Measuring Errors and Adverse Events in Health Care

Eric J. Thomas, MD, MPH, Laura A. Petersen, MD, MPH

In this paper, we identify 8 methods used to measure errors and adverse events in health care and discuss their strengths and weaknesses. We focus on the reliability and validity of each, as well as the ability to detect latent errors (or system errors) versus active errors and adverse events. We propose a general framework to help health care providers, researchers, and administrators choose the most appropriate methods to meet their patient safety measurement goals.

KEY WORDS: **medical error; adverse events; patient safety; measurement.**
J GEN INTERN MED 2003;18:61–67.

Researchers and quality improvement professionals in the United States,[1] Great Britain,[2] and elsewhere are mobilizing to reduce errors and adverse events in health care. But which measurement methods should be used to demonstrate reductions in error and adverse event rates, and what are their strengths and weaknesses? Recent articles about errors and adverse events,[3–9] have only superficially addressed measurement issues. Other articles have debated the applicability of principles of evidence-based medicine to the study of patient safety[10,11] but have not provided a broad framework for understanding measurement issues.

Our goal is to present a conceptual model for measuring latent errors, active errors, and adverse events. In addition, we identify some of the error and adverse event measurement methods commonly used in health care, and discuss the strengths and weaknesses of these methods. Researchers and individuals interested in quality improvement can use our model to help select the most appropriate measurement methods to meet their goals and to evaluate the methods used by others.

We have not conducted an exhaustive review of the literature to identify all error and adverse event measurement methods used in health care, nor do we attempt to review all methods used in other industries. Instead, we discuss what we believe to be commonly used methods for measuring errors and adverse events in health care. We believe that measurement methods we do not discuss, or ones developed in the future, can be placed into our conceptual model. Hopefully, the model will assist researchers and quality improvement professionals in understanding the strengths and weaknesses of various measurement methods and in choosing the appropriate measurement method(s) for their goals.

Received from The Center for Clinical Research and Evidence Based Medicine, Division of General Medicine, and Department of Medicine at The University of Texas Houston Medical School (EJT), Houston Tex; and the Houston Center for Quality of Care and Utilization Studies, a Health Services Research and Development Center of Excellence, Houston VA Medical Center, and Section for Health Services Research, Baylor College of Medicine (LAP), Houston, Tex.

Address correspondence and requests for reprints to Dr. Thomas: The University of Texas Houston Medical School, 6431 Fannin MSB 1.122, Houston TX 77030 (e-mail: eric.thomas@uth.tmc.edu).

DEFINITIONS

In this paper, we use the phrase *errors and adverse event*s to encompass a number of commonly used terms. We use the word *error* to include terms such as mistakes, close calls, near misses, active errors, and latent errors. The term *adverse events* includes terms that usually imply patient harm, such as medical injury and iatrogenic injury. We believe the phrase *errors and adverse events* is useful for this paper because *errors*, as defined by Reason,[12] do not necessarily harm patients, whereas the term *adverse event* does imply harm. Together, the phrase *errors and adverse events* encompasses most terms pertinent to patient safety.[8]

"Measurements describe phenomena in terms that can be analyzed statistically,"[13] and they should be both precise (free of random measurement error) and accurate (free of systematic measurement error). Readers should recall that *precise* is a synonym for *reliable*, and *accurate* a synonym for *valid*. Unfortunately, measuring errors and adverse events is more difficult than measuring many other health care processes or outcomes because errors and adverse events need to be understood in the context of the systems within which they occur. Based upon research in health care,[14] as well as in other fields,[15] an individual error or adverse event is usually the result of numerous *latent errors* in addition to the *active error* committed by a practitioner.

Latent errors include system defects such as poor design, incorrect installation, faulty maintenance, poor purchasing decisions, and inadequate staffing. These are difficult to measure because they occur over broad ranges of time and space and they may exist for days, months, or even years before they lead to a more apparent error or adverse event directly related to patient care. Active errors occur at the level of the frontline provider (such as administration of the wrong dose of a medication) and are easier to measure because they are limited in time and space. We propose that because latent errors occur over broad temporal and geographical ranges, and because active errors are more constrained in this regard, some measurement methods are best for latent errors and others for active errors.

