The Outcome of the NVASRS

Based on data from 117,000 major operations collected prospectively by a dedicated nurse in each of 44 VA medical centers, the NVASRS developed predictive risk adjustment models for 30-day mortality and morbidity in

### Table 1. National Veterans Affairs Surgical Risk Study Expert Advisory Panel (1991-1994)

<table>
<thead>
<tr>
<th>Surgeons</th>
</tr>
</thead>
<tbody>
<tr>
<td>J. Bradley Aust, MD, San Antonio, Tex</td>
</tr>
<tr>
<td>Paul Ebert, MD, Chicago, Ill</td>
</tr>
<tr>
<td>John Mannick, MD, Boston, Mass</td>
</tr>
<tr>
<td>Health services researchers</td>
</tr>
<tr>
<td>Liza Iezzoni, PhD, Boston</td>
</tr>
<tr>
<td>Barbara J. McNeil, MD, PhD, Boston</td>
</tr>
<tr>
<td>Epidemiologists and statisticians</td>
</tr>
<tr>
<td>L. Richard Smith, MD, Durham, NC</td>
</tr>
<tr>
<td>J. William Thomas, PhD, Ann Arbor, Mich</td>
</tr>
</tbody>
</table>

### Figure 1. The National Surgical Quality Improvement Program (NSQIP) is based on the assumption, forwarded by Iezzoni,7 that outcomes in surgery are determined by the sum of patients' risk factors, effectiveness or quality of care, and random variation. By providing a database to account for patients' risk factors and outcomes, and analytic tools for risk adjustment and to account for random variation, the NSQIP is able to equate the measurement of outcome with the measurement of quality of care. Adapted with permission from Mosby Inc.7

#### Materials and Methods

Data Acquisition and Transmission

At each VA medical center that performs major surgery, a clinical surgical nurse reviewer oversees the prospective collection of preoperative, intraoperative, and 30-day outcome data on all major operations into a special module of the VA's electronic clinical record (Table 2). In some high-volume hospitals, data collection is limited to the first 36 major operations in an 8-day cycle, with each cycle starting on a different day of the week.
produce a representative sample. About 115,000 cases are accrued annually. By the end of fiscal year 2000, the NSQIP database contained 727,447 complete operations in 9 major specialties (Table 3). Forty-five days after surgery, data are transmitted to 1 of 2 data coordinating centers. The noncardiac surgery data are transmitted to the Hines VA Cooperative Studies Program Coordinating Center, and the cardiac surgery data are transmitted to the Center for Continuous Improvement in Cardiac Surgery at the Denver VA Medical Center, Denver, Colo. At the coordinating center, the data are passed through an editing program. Potential errors in the data are transmitted back to the participating sites via query reports for resolution. Annually, the NSQIP data are passed through the VA’s Beneficiary Identification and Records Locator Subsystem, an administrative database of all veterans’ benefits, which includes survival information. The database, used by the NSQIP for the assertion of interrater reliability, is more than 95% complete with regard to a veteran’s vital status.

Predictive Modeling and Risk Adjustment

Annually, the noncardiac surgery data are subjected to logistic regression analysis to identify the independent predictors of 30-day mortality, 30-day morbidity, and postoperative length of hospital stay. Separate models are created for all operations and for each of the 8 noncardiac surgical specialties. These models are predictive, with high C indexes.4,5 Table 4 lists, in order of decreasing importance, the top 8 variables predictive of 30-day mortality for all operations, as they were calculated over time. The rank order of these variables did not change much over time, with the preoperative serum albumin, the American Society of Anesthesiologists class, and the presence of disseminated cancer remaining as the most important risk factors throughout. The β coefficient generated by the model for each predictive variable allows for the calculation of a population-specific expected outcome that reflects the severity of illness of that population. The risk-adjusted outcome for that specific population is expressed as the O/E ratio, where O represents the total number of observed events (deaths or complications) and E, the number of events that is expected on the basis of the compendium of the preoperative risk factors prevalent in that population. Figure 2 shows the O/E ratio and the 90% confidence interval for the 30-day mortality of all noncardiac operations at each hospital participating in the NSQIP in fiscal year 2000. Statistically significant high outlier hospitals, ie, hospitals in which the confidence interval is above 1.0, are indicated by an asterisk; statistically significant low outlier hospitals, ie, hospitals in which the confidence interval is below 1.0, are indicated by a dagger. High outlier hospitals have mortality rates that are significantly higher than what is expected based on the severity of illness of their respective patient populations, while low outlier hospitals have mortality rates that are significantly lower than what is expected based on the severity of the patients’ preoperative risk factors. The validation study, which was conducted as part of the NVASRS, has validated the concept that high outlier hospitals are more likely to have inferior structures and processes of care, and that low outlier hospitals are more likely to have superior structures and processes.5,9,10

