# Quality Assessment in Surgery

## *Riding a Lame Horse*

*Daniel Dindo, MD, Dieter Hahnloser, MD, and Pierre-Alain Clavien, MD, PhD, FACS*

**Background:** Quality assessment in surgery is paramount for patients and health care providers. In our center, quality assessment is based on the recording of preoperative risk factors of each patient and a well-established grading system to track complications. Our prospective quality database is administrated by residents. However, the validity of such data collection is unknown.

**Methods:** To evaluate the validity of the recorded data, a specially trained study nurse audited our prospective quality database over a 6-month period. In the first 3 months, the audit was done in an undisclosed manner. Then, the audit was disclosed to the residents who were again subjected to a teaching course. Thereafter, the audit was continued in a disclosed manner for another 3 months, and data were compared between the 2 periods. Furthermore, we inquired about the strategies to assess surgical quality in 108 European medical centers.

**Results:** Surprisingly, residents failed to report most complications; 80% (164/206) and 79% (275/347; $P = 0.27$) of the negative postoperative events were not recorded during the first and the second period, respectively. When captured, however, grading of complications was correct in 97% of the cases. Moreover, comorbidities were incorrectly assessed in 20% of the patients in the first period and in 14% thereafter ($P = 0.07$). The survey disclosed that residents and junior staff are responsible of recording surgical outcome in 80% of the participating European centers.

**Conclusions:** Recording of outcome by surgical residents is unreliable, despite active and focused training. Hence, surgery should be evaluated by dedicated personnel.

(*Ann Surg* 2010;251: 766–771)



Reliable outcome reporting is essential to assess and compare quality in surgery. Although there is a wide range of outcome measures used in medical published data, there is currently no consensus on which data should be collected for quality assessment, as well as how and by whom quality should be measured. Repeatedly, significant variance in morbidity has been reported hampering interpretation of surgical outcome reports. For example, morbidity after pancreas surgery (Whipple's procedure) was reported to range between 18% and 72%.[1,2] Differences in case-mix,[3] outcome definitions,[4] and observation time[5] are responsible for the diverging results in surgical literature. Remarkably, the reliability of data collections has rarely been studied as potential explanation of differences in reported outcome.

From the Department of Surgery, University Hospital, Raemistr. 100, CH-8091 Zurich, Switzerland.

This study was presented at the 15th annual meeting of the European Surgical Association (ESA), April 18-19, 2008, Venice, Italy.

Reprints: Pierre-Alain Clavien, MD, PhD, FACS, Department of Surgery, University Hospital, Raemistr. 100, 8091 Zurich, Switzerland. E-mail: clavien@access.uzh.ch.



ISSN: 0003-4932/10/25104-0766
DOI: 10.1097/SLA.0b013e3181d0d211

In the United States, large databases such as the National Surgical Quality Improvement Program (NSQIP) were established in Veteran Affairs Hospitals to record risk-adjusted surgical outcome, to rate hospital quality, and to benchmark surgical performance. By tracking hospital's performance, surgical morbidity and mortality could be decreased.[6] However, the annual costs for data collection and the fear of medicolegal consequences have hampered the introduction of this or similar programs beyond selected medical centers.[7] Instead, administrative databases are being increasingly used in outcome research.[8,9] However, such administrative quality databases are often unreliable[9,10] as they have shown significant variance compared with clinical databases.[9–12] Clinical databases, that is, those designed and controlled by the physicians, on the other hand, might underreport complications,[13] and their validity is unknown. Nevertheless, such clinical databases are currently used for benchmarking and marketing in many hospitals.

In our center, we have developed a particular interest in outcome research for many years.[14–20] In 2004, we proposed a classification of surgical complications allowing objective assessment of adverse events,[15] usually reported as "the Clavien classification," or the "Clavien–Dindo classification," which is now well implemented in our center as a means of quality control. This classification is also increasingly used in the surgical literature, with currently more than 330 citations since its introduction (available at: www.surgicalcomplication.info).[21–34] Ranking of each complication in our center is presented at the weekly Morbidity and Mortality conference, and the outcome of each patient is recorded electronically in a clinical database that is administrated by residents. Because the reliability of such clinical databases is largely unknown, we designed a study to audit our outcome database using a specially trained study nurse. The main purpose of this study was to assess the reliability of our residents in tracking complications after surgery. Furthermore, we inquired through an international survey, targeted on members of 2 leading European surgical societies, how other centers assess their quality.