Importantly, some methods are able to detect both latent and active errors. But we believe that when comparing those methods to each other they each have relative strengths and weaknesses for detecting latent versus active errors. For example, morbidity and mortality conferences (with or without autopsy results), malpractice claims analysis, and error reporting systems all have the ability to detect latent errors, active errors, and adverse events. However, their strength is in detecting latent errors, when compared to methods such as direct observation of patient care and prospective clinical surveillance for predefined adverse events. The reason is that they include information from multiple providers who were involved in the care over different times and physical locations. This increases the likelihood that latent errors may be detected. In contrast, methods such as direct observation and prospective clinical surveillance tend to focus on a few providers in a relatively limited time period, decreasing the likelihood of detecting latent errors.

Before discussing those points in more detail, we review the strengths and weaknesses of 8 measurement methods that have been used to measure errors and adverse events in health care (Table 1).

METHODS USED TO MEASURE ERRORS AND ADVERSE EVENTS

Morbidity and Mortality Conferences and Autopsy

Morbidity and mortality (M and M) conferences with or without autopsy results have had a central role in surgical training for many years.[16] The goal of these conferences is to learn from surgical errors and adverse events and thereby educate residents and improve the quality of care. The ability of M and M conferences to improve care has not been proven, but there is a strong belief in its effectiveness. In fact, the Accreditation Council for Graduate Medical Education requires surgery departments to conduct weekly M and M conferences,[17] and faculty and residents have

Table 1. Advantages and Disadvantages of Methods Used to Measure Errors and Adverse Events in Health Care

Error Measurement Method	Examples*	Advantages	Disadvantages
Morbidity and mortality conferences and autopsy	16–21	Can suggest latent errors Familiar to health care providers and required by accrediting groups	Hindsight bias Reporting bias Focused on diagnostic errors Infrequently and nonrandomly utilized
Malpractice claims analysis	25–28	Provides multiple perspectives (patients, providers, lawyers) Can detect latent errors	Hindsight bias Reporting bias Nonstandardized source of data
Error reporting systems	29–35	Can detect latent errors Provide multiple perspectives over time Can be a part of routine operations	Reporting bias Hindsight bias
Administrative data analysis	36–40	Utilizes readily available data Inexpensive	May rely upon incomplete and inaccurate data The data are divorced from clinical context
Chart review	41–44	Utilizes readily available data Commonly used	Judgements about adverse events not reliable Expensive Medical records are incomplete Hindsight bias
Electronic medical record	45, 46	Inexpensive after initial investment Monitors in real time Integrates multiple data sources	Susceptible to programming and/or data entry errors Expensive to implement Not good for detecting latent errors
Observation of patient care	47–50	Potentially accurate and precise Provides data otherwise unavailable Detects more active errors than other methods	Expensive Difficult to train reliable observers Potential Hawthorne effect Potential concerns about confidentiality Possible to be overwhelmed with information Potential hindsight bias Not good for detecting latent errors
Clinical surveillance	53, 54	Potentially accurate and precise for adverse events	Expensive Not good for detecting latent errors

** The numbers refer to the references.*

positive attitudes about the effectiveness of M and M conferences.[18]

Regarding autopsy, some studies have found that potentially fatal misdiagnoses occur in 20% to 40% of cases.[19–21] When coupled with a review of the medical record and discussions with the providers who cared for the patient as part of M and M conferences, we believe that autopsies become a rich source of information that can illuminate errors that lead to misdiagnosis. Although both active and latent errors may be identified by autopsy and M and M conferences, relative to other methods discussed below, the strength of this method is in illuminating latent errors. However, autopsy rates and the total number of cases discussed at M and M conferences are too low to measure incidence or prevalence rates of errors and adverse events. These methods of detecting errors and adverse events are at best a case series, the weakest form of study design.[22]

Autopsies and M and M conferences, like other methods discussed here, are also limited by hindsight bias.[23] For example, when asked to judge quality of care in groups of cases that were identical regarding the process of care but varied only by the outcome, anesthesiologists consistently rated care in the context of a bad outcome as substandard, and care with a good outcome as neutral, even though the processes of care were identical.[24] However, this bias may be avoided by using evaluators blinded to outcome.

