Series Editor: Kamal M.F. Itani, MD

Approaches to Assessing Surgical Quality of Care

Alexandra L.B. Webb, MD
Aaron S. Fink, MD

Historically, measuring the quality of surgical care had been limited to 30-day mortality and morbidity rates. With the increasing complexity of surgical care and greater public demand for information about the quality of their health care, the development of more sophisticated metrics for assessing quality has become increasingly important. Systematic efforts to improve quality of care began as early as the 1960s, when Donabedian[1,2] developed a conceptual framework for defining and assessing quality of care. In this framework, he identified 3 basic components central to the quality of health care: structure, process, and outcome. Although each of these elements can be assessed individually, Donabedian emphasized that proper integration of all 3 elements is critical. Quality measures based on this triad have been an integral part of surgical quality improvement initiatives. Understanding these elements allows closer examination of various quality measures. This article, which is the sixth in a series addressing recent evidence-based recommendations for improving the quality and safety of surgical care, provides an overview of the structure-process-outcome framework of quality of care, with a focus on the advantages and disadvantages of each element.

TAKE HOME POINTS

- Quality of care can be assessed using structural, process, or outcome measures. Although each of these elements can be assessed individually, proper integration of all 3 elements is critical.
- Structural measures, such as procedure volume, have been used frequently to indirectly assess surgical quality.
- Process measures are used to measure the quality of care provided by individual physicians as well as individual hospitals.
- The National Surgical Quality Improvement Program is a risk-adjusted measurement tool that is used in the Veterans Affairs system and is gaining widespread acceptance in the private sector to measure and directly compare surgical outcomes.
- It is important to tailor the type of quality of care measure to the specific characteristics of the surgical procedure to be assessed in order to obtain optimal results.

STRUCTURAL MEASURES

Structural measures assess the setting of care or the health care system itself. Donabedian's definition of structure includes material resources, human resources, and organizational structure.[1,2] Examples of structural measures include procedure volume, subspecialty training, presence of "closed" intensive care units, nurse-to-bed ratios, and presence of certain technology or equipment. The primary advantage of targeting structural variables as a measure of quality revolves around the expediency of such measures. Structural variables can often be reviewed rapidly, as these data are usually readily available within administrative databases; as such, analyses of large numbers of patients or hospitals may be relatively straightforward. However, most structural variables can only be examined in observational studies. Such studies often engender uncertainty as to whether the measured differences represent actual differences in surgical quality. In addition, structural measures allow generalizations regarding large groups

Dr. Itani is a professor of surgery, Boston University; an associate chief of surgery, Boston Medical Center and Brigham & Women's Hospital; and chief of surgery, Boston VA Health Care System, Boston, MA. Dr. Webb is an assistant professor, Division of Gatrointestinal and General Surgery, Emory University School of Medicine; and chief of general surgery, Atlanta VA Medical Center, Atlanta, GA. Dr. Fink is a professor of surgery, Emory University School of Medicine, and chief of surgery, Atlanta VA Medical Center, Atlanta, GA.

of providers or hospitals but do not measure the quality of care provided by an individual practitioner or a single hospital. Finally, many structural variables are not easily actionable.[3] Therefore, structural measures may be useful for quality assessment, but they may not be useful in quality improvement. A structural measure that has been used frequently in surgical quality assessment in hospitals is volume for a given procedure. An example illustrating the utility and the pitfalls of structural measures is provided in the following section.

Volume-Outcome Relationship

Beginning in the 1980s, several reports appeared suggesting that outcomes in high-volume hospitals were better than outcomes in centers with lower clinical volumes.[4–10] In 2002, Birkmeyer et al[11] surveyed the national Medicare claims database and the Nationwide Inpatient Sample for 30-day mortality rates following various cardiovascular and oncologic procedures. They reported a correlation between increased hospital volume and lower mortality rates for all procedures examined. Subsequently, the validity of this association was questioned.[12–17] Most of the data supporting the volume-outcome association have been derived from administrative data that were not adjusted adequately for the case mix.[12] Thus, the reported differences may have been misleading. Indeed, several studies have demonstrated that the apparent volume-mortality relationships disappear when adequate risk adjustment is performed.[16,18] For example, when the relationship between risk-adjusted mortality and institutional surgical volume was examined in the Veterans Affairs (VA) health system, no association between lower mortality rates and increased hospital volume was found for abdominal aortic aneurysm (AAA) repair, infrainguinal vascular reconstruction, carotid endarterectomy (CEA), lung resection, cholecystectomy, colectomy, and total hip arthroplasty.[16] For certain high-risk procedures such as esophagectomy and pancreatic resection, however, mortality rates appear to be significantly lower at high-volume centers, even after the data are adjusted for risk.[11,19]

Other concerns have been raised regarding the volume-outcome relationship. Much of the published literature on this topic consists of cross-sectional studies, which provide a "snapshot" of an outcome at a given point in time. Longitudinal studies might be more appropriate, as they would be able to measure changes in outcomes at a given hospital over time and determine whether such changes were associated with changes in its procedural volume.[12] Additionally, low-volume centers often have falsely elevated mortality rates due to their low case volume. A longitudinal study might provide large case numbers, thus better defining the actual impact of the volume itself—as well as other structural or procedural factors—that contribute to the observed relationship.

