Journal of Evaluation in Clinical Practice, **11**, 3, 209–217

Measures of surgical quality: what will patients know by 2005?

Michael S. Broder MD MSHS,[1] Lisa Payne-Simon MPH[2] and Robert H. Brook MD ScD[3,4]
[1]Partnership for Health Analytic Research, LLC, Beverly Hills, CA, USA
[2]California HealthCare Foundation, San Francisco, CA, USA
[3]UCLA Center for the Health Sciences, Los Angeles, CA, USA
[4]RAND Health, Santa Monica, CA, USA

Correspondence
Dr Michael S. Broder
Partnership for Health Analytic Research, LLC
369 S. Doheny Dr, #169
Beverly Hills, CA 90211
USA
E-mail: mbroder@ucla.edu

Keywords: evaluation, health care quality, quality indicators, surgical procedures

Accepted for publication:
5 April 2004

Abstract

Rationale, aims and objectives Many objective measures rating quality of doctors, hospitals, and medical groups are publicly reported. Surgical patients may have more opportunity to use quality measures than other types of patients to guide their choice of provider. If surgical patients are able to choose higher quality providers, overall surgical quality might increase. **Objective** To determine what procedure-specific measures of surgical quality are available to consumers facing surgery in California and what new measures will be available by 2005. **Methods** We searched for and surveyed organizations publicly reporting data on health care quality in California. We asked about current quality measures and new measures set for public release by 2005. Included measures had to be procedure-specific and results separated by hospital. The main outcome measures were the number of quality measures; conceptual aspect of quality measured; and type of risk-adjustment used. **Results** Eighteen organizations publicly report any health care quality measures in California. These organizations report 333 measures, of which 32 (10%) are procedure-specific measures of surgical quality. There is at least one quality measure for 21 different procedures; these procedures account for 14% of all major operations. Three new measures will be released by 2005. **Conclusions** Californians facing surgery have limited information regarding quality of their care; few new measures are planned. Eighty-six per cent of patients would find no quality measures related to planned procedures. Public release of performance data is unlikely to improve the quality of health care unless the number and comprehensiveness of measures increase dramatically.

Introduction

Many industrialized countries are undertaking efforts to publicly report data on health care quality (Department of Health 2000; Australian Council for Safety and Quality in Health care 2002). In the United States, some measures of the quality of plans, hospitals, medical groups, and individual physicians have been publicly reported for several years. Data suggest that consumers may use these measures to guide some of their choices (Chassin *et al*. 1996). However, in some cases, patients may be unable to use these data to choose a health care provider. For example, a person requiring trauma care has little opportunity to make provider choices. For some conditions, consumers may value familiarity with the provider more than a quality score (Berry *et al*. 2001).

Surgical patients may be ideal candidates to use quality data in selecting a provider because many

have time to plan their care. They might also be willing to travel to a less convenient location for surgery because surgical care is typically time-limited. Americans undergo more than 41 million major surgical procedures every year, creating a large group of people potentially interested in measures of surgical quality (Vital and Health Statistics 2001).

The United States has been at the forefront of the quality report card movement (Marshall *et al*. 2000). We examined the adequacy of the currently reported data on surgical quality in one large US state. We determined what procedure-specific quality data will be available to California patients facing non-obstetric surgery in 2005, and whether patients could use this information to determine which provider, plan, or hospital would provide the best care. We used California-specific data because California, with 12% of the total US population, is the most populous state in the nation and has the largest number of surgical procedures of any state (Giooum *et al*. 1996). California is also one of 17 states that publicly release an all-payer hospital database, which has been used to examine hospital quality (California Office of Statewide Health Planning and Development 2003).

Methods

There is no central repository of quality of care reports in California, nor is there a regulatory or licensing body that maintains a comprehensive listing of data sources. We began with a list of such measures, reported by a diverse group of organizations, that were available in California in 1998 (California Healthcare Foundation, unpublished manuscript). Using this list as a base, we developed a survey to elicit updated information from each of the organizations that provided data in 1998. We also asked each 1998 source to identify organizations that might be reporting quality data by 2005.

Where possible, we contacted these organizations and administered our survey. We also searched the Internet for sources of quality data on California health care providers. Our approach was designed to find a wide variety of information, all of which a determined patient would be able to find. We did not include measures that are not regularly updated or were reported only once.