Malpractice Claims Analysis

Analysis of medical malpractice claims files is another method used to identify errors and adverse events.[25,26] The medical records, depositions, and court testimony that comprise claims files are a large pool of data that investigators and clinicians could use to qualitatively analyze medical errors. There are approximately 110,000 claims received each year by the 150 medical malpractice insurers in the United States[27] and analyses of malpractice claims files have led to important patient safety standards in anesthesia, such as use of pulse oximetry.[28] Relative to other methods, the strength of claims file analysis lies in its ability to detect latent errors, as opposed to active errors and adverse events. This powerful example of the utility of malpractice claims analysis is balanced by several limitations. Claims are a series of highly selected cases from which it is difficult to generalize. Also, malpractice claims analysis is subject to hindsight bias as well as a variety of other ascertainment and selection biases, and the data present in claims files is not standardized. Finally, although malpractice claims files analysis may identify potential causes of errors and adverse events that may be addressed and studied, the claims files themselves cannot be used to estimate the incidence or prevalence of errors or adverse events or the effect of an intervention to decrease errors and adverse events.

Error Reporting Systems

Errors witnessed or committed by health care providers may be reported via structured data collection systems. Numerous reporting systems exist in health care and other industries[29] and their use was strongly endorsed by the Institute of Medicine report.[2] Reporting systems, including surveys of providers[30] and structured interviews, are a way to involve providers in research and quality improvement projects.[31]

Analysis of error reports may provide rich details about latent errors that lead to active errors and adverse events. But error reporting systems alone cannot reliably measure incidence and prevalence rates of errors and adverse events because numerous factors may affect whether errors and adverse events are reported. Providers may not report errors because they are too busy, afraid of lawsuits, or worried about their reputation. High reporting rates may indicate an organizational culture committed to identifying and reducing errors and adverse events rather than a truly high rate.[32] Despite these limitations, error reporting systems can identify errors and adverse events not found by other means, such as chart reviews,[33] and can thereby be used in efforts to improve patient safety.[34,35]

Administrative Data Analysis

Theoretically, administrative or billing data might provide an attractive source of data for measuring errors and adverse events. However, administrative data may be incomplete and subject to bias from reimbursement policies and regulations that provide incentives to code for conditions and complications that increase payments to hospitals ("DRG creep").[36] The Complication Screening Program (CSP) is designed to identify preventable complications of hospital care using hospital discharge data,[37] but the utility of the CSP for detecting errors and adverse events is unclear.[38,39] Other screening methods, including those available through hospital billing data, do not identify a high percentage of adverse events.[40] Despite these limitations, administrative data is in our opinion less susceptible to the ascertainment and selection biases that limit autopsies, morbidity and mortality conferences, malpractice claims analyses, and incident reporting systems.

Chart Review

Large, population-based chart reviews have been the foundation of research into errors and adverse events,[41] and the continuing usefulness of chart review is demonstrated by the Center for Medicare and Medicaid Services' recent decision to use chart review as a methods for monitoring patient safety for its beneficiaries throughout the country. Despite advances in the science of medical record review,[42] there are many flaws in this methodology. Judgments about the presence of adverse events by chart reviewers are known to have only low to moderate reliability (precision).[43] Another limitation of chart review

is incomplete documentation in the medical record.[44] In our experience, incomplete documentation can affect the ability to detect both latent errors and active errors that may lead to adverse events, a weakness that may be addressed by combining chart review with provider reporting.[33]

Electronic Medical Record Review

Reviewing the electronic medical record may improve detection of errors and adverse events by monitoring in "real time" and by integrating multiple data sources (e.g., laboratory, pharmacy, billing). Use of computers to search the electronic medical record[45] can find errors and adverse events not detected by traditional chart review or provider self-reporting.[46]

While more complete than hand written records (because of the ability to integrate multiple data sources), the data in the electronic medical record are still entered by humans and therefore prone to error and bias. Also, no standardized method exists to search for errors and adverse events. The reliability and validity of such measurement tools is unknown and deserves further study. The initial cost of these systems remains another significant limitation.[5] But as hospital medication delivery systems and laboratory reporting systems become more integrated with computers, structured review of the electronic medical record is likely to be an important method for measuring errors and adverse events and has the potential to be both accurate and precise.