## METHODS

### Study Design

The study was approved by the ethics committee at the University of Zurich, and was registered at clinicaltrials.gov (NCT 00548535). The study was divided into 2 periods each lasting 3 months (Table 1). In the first period, a specially trained study nurse, not being involved in the primary care process of the patients, audited our outcome data as recorded by residents in an undisclosed manner. In this database, all complications are recorded and graded according to the Clavien–Dindo classification,[15] as well as all comorbidities using the Charlson Risk Index (CRI[35]; see later). The focus of the audit was laid on the completeness and accuracy of the recording of postoperative complications and their respective grading, and on the accuracy of the recorded CRI.[35] Inconsistencies between recorded and audited data were evaluated by the study nurse and the authors. We defined this approach to be the "gold-standard."

**TABLE 1.** Study Design

| | Period 1 | | Period 2 |
|---|---|---|---|
| Task | Undisclosed audit | Analysis of data<br>Disclosure and discontinuation of audit<br>Discussion of performance<br>Renewal of teaching | Restarting disclosed audit |
| Action | Revision of resident-administered database by study nurse<br>-Assessment of complications<br>-Accuracy of complication grading[15]<br>Accuracy of Charlson Risk Index[30] | Residents were individually contacted and performance was displayed<br>-Illustration of Charlson Risk Index[30] by examples<br>-Illustration of Clavien–Dindo Complication Classification[15] by examples<br>-Designation of an outcome adept to be contacted in case obscurities | Same as in period 1 |

After the first period, the results of the audit were disclosed to the residents, and they were informed individually about their performance. During this time, the audit was discontinued. During this phase, the authors held an additional teaching course compulsory for all residents. The importance of reliable quality assessment was again emphasized and the CRI[35] as well as the Clavien–Dindo complication classification[15] were illustrated by examples. In the second period, the audit was renewed and continued for another 3 months in a disclosed manner.

Finally, a survey was sent to members of 2 major European surgical societies, the European Surgical Association (ESA) and the European Hepato-Pancreato-Biliary Association (EHPBA). Of note, the members of the ESA include the continent's most prominent surgeons leading academic medical institutions. Members were contacted by email by the secretary general of the respective societies. The survey focused on the local practice in capturing surgical quality (see the proceeding paragraphs).

## Prospective Database

Each complication including its respective grading and the comorbidities as assessed by the Charlson Risk Index (CRI)[35] was entered in the database by residents responsible for the primary care of the patients. The Charleston Risk Index (CRI) ranges from 0 to 37 points. Mortality after surgery has been shown to be doubled in patients with a score $\geq$3 points compared with those displaying no comorbidity.[36,37]

Complications are stratified by seriousness and classified into 5 grades.[15] Briefly, grade I includes complications without need for special therapy, such as a small hematoma or a small pneumothorax. Grade II complications need pharmacological treatment (eg, a pneumonia requiring antibiotics). Complications requiring endoscopic, radiologic, or surgical intervention are classified as grade III, either IIIa without or IIIb with the need for general anesthesia. Grade IV complications are life-threatening complications requiring management in an intermediate or intensive care unit; one organ dysfunctions are classified as IVa, whereas multiorgan dysfunctions are grade IVb complications. Death of the patient is considered as a grade V complication. In the database, explanations regarding this classification are listed with the frequent examples to facilitate the data collection. In addition, all complications and their classification are listed and briefly discussed in our weekly Morbidity and Mortality conference.

The extensiveness of surgery is also recorded for every procedure according to a previously described system[38]; type I surgery comprises of procedures without opening of the abdominal cavity (eg, inguinal hernia operation), type II procedures correspond to abdominal operations (eg, gallbladder and colon surgery) without retroperitoneal and liver surgery, and major abdominal operations involving retroperitoneal organs (pancreas, rectum, esophagus), and liver surgery are defined as type III surgery.

## International Survey

The survey inquired (*a*) whether all postoperative complications and comorbidities are assessed, (*b*) whether complications/comorbidities are assessed by diagnosis, and/or by using a grading system, and (*c*) by whom (students, residents, junior staff, consultants, study nurses) data are collected. A total of 604 different centers were contacted by emails.