Our survey elicited information about each quality measure, including the health system sector to which it applied; how the performance data were obtained; and what disease, procedure, or condition the item covered. Where possible we determined the populations to which the measure applied. The survey also asked how the data were publicized, what the time lag was from collection to release, and whether there were new measures under consideration for public release by 2005. We excluded obstetric procedures because obstetric providers are not usually selected with the expectation that they will be performing surgery.

We obtained data on frequency of procedures from the National Center for Health Statistics, the United Network for Organ Sharing, and California's Office of Statewide Health Planning and Development, which publishes a database of all California hospital discharges (Vital and Health Statistics 2001; California Office of Statewide Health Planning and Development 2003; Scientific Registry of Transplant Recipients 2003).

Results

In 2002, we contacted more than 20 organizations; 18 publicly reported 333 measures of health care quality in California. Thirty-two (10%) of these were procedure-specific measures of surgical quality, which rated quality for 21 different procedures. All surgical quality measures described care at the hospital level; none reported information about specific health plans, medical groups, or surgeons. The 10 most common surgical procedures account for 22% of the procedures done in California; six of the 10 most common non-obstetric procedures had no associated quality measures.

The procedures for which measures were available include eight types of solid organ transplant, six cardiovascular surgeries, a group of interventional cardiology procedures, several orthopaedic procedures, one urologic procedure (prostatectomy), and surgery for oesophageal cancer. These procedures accounted for 331 693 (14%) of the 2381 601 procedures performed in California in 1999.

We divided the measures into structure, process, and outcome measures using definitions described by Donabedian (1980). Twelve procedures had associ-

ated structural measures. There were no process measures, and 20 procedures had associated outcome measures (Fig. 1). The one reported measure of health care structure was the number of procedures performed annually at a given hospital. This number was reported for 12 procedures (abdominal aortic aneurysm surgery, coronary bypass surgery, carotid endarterectomy, oesophageal cancer surgery, and eight types of solid organ transplant). These 12 procedures accounted for 74 176 surgeries in 1999, 3% of the California total for that year. No surgeon-specific volume measures were reported.

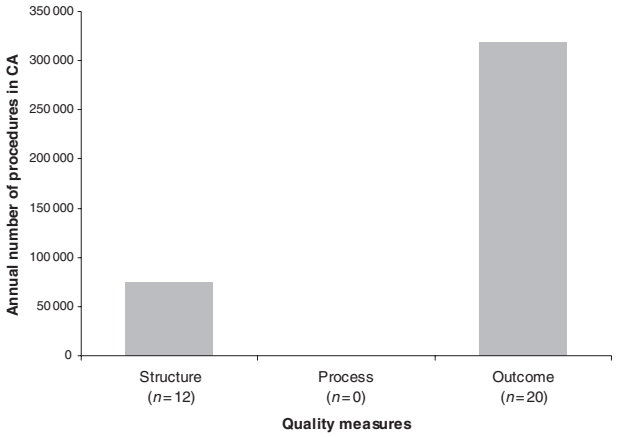

Figure 1 Annual number of surgical procedures by type of quality measure.

Outcome measures were of two types: death rates, and combined death and complication rates. Individual hospital death rates were reported for each type of solid organ transplant and for Coronary Artery Bypass Graft (CABG). For 11 procedures, in-hospital death rates and major complication rates were combined to produce a quality score. Major complications, defined by ICD-9CM codes, include such things as complication of an orthopaedic implant, stroke, cardiac arrest, excessive bleeding, and some infections. The nine procedures for which death rates were reported accounted for 58 716 surgical procedures in California in 1999 (2% of the total); combined morbidity and mortality measures covered 316 346 procedures (13% of the total).