Similar to hospital volume, individual surgeon volume has also been linked to mortality.[20,21] When surgeon volume is examined, the assumption is that high-volume surgeons possess superior technical skill and decision-making ability, accounting for improved outcomes.[15] Two hypotheses are frequently offered in support of this relationship.[22] The first is the "practice makes perfect" hypothesis. High volume presumably leads to more experience, better surgical judgment, and superior technical skills, thus improving quality. The second hypothesis, "selective referral," suggests that providers with favorable outcomes and reputations receive more patient referrals, thus increasing their volume. However, it is difficult to obtain meaningful volume data for an individual provider, as the numbers for most procedures are too low to measure outcomes quarterly or even annually. If outcomes are measured over a longer period of time, their merit may be questioned, as operative and clinical practice may have changed during the time period in which the outcomes are being measured. A "real-time" analysis is necessary in order for the measure to be useful in rewarding individual performance or for quality improvement initiatives.

As noted earlier, the volume-outcome relationship has been frequently used to assess quality of surgical care; it also has been used as the impetus for surgical quality improvement initiatives. In 2000, the Leapfrog group, a coalition of more than 150 large public and private health care purchasers, was formed to address concerns about preventable errors in medical and surgical care.[23,24] Leapfrog member employers agreed to purchase health care from institutions that met a set of standards that they had developed. One of these standards was evidence (volume)-based hospital referral (EHR). The EHR measure was adopted due to a Birkmeyer et al[25] paper, which estimated that referral of patients needing high-risk procedures to high-volume hospitals would save 2581 lives annually.

For certain procedures, the Leapfrog volume standards were initially set to such high levels (ie, 500 cases/yr for coronary artery bypass graft [CABG], 30 cases/yr for AAA repair, 100 cases/yr for CEA, and 7 esophagectomies/yr) that they proved to be largely unattainable in many hospitals. Increasing evidence also diminished the importance of volume as a sole predictor of mortality for other procedures.[26] For example, when mortality rates of Medicare patients who underwent

CEA were compared, no difference was found between hospitals with very low procedure volume and hospitals with very high procedure volume.[11] As a result, CEA was removed from Leapfrog's EHR program in 2003. In addition, Leapfrog decreased the volume thresholds for CABG and added process and outcome measures in most states. CABG volume criteria were dropped entirely in states with risk-adjusted mortality databases. Volume thresholds were increased for esophagectomy and AAA repair, and new volume-based referral criteria were established for pancreatectomy. Despite these changes, many hospitals will still be unable to meet the volume criteria.[26] For example, 3340 (62%) patients undergoing pancreatic surgery annually in the United States had their surgery performed at centers not meeting the Leapfrog volume criterion of 11 cases per year. Likewise, 148,508 (39%) of patients undergoing CABG were treated at centers that did not meet the Leapfrog volume criterion of 500 cases per year.[10] As such, concerns have been raised regarding the wisdom of attempting to redistribute such a large number of patients to higher volume hospitals.[27]

Structural measures such as procedure volume provide an indirect measure of surgical quality. In some cases, procedure volume may provide an adequate proxy for outcome, as is seen with esophagectomy and pancreatectomy. For other procedures, the relationship between procedure volume (either for the hospital or the surgeon) and outcome is much less clear.

PROCESS MEASURES

Process measures assess the activities performed when health care professionals provide care to patients. These measures address whether "good" medical care has been provided. Such measures include the completeness of clinical history, physical examination, and diagnostic testing; justification of diagnosis and therapy; technical competence in performing diagnostic and therapeutic procedures; evidence of preventive management; coordination and continuity of care; and acceptability of care to the recipient.[1] These measures are useful because they assess care that patients actually receive.[3] Theoretically, if the proper actions are chosen, greater compliance with these actions should lead to improved outcomes. Many process measures were actually selected following study of care processes at institutions with excellent outcomes. It is assumed that these processes are the cause of the improved outcomes at these institutions.

There are several advantages and disadvantages to using process measures. Information about care received by an individual patient is readily available in the medical record, allowing prompt remedial action if needed.[2] Additionally, these measures are usually actionable and lend themselves well to quality improvement initiatives.[3] Often, they are amenable to study in randomized trials, thus providing a high level of evidence for their effectiveness. These measures also have a certain degree of face validity and are often perceived by providers to be more "fair" than structural measures.

Often lacking in studies of process measures, however, is a firm evidential link to patient outcomes. When process measures are introduced despite a lack of evidence for their effectiveness, the result may be a measure embraced by the public but not supported by care providers.[28,29] For example, extensive data regarding perioperative medical management of surgical patients exist; however, there is a paucity of clinical data linking specific operative technique to patient outcome. (Exceptions include total mesorectal excision in patients with rectal cancer,[30] CABG,[31] and CEA.[32]) Likewise, data demonstrating benefit are lacking for many other common practices of perioperative surgical care, such as the need for and proper method for preoperative bowel preparation.[33] A final disadvantage to using process measures involves the difficulty in identifying patient populations eligible for a given intervention. For example, when examining postoperative discontinuation of prophylactic antibiotics, it is important to exclude patients receiving therapeutic antibiotics. The following section discusses 2 surgical quality improvement initiatives that illustrate the advantages and disadvantages of using process measures.