## Statistical Analyses

Chi square test, Mann-Whitney *U* test, and student's *t* test were used where appropriate. Normally distributed data are given as mean ($\pm$ standard deviation) and median (range) is given for not normally distributed data. For all analyses, $P < 0.05$ was considered significant. Statistical package for the Social Sciences (SPSS, version 12.0, Chicago, IL) was used for all analyses.

# RESULTS

## Which Patients Were Analyzed?

A total of 305 patients were included during the first period and 447 patients during the second period. There were 335 type I operations (surgery without opening of abdominal cavity), 319 type II (abdominal surgery without retroperitoneal and liver surgery), and 98 type III procedures (retroperitoneal and liver surgery). There was no significant difference between the 2 study periods regarding these types of operation, age, and gender of the patients, ASA, body mass index, and length of hospitalization (Table 2).

**TABLE 2.** Patient Demographics

| | First Period n = 305 | Second Period n = 447 | *P* |
|---|---|---|---|
| Age (yr) | 50 (18–90) | 51 (18–86) | 0.71 |
| Gender (female; %) | 47.2 | 47.5 | 0.92 |
| BMI (kg/m$^2$) | 24.9 (13.1–58.0) | 25·0 (11.2–60.6) | 0.47 |
| ASA 3/4 (%) | 31.8 | 36.3 | 0.06 |
| Type of surgery | | | 0.10 |
| Type I (%) | 44 | 45 | |
| Type II (%) | 45 | 39 | |
| Type III (%) | 11 | 16 | |
| Length of hospitalization (d) | 5 (1–171) | 5 (1–89) | 0.63 |

## Did Residents Adequately Assess Outcome Data and Comorbidities?

A total of 206 complications occurred during the first period and 80% of these complications were not recorded (162/206). Of grade I complications (without need for further treatment), 94% (110/117) were not recorded by residents, of grade II complications (requiring drugs) 54% (30/56), and of grade III complications (requiring surgical, endoscopic, or radiologic intervention) 71% (22/31). Grade IV (requiring intensive care; n = 1) and grade V complications (death of the patient, n = 1) were not documented in the database (Fig. 1). However, when recorded, all complications (n = 44) were correctly graded.

During the first period, 57% (173/305) of patients had comorbidities (CRI > 0). In 20% of patients (62/305), comorbidities were not correctly assessed. Eighteen percent of the patients had a CRI ≥3, whereas the residents recorded a CRI ≥3 in 16% of the patients ($P = 0.52$). Of note, we found no difference between junior (<3 years of training) and senior residents (>3 years of experience) in respect to the quality of the complication and comorbidity assessment ($P = 0.54$ and $P = 0.38$, respectively).