Feedback to Providers and Managers

The quality improvement function of the NSQIP is effected primarily through feedback of information to the providers and managers in the field. The various formats of this feedback include the following.

(1) A comprehensive chief of surgery annual report that is specifically prepared for each surgical center. The report contains the risk-adjusted outcomes of all the participating hospitals, but the identities of these hospitals are blinded by a code that is changed annually. Each surgical center is privy only to its own hospital code. Site and specialty-specific data are provided that enable each center to compare the preoperative risk profiles and the outcomes of its patients with those of other peer hospitals and with national averages.

(2) Periodic assessment of the performance of high and low outlier institutions. This assessment is performed annually by the executive committee at a 2-day meeting in which the performance of each of the participating hospitals during a 4-year period is reviewed. The executive committee communicates to the field certain levels of concern about the high outlier status of various surgical centers and subspecialties within these centers.

Table 2. Elements in the National Surgical Quality Improvement Program Database

<table>
<thead>
<tr>
<th>Preoperative data</th>
<th>10 Demographic variables</th>
<th>30 Clinical variables</th>
<th>12 Laboratory variables</th>
<th>Intraoperative data</th>
<th>15 Clinical variables</th>
<th>Postoperative data</th>
<th>10 Laboratory variables</th>
<th>30-Day postoperative mortality</th>
<th>21 Categories of 30-day postoperative morbidity</th>
<th>Length of hospital stay</th>
</tr>
</thead>
</table>

Table 3. National Surgical Quality Improvement Program Database as of September 30, 2000*

<table>
<thead>
<tr>
<th>Subspecialty</th>
<th>FY 1998</th>
<th>FY 1999</th>
<th>FY 2000</th>
<th>Total</th>
<th>All Phases</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>General</td>
<td>31,437</td>
<td>32,163</td>
<td>31,201</td>
<td>199,519</td>
<td>27.4</td>
<td></td>
</tr>
<tr>
<td>Orthopedics</td>
<td>20,286</td>
<td>20,531</td>
<td>20,524</td>
<td>138,470</td>
<td>19.0</td>
<td></td>
</tr>
<tr>
<td>Urology</td>
<td>13,251</td>
<td>13,729</td>
<td>13,133</td>
<td>99,507</td>
<td>13.7</td>
<td></td>
</tr>
<tr>
<td>Vascular</td>
<td>10,281</td>
<td>10,753</td>
<td>9,905</td>
<td>75,963</td>
<td>10.4</td>
<td></td>
</tr>
<tr>
<td>Cardiac surgery</td>
<td>7,481</td>
<td>7,214</td>
<td>8,281</td>
<td>67,519</td>
<td>9.3</td>
<td></td>
</tr>
<tr>
<td>Neurosurgery</td>
<td>5,922</td>
<td>5,943</td>
<td>5,549</td>
<td>44,066</td>
<td>6.1</td>
<td></td>
</tr>
<tr>
<td>Otolaryngology</td>
<td>5,483</td>
<td>5,193</td>
<td>3,960</td>
<td>37,743</td>
<td>5.2</td>
<td></td>
</tr>
<tr>
<td>Thoracic</td>
<td>3,725</td>
<td>3,560</td>
<td>3,419</td>
<td>27,836</td>
<td>3.8</td>
<td></td>
</tr>
<tr>
<td>Plastic</td>
<td>2,210</td>
<td>2,246</td>
<td>2,162</td>
<td>17,257</td>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>2,797</td>
<td>2,921</td>
<td>2,366</td>
<td>19,567</td>
<td>2.7</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>102,873</td>
<td>104,253</td>
<td>100,440</td>
<td>727,447</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

*Data are given as number of operations unless otherwise indicated. FY indicates fiscal year.
The committee also communicates praise and provides certificates of commendation to centers and surgical specialties within these centers that demonstrate improved or superior performance, as evidenced by low O/E ratios.