All outcome measures were risk adjusted (Table 1). For solid organ transplants and for CABG,

Table 1 Surgical quality measure, procedure, and risk adjustment method

Measure	*Procedure*	*Risk adjustment method*
Structure		
Annual hospital volume	Repair abdominal aortic aneurysm	None
	Coronary bypass	None
	Carotid endarterectomy	None
	Oesophageal cancer surgery	None
	Solid organ transplant (eight types)	None
Process		
None		
Outcome		
Death rate	Coronary bypass	Medical record abstraction
	Solid organ transplant (eight types)	Medical record abstraction
Combined in-hospital death and complication rate	Percutaneous Transluminal Coronary Angioplasty	Administrative data
	Spinal fusion	Administrative data
	Back and neck surgery (other than fusion)	Administrative data
	Coronary bypass surgery	Administrative data
	Knee replacement (primary)	Administrative data
	Hip replacement (primary)	Administrative data
	Prostatectomy	Administrative data
	Peripheral Vascular Interventional Procedures	Administrative data
	Peripheral vascular bypass and revision	Administrative data
	Valve replacement	Administrative data
	Abdominal aortic aneurysm resect/replace	Administrative data

death rates were adjusted using data abstracted from medical records. Examples of medical record data used in adjustments for organ transplants include cold ischaemia time, medical condition at transplant (i.e. in hospital, in ICU, not hospitalized), and degree of HLA mismatch. For CABG, adjusters include New York Heart Association class, serum creatinine, and degree of coronary artery stenosis.

The remainder of the measures relied on administrative risk adjustment models, using ICD-9CM discharge codes to control for patient characteristics (Table 1). The risk adjustment process was reflected in reported death rates by providing observed and expected death rates. A lower observed than expected mortality indicates better quality and a higher observed to expected ratio worse quality. For quality scores (which integrate death and complication rates), risk adjustment was reflected by grouping hospitals into three categories. The category 'best' means that risk-adjusted outcomes were statistically significantly better than expected, 'as expected' means no statistically significant difference between observed and expected, and 'poor' means risk-adjusted outcomes are statistically significantly worse than expected (Healthgrades, Inc., copyright 1999–2004).

The 32 surgical quality measures we found were reported by five different organizations. There was some joint and duplicate reporting. For example, the Pacific Business Group on Health and the Office of Statewide Health Planning and Development (OSHPD) jointly report data on coronary artery bypass mortality.

Most current measures reflected care delivered between 2 and 5 years ago because measures other than annual volumes are based on reporting periods of 2–3 years. The shortest time lag between data collection and reporting was 6 months (for the eight solid organ transplant volumes). The remaining 24 measures were released between 18 and 36 months after the corresponding data collection ended.

Two groups of measures relied on voluntary reporting. As part of a statewide program, 79 California hospitals (out of 118 performing the procedure) voluntarily reported 41 data elements (including outcomes and risk factors) on each patient undergoing CABG. These elements were used to calculate quality measures related to coronary artery bypass surgery (Damberg *et al*. 2001). Hospitals performing transplants voluntarily reported organ-specific data for the 16 transplant outcome measures to the United Network for Organ Sharing (Scientific Registry of Transplant Recipients 2003).

The remainder of the measures relied on data reported under statutory or regulatory requirement. Hospitals receiving Medicare payments must report in-hospital death and complication rates for Medicare patients having a variety of procedures, and in California a state agency releases similar data for all health care payers. Complications reported in this database are limited to those with ICD-9-CM codes.

Three new measures are expected by 2005, and one current measure will change to mandatory reporting. By that date, Healthgrades will add a hospital-level measure related to carotid endarterectomy and OSHPD will begin to release hospital-specific data on mortality after hip fracture. Recent legislation requires that by 2004, all California hospitals performing CABG publicly report risk-adjusted mortality (e.g. reporting this measure will no longer be voluntary). By 2005 these hospitals must also report surgeon-specific mortality data, and OSHPD will publicly release these data (Table 2).

Discussion

We searched an existing database, surveyed organizations that report health care quality, and searched the Internet for measures related to surgical quality of care in California. We found relatively little data to aid a conscientious consumer interested in using objective quality measures to plan surgery, and the existing data had significant shortcomings. There were no data related to specific surgeons or medical groups; all measures were reported at the hospital level only. The majority of measures relied on administrative data for risk adjustment. There were no procedure-specific functional assessments or procedure-specific satisfaction scores reported (the Patient Evaluation of Performance in California survey assesses satisfaction with surgical care by hospital, but it is not procedure specific) (California HealthCare Foundation 2002). Few new measures will be added to the current set in the next several years.