Use in Surgical Quality Improvement Initiatives

Process measures have garnered a great deal of attention in recent years. Since process measures are easy to measure and track, they are the focus of several national quality improvement initiatives. The Surgical Infection Prevention Project (SIP) was developed as a joint effort of the Centers for Disease Control and Prevention (CDC) and the Centers for Medicare and Medicaid Services (CMS).[34] SIP identified 3 perioperative process measures expected to decrease rates of surgical site infection: SIP 1 requires administration of prophylactic antibiotics within 60 minutes prior to surgical incision; SIP 2 provides guidelines for selecting the appropriate antibiotic for a given procedure; and SIP 3 requires timely discontinuation of prophylactic antibiotics (usually 24 hr postoperatively) in an effort to decrease antibiotic resistance. It has been clearly demonstrated that quality improvement initiatives promoting these practices can improve adherence to these guidelines. In a study of 35,543 patients at 44 hospitals, SIP 1

Table 1. Surgical Care Improvement Project Process and Outcome Measures*

Cardiac (CARD)
SCIP CARD 2: Surgery patients on a β-blocker prior to arrival who received a β-blocker during the perioperative period
SCIP CARD 3: Intra- or postoperative acute myocardial infarction diagnosed during index hospitalization and within 30 days of surgery†
Global
SCIP Global 1: Mortality within 30 days of surgery
SCIP Global 2: Readmission within 30 days of surgery
Infection (INF)
SCIP INF 1: Prophylactic antibiotic received within 1 hr prior to surgical incision
SCIP INF 2: Prophylactic antibiotic selection for surgical patients
SCIP INF 3: Prophylactic antibiotics discontinued within 24 hr after surgery end time (48 hr for cardiac patients)
SCIP INF 4: Cardiac surgery patients with controlled 6 AM postoperative serum glucose
SCIP INF 5: Postoperative wound infection diagnosed during index hospitalization†
SCIP INF 6: Surgery patients with appropriate hair removal
SCIP INF 7: Colorectal surgery patients with immediate postoperative normothermia
End-stage renal disease (ESRD)
SCIP VA 1: Proportion of permanent hospital ESRD vascular access procedures that are autogenous arteriovenous fistulas (to be derived from administrative data)
Venous thromboembolism (VTE)
SCIP VTE 1: Surgery patients with recommended VTE prophylaxis ordered
SCIP VTE 2: Surgery patients who received appropriate VTE prophylaxis within 24 hr prior to surgery to 24 hr after surgery
SCIP VTE 3: Intra- or postoperative pulmonary embolism diagnosed during index hospitalization and within 30 days of surgery†
SCIP VTE 4: Intra- or postoperative deep vein thrombosis diagnosed during index hospitalization and within 30 days of surgery†

Adapted from MedQIC. Measures: Surgical Care Improvement Project. Available at www.medqic.org/dcs/ContentServer?cid=1137346750659&pagename=Medqic%2FMeasure%2FMeasuresHome&parentName=TopicCat&level3=Measures&c=MQParents. Accessed 8 Jan 2008.

*Measures addressing respiratory interventions are currently in development.

†Denotes an outcomes measure. Currently, the outcome measures are still in the development stage.

compliance improved from 56% at baseline to 95% after 1 year following implementation. SIP 2 compliance likewise improved from 92.6% to 95%, and SIP 3 compliance improved from 40% to 85%. Additionally, the overall rate of surgical site infection decreased by 27% during this time, suggesting a direct link between improvement in these processes of care and actual patient outcomes.[35] However, a recent study using risk-adjusted data did not find a statistically significant relationship between surgical site infection rates and timely administration of prophylactic antibiotics.[36]

The Surgical Care Improvement Project (SCIP) was introduced in 2004 as a collaborative effort endorsed by 10 major national organizations, including the American College of Surgeons (ACS), the CMS, and the Joint Commission on Accreditation of Healthcare Organizations (JCAHO). This project's stated goal is to achieve a 25% reduction in all surgical complications by 2010. The project focuses on 4 areas: surgical infection prevention (including the previously mentioned SIP measures focusing on appropriate antibiotic prophylaxis), cardiac event prevention, deep venous thrombosis prophylaxis, and ventilator-associated pneumonia prevention (**Table 1**). At present, only the process measures have been formally introduced, but several direct outcome measures are in development.[37]

The national reporting of compliance with SIP and SCIP measures has focused attention on processes of care at most hospitals. Many hospitals have demonstrated improvement in adherence to SIP guidelines as a result of this focus.[35,36] However, few institutions have demonstrated a correlation between improved adherence to these measures and improved patient outcomes.[35] There has been much debate regarding the evidence base for many of the measures, resulting in their withdrawal or modification. It is clear, however, that the SCIP initiative will continue to expand its purview and similar programs will be developed in an effort to monitor and compare hospital and provider quality.

OUTCOME MEASURES

Outcome measures assess the effect of care on the health status of patients and populations.[2] As noted earlier, quality measurement for surgical outcomes traditionally focused on direct measurement of surgical outcomes. As a result, average 30-day mortality and morbidity rates have been established for most common procedures. Other patient outcomes that have been studied in order to assess surgical quality include length of stay, readmission rates, patient satisfaction, health-related quality of life, cost-effectiveness, and resource utilization. These measures usually afford excellent face validity, as most people consider improving patient outcomes to be the main goal of surgical practice. Surgeons more readily accept such measures than indirect measures of surgical care. Additionally, there is evidence that

merely measuring these outcomes may have the effect of improving them (the "Hawthorne effect").[4]