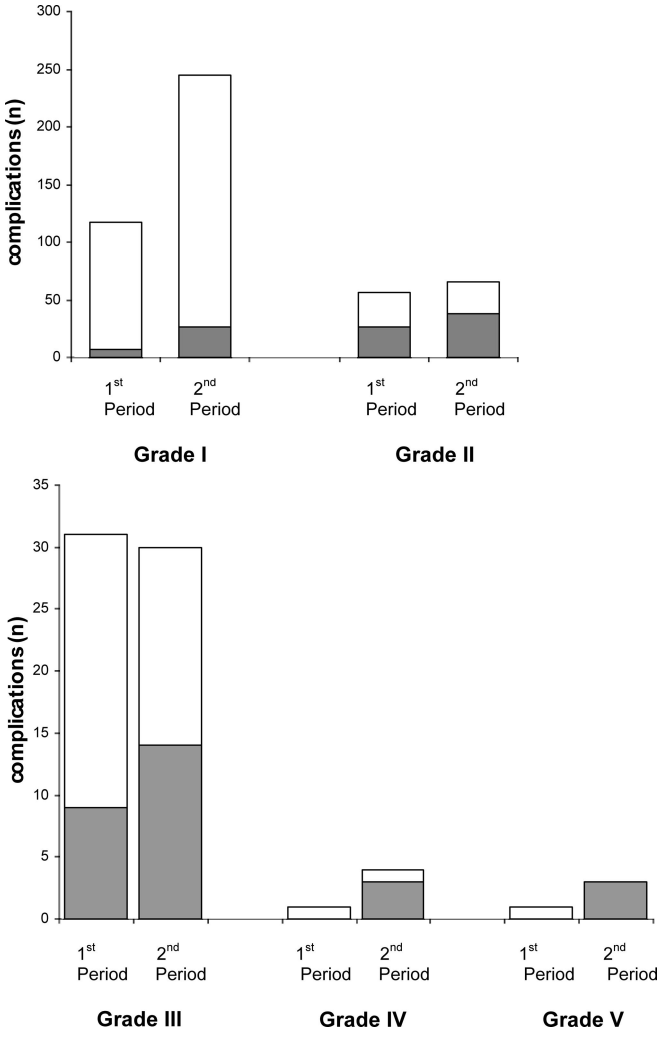


**FIGURE 1.** Numbers of recorded and missed complications during the first period of the study (undisclosed data revision) and during the second period (after disclosure of data revision) (□missed data; ■ recorded data).

## Did Assessment of Outcome Data and Comorbidities Improve After Teaching?

In the second period, 347 complications occurred. Surprisingly, quality recording did not improve since only 75 complications were recorded. Of the complications, 79% were still not or not correctly assessed (275/347; $P = 0.27$ compared with the first period); 89% (219/245; $P = 0.11$) of grade I complications were not or not properly recorded, 59% of grade II complications (38/65; $P = 0.36$), 47% of grade III complications (14/30; $P = 0.05$), and one-fourth grade IV complications (1/4; $P = 0.60$). All grade V complications were recorded (3/3; $P = 0.25$) (Fig. 1). Focusing on clinically relevant complications (grade II and higher), there was a marginal improvement in the second period with 52% (53/102) of the complications that were missed compared with 61% (54/89) in the first period. However, this difference did not reach statistical significance ($P = 0.18$). Of note, only 3 of the 75 (4%) recorded complications were incorrectly graded.

In 55% of the patients, comorbidities were recorded in the second period of the study (246/447). In 14% (64/447) of the patients, comorbidity recording was incorrect ($P = 0.07$ compared with the first period). Of all patients, 23% had a CRI ≥3, while the residents did only record such a CRI in 18% of the patients ($P = 0.08$); this discrepancy was not significantly different compared with the first study period ($P = 0.98$).

Again, the level of experience (<3 and >3years of surgical training, respectively) did not significantly influence the quality of data collection ($P = 0.76$ for complication recording and $P = 0.40$ for assessment of comorbidity, respectively).

## How Are Complications and Comorbidities Recorded in Other Centers?

Surgeons from 108 European centers answered the survey. The response rate to the questionnaire was 18% (108/604). Of the 135 active members of the ESA, 52 answered the survey (response rate of 39%) and of the 469 members of the EHPBA, 56 answered the survey (response rate of 12%). A total of 63 answers were obtained from Central Europe (Germany, United Kingdom, France, the Netherlands, Belgium, Austria, Denmark, Switzerland), 22 from Southern Europe (Italy, Spain, Greece), 12 from Scandinavia (Norway, Sweden, Finland), and 11 from Eastern Europe (Turkey, Hungary, Czech Republic, Romania, Poland, Russia).

Eighty-one percent (87/108) of the European centers claimed that they routinely record all their complications. Of those, only in one-third of the centers (25/87) a grading system such as our 5-scale classification[15] or another local system, was used to assess the seriousness of the complications. In two-thirds of the centers (57/87; 66%), which assess complications on a routine basis, complications are tracked by residents, and in 11% (10/87) by junior staff (Table 3).

The recording of comorbidity allowing risk-adjustment of outcome data is routinely performed in only 57% (62/108) of the participating centers. In only 43% (27/62), a grading system, such as Physiologic and Operative Severity Score for the enUmeration of Mortality and morbidity (POSSUM), American Society of Auesthesiologists Score (ASA), or CRI, is used. Of note, data are collected by residents and junior staff in 78% (48/62) of the centers (Table 3).