(3) Provision of self-assessment tools. The NSQIP has developed instruments that it has made available to providers and managers to help them assess the strengths and weaknesses of their respective programs, particularly when the NSQIP reports show these programs to be high outliers in their respective risk-adjusted 30-day mortality or morbidity rates.

(4) Structured site visits for the assessment of data quality and specific performance. Providers and managers who face potential problems with the outcomes of their respective surgical services have the option to invite the NSQIP to conduct structured site visits to help identify and address deficiencies in the quality of care delivered by these services. These site visits are conducted in 2 stages. In the first stage, experienced NSQIP nurses conduct a thorough audit to assess the reliability of the data collection at the site. If no problems with the data are identified, a second site visit is conducted by a team composed of an experienced surgeon (who usually heads the team), a health services specialist, an anesthesiologist, and a critical care nurse. The site visit teams are only consultative to the providers and managers. After the site visit, the team issues a confidential detailed report of its findings and recommendations, which is distributed to the senior management and the surgery leadership at the site of the visit.

(5) Identification and dissemination of best practices. Hospitals that have significantly reduced their O/E ratios, and hospitals that have maintained consistently significantly low O/E ratios, are encouraged to report back to the NSQIP the methods and procedures that they have used to improve or sustain good risk-adjusted outcomes. The feedback from these hospitals, along with relevant knowledge gained from specific NSQIP site visits and from investigations conducted on the NSQIP database, are published regularly in the NSQIP’s annual report to the hospitals. Dissemination of “good practices”

Table 4. Order of Entry for the Most Predictive Preoperative Risk Factors of 30-Day Postoperative Morbidity (All Operations Models)*

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>Phase 1</th>
<th>Phase 2</th>
<th>FY 1996</th>
<th>FY 1997</th>
<th>FY 1998</th>
<th>FY 1999</th>
<th>Mean Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serum albumin, g/dL</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1.0</td>
</tr>
<tr>
<td>American Society of Anesthesiologists class</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2.0</td>
</tr>
<tr>
<td>Disseminated cancer</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3.0</td>
</tr>
<tr>
<td>Emergency operation</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>7</td>
<td>5.0</td>
</tr>
<tr>
<td>Age</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>4.7</td>
</tr>
<tr>
<td>Serum urea nitrogen &gt;40 mg/dL (&gt;14.3 mmol/L)</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>9</td>
<td>12</td>
<td>5</td>
<td>7.5</td>
</tr>
<tr>
<td>Do not resuscitate</td>
<td>7</td>
<td>10</td>
<td>8</td>
<td>4</td>
<td>11</td>
<td>10</td>
<td>8.3</td>
</tr>
<tr>
<td>Complexity score</td>
<td>12</td>
<td>9</td>
<td>10</td>
<td>13</td>
<td>14</td>
<td>14</td>
<td>12.0</td>
</tr>
<tr>
<td>Aspartate aminotransferase &gt;40 U/L</td>
<td>11</td>
<td>8</td>
<td>9</td>
<td>17</td>
<td>28</td>
<td>13</td>
<td>14.3</td>
</tr>
<tr>
<td>Weight loss &gt;10%</td>
<td>9</td>
<td>13</td>
<td>14</td>
<td>10</td>
<td>7</td>
<td>8</td>
<td>10.2</td>
</tr>
<tr>
<td>Functional status</td>
<td>24</td>
<td>6</td>
<td>11</td>
<td>8</td>
<td>6</td>
<td>6</td>
<td>10.2</td>
</tr>
<tr>
<td>White blood cell count &gt;11.0 × 10^3/µL</td>
<td>16</td>
<td>14</td>
<td>15</td>
<td>11</td>
<td>9</td>
<td>11</td>
<td>12.7</td>
</tr>
</tbody>
</table>

*C index for all operations models = 0.89 – 0.91. FY indicates fiscal year.
in this manner is a critical component of the quality improvement initiatives of the NSQIP.