Direct measurement of surgical outcomes may be complicated by sample size concerns. Thus, it may be difficult to obtain meaningful data for surgeon- or procedure-specific mortality at a given hospital. For surgeon-specific outcomes to be meaningful, a procedure should be performed with reasonable frequency and should impose a significant risk profile. Another disadvantage is that by their very nature outcome measures are delayed events. It can be difficult to obtain information about outcomes that occur after care is completed, especially after hospital discharge.[2] Once an outcome has occurred, it also may be too late to improve quality of care given to that individual patient, although lessons learned may be used to prevent similar events from occurring in the future. Outcome assessments also have the unique attribute of reflecting all contributions to care, including structural and process issues as well as patient contributions. Thus, outcome measures can measure successes or failures at all points in the system. Nonetheless, it can be difficult to trace an adverse outcome to the complex sequence of events resulting in a given outcome.[2]

An important consideration involved with the assessment of outcomes is data comparison. When individual or institutional outcomes are compared, divergent risk profiles of the patient populations in question compromise the quality of the comparison.[38,39] Without proper risk adjustment (**Figure 1**), physicians and institutions caring for the sickest patients could be inappropriately identified as having poor outcomes.[39] Conceivably, publication of unadjusted outcome data could result in the denial of care to the sickest patients for fear of making the institution or the individual practitioner appear inferior. Initially, most risk-adjustment strategies were based on the use of administrative data in an effort to contain costs.[39] However, risk factor information obtained in this manner is of questionable accuracy, compromising the quality of the subsequent risk adjustment. Hence, risk-adjustment initiatives that use clinical information have been developed and appear to have achieved higher levels of accuracy.[39–42] Outcome registries have been used with increasing frequency for surgical quality assessment. The following section discusses the role of these registries in quality assessment as well as their benefits and disadvantages.

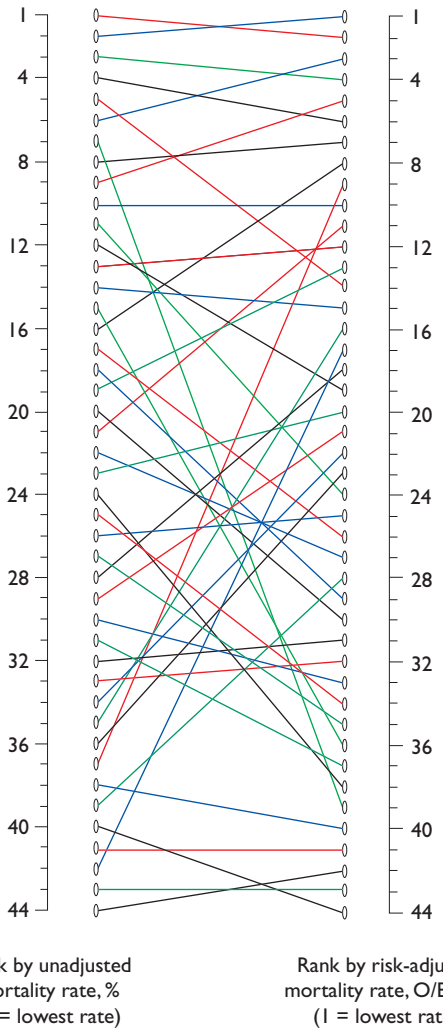

Figure 1. Changes in rank after risk adjustment for 30-day mortality. The left column shows the rank of the 44 hospitals by unadjusted postoperative mortality rate, with the hospital listed in the first position having the lowest rate. The right column shows the rank of the same hospitals based on their risk-adjusted observed-to-expected (O/E) ratio for postoperative mortality. Each hospital is connected with a line that demonstrates the change in rank order after risk adjustment. (Adapted from Khuri SF, Daley J, Henderson W, et al. Risk adjustment of the postoperative mortality rate for the comparative assessment of the quality of surgical care: results of the National Veterans Affairs Surgical Risk Study. J Am Coll Surg 1997;185:324. Copyright 1997, with permission from Elsevier.)

Outcome Registries

Outcome registries refer to data collected on procedure outcomes for a given institution or system. These registries allow comparisons to be made between individual practitioners and/or institutions. The value of these registries is dependent on the types of data collected and on the methods used for comparing these data. Multiple administrative databases exist, such as the Medicare claims database. This database is composed of data collected by nonclinical personnel for the

Table 2. Postoperative Outcomes Recorded in the National Surgical Quality Improvement Program Database

30-Day postoperative mortality
30-Day postoperative morbidity
Wound classification (superficial, deep incisional, or organ/space)
Wound disruption
Pneumonia
Unplanned intubation
Pulmonary embolism
On ventilator > 48 hr
Progressive renal insufficiency
Acute renal failure
Urinary tract infection
Central nervous system occurrences
Cerebrovascular accident
Coma > 24 hr
Peripheral nerve injury
Cardiac arrest requiring cardiopulmonary resuscitation
Myocardial infarction
Bleeding > 4 units red blood cells
Graft/prosthesis/flap failure
Deep venous thrombosis/thrombophlebitis
Systemic sepsis/septic shock
Length of hospital stay

Data from Khuri SF, Daley J, Henderson W, et al. The Department of Veterans Affairs' NSQIP: the first national, validated, outcome-based, risk-adjusted, and peer-controlled program for the measurement and enhancement of the quality of surgical care. National VA Surgical Quality Improvement Program. Ann Surg 1998;228:491–507.

purpose of billing. Data collected include basic demographic data as well as comorbidities, length of hospital stay, mortality, and complications. The accuracy of data collected in this manner has been questioned. Other clinical outcome registries contain various types of data collected by clinicians. Since the purpose of these registries is research or quality improvement, the data collected usually include more clinical factors. Also, if the data are prospectively collected or collected by clinical personnel, they are generally of higher quality.