## DISCUSSION

Quality control in surgery is gaining increasing attention among health care professionals, the public, and the media. Tracking hospital performance is an essential step to reduce surgical morbidity and mortality, but only reliable data may enable conclusive comparison of outcome in a single center over time or among different centers. Our survey disclosed that in most of the partici-

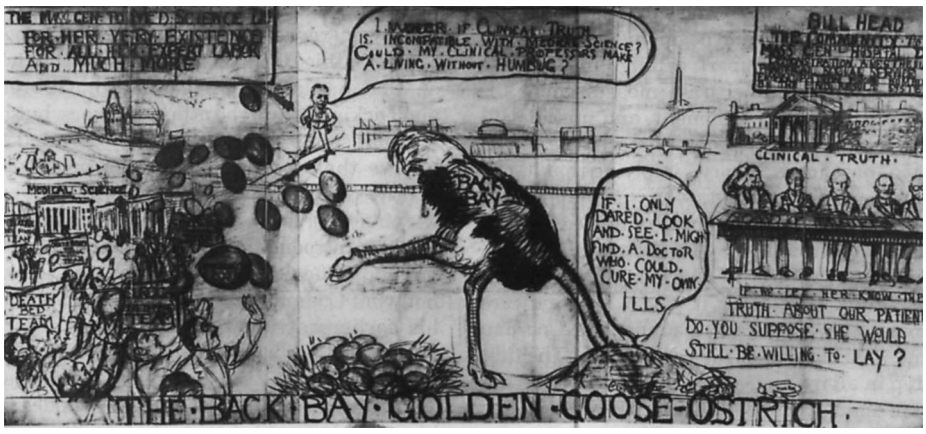


**FIGURE 2.** Cartoon displayed by Codman at the Meeting of the Surgical Section of the Suffolk District Medical Society on January 6, 1915, where he spoke about hospital efficiency. Courtesy of the Harvard Medical Library in the Francis A. Countway Library of Medicine.

pating European centers, quality data are gathered by residents. This situation is worrisome because we demonstrate that outcome assessment by residents is largely unreliable. Therefore, self-reported quality reports in Europe are likewise largely inaccurate. The interpretation of surgical quality reports is further hampered since more than half of the centers do not adjust their outcome data for the case-mix, as they do not record comorbidities.

In this study, we could highlight that data collection by residents is not suitable for quality control. Strikingly, the reliability of the collected data did not improve despite of the additional teaching, and despite of the disclosure of the audit. The reason for this dramatic underreporting of complications is likewise multifactorial. First, recording of outcome data is time-consuming and might therefore be neglected by the residents. The restriction of the weekly working hours (restricted to 50 hours in Switzerland, and to less in many other European countries) may also significantly impede reliable data collection. For example, information regarding outcome might be lost because of transitions of care that are adherent to the restriction of the working time. Second, lack of incentives may partly explain the insufficient data collection by the residents; comprehensive data collection is not rewarded and on the other hand, unreliable data collection has no negative consequences. This lack of motivation also holds for the hospital as a whole, since there is no apparent monetary benefit for the institution to collect such data. In contrast, the additional load for data collection is not reimbursed by payers in most systems. The third explanation might be that surgeons in general are keener on focusing on their core business, the work in the operating room, than on administrative duties, which may unveil their own complication rates possibly pointing out poor quality.

Although data collectors have been suggested to be a crucial factor for the quality of registers in surgery in some reports,[13,39] little is known about their effect on the reliability of outcome assessment. From experience in North America, there is evidence that nonclinicians are better data collectors than clinicians[40]; this is strongly supported by our study. Morbidity and mortality, as identified by physicians, was also considerably underreported in another study when compared with NSQIP data.[41] In that study, 1 of 2 deaths were not registered and the physicians did not identify 3 of 4 complications. A similar rate of unreported complications was also documented in another study.[42] Based on our survey, we might assume that surgical quality is tracked by clinicians in more than 90% of the medical centers in Europe with most of the of the centers relying on residents for their data collection. Taking into account the data of our study, the situation concerning surgical quality assessment in Europe might be alarming.

Reliable outcome data are crucial to improve surgical performance, to gain data for benchmarking, to assess the safety of new procedures and devices, and to inform the patients about the risk of surgery. For example, performance data are reported to the public in many states in the United States[43] and the United Kingdom (eg, available at: www.scts.org) to provide objective data for the patients to select the "best" hospital for their conditions. Health policy makers point out that the availability of comparative data on individual hospital's and physician's performance may contribute to limit the costs of health care, while improving quality.[44] In addition, insurers actively look for reliable outcome data for selective contracting to steer patients to hospitals and to surgeons, which have the best results.[45]

Quality registers are tools for detecting variations in healthcare,[46] but reliable data are fundamental to reach this purpose.[13] In the United States, large clinical databases have been established to portray performance of health care providers. The NSQIP data are gathered from dedicated, trained surgical nurses in more than 150 hospitals.[47] This and other large databases[48] are continuously audited on many levels for completeness and accuracy. The reliability of such data has recently been demonstrated with an error rate of less than 1% for major complications compared with the source data.[49] However, costs are a major drawback for many centers to join such large registries. For example, the required financial commitment to participate in the NSQIP is $100,000 annually.[7] Therefore, most US centers rely on internal complication assessment for quality assurance and clinical research, and do not participate in audited centralized or collaborative systems. In Europe, large databases similar to the NSQIP program are not yet available.

