Analysis of errors reported by surgeons at three teaching hospitals

Atul A. Gawande, MD, MPH, Michael J. Zinner, MD, David M. Studdert, LLB, ScD, MPH, *and* **Troyen A. Brennan, MD, JD, MPH,** *Boston, Mass*

Background. *Little is known of the factors that underlie surgical errors. Incident reporting has been proposed as a method of obtaining information about medical errors to help identify such factors.*
Methods. *Between November 1, 2000, and March 15, 2001, we conducted confidential interviews with randomly selected surgeons from three Massachusetts teaching hospitals to elicit detailed reports on surgical adverse events resulting from errors in management ("incidents"). Data on the characteristics of the incidents and the factors that surgeons reported to have contributed to the errors were recorded and analyzed.*
Results. *Among 45 surgeons approached for interviews, 38 (84%) agreed to participate and provided reports on 146 incidents. Thirty-three percent of incidents resulted in permanent disability and 13% in patient death. Seventy-seven percent involved injuries related to an operation or other invasive intervention (visceral injuries, bleeding, and wound infection/dehiscence were the most common subtypes), 13% involved unnecessary or inappropriate procedures, and 10% involved unnecessary advancement of disease. Two thirds of the incidents involved errors during the intraoperative phase of surgical care, 27% during preoperative management, and 22% during postoperative management. Two or more clinicians were cited as substantially contributing to errors in 70% of the incidents. The most commonly cited systems factors contributing to errors were inexperience/lack of competence in a surgical task (53% of incidents), communication breakdowns among personnel (43%), and fatigue or excessive workload (33%). Surgeons reported significantly more systems failures in incidents involving emergency surgical care than those involving nonemergency care (*$P < .001$*).*
Conclusions. *Subjective incident reports gathered through interviews allow identification of characteristics of surgical errors and their leading contributing factors, which may help target research and interventions to reduce such errors. (Surgery 2003;133:614-21.)*

From Brigham and Women's Hospital and Harvard School of Public Health, Boston, Mass

STUDIES HAVE ESTIMATED that one half to two thirds of hospital adverse events are attributable to surgical care.[1-3] More than half of these events appear preventable.[3-5] However, little is known about the human and systems factors that underlie such errors in surgery.

The surgical management of disease is complex and difficult. Observers describe a large variety of organizational and human factors that contribute to poor surgical outcomes, including surgeon inexperience,[6-8] low hospital volume for an operation,[9-11] excessive workload,[12] fatigue,[13] poor technology,[14] insufficient supervision of trainees,[15] inadequate hospital systems,[16] poor communication among staff,[17] time of day,[14] and bureaucracy or administrative failures.[18] To target interventions and policies, clinicians, administrators, and regulators have sought to identify which of these factors are implicated most frequently in surgical error.

However, determining the relative importance of the various factors has proven extremely difficult. Chart review and observational studies have not revealed sufficiently detailed information about a large enough number of events to discern the underlying patterns.[3,19-22] By contrast, in-depth investigations of incidents involving error have allowed identification of "root causes" in individual cases,[23] but have been too time and labor intensive to replicate on a larger scale. Recently, policymakers in the United States and elsewhere have advocated both mandatory and voluntary error reporting systems to collect the desired information.[22] These reporting systems face significant operational challenges, particularly in obtaining more than sporadic participation by clinicians.

Accepted for publication March 7, 2003.

Funding provided by the Warren Whitman Richardson Fund at Harvard Medical School

Reprint Requests: Atul A. Gawande, MD, MPH, Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02155.

Surgery 2003;133:614-21.

0039-6060/2003/$30.00 + 0

doi:10.1067/msy.2003.169

Confidential interviews designed to collect incident reports from single participants in mishaps is one relatively simple approach that has been successful in anesthesia and other fields in allowing identification of patterns in causation.[24,25] To test the feasibility and usefulness of this method in the field of surgery, we undertook a study to gather and analyze reports on adverse events resulting from errors in surgical care through interviews with surgeons at three hospitals.