Early outcome registries. Beginning in 1986, the Health Care Financing Administration (HCFA) released annual unadjusted hospital mortality statistics based on Medicare claims data.[43] Many other organizations followed suit and reported this information (eg, the JCAHCO reports on hospital mortality rates). Amid great criticism regarding the credibility of this unadjusted administrative data, HCFA ultimately ceased publishing reports on hospital mortality rates.[44,45]

The first large-scale registries to focus on direct measurement of surgical outcomes were the cardiac surgery registries in New York[46] and Northern New England in the 1980s.[47] Since that time, several other outcome registries have been implemented in other surgical fields such as oncology[48] and bariatrics.[49] Many of these registries offer incomplete or no risk adjustment.

National Surgical Quality Improvement Program (NSQIP). The Department of Veterans Affairs organized the National VA Surgical Risk Study (NVASRS) between 1991 and 1993.[50] This program was developed in response to Public Law 99-16, which was enacted by the US Congress in December 1985. This law mandated that surgical outcomes of VA hospitals be compared with those of private hospitals.[51] These initial comparisons were performed with limited risk adjustment.[52] Since the patient population at VA hospital centers was claimed to be older and sicker than that in most private hospitals, the validity of unadjusted or minimally adjusted comparisons was questioned.

The purpose of NVASRS was to develop risk-adjustment methodology that would permit valid comparisons between institutions with divergent patient populations. The NVASRS successfully developed such a model, using a nurse reviewer to collect preoperative, intraoperative, and postoperative outcome data on patients undergoing noncardiac operations at 44 VA hospitals. The outcomes measured were 30-day mortality and 21 major comorbidities. With the success demonstrated by the NVASRS, the VA utilized the resultant methodology and established the NSQIP in 1994.[53–58] Like NVASRS, NSQIP employs specially trained, dedicated nurses to collect data on preoperative and intraoperative factors such as patient comorbidities (eg, diabetes mellitus) and procedure type, respectively, as well as postoperative occurrences in patients undergoing noncardiac procedures at VA hospitals. The outcomes measured include all-cause 30-day mortality as well as major comorbidities that are aggregated into 21 different groups (**Table 2**).[55] Standard definitions are used for each variable, thus minimizing intercenter variability. These data are used to calculate predicted morbidity and mortality rates, which are then compared with measured morbidity and mortality rates. The latter comparisons are then expressed as a ratio of observed-to-expected-morbidity or mortality (O/E ratio). O/E ratios for each institution are compared to identify institutions with values significantly higher (high outlier) or lower (low outlier) than the average.

This methodology was validated when site visits to high- and low-outlier sites confirmed the presence of superior structural components and care processes at

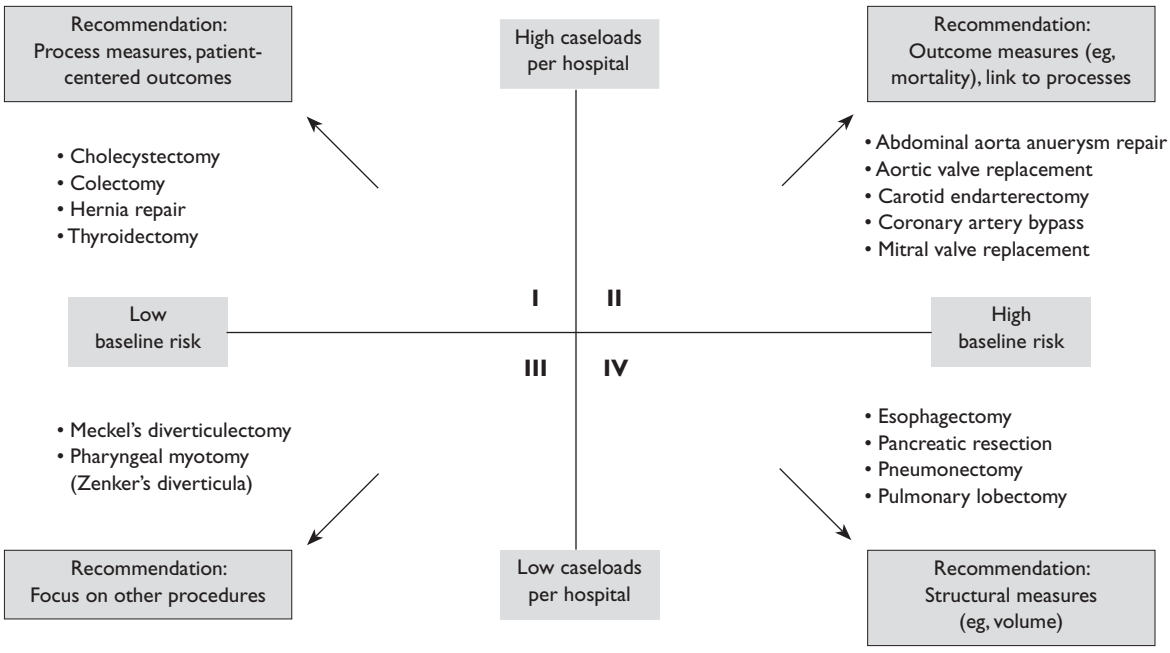

Figure 2. Recommendations for when to focus on structure, process, or outcome measures. (Adapted from Birkmeyer JD, Dimick JB, Birkmeyer NJ. Measuring the quality of surgical care: structure, process, or outcomes? J Am Coll Surg 2004;198:631. Copyright 2004, with permission from Elsevier.)