Some limitations of this study should be discussed. First, teaching of the residents was not undertaken by independent outcome specialists, but by the authors themselves. However, the authors are experienced in outcome research,[14–20] as they have defined the Clavien complication classification,[15] which is used in their quality database. The observation that the complications were correctly graded in almost all cases, also in the initial phase of the study, demonstrates that the residents were well trained in the assessment of surgical outcome. Second, the value of scheduled teaching sessions to improve reliability of residents in collecting data might be doubtful, as suggested in a study in cardiac surgery showing that a large clinical database could only be marginally improved by teaching and feedback mechanisms.[50] Third, the reteaching consisted only of a single session focusing on the importance of consistent data collection for quality assessment and later studies. Whether longer or serial courses would have been more effective is questionable. We rather believe that the poor reliability

**TABLE 3.** Data Collectors of Complications and Comorbidities

| | Complications | | Comorbidities | |
|---|---|---|---|---|
| | ESA n = 52 | EHPBA n = 56 | ESA n = 52 | EHPBA n = 56 |
| Study nurse/data manager (%) | 3 | 9 | 7 | 8 |
| Resident (%) | 72 | 63 | 76 | 64 |
| Junior staff/fellow (%) | 5 | 14 | 7 | 16 |
| Attending (%) | 20 | 8 | 10 | 8 |
| Head (%) | 0 | 6 | 0 | 4 |

Data are based on the international survey (n = 108).

ESA indicates European Surgical Association; EHPBA, European Hepato-Pancreatic-Biliary Association.

of residents in recording this data relies on external factors, as discussed earlier, such as the restriction of working hours, which is regulated through European laws, and therefore, active in most countries in this area. In addition, whether additional incentives (eg, financial) or punishment (eg, restriction to the operating room) may effect on the resident is unknown, but those strategies are unlikely to be accepted in many systems. Finally, the findings of the survey may not be applicable to other nonparticipating European centers because of the rather low overall response rate (18%); however, the response rate was much better for the ESA (39%) that includes mostly surgeons from academic centers, often chairs of large surgical services. To explain the overall poor response rate in nonacademic centers, we are tempted to speculate that those centers might have been reluctant to disclose the lack of quality control in their places since the survey was not anonymous. If this assumption is correct, then the situation in Europe concerning insufficient quality control might be even worse than revealed by the survey.

In summary, this study points to major problems in recording negative outcome in surgery in Europe relying on residents who are restricted with their working time and might also be unmotivated to collect reliable data. We would like to conclude with the maxim from one of the fathers of surgical outcome research, Dr. Ernest Amory Codman (1869–1940), which unfortunately is still current today: "Many surgical leaders behave as ostriches, but it is time to get the head out of the sand" (Fig. 2). Indeed, we urgently need to develop new strategies to assess surgical quality using dedicated and appropriately trained personnel. Although it is important that residents understand how to assess and document quality of care issues, quality assessment must not be at their sole responsibility. Furthermore, institutions must be audited for the accuracy of their databases. Until this is achieved, quality assessment will remain a lame horse that will not bring us forward to better surgical quality and finally, we will lose credibility among other colleagues, health care providers, and payers.