**Investigations Using the Database**

Using policies modeled after those of the VA Cooperative Studies Program, the NSQIP has established an infrastructure that allows VA surgeons and researchers to query the database through hypothesis-driven proposals that are submitted to and peer reviewed by the executive committee. Once a study is approved by the NSQIP executive committee and the investigators’ local institutional review board, a statistician from the NSQIP Data Coordinating Center at Hines VA Cooperative Studies Program Coordinating Center is assigned to it, and the data are analyzed either at the Data Coordinating Center or at the respective principal investigator’s facility. The executive committee monitors the scientific quality and integrity of these studies and must review and approve all scientific presentations and publications that emanate from these studies before submission. The NSQIP publications to date have included 28 papers published in peer-reviewed journals, \(^2\)-\(^5\),\(^8\)-\(^29\) 24 papers in preparation for publication in peer-reviewed journals, and 11 invited publications and book chapters.\(^7\),\(^30\)-\(^40\) There are 59 research groups that are querying the database about hypotheses related to different surgical issues and conditions.

**THE ESSENTIALS**

**Quality Improvement**

The NSQIP is first and foremost a quality improvement program (Figure 3). The validity of its outcome-based methods in assessing the quality of surgical care has been established. However, the NSQIP is not a punitive program for the purpose of identifying “bad apples.” On the contrary, its primary focus is to provide the surgeons and managers in the field with reliable information, benchmarks, and consultative advice that will guide them in assessing and continually improving their local processes and structures of care. Chiefs of surgery in the VA have come to find value in the program and have learned to use the data and information contained in the annual report to identify and improve deficiencies that otherwise would have remained unnoticed.\(^4\)\(^1\) They have benefited from the various site visits that the NSQIP has organized, and from learning how other colleagues have addressed specific problems through the section of the annual report that is dedicated to the dissemination of best practices. Since the inception of the NSQIP data collection in 1991, the 30-day mortality of major surgery in the VA has decreased by 27% and the 30-day morbidity has decreased by 45% (Figure 4). The value that VA surgeons find in the NSQIP is the biggest asset of the program and the only guarantee of its continued effectiveness.

**Reliable Data**

The bases for any successful quality improvement process are reliable data. Driving the NSQIP is a compulsion for assurance of data reliability. Only data that could be reliably collected in pilot studies were incorporated into the data collection instruments. The data collection process in each hospital is attended to by a nurse reviewer, capable of making clinical judgments about definitions and outcomes. Nurses are trained on a uniform set of definitions that is continually improved to eliminate ambiguity and confusion. Interrater reliability and

---

**Figure 3.** The Veterans Affairs National Surgical Quality Improvement Program (NSQIP) collects data from each of its participating sites, ascertains their cleanliness and reliability, and processes them into comparative risk-adjusted outcomes. These outcomes, along with other information and tools, are continuously fed back to the participating centers for the primary purpose of achieving quality improvement (QI). Adapted with permission from Mosby Inc.\(^7\)

**Figure 4.** The 30-day postoperative mortality (A) and 30-day postoperative morbidity (B) for all major operations performed in the Department of Veterans Affairs hospitals throughout the duration of the National Surgical Quality Improvement Program data collection process. A 27% decrease in the mortality and a 45% decrease in the morbidity were observed in the face of no change in the patients’ risk profiles. FY indicates fiscal year.
core competencies are assessed annually at a nurse reviewers’ national meeting. The nurse reviewers share their questions and concerns with regional and national nurse coordinators through periodic communication, including monthly conference calls. The robustness of the Veterans Health Information Systems and Technology Architecture, the VA’s electronic medical record infrastructure, allows for electronic assessment of interrater reliability and for direct transmission of the laboratory data from the respective medical center laboratory to the coordinating center at Hines, thereby saving the nurse reviewer time and improving the accuracy of the data collection process.

As part of ascertaining the reliability of the data in the development of outcome-based risk adjustment models, the NVASRS had to develop and validate certain quantitative scores, most important of which were a complexity score and a morbidity score. Panels of specialists in 8 surgical specialties were assembled to devise, using a common set of guidelines, a technical complexity score for each of the more than 3500 Current Procedural Terminology codes in the database. Several scores were also developed to provide a measure of morbidity and were compared with each other. The morbidity score that was decided on, a dichotomous variable based on the presence or absence of 1 or more complications, was chosen because of its simplicity, ease of use, and the good correlation it exhibited with the other scores.