Table 2 Organizations reporting procedure-specific surgical quality measures in California

Reporting organization	*Publicly reported measures*	*Additional measures to be reported by 2005*
United Network for Organ Sharing (UNOS)	16	0
Healthgrades	10	1
Pacific Business Group on Health (PBGH)	5*	0
Office of Statewide Health Planning and Development (OSHPD)	0	2
Consumers Checkbook	1	0
Total unique measures	32	3

*Excludes measures reported jointly with UNOS and OSHPD.

Overall, only 14% of surgical procedures performed in the state had associated measures. Many common procedures, including six of the 10 most common major non-obstetric surgical procedures performed in California, were not represented by any quality measures. Only three new measures are expected over the next 3 years.

Using Donebedian's conceptual model, we found measures of surgical quality pertaining to structure and outcome of care. Examples of structural measures include the type of specialists performing surgery, whether residents or trainees are involved in the procedure, and the number of procedures performed at a particular institution. Of these, only numbers of procedures are currently publicly reported in an organized fashion. Procedure-specific measures of hospital volume account for 12 of the 32 surgical quality measures reported in California.

A large body of research has examined the relationship between mortality and surgical volume (Birkmeyer *et al*. 2003). For some complex procedures that are infrequently performed, researchers have found an inverse association between volume and mortality (Begg *et al*. 1998; Dudley *et al*. 2000; Epstein 2002). Birkmeyer and colleagues recently found a statistically significant inverse relationship between hospital volume and mortality for 14 cardiovascular or cancer-related procedures (Birkmeyer *et al*. 2002). However, this association has not been consistently found in studies of the mildly to moderately complex procedures that comprise the majority of surgeries performed (Khuri *et al*. 1999). Even when the hospital-volume–mortality link exists, its explanatory power may be limited. For example, for some procedures volume explains as little as 18% of the variation in mortality (Hughes *et al*. 1987). Physician volume may also predict morbidity and mortality; some studies have found such a link, others have not (Hughes *et al*. 1987; Pellegrini *et al*. 1989; Mcgrath *et al*. 2000; Birkmeyer *et al*. 2003). If a consumer uses the available data to select a high-volume hospital, he or she cannot be guaranteed a high-volume surgeon.

For individuals facing a common procedure of low-to-moderate complexity, it is far from clear that selecting a high volume hospital or surgeon will improve morbidity or reduce mortality. Hospital volume may be more useful for making policy decisions regarding regionalization of very complex, infrequently performed procedures than for individual patient decision making (Birkmeyer *et al*. 2002; Epstein 2002). Dudley calculated that regionalizing surgery for 11 conditions could potentially reduce mortality for these procedures by 13–37% (Dudley *et al*. 2000). The Leapfrog Group, a US business group dedicated to improving health care quality, has proposed volume standards for five procedures, and its members intend to preferentially contract with health care organizations meeting those standards (Birkmeyer *et al*. 2000).

Structural factors other than volume have also been linked to differences in quality. Several studies have found lower mortality in university teaching hospitals, compared with affiliated teaching and non-teaching hospitals, although morbidity and length of stay may be greater (Hartz *et al*. 1989; Hutter *et al*. 2000; Yuan *et al*. 2000; Khuri *et al*. 2001). Whether a hospital is part of an academic medical centre or has a teaching program is publicly available information and could be useful to consumers. But structural

measures such as procedure volumes do not give a complete picture of surgical quality (Begg *et al*. 1998). More direct measures of quality include measures of what happens to patients (outcome measures) and assessments of what surgeons actually do to their patients (process measures).

Risk adjustment to account for inherent differences among patient populations is a critical step in developing sound outcome measures for surgery (Daley *et al*. 2001), and all reported outcomes were risk adjusted. There are many ways to risk adjust for surgical outcomes, and the method chosen may influence what is found: two studies using clinical risk factors did not find a link between a hospital's volume and its risk-adjusted mortality rate (Khuri *et al*. 1999; Damberg *et al*. 2001).

The most accurate methods of risk adjustment involve collecting clinical data. Systems such the US Veterans Administration (VA) National Surgical Quality Improvement Project (NSQIP) and the Physiological and Operative Severity Score for the enUmeration of Mortality and Morbidity (POSSUM) system from the UK are among the most accurate methods identified to date (Khuri *et al*. 1997; Copeland 2002). These data collection programs are labour intensive. NSQIP, for example, relies on a full-time nurse abstractor at each VA hospital.