Observation of Patient Care

Observing or videotaping actual patient care may be good for measuring active errors. Observation has been used in operating rooms,[47] intensive care units,[48] surgical wards,[49] and to assess errors during medication administration.[50] These studies found many more active errors and adverse events than previously documented, again highlighting the limitations of the other measurement methods described above.

Direct observation is limited by practical and methodologic issues. First, confidentiality is a concern because data could potentially be used by supervisors to punish employees, or could be obtained by medical malpractice plaintiff attorneys. Second, direct observation requires time-intensive training of observers to ensure reliability (precision). Third, if the observers are not or cannot be blinded to patient outcome, hindsight bias may be present. Fourth, observation of care focuses on the "sharp end" or on the providers instead of the entire system of delivery.[51] As we have noted, most errors and adverse events are the result of many antecedent latent errors that are usually not observable (e.g., equipment purchasing decisions) while watching patient encounters at one point and place in time. This limitation can be addressed by interpreting data from direct observation in the context of data derived from incident reports, claims files analysis, and autopsies. Finally, the Hawthorne effect,[52] which occurs when individuals alter their normal behavior because they are being observed, is also a limitation.

Clinical Surveillance

Potentially, the most precise and accurate method of measuring adverse events is exemplified by many of the studies that comprise the large literature on postoperative complications.[53] For example, the measurement of postoperative myocardial infarctions may include administration of electrocardiograms and measurement of cardiac enzymes in a standardized manner to all patients in a prospective cohort study.[54] Other adverse events could be measured in a similar fashion.

We believe that this type of active, prospective surveillance typical of classical epidemiologic studies is ideal for assessing the effectiveness of specific interventions to decrease explicitly defined adverse events. However, because clinical surveillance tends to focus on specific events in a focused time and place, we believe it provides relatively less contextual information on the latent errors that cause adverse events, and furthermore, may be costly.

DISCUSSION

Figure 1 illustrates our proposal for a general framework to help select error and adverse event measurement methods. Health care providers, researchers, administrators, and policy makers may find it useful to see these methods as existing on a continuum that illustrates the relative utility of each method for measuring latent errors as compared with active errors and adverse events (Fig. 1). On the left of this continuum are methods that capture the rich contextual issues that surround errors and adverse events and thereby allow detection of the latent errors that lead to them. These include medical malpractice claims file analysis, incident report analysis, morbidity and mortality conferences, and autopsies.

Using methods like these to identify latent errors has helped improve patient safety in areas like anesthesiology and pharmacy. For example, claims file analyses led to implementation of pulse oximetry in anesthesia,[28] and

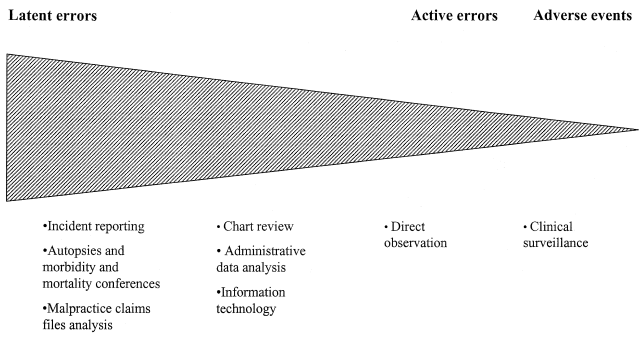

FIGURE 1. The relative utility of methods for measuring latent errors, active errors, and adverse events.

incident reports have led to pharmacy practices such as removal of concentrated potassium chloride from nursing units, or requiring the use of leading zeros when writing medication doses.

Although these methods can provide important and actionable information about systems, they also have weaknesses. They are incapable of providing error or adverse event rates because they are imprecise, primarily because of the various factors that influence whether an error or adverse event leads to a claim, incident report, or autopsy. Therefore, they should be used sparingly, if at all, to assess the efficacy of interventions to improve patient safety. Instead, they can identify the latent errors that need to be addressed. The efficacy of interventions to address these errors and the related active errors and adverse events can be determined with more precise methods to assess baseline rates of errors and adverse events and the efficacy of interventions.