The NSQIP assigns high and low outlier status to specific surgical services and divisions, based on risk-adjusted outcomes. As such, it provides a statement on the quality of care provided by these surgical groups. Proper risk adjustment is crucial if hospitals or services are to be singled out as high or low outliers. In the absence of proper risk adjustment, a 60% error rate can be encountered in assigning an outlier status to a specific hospital or surgical service. Therefore, data reliability, combined with valid risk-adjustment models, is imperative to a proper comparative assessment of the quality of surgical care between various institutions.

Setting Standards of Surgical Care

The NSQIP database provides the basis for evidence-based policies and standards of care in surgery. For example, one of the cost cutting measures that was contemplated by VA managers in the mid 1990s was the closure of surgical services that performed a low volume of major surgery, on the assumption that surgical outcomes in low-volume hospitals were not as good as outcomes in high-volume hospitals. An NSQIP study was conducted to address this issue. It found no relationship in the VA between volume and risk-adjusted outcome of surgery in 8 major common operations. The study confirmed that the volume of surgery in the VA could not be used as a surrogate for quality and found no merit in setting up VA standards or policies that are based on volume of surgery alone. The relationship between volume and outcome of surgery is subject of a national debate that has been fueled by the Leapfrog Group of employers’ attempt to set standards of care that are based in part on the volume of surgery performed per surgical center. In the absence of a national database similar to that of the VA’s NSQIP, it will be difficult for the surgical community to reach evidence-based consensus on standards, policies, and performance measures.

Emphasis on Systems, Not Providers

The compendium of studies and site visits performed by the NSQIP to date have underscored the fact that quality of surgical care is primarily a function of well-coordinated systems of care. The identification of breakdowns in systems of care and the attempts to restore and improve these systems are the bases for the quality improvement initiative in the NSQIP. Single providers of care, such as the attending surgeon, contribute to the quality of the systems at their respective institutions, but their role in determining the overall quality of these systems is just one of many determinants of quality on the surgical service. Therefore, in the NSQIP, more emphasis is made on the system than on the provider, and provider-specific data are not transmitted to the central database. There are several reasons why the NSQIP discourages uploading provider-specific information into its national database. The first is purely statistical; the average surgeon does not perform enough major operations annually to provide a sample size adequate for statistical analysis. The second and more important reason is that one cannot separate the outcome-based quality of care rendered by a specific surgeon from that of the outcome-based quality of care rendered by his or her institution. If systems of care are poor at a specific institution, even the most competent surgeon can have poor outcomes. Likewise, a mediocre surgeon can have excellent outcomes if he or she is functioning in an environment with excellent systems of care. Last, and most important, surgeons need to buy into the NSQIP if it is to achieve success in its primary mission of improving the quality of surgical care. Directing the national focus to the comparative “performance” of individual surgeons is sure to alienate the most important constituency of the NSQIP, a lesson learned more than a decade ago when the comparative outcomes of cardiac surgeons in the state of New York were publicized on the pages of local and national newspapers.

THE FUTURE

Applicability to the Private Sector

Is the VA NSQIP applicable outside the VA? Three reasons prompted the NSQIP to address this question in 1999. (1) Several non-VA academic surgical centers, which came to know about the NSQIP through their respective affiliations with VA hospitals, expressed interest in joining the NSQIP if the option were made available to them. (2) Making the NSQIP available to the private sector would satisfy, for the first time, the congressional mandate that the VA compare its risk-adjusted surgical outcomes with those of the private sector. (3) The VA leadership and the NSQIP executive committee were eager to share with the rest of the surgical community the value that they recognized in the NSQIP as a repository.
Efforts have already begun in the NSQIP to evaluate availability to the surgical community nationwide. Of the NSQIP the efforts to make this program available to non-VA hospitals? (2) Are the models predictive of surgical outcomes of VA patients, developed by the NSQIP, applicable to non-VA patient populations? Three non-VA medical centers participated in the private sector initiative: the University of Michigan Medical Center, Ann Arbor; Emory University Hospital, Atlanta, Ga; and University of Kentucky Chandler Medical Center, Lexington. Using processes and definitions used by the VA NSQIP, a trained nurse in each of these 3 medical centers collected data on all major general and vascular operations and entered them into a specially developed secure Internet Web site. The data were then automatically transmitted to the NSQIP Data Coordinating Center at Hines, where they were edited for outlying or inconsistent values and analyzed. Analysis of the results of the private sector initiative after the first year of complete data collection showed that the VA NSQIP data collection and transmission methods were applicable to all 3 non-VA sites. The predictive and risk adjustment models were also equally applicable to the VA and the non-VA populations. These preliminary observations were encouraging enough for the VA leadership and the NSQIP executive committee to incorporate into the future plans of the NSQIP the efforts to make this program available to the surgical community nationwide.