low-outlier sites and the converse at high-outlier sites.[59] Further, when the NSQIP risk-adjustment methodology was compared with risk adjustment based on administrative data (from the patient treatment file at VA hospitals), the latter were found to afford poor sensitivity and predictive value.[60] As might be expected, the NSQIP clearly offers superior risk adjustment when compared with other commonly used risk-adjustment scales.[42]

NSQIP represents the first large-scale program to assess risk-adjusted, hospital-specific mortality and morbidity rates for most surgical procedures and subspecialties. In 2002, application of the NSQIP was shown to be feasible in the private sector.[61,62] Thus, the ACS adopted the NSQIP in 2004 and began to further promote its use in the private sector as the ACS-NSQIP. With NSQIP methodology in place at both VA and private hospitals, the initial congressional mandate to compare their surgical outcomes can be fulfilled.[63–66]

The major obstacle to widespread implementation of the NSQIP is cost, which has been estimated at $38 per operation.[55] The main expense is the requirement for a dedicated clinical reviewer at each site. Also, while reliable risk-adjusted outcome data is essential in any quality measurement endeavor, it must be coupled with an understanding of the structure and processes of care implicit in those outcomes.[2]

Thus, outcome measures such as outcome registry databases provide a direct measure of surgical quality. The NSQIP, developed by the VA and expanded to the private sector, provides a validated tool for the measurement and comparison of risk-adjusted surgical outcomes. Although it is more costly, its benefits have been clearly demonstrated. The data generated by the NSQIP are useful for guiding quality improvement initiatives.

SELECTING QUALITY MEASURES

In order to comprehensively measure the quality of surgical care, structure, process, and outcome measures are all necessary. As Donabedian asserts, "good structure increases the likelihood of good process, and good process increases the likelihood of a good outcome."[2] Such a combined approach builds on the strength of each measure and mitigates their weaknesses. As our health care delivery system increases in complexity, so should our quality assessment tools.

Birkmeyer et al[3] propose that the procedure itself be the primary factor in determining which approach to quality measurement should be used. They recommend stratifying procedures by risk and by volume (**Figure 2**). High-risk, infrequently performed procedures (eg, esophagectomy, pancreatectomy) may best be assessed via procedure volume, which may be the only practical approach for this group. High-risk,

frequently performed procedures (eg, CABG) may be best examined by risk-adjusted outcome measures, such as the NSQIP or its cardiac surgery counterpart, the CICSP. Low-risk, high-volume procedures (eg, cataract surgery, inguinal hernia repair) are more difficult to evaluate. Due to low mortality risk, neither volume nor direct measures of morbidity and mortality are likely to be informative. Therefore, either process measures or measures of outcomes other than morbidity and mortality (eg, functional health status) should be considered in this group. Procedures with low risk and low volume should probably not be a major focus for quality measurement at this time. Birkmeyer et al's approach is appealing as it promotes a "right tool for the right job" philosophy and recognizes the role for multiple complementary strategies in the measurement of surgical quality.

FUTURE DIRECTIONS

Clearly, there is room for improvement in the accurate measurement of surgical quality. Improvements are needed in the quality of data collection, including a decreased reliance on administrative data. As development of process measures continues, more data supporting their effectiveness are needed. Focus on patient-centered outcomes should also be increased, with the goal of using these outcomes to drive future quality improvement initiatives. Measures of surgical decision-making are also needed, including measures of how well patient preferences are incorporated in these decisions. **HP**

Corresponding author: Aaron S. Fink, MD, Atlanta VA Medical Center, 1670 Clairmont Road (112), Decatur, GA 30033; Aaron.Fink@VA.gov.