For example, at the far right of the continuum are prospective clinical surveillance and direct observation. These methods can provide precise and accurate estimates of error and adverse event rates in a prospective fashion, and are thus suited to measure incidence, prevalence, and the impact of interventions. However, we believe that direct observation and clinical surveillance alone have relatively limited ability to measure latent errors because such errors may have occurred in a different time or place than is being observed.

The usefulness of our conceptual model is supported by its ability to explain and place in context debates about improving patient safety. For example, indirect support for our model (Fig. 1) comes from applying it to two recent articles that discussed the merits of using evidence-based medicine (EBM) principals to evaluate patient safety practices. Leape et al. argued that exclusive reliance on EBM would result in missed opportunities for improving patient safety because many patient safety practices that are believed to be effective have not been, and cannot be, assessed with randomized controlled trials. Leape et al.[10] gave examples of patient safety practices such as pharmacy-based intravenous admixture systems, unit dose dispensing of medications, removal of concentrated potassium chloride from nursing units, and various anesthesia safety practices that are believed to improve patient safety but do not meet EBM criteria for acceptance. These practices have come into existence because of errors and adverse events that were initially detected by methods on the left side of our figure (malpractice claims analysis and incident reporting). These methods allowed not only for the adverse event to be detected, but for the latent errors that led to the adverse event (e.g., the presence of concentrated potassium chloride on nursing units) to be detected.

In a companion piece, Shojania et al.[11] urged more reliance on EBM principals to determine whether patient safety practices were effective. They identified certain practices that have been tested in clinical trials (for example, perioperative β blocker use and thromboembolism prophylaxis). These types of patient safety practices have been tested by using measurement methods on the right side of our figure (e.g., prospective clinical surveillance to detect postoperative thromboembolic disease or cardiac complications).

Our model illustrates how the contrasting perspectives of Leape and Shojania originate in part from reliance on different measurement methods that vary in their precision, accuracy, and ability to detect latent errors versus active errors and adverse events. Leape urges us to use the methods on the left of our figure because of their ability to detect very important latent errors. Shojania favors reliance on measurement methods on the right because of their ability to provide precise and accurate measurements. Our model accommodates both of these approaches and suggests that they exist on a continuum.

Our model also suggests that a comprehensive monitoring system for patient safety might include combinations of the measurement methods we discussed. For example, ongoing incident reporting, autopsies, morbidity and mortality conferences, and malpractice claims file analysis could be used to identify latent errors and some active errors and adverse events. These methods would not be used to calculate rates, but rather to direct subsequent projects that would use chart review, direct observation, or prospective clinical surveillance to measure explicitly defined errors and adverse events. Combining different measurement methods has been used successfully by hospital epidemiologists to detect nosocomial infections.[55]

One primary goal of health care is to "do no harm." Understanding the relative strengths and weaknesses of the error and adverse event measurement methods discussed here can help investigators, clinicians, administrators, and policy makers meet this goal.

The authors thank Jon E. Tyson, MD, Robert Helmreich, PhD, and Marshall Chin, MD for their helpful comments on earlier versions of this paper.

Dr. Thomas is a Robert Wood Johnson Foundation Generalist Physician Faculty Scholar. Dr. Petersen was an awardee in the Research Career Development Award Program of the VA HSR&D Service (grant RCD 95-306) and is a Robert Wood Johnson Foundation Generalist Physician Faculty Scholar.

The authors have no potential conflicts of interest. The authors' funding agencies had no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

REFERENCES

1. Kohn LT, Corrigan JM, Donaldson MS, eds. To Err is Human. Washington DC: National Academy Press, 1999.
2. Expert Group on Learning from Adverse Events in the NHS. An Organisation With a Memory. London: Stationery Office; 2000.