## REFERENCES


1. Trede M, Schwall G, Saeger HD. Survival after pancreatoduodenectomy. 118 consecutive resections without an operative mortality. *Ann Surg*. 1990;211:447–458.
2. Seiler CA, Wagner M, Sadowski C, et al. Randomized prospective trial of pylorus-preserving vs. classic duodenopancreatectomy (Whipple procedure): initial clinical results. *J Gastrointest Surg*. 2000;4:443–452.
3. Martin RC II, Brennan MF, Jaques DP. Quality of complication reporting in the surgical literature. *Ann Surg*. 2002;235:803–813.
4. Bruce J, Russell EM, Mollison J, et al. The measurement and monitoring of surgical adverse events. *Health Technol Assess*. 2001;5:1–194.
5. Macintyre IM, Akoh JA. Improving survival in gastric cancer: review of operative mortality in English language publications from 1970. *Br J Surg*. 1991;78:771–776.
6. Rowell KS, Turrentine FE, Hutter MM, et al. Use of national surgical quality improvement program data as a catalyst for quality improvement. *J Am Coll Surg*. 2007;204:1293–1300.
7. Miller DC, Filson CP, Wallner LP, et al. Comparing performance of morbidity and mortality conference and national surgical quality improvement program for detection of complications after urologic surgery. *Urology*. 2006;68:931–937.
8. Fiscella K, Franks P, Meldrum S, et al. Racial disparity in surgical complications in New York state. *Ann Surg*. 2005;242:151–155.
9. Bertges DJ, Shackford SR, Cloud AK, et al. Toward optimal recording of surgical complications: concurrent tracking compared to the discharge data set. *Surgery*. 2007;141:19–31.
10. 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;194:257–266.
11. Mack MJ, Herbert M, Prince S, et al. Does reporting of coronary artery bypass grafting from administrative databases accurately reflect actual clinical outcomes? *J Thorac Cardiovasc Surg*. 2005;129:1309–1317.
12. Shahian DM, Silverstein T, Lovett AF, et al. Comparison of clinical and administrative data sources for hospital coronary artery bypass graft surgery report cards. *Circulation*. 2007;115:1518–1527.
13. Gunnarsson U, Seligsohn E, Jestin P, et al. Registration and validity of surgical complications in colorectal cancer surgery. *Br J Surg*. 2003;90:454–459.
14. Dindo D, Clavien PA. What is a surgical complication? *World J Surg*. 2008;32:942–944.
15. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann Surg*. 2004;240:205–213.
16. Clavien PA, Dindo D. Surgeon's intuition: is it enough to assess patients' surgical risk? *World J Surg*. 2007;31:1909–1911.
17. Clavien PA, Sanabria JR, Mentha G, et al. Recent results of elective open cholecystectomy in a North American and a European center. Comparison of complications and risk factors. *Ann Surg*. 1992;216:618–626.
18. Clavien PA, Sanabria JR, Strasberg SM. Proposed classification of complications of surgery with examples of utility in cholecystectomy. *Surgery*. 1992;111:518–526.
19. Clavien PA, Camargo CA Jr, Croxford R, et al. Definition and classification of negative outcomes in solid organ transplantation. Application in liver transplantation. *Ann Surg*. 1994;220:109–120.
20. Clavien P.-A., de Oliveria M, et al. The Clavien–Dindo Classification of surgical complications: five-year experience. *Ann Surg*. 2009;250:187–196.
21. Laurent C, Leblanc F, Bretagnol F, et al. Long-term wound advantages of the laparoscopic approach in rectal cancer. *Br J Surg*. 2008;95:903–908.
22. Lagarde SM, Reitsma JB, Maris AK, et al. Preoperative prediction of the occurrence and severity of complications after esophagectomy for cancer with use of a nomogram. *Ann Thorac Surg*. 2008;85:1938–1945.
23. Tarantino I, Hetzer FH, Warschkow R, et al. Local excision and endoscopic posterior mesorectal resection versus low anterior resection in T1 rectal cancer. *Br J Surg*. 2008;95:375–380.
24. Patel S, Cassuto J, Orloff M, et al. Minimizing morbidity of organ donation: analysis of factors for perioperative complications after living-donor nephrectomy in the United States. *Transplantation*. 2008;85:561–565.
25. Permpongkosol S, Link RE, Su LM, et al. Complications of 2,775 urological laparoscopic procedures: 1993 to 2005. *J Urol*. 2007;177:580–585.
26. Yi NJ, Suh KS, Cho JY, et al. Three-quarters of right liver donors experienced postoperative complications. *Liver Transpl*. 2007;13:797–806.
27. Tamura S, Sugawara Y, Kaneko J, et al. Systematic grading of surgical complications in live liver donors according to Clavien's system. *Transpl Int*. 2006;19:982–987.
28. Khalaf H, Al-Sofayan M, El-Sheikh Y, et al. Donor outcome after living liver donation: a single-center experience. *Transplant Proc*. 2007;39:829–834.
29. Patel S, Orloff M, Tsoulfas G, et al. Living-donor liver transplantation in the United States: identifying donors at risk for perioperative complications. *Am J Transplant*. 2007;7:2344–2349.
30. Tefekli A, Ali Karadag M, Tepeler K, et al. Classification of percutaneous nephrolithotomy complications using the modified clavien grading system: looking for a standard. *Eur Urol*. 2008;53:184–190.
31. Vollmer CM Jr, Pratt W, Vanounou T, et al. Quality assessment in high-