Administrative data, while easier to collect, have not been as useful for predicting surgical outcomes (Best *et al*. 2002). Current coding practices make it impossible to separate pre-existing conditions from post-surgical complications. (California has recently introduced a sixth digit into its hospital discharge coding that asks whether the condition coded was present on admission or occurred during hospitalization. If the reliability and validity of that code can be demonstrated, then this limitation of administrative data may be overcome.) In addition, between-hospital variation in the conscientiousness of coding may affect quality scores that are risk-adjusted using administrative data. More thorough coding of comorbid conditions might increase the expected mortality rate for a procedure; more thorough coding of complications might increase the observed complication rate. We found that only those measures involving voluntary reporting (i.e. organ transplant and CABG) currently use clinical data for risk adjustment. The remainder use administrative data. This will change when, in 2005, California's mandatory CABG reporting system becomes active, as it requires clinical data collection.

Process quality measures may be more sensitive to differences in quality than are outcome measures. A smaller sample is needed to calculate statistically valid measures because process measures can focus on care that all or nearly all patients with a particular condition should have (e.g. whether patients get appropriate prophylactic antibiotics before surgery). Because smaller samples are adequate, process measures may be used to examine care delivered by individual physicians, rather than by groups of physicians. Furthermore, because process measures address what is actually done, they are less confounded by patients' disease characteristics and do not require extensive risk adjustments. Strong links between processes and outcomes have been documented for some conditions, allowing these measures to be used as proxies for outcomes of interest (Adler *et al*. 2000).

Despite these theoretical advantages, we found no publicly reported measures of process quality for surgical care. In many other areas of care, process measures predominate. For example, the Center for Medicare and Medicaid Services (CMS) Diabetes Quality Improvement Project and the National Committee on Quality Assurance Health Employer Data and Information Set (HEDIS) both use process measures almost exclusively. Efforts to better understand surgical quality, such as the VA NSQIP, are underway and may make measuring the quality of surgical processes more feasible in the future (Khuri *et al*. 2002).

How can consumers use the available data to better inform their choices about surgical care? They could use information that relates to overall care, such as whether a hospital is accredited by the Joint Commission, or how it scores on measures of patient satisfaction. However, a recent survey suggests that patients would prefer information related to their specific condition (Berry *et al*. 2001). For 86% of surgical patients in California, this type of condition-specific information does not exist. These patients can only rely on indirect measures, such as whether or not a hospital has a surgical training program, or on the results of patient experience surveys such as the Patients' Evaluation of Performance in California.

Even if a consumer has a condition for which data are reported, he or she will still face difficulties in using the data effectively. Until 2005, when surgeon-level data on coronary artery bypass become public in California, all surgical quality measures provide data at the hospital level only.

For most common procedures, morbidity and mortality are low and consumers might be better served by data on satisfaction or functional outcomes. Many such procedure-specific measures exist (e.g. the Western Ontario and McMaster University Osteoarthritis Index for functional outcomes after knee or hip replacement, and the Duke Activity Status Index for CABG), but we found no public reporting of these types of measures (Bellamy *et al*. 1988; Hlatky *et al*. 1989). Furthermore, none of the organizations surveyed had plans to release such measures by 2005.

Many procedures have only one related quality measure. For those few procedures that have more than one available measure, a consumer could try to integrate the procedure-specific results with more general information to arrive at a decision. That is, a consumer could try to have his or her procedure at an academic medical centre that has a high surgical volume and low risk-adjusted morbidity and mortality. For procedures with only one measure, it would seem reasonable to give weight to that measure, despite the limitations of administrative risk adjustment and of using volume as a proxy for quality. For the consumer who finds discordance among available measures, little guidance can be offered. No data exist to decide whether outcome measures that are risk-adjusted with administrative data are more valid indicators of overall surgical quality than are, for example, structural measures (e.g. volume).

We found only three new measures proposed for release by 2005. Quality measures take several years from conception to implementation, and any new measures to be publicly available by 2005 would have to be in the advanced stages of development at the time of our survey.

Our data have several shortcomings. We collected data in 2002 and have not repeated our survey. We focused our search on California, and our findings may not be representative of other states or of efforts at the national level. California makes public an all-payer hospital database that the majority of other states do not. In addition, because of high managed-care penetration, California may be a more 'mature' market in terms of health care quality measures. Both of these factors might increase the relative number of measures reported in California compared with other states.