METHODS

Overview and definitions. Interviews lasting approximately 1 hour were conducted with attending surgeons, senior surgical residents, and surgical fellows at three Massachusetts teaching hospitals. We used the interviews to elicit reports on any of the interviewees' cases in which: (1) an adverse event occurred, meaning an injury involving disability (temporary or permanent) or death that resulted from medical management, as opposed to disease;[26] (2) the surgeon was personally aware of the circumstances that led to the adverse event; and (3) the surgeon judged the adverse event to be partly or wholly the result of an error in management. Following previous literature, we defined an error as a mistake: either a failure of a planned action to be completed as intended (ie, an error of execution), or the use of a wrong plan to achieve an aim (ie, an error of planning).[27] Events meeting all three criteria for inclusion were termed "incidents."

Identification of interviewees and incidents. We sought 100 incident reports for analysis and anticipated two to three error reports per interviewee. We randomly selected 45 surgeons from a list of all senior surgical trainees and clinically active surgeons (in cardiac, general, thoracic, transplant, trauma, and vascular surgery) at three Massachusetts teaching hospitals (n = 75). Thirty-two selected surgeons were faculty and 13 were fellows or residents in the final 2 years of training. Interviews were conducted by a single surgeon-interviewer (AAG) between November 1, 2000, and March 15, 2001.

We ensured the confidentiality of the interviewees and the information they conveyed by maintaining the anonymity of the surgeon, colleagues, patients, and facilities in all records and by obtaining a federal certificate of confidentiality. All interviewees provided signed informed consent to participate. The study was approved by the institutional review board at the Brigham and Women's Hospital.

We identified incidents in two ways. First, we reviewed with surgeons all their cases reported for weekly morbidity and mortality (M&M) conferences during the previous 6 months. At two of the three institutions, M&M cases were reported by residents alone. One of the three institutions also gathered cases using administrative data. Only cases that met the three inclusion criteria were retained in our sample. Second, because only a subset of adverse events involving error are reported at M&M conferences,[28,29] we included any reports of other adverse events volunteered by the interviewees that met the inclusion criteria.

We anticipated that a single incident could involve multiple errors. If multiple errors occurred in the care of one patient and involved separate circumstances with separate adverse outcomes, they were analyzed as separate incidents. Also, several incidents were reported by more than one surgeon. To have comparable data from each case, only the data reported by the first surgeon interviewed were included in the analysis. However, data provided by subsequent surgeons were used to test agreement among interviewees' reports.

Incident analysis. The interviews were then structured to gather detailed information on each incident following critical incident methods established for analyzing military, airline, and other disasters,[30] and previously used in studying critical medical incidents.[25,31] Surgeons were asked to provide an open-ended description of each incident, and the factors they recalled that contributed to the errors in care associated with the incident. The interviewer and surgeon then reviewed the recounted events in detail, refining the description of the circumstances involved. Finally, the interviewer queried the surgeons about the role of 15 possible contributing factors.

To identify these contributing factors, we followed the error analysis framework of Vincent et al,[32] refining our categories based on interviews with, and reviews by, surgeons and surgical nurses.[32] The final 15 categories of factors were: (1) interruption/distraction; (2) ergonomic problems (such as lighting, space, noise); (3) technology failure; (4) fatigue (from the length or lateness of duty); (5) excessive workload/inadequate staffing (for the tasks required at a given time); (6) breakdown in accurate transmission of necessary information (communication) between personnel; (7) inappropriate protocols; (8) lack of supervision of trainees; (9) lack of experience with or competence at a task; (10) administrative complexity/bureaucracy; (11) emergent versus elective setting; (12) time of day; (13) failure of vigilance; (14) failure of memory; and (15) error in judgment. In addition, for errors that occurred in the operating room, the time of the

incident was confirmed by checking the operative schedule. Length of time on duty was also confirmed by checking with the surgeon's operative and call schedules.

The interviewer recorded the information reported on each incident on a worksheet during the interview[24] and subsequently completed three comprehensive data forms: one recording the details of the injury itself; one the clinical circumstances of the incident (including location, timing, and who was involved); and one the human and systems factors that contributed to the error. The surgeon's judgment of whether a given factor (eg, excessive workload/inadequate staffing) contributed to error was recorded on a five-point Likert scale (ranging from "highly unlikely" to "somewhat likely" to "highly likely").

Data analyses. The surgeons' judgments about contributing factors were analyzed as binary variables. Factors judged at least "somewhat likely" to have contributed