REFERENCES

1. Donabedian A. Evaluating the quality of medical care. 1966. Milbank Q 2005;83:691–729.
2. Donabedian A. The quality of care. How can it be assessed? JAMA 1988;260:1743–8.
3. Birkmeyer JD, Dimick JB, Birkmeyer NJ. Measuring the quality of surgical care: structure, process, or outcomes? J Am Coll Surg 2004;198:626–32.
4. Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. N Engl J Med 1979;301:1364–9.
5. Hannan EL, O'Donnell JF, Kilburn H Jr, et al. Investigation of the relationship between volume and mortality for surgical procedures performed in New York State hospitals. JAMA 1989;262:503–10.
6. Begg CB, Cramer LD, Hoskins WJ, Brennan MF. Impact of hospital volume on operative mortality for major cancer surgery. JAMA 1998;280:1747–51.
7. Flood AB, Scott WR, Ewy W. Does practice make perfect? Part I: The relation between hospital volume and outcomes for selected diagnostic categories. Med Care 1984;22:98–114.
8. Dudley RA, Johansen KL, Brand R, et al. Selective referral to high-volume hospitals: estimating potentially avoidable deaths. JAMA 2000;283:1159–6.
9. Halm EA, Lee C, Chassin MR. How is volume related to quality in health care? A systematic review of the research literature. Ann Intern Med 2002; 137:511–20.
10. Birkmeyer JD, Lucas FL, Wennberg DE. Potential benefits of regionalizing major surgery in Medicare patients. Eff Clin Pract 1999;2:277–83.
11. Birkmeyer JD, Siewers AE, Finlayson EV, et al. Hospital volume and surgical mortality in the United States. N Engl J Med 2002;346:1128–37.
12. Sowden AJ, Sheldon TA. Does volume really affect outcome? Lessons from the evidence. J Health Serv Res Policy 1998;3:187–90.
13. Betensky RA, Christian CK, Gustafson ML, et al. Hospital volume versus outcome: an unusual example of bivariate association. Biometrics 2006;62: 598–604.
14. Christian CK, Gustafon ML, Betensky RA, et al. The Leapfrog volume criteria may fall short in identifying high-quality surgical centers. Ann Surg 2003;238:447–57.
15. Christian CK, Gustafson ML, Betensky RA, et al. The volume-outcome relationship: don't believe everything you see [published erratum appears in World J Surg 2006;30:2085]. World J Surg 2005;29:1241–4.
16. Khuri SF, Daley J, Henderson W, et al. Relation of surgical volume to outcome in eight common operations: results from the VA National Surgical Quality Improvement Program. Ann Surg 1999;230:414–32.
17. Khuri SF, Henderson WG. The case against volume as a measure of quality of surgical care. World J Surg 2005;29:1222–9.
18. Shroyer AL, Marshall G, Warner BA, et al. No continuous relationship between Veterans Affairs hospital coronary artery bypass grafting surgical volume and operative mortality. Ann Thorac Surg 1996;61:17–20.
19. Birkmeyer JD, Dimick JB, Staiger DO. Operative mortality and procedure volume as predictors of subsequent hospital performance. Ann Surg 2006; 243:411–7.
20. Dimick JB, Birkmeyer JD, Upchurch GR Jr. Measuring surgical quality: what's the role of provider volume? World J Surg 2005;29:1217–21.
21. Neumayer LA, Gawande AA, Wang J, et al; CSP #456 Investigators. Proficiency of surgeons in inguinal hernia repair: effect of experience and age. Ann Surg 2005;242:344–52.
22. Luft HS, Hunt SS, Maerki SC. The volume-outcome relationship: practice-makes-perfect or selective-referral patterns? Health Serv Res 1987;22:157–82.
23. Shaller D, Sofaer S, Findlay SD, et al. Consumers and quality-driven health care: a call to action. Health Aff (Millwood) 2003;22:95–101.
24. Milstein A, Galvin RS, Delbanco SF, et al. Improving the safety of health care: the leapfrog initiative [published erratum appears in Eff Clin Pract 2001;4:94]. Eff Clin Pract 2000;3:313–6.
25. Birkmeyer JD, Finlayson EV, Birkmeyer CM. Volume standards for high-risk surgical procedures: potential benefits of the Leapfrog initiative. Surgery 2001;130:415–22.
26. Birkmeyer JD, Dimick JB. Potential benefits of the new Leapfrog standards: effect of process and outcomes measures. Surgery 2004;135:569–75.
27. Ward MM, Jaana M, Wakefield DS, et al. What would be the effect of referral to high-volume hospitals in a largely rural state? J Rural Health 2004;20: 344–54.
28. Bradley EH, Herrin J, Elbel B, et al. Hospital quality for acute myocardial infarction: correlation among process measures and relationship with short-term mortality. JAMA 2006;296:72–8.
29. Devereaux PJ, Beattie WS, Choi PT, et al. How strong is the evidence for the use of perioperative beta blockers in non-cardiac surgery? Systematic review and meta-analysis of randomised controlled trials. BMJ 2005;331:313–21.
30. Hermanek P. Impact of surgeon's technique on outcome after treatment of rectal carcinoma. Dis Colon Rectum 1999;42:559–62.
31. Peterson ED, DeLong ER, Jollis JG, et al. The effects of New York's bypass surgery provider profiling on access to care and patient outcomes in the elderly. J Am Coll Cardiol 1998;32:993–9.
32. Hannan EL, Popp AJ, Feustel P, et al. Association of surgical specialty and processes of care with patient outcomes for carotid endarterectomy. Stroke 2001;32:2890–7.
33. Itani KM, Wilson SE, Awad SS, et al. Polyethylene glycol versus sodium phosphate mechanical bowel preparation in elective colorectal surgery. Am J Surg 2007;193:190–4.
34. Bratzler DW, Houck PM; Surgical Infection Prevention Guideline Writers Workgroup. Antimicrobial prophylaxis for surgery: an advisory statement from the National Surgery Infection Prevention Project. Am J Surg 2005; 189:395–404.
35. Dellinger EP, Hausmann SM, Bratzler DW, et al. Hospitals collaborate to decrease surgical site infections. Am J Surg 2005;190:9–15.
36. Hawn MT, Itani K, Gray S, et al. Is timely administration of prophylactic antibiotics a quality measure for surgery [abstract]? Presented at the 119th Annual Southern Surgical Association Meeting; 2007 Dec 2–5; Hot Springs, VA.
37. MedQIC. SCIP project information. Available at www.medqic.org/scip.