The Big Picture

Although the NSQIP has achieved a good measure of success in its efforts to use outcomes as a means for measuring and improving the quality of surgical care in the VA, the NSQIP is still a long way from realizing the vision conceived by its founders. In its efforts to achieve quality improvement, the NSQIP today uses 2 measurable outcomes of surgery: postoperative mortality and postoperative morbidity (Figure 5A). There are other dimensions of surgical outcome that could be incorporated into the NSQIP to more thoroughly assess the quality of surgical care. The most important of these are long-term survival, functional outcomes, quality of life, and patients' satisfaction. Efforts have already begun in the NSQIP to evaluate available tools for the assessment of these outcomes, and to develop and validate new tools for analyses in which current tools are inadequate. More important, the NSQIP does not plan to limit its measures of quality to measures of outcome alone. It is important that process measures also be incorporated into the assessment of the quality of surgical care. However, only process measures that are demonstrated to affect outcome should be used for this purpose. Therefore, an important goal for the NSQIP is the identification of process measures in surgery that directly affect outcome, and the validation of these measures as tools in the assessment of the quality of surgical care. Cost is another dimension of health care that cannot be ignored in the assessment of the overall quality of care. Although cost can be regarded as an important outcome of health care, it is the relationship of cost to outcome that defines quality in surgery. The NSQIP is in a unique position to develop tools for the assessment of cost-effectiveness, the lowest cost for the best outcome. The use of process and cost measures in the assessment of the quality of health care also requires proper risk adjustment, a unique feature of the NSQIP. Hence, the vision for the NSQIP of tomorrow is a program that accurately quantifies patients' risk and achieves continuous quality improvement through the proper assessment of risk-adjusted outcomes, outcome-related structures and processes, and outcome-related costs (Figure 5B).

We acknowledge the continuous efforts of all the participants in the NSQIP, particularly the chiefs of surgery and the clinical nurse reviewers at each VA medical center. We also acknowledge the members of the NSQIP executive committee, who provide continuous leadership, guidance, and vision to the quality improvement initiatives described in this article. The editorial help of Nancy Healey and her assistance in the preparation of the manuscript are gratefully acknowledged.

Corresponding author and reprints: Shukri F. Khuri, MD, Veterans Affairs Boston Healthcare System, 1400 VFW Parkway, West Roxbury, MA 02132 (e-mail: shukri.khuri@med.va.gov).

REFERENCES

24. Collins TC, Johnson M, Khuri SF, Henderson WG, Daley J. Lower extremity non-
22. Weaver F, Hynes D, Goldberg JM, Khuri S, Daley J, Henderson W. Hysterec-
14. Harpole DH, DeCamp MM, Daley J, et al. Prognostic models of 30-day mortal-
16. Arozullah AM, Ferreira MR, Bennett RL, et al. Racial variations in the use of lapa-
rosopic cholecystectomy in the Department of Veterans Affairs medical sys-
20. Charns MP, Young GJ, Henderson WG, Khuri SF. Daley J. Coordination and pa-
25. Charns MP, Young GJ, Henderson WG, Khuri SF. Daley J. Coordination and pa-
27. Charns MP, Young GJ, Henderson WG, Khuri SF. Daley J. Coordination and pa-
29. Charns MP, Young GJ, Henderson WG, Khuri SF. Daley J. Coordination and pa-
31. Charns MP, Young GJ, Henderson WG, Khuri SF. Daley J. Coordination and pa-