3. Bates DW, Gawande AA. Error in medicine: what have we learned. Ann Intern Med. 2000;132:763–7.
4. Weingart SN, Wilson RM, Gibberd RW, Harrison B. Epidemiology of medical error. BMJ. 2000;320:774–7.
5. Brennan TA. The Institute of Medicine report on medical errors—could it do harm? N Engl J Med. 2000;342:1123–5.
6. McDonald CJ, Weiner M, Hui SL. Deaths due to medical errors are exaggerated in Institute of Medicine report. JAMA. 2000;284: 93–5.
7. Sox HC, Woloshin S. How many deaths are due to medical error? Getting the number right. Eff Clin Pract. 2000;3:277–83.
8. Hofer TP, Kerr EA. What is an error? Eff Clin Pract. 2000;3:261–9.
9. Wears RL, Janiak B, Moorehead JC, et al. Human error in medicine: promise and pitfalls, part 2. Ann Emerg Med. 2000;36:142–4.
10. Leape LL, Berwick DM, Bates DW. What practices will most improve patient safety? Evidence-based medicine meets patient safety. JAMA. 2002;288:501–7.
11. Shojania KG, Duncan BW, McDonald KM, Wachter RM. Safe but sound. Patient safety meets evidence-based medicine. JAMA. 2002;288:508–13.
12. Reason J. Human Error. Cambridge: Cambridge University Press; 1990.
13. Hulley SB, Martin JN, Cummings SR. Planning the measurements: precision and accuracy. In: Designing Clinical Research: An Epidemiologic Approach. Hulley SB, Cummings SR, eds. Philadelphia: Lippincott; 2001:37.
14. Leape LL, Bates DW, Cullen DJ. Systems analysis of adverse drug events. ADE prevention study. JAMA. 1995;274:35–43.
15. Perrow C. Normal Accidents: Living With High Risk Technologies. New York: Basic Books; 1984.
16. Gordon L. Gordon's Guide to the Surgical Morbidity and Mortality Conference. Philadelphia: Hanley and Belfus; 1994.
17. Accreditation Council for Graduate Medical Education. Essentials and Information Items. Graduate Medical Education Directory. Chicago Ill: American Medical Association; 1995.
18. Harbison S, Regehr G. Faculty and resident opinions regarding the role of Morbidity and Mortality Conference. Am J Surg 1999;177: 136–9.
19. Goldman L, Sayson R, Robbins S, Cohn LH, Bettmann M, Weisberg M. The value of the autopsy in three medical eras. N Engl J Med. 1983;308:1000–5.
20. Cameron HM, McGoogan E. A prospective study of 1152 hospital autopsies, I: inaccuracies in death certification. J Pathol. 1981; 133:273–83.
21. Anderson RE, Hill RB, Key CR. The sensitivity and specificity of clinical diagnostics during five decades: toward an understanding of necessary fallibility. JAMA. 1989;261:1610–17.
22. Sackett DL, Haynes RB, Guyatt GH, Tugwell P. Clinical Epidemiology. Boston: Little Brown, and Company; 1991.
23. Fischoff B. Hindsight does not equal foresight: the effect of outcome knowledge on judgment under uncertainty. J Exp Psychol Hum Percept Perform. 1975;1:288–99.
24. Caplan RA, Posner KL, Cheney FW. Effect of outcome on physician judgments of appropriateness of care. JAMA. 1991; 265:1957–60.
25. Cheney FW, Posner K, Caplan RA, Ward RJ. Standard of care and anesthesia liability. JAMA. 1989;261:1599–1603.
26. Rolph JE, Kravitz RL, McGuigan K. Malpractice claims data as a quality improvement tool. II. Is targeting effective? JAMA. 1991; 266:2093–7.
27. Weiler PC. Medical Malpractice on Trial. Cambridge, Mass: Harvard University Press; 1991.
28. Eichhorn JH, Cooper JB, Cullen DJ, Maier WR, Philip JH, Seeman RG. Standards for patient monitoring during anesthesia at Harvard Medical School. JAMA. 1986;256:1017–20.
29. Barach P, Small SD. Reporting and preventing medical mishaps: lessons from non-medical near miss reporting systems. BMJ. 2000;320:759–63.
30. Sexton JB, Thomas EJ, Helmreich RL. Error, stress, and teamwork in medicine and aviation: cross-sectional surveys. BMJ. 2000;320: 745–9.
31. Brennan TA, Lee TH, O'Neil AC, Petersen LA. Integrating providers into quality improvement: a pilot project at one hospital. Qual Manag Health Care. 1992;1:29–35.