Our method of surveying existing reporting organizations and searching the Internet for quality measures may have missed some measures. Our goal was to identify the data a conscientious person would find as he or she looked for health care information, and we conducted a more thorough search than even the most dedicated person might be expected to perform. We found no data on surgeon quality and only limited data on hospital quality for surgical services.

By 2005, the majority of California consumers facing surgery will be unable to find information to guide their decisions. Those consumers that do find information may not find it useful, particularly for procedures with low morbidity and mortality. Unless a concerted effort is undertaken to increase the number of conditions for which quality data are available, and to make available reliable measures of functional outcomes, most consumers will not be able to choose surgical providers based on important measures of quality. Thus, in the near term, improvements in surgical quality in California will need to depend on stimuli other than public release of data.

Acknowledgements

This work was supported in part by a grant from the California HealthCare Foundation and by the UCLA Building Interdisciplinary Research Careers in Women's Health Program.

References

Adler A.I., Stratton I.M., Neil H.A.W., Yudkin J.S., Matthews D.R., Cull C.A., Wright A.D., Turner R.C. & Holman R.R. (2000) Association of systolic blood pressure with macrovascular and microvascular complications of Type 2 diabetes: prospective observational study. *British Medical Journal* **321**, 412–419.

Australian Council for Safety and Quality in Health Care (2002) *Safety Through Action: Improving Patient Safety in Australia.* The Third Report to the Australian Health Ministers Conference. California Healthcare Foundation, Oakland, CA.

Begg C.B., Cramer L.D., Hoskins W.J. & Brennan M.F. (1998) Impact of hospital volume on operative mortality for major cancer surgery. *Journal of the American Medical Association* **280**, 1747–1751.

Bellamy N., Buchanan W.W., Goldsmith C.H., Campbell J. & Stitt L.W. (1988) Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *Journal of Rheumatology* **15** (12), 1833–1840.

Berry S., Spranca M. & Brown J. (2001) *Consumers and Health Care Quality Information: Need, Availability, Utility.* California Health Care Foundation, Oakland, CA.

Best W., Khuri S., Phelan M., Hur K., Henderson W., Demakis J. & Daley J. (2002) Identifying patient preoperative risk factors and postoperative adverse events in administrative databases. *Journal of the American College of Surgeons* **194**, 257–266.

Birkmeyer J.D., Birkmeyer C.M. & Wennberg D.E. (2000) *Leapfrog Standards: Potential Benefits of Universal Adoption.* The Leapfrog Group, Washington, DC.

Birkmeyer J.D., Siewers A.E., Finlayson E.V.A., Stukel T.A., Lucal F.L., Batista I., Welch G.H. & Wennberg D.E. (2002) Hospital volume and surgical mortality in the United States. *New England Journal of Medicine* **346**, 1128–1137.

Birkmeyer J.D., Stukel T.A., Siewers A.E., Goodney P.P., Wennberg D.E. & Lucas F.L. (2003) Surgeon volume and operative mortality in the United States. *New England Journal of Medicine* **349**, 2117–2127.

California HealthCare Foundation (2002) *What Patients Think of California Hospitals: Patients' Evaluation of Performance (PEP-C) Policy Report, The Quality Initiative* April 2002. Office of Statewide Health Planning and Development, Sacramento, CA.

California Office of Statewide Health Planning and Development (2003) Available at: http://www.oshpd.ca.gov/HQAD/HIRC/index.htm. Updated October 2003. Accessed February 18 2004.

Chassin M.R., Hanna E.L. & DeBuono B.A. (1996) Benefits and hazards of reporting medical outcomes publicly. *New England Journal of Medicine* **334**, 394–398.

Copeland G. (2002) The POSSUM system of surgical audit. *Archives of Surgery* **137**, 15–19.

Daley J., Henderson W.G. & Khuri S.F. (2001) Risk adjusted surgical outcomes. *Annual Review of Medicine* **52**, 275–287.

Damberg C.L., Chung R.E. & Steimle A. (2001) *The California Report on Coronary Artery Bypass Graft Surgery: 1997–1998 Hospital Data, Summary Report.* Pacific Business Group on Health and the California Office of Statewide Health Planning and Development, San Francisco, CA. (July 2001)

Department of Health (2000) *The NHS Plan: a Plan for Investment, a Plan for Reform*. Department of Health, London.

Donabedian A. (1980) *The Definition of Quality and Approaches to Its Assessment*. pp. 79–128. Health Administration Press, Ann Arbor, MI.

Dudley R.A., Johansen K.L., Brand R., Rennie D.J. & Milstein A. (2000) Selective referral to high volume hospitals: estimating potentially avoidable deaths. *Journal of the American Medical Association* **283**, 1159–1166.

Epstein A. (2002) Volume and outcome – it's time to move ahead. *New England Journal of Medicine* **346**, 1161–1163.

Giooum B.S., Graves E.J. & Kozak L.J. (1996) Trends in hospital utilization: United States, 1988–92. National Center for Health Statistics. Vital Health Stat, Series 13 No. 124.

Hartz A., Krakauer H., Kuhn E., *et al.* (1989) Hospital characteristics and mortality rates. *New England Journal of Medicine* **321**, 1720–1725.

Healthgrades. Inc., (Copyright 1999–2004) *The Healthcare Quality Experts.* Hospital Report Card Grading Methodology. Available at: http://www.healthgrades.com/public/index.cfm?fuseaction=mod&type=content&modact=Hrc_Methodology

Hlatky M., Boineau R., Higginbotham M., Lee K.L., Mark D.B., Califf R.M., Cobb F.R. & Pryor D.B. (1989) A brief self-administered questionnaire to determine functional capacity (the Duke Activity Status Index). *American Journal of Cardiology* **64**, 651–654.

Hughes R.G., Hunt S.S. & Luft H.S. (1987) Effects of surgeon volume and hospital volume on quality of care in hospitals. *Medical Care* **25**, 489–503.

Hutter M.M., Glasgow R.E. & Mulvihill S.J. (2000) Does the participation of a surgical trainee adversely impact patient outcomes? A study of major pancreatic resections in California. *Surgery* **128**, 286–292.

Khuri S.F., Daley J. & Henderson W.G. (2002) The comparative assessment and improvement of quality of surgical care in the Department of Veterans Affairs. *Archives of Surgery* **137**, 20–27.

Khuri S., Daley J., Henderson W., *et al.* (1997) Risk adjustment of the postoperative mortality rate for the comparative assessment of the quality of surgical care. *Journal of the American College of Surgeons* **185**, 315–327.

Khuri S.F., Daley J., Henderson W., *et al.*, and the Participants in the VA National Surgical Quality Improvement Program (1999) Relation of surgical volume to outcome

in eight common operations – results from the VA National Surgical Quality Improvement Program. *Annals of Surgery* **230** (3), 414–432.

Khuri S., Najjar S., Daley J., *et al.* (2001) Comparison of surgical outcomes between teaching and nonteaching hospitals in the Department of Veterans Affairs. *Annals of Surgery* **234**, 370–383.

Marshall M.N., Shekelle P.G., Leatherman S. & Brook R.H. (2000) The public release of performance data: what do we expect to gain? A review of the evidence. *Journal of the American Medical Association* **283** (14), 1866–1867.

Mcgrath P.D., Wennberg D.E., Dickens J.D., Siewers A.E., Lucas F.L., Malenka D.J., Kellet M.A. & Ryan T.J. (2000) Relation between operator and hospital volume and outcomes following percutaneous coronary interventions in the area of the coronary stent. *Journal of the American Medical Association* **284**, 3139–3144.

Pellegrini C.A., Heck C.F., Raper S. & Way L.W. (1989) An analysis of the reduced morbidity and mortality rates after pancreaticoduodenectomy. *Archives of Surgery* **124**, 778–781.

Scientific Registry of Transplant Recipients (2003) Available at: http://www.ustransplant.org/programs-report.pdf

Vital and Health Statistics (2001) Advance data from the National Health Interview Survey. *Centers for Disease Control and Prevention.* National Center for Health Statistics, Advance Data No 319, Hyattsville, MD.

Yuan Z., Cooper G.S., Einstadter D., Cebul R.D. & Rimm A.A. (2000) The association between hospital type and mortality and length of stay: a study of 16.9 million hospitalized Medicare beneficiaries. *Medical Care* **38** (2), 231–245.