Accessed 10 Dec 2007.
38. Iezzoni LI. An introduction to risk adjustment. Am J Med Qual 1996;11: s8–11.
39. Iezzoni LI. The risks of risk adjustment. JAMA 1997;278:1600–7.
40. Iezzoni LI. Asessing quality using administrative data. Ann Intern Med 1997;127(8 Pt 2):666–74.
41. Iezzoni LI, Ash AS, Shwartz M, et al. Predicting who dies depends on how severity is measured: implications for evaluating patient outcomes. Ann Intern Med 1995;123:763–70.
42. Atherly A, Fink AS, Campbell DC, et al. Evaluating alternative risk-adjustment strategies for surgery. Am J Surg 2004;188:566–70.
43. Bowen OR, Roper WL. Medicare hospital mortality information, 1986. Washington (DC): US Government Printing Office; Health Care Financing Administration; 1987; HCFA Publication No 01–002.
44. Rosen HM, Green BA. The HCFA excess mortality lists: a methodological critique. Hosp Health Serv Adm 1987;32:119–27.
45. Dubois RW. Inherent limitations of hospital death rates to assess quality. Int J Technol Assess Health Care 1990;6:220–8.
46. Hannan EL, Kilburn H Jr, O'Donnell JF, et al. Adult open heart surgery in New York State. An analysis of risk factors and hospital mortality rates. JAMA 1990;264:2768–74.
47. O'Connor GT, Plume SK, Olmstead EM, et al. A regional intervention to improve the hospital mortality associated with coronary artery bypass graft surgery. The Northern New England Cardiovascular Disease Study Group. JAMA 1996;275:841–6.
48. Surveillance Epidemiology and End Results (SEER) database. Available at www.seer.cancer.gov. Accessed 10 Dec 2007.
49. Hutter MM, Crane M, Keenan M, et al. Data collection systems for weight loss surgery: an evidence-based assessment. Obes Res 2005;13:301–5.
50. Khuri SF, Daley J, Henderson W, et al. The National Veterans Administration Surgical Risk Study: risk adjustment for the comparative assessment of the quality of surgical care. J Am Coll Sur 1995;180:519–31.
51. Veterans Administration Health-Care Amendments of 1985. Pub. L. No. 99 Stat. 941 (3 Dec 1985).
52. Stremple JF, Bross DS, Davis CL, McDonald GO. Comparison of postoperative mortality in VA and private hospitals. Ann Surg 1993;217:277–85.
53. Daley J, Khuri SF, Henderson W, et al. Risk adjustment of the postoperative morbidity rate for the comparative assessment of the quality of surgical care: results of the National Veterans Affairs Surgical Risk Study. J Am Coll Surg 1997;185:328–40.
54. Khuri SF, Daley J, Henderson W, et al. Risk adjustment of the postoperative mortality rate for the comparative assessment of the quality of surgical care: results of the National Veterans Affairs Surgical Risk Study. J Am Coll Surg 1997;185:315–27.
55. Khuri SF, Daley J, Henderson W, et al. The Department of Veterans Affairs' NSQIP: the first national, validated, outcome-based, risk-adjusted, and peer-controlled program for the measurement and enhancement of the quality of surgical care. National VA Surgical Quality Improvement Program. Ann Surg 1998;228:491–507.
56. Khuri SF, Daley J, Henderson WG. The comparative assessment and improvement of quality of surgical care in the Department of Veterans Affairs. Arch Surg 2002;137:20–7.
57. Khuri SF. The NSQIP: a new frontier in surgery. Surgery 2005;138:837–43.
58. Khuri SF. Quality, advocacy, healthcare policy, and the surgeon. Ann Thorac Surg 2002;74:641–9.
59. Daley J, Forbes MG, Young GJ, et al. Validating risk-adjusted surgical outcomes: site visit assessment of process and structure. National VA Surgical Risk Study. J Am Coll Surg 1997;185:341–51.
60. Best WR, Khuri SF, Phelan, M, et al. Identifying patient preoperative risk factors and postoperative adverse events in administrative databases: results from the Department of Veterans Affairs National Surgical Quality Improvement Program. J Am Coll Surg 2002;194:257–66.
61. Fink AS, Campbell DA Jr, Mentzer RM Jr, et al. The National Surgical Quality Improvement Program in non-veterans administration hospitals: initial demonstration of feasibility. Ann Surg 2002;236:344–54.
62. Khuri S, Henderson W, Daley J, et al. Successful implementation of the Department of Veterans Affairs' NSQIP in the private sector: the patient safety in surgery study. Ann Surg. In press 2008.
63. Fink AS, Hutter MM, Campbell DA Jr, et al. Comparison of risk-adjusted 30-day postoperative mortality and morbidity in Department of Veterans Affairs hospitals and selected university medical centers: general surgical operations in women. J Am Coll Surg 2007;204:1127–36.
64. Henderson WG, Khuri SF, Mosca C, et al. Comparison of risk-adjusted 30-day postoperative mortality and morbidity in Department of Veterans Affairs hospitals and selected university medical centers: general surgical operations in men. J Am Coll Surg 2007;204:1103–14.
65. Hutter MM, Lancaster RT, Henderson WG, et al. Comparison of risk-adjusted 30-day postoperative mortality and morbidity in Department of Veterans Affairs hospitals and selected university medical centers: vascular surgical operations in men. J Am Coll Surg 2007;204:1115–26.
66. Johnson RG, Wittgen CM, Hutter MM, et al. Comparison of risk-adjusted 30-day postoperative mortality and morbidity in Department of Veterans Affairs hospitals and selected university medical centers: vascular surgical operations in women. J Am Coll Surg 2007;204:1137–46.