32. Edmonson AC. Learning from mistakes is easier said than done: group and organizational influences on the detection and correction of human error. J Appl Behav Sci. 1996;32:5–28.
33. O'Neil AC, Petersen LA, Cook EF, Bates DW, Lee TH, Brennan TA. A comparison of physicians self-reporting to medical record review to identify medical adverse events. Ann Intern Med. 1993;119: 370–6.
34. Petersen LA, Brennan TA, O'Neil AC, Cook EF, Lee TH. Does house staff discontinuity of care increase the risk for preventable adverse events? Ann Intern Med. 1994;121:866–72.
35. Petersen LA, Orav EJ, Teich JM, O'Neil AC, Brennan TA. Using a computerized sign-out to improve continuity of inpatient care and prevent adverse events. Jt Comm J Qual Improv. 1998;24: 77–87.
36. Iezzoni LI. Assessing quality using administrative data. Ann Intern Med. 1997;127:666–74.
37. Iezzoni LI. Identifying complications of care using administrative data. Med Care. 1994;32:700–15.
38. Iezzoni LI, Davis RB, Palmer RH, et al. Does the Complications Screening Program flag cases with process of care problems? Using explicit criteria to judge processes. Int J Qual Health Care. 1999; 11:107–18.
39. Weingart SN, Iezzoni LI, Davis RB, et al. Use of administrative data to find substandard care: validation of the complications screening program. Med Care. 2000;38:796–806.
40. Bates DW, O'Neil AC, Petersen LA, Lee TH, Brennan TA. Evaluation of screening criteria for adverse events in medical patients. Med Care. 1995;33:452–62.
41. Brennan TA, Leape LL, Laird NM, et al. Incidence of adverse events and negligence care in hospitalized patients. N Engl J Med. 1991; 324:370–6.
42. Ashton CM, Kuykendall DH, Johnson ML, Wray NP. An empirical assessment of the validity of explicit and implicit process of care criteria for quality assessment. Med Care. 1999; 37:798–808.
43. Localio AR, Lawthers A, Brennan TA. Identifying adverse events caused by medical care: degree of physician agreement in retrospective chart review. Ann Intern Med. 1996;125: 457–64.
44. Luck J, Peabody JW, Dresselhaus TR, Lee M, Glassman P. How well does chart abstraction measure quality? A prospective comparison of standardized patients with the medical record. Am J Med. 2000;108:642–9.
45. Classen DC, Pestotnik SL, Evans RS, Burke JP. Computerized surveillance of adverse drug events in hospital patients. JAMA. 1991;266:2847–51.
46. Jha AK, Kuperman GJ, Teich JM, et al. Identifying adverse drug events: development of a computer-based monitor and comparison to chart-review and stimulated voluntary report. J Am Med Inform Assoc. 1998;5:305–14.
47. Helmreich RL, Schaefer HG. Team performance in the operating room. In: Bogner MS, ed. Human Error in Medicine. Hillsdale NJ: Lawrence Erlbaum Associates; 1994:225–53.
48. Donchin Y, Gopher D, Olin M, et al. A look into the nature and causes of human errors in the intensive care unit. Crit Care Med. 1995;23:294–300.
49. Andrews LB, Stocking C, Krizek T, et al. An alternative strategy for studying adverse events in medical care. Lancet. 1997;349: 309–13.
50. Barker KN. Data collection techniques: observation. Am J Hosp Pharm. 1980;37:1235–43.
51. Cook RI, Woods DD. Operating at the sharp end: the complexity of

human error. In: Bogner MS, ed. Human Error in Medicine. Hillsdale NJ: Lawrence Erlbaum Associates; 1994.
52. Last JM. A Dictionary of Epidemiology. New York: Oxford University Press; 1995.
53. Mangano DT, Goldman L. Preoperative assessment of patients with known or suspected coronary disease. N Engl J Med. 1995;333: 1750–6.
54. Lee TH, Marcantonio ER, Mangione CM, et al. Derivation and prospective validation of a simple index for prediction of cardiac risk of major noncardiac surgery. Circulation. 1999; 100:1043–9.
55. Gaynes RP, Horan TC. Surveillance of nosocomial infections. In: Mayhall GC, ed, Hospital Epidemiology and Infection Control. Philidelphia: Lippincott Williams and Wilkins; 1999.