acuity surgery: volume and mortality are not enough. *Arch Surg*. 2007;142: 371–380.

32. Atug F, Castle EP, Srivastav SK, et al. Prospective evaluation of concomitant lymphadenectomy in robot-assisted radical prostatectomy: preliminary analysis of outcomes. *J Endourol*. 2006;20:514–518.
33. Kaul S, Savera A, Badani K, et al. Functional outcomes and oncological efficacy of Vattikuti Institute prostatectomy with veil of aphrodite nerve-sparing: an analysis of 154 consecutive patients. *BJU Int*. 2006;97:467–472.
34. Gonzalgo ML, Pavlovich CP, Trock BJ, et al. Classification and trends of perioperative morbidities following laparoscopic radical prostatectomy. *J Urol*. 2005;174:135–139; discussion 139.
35. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40:373–383.
36. Birim O, Kappetein AP, Takkenberg JJ, et al. Survival after pathological stage IA nonsmall cell lung cancer: tumor size matters. *Ann Thorac Surg*. 2005;79:1137–1141.
37. Rieker RJ, Hammer E, Eisele R, et al. The impact of comorbidity on the overall survival and the cause of death in patients after colorectal cancer resection. *Langenbecks Arch Surg*. 2002;387:72–76.
38. Dindo D, Muller MK, Weber M, et al. Obesity in general elective surgery. *Lancet*. 2003;361:2032–2035.
39. Maudsley G, Williams EM. What lessons can be learned for cancer registration quality assurance from data users? Skin cancer as an example. *Int J Epidemiol*. 1999;28:809–815.
40. Russell EM, Bruce J, Krukowski ZH. Systematic review of the quality of surgical mortality monitoring. *Br J Surg*. 2003;90:527–532.
41. Hutter MM, Rowell KS, Devaney LA, et al. Identification of surgical complications and deaths: an assessment of the traditional surgical morbidity and mortality conference compared with the American College of Surgeons-National Surgical Quality improvement program. *J Am Coll Surg*. 2006;203: 618–624.
42. Feldman L, Barkun J, Barkun A, et al. Measuring postoperative complications in general surgery patients using an outcomes-based strategy: comparison with complications presented at morbidity and mortality rounds. *Surgery*. 1997;122:711–719; discussion 719–720.
43. Dimick JB, Welch HG, Birkmeyer JD. Surgical mortality as an indicator of hospital quality: the problem with small sample size. *JAMA*. 2004;292:847–851.
44. Daley J, Henderson WG, Khuri SF. Risk-adjusted surgical outcomes. *Annu Rev Med*. 2001;52:275–287.
45. Birkmeyer NJ, Share D, Campbell DA Jr, et al. Partnering with payers to improve surgical quality: the Michigan plan. *Surgery*. 2005;138:815–820.
46. Berwick DM. Controlling variation in health care: a consultation from Walter Shewhart. *Med Care*. 1991;29:1212–1225.
47. Fink AS, Itani KM, Campbell DC Jr. Assessing the quality of surgical care. *Surg Clin North Am*. 2007;87:837–852, vi.
48. Ferguson TB Jr, Dziuban SW Jr, Edwards FH, et al. The STS national database: current changes and challenges for the new millennium. Committee to establish a national database in cardiothoracic surgery, the society of thoracic surgeons. *Ann Thorac Surg*. 2000;69:680–691.
49. Herbert MA, Prince SL, Williams JL, et al. Are unaudited records from an outcomes registry database accurate? *Ann Thorac Surg*. 2004;77:1960–1964; discussion 1964–1965.
50. Fine LG, Keogh BE, Cretin S, et al. How to evaluate and improve the quality and credibility of an outcomes database: validation and feedback study on the UK cardiac surgery experience. *BMJ*. 2003;326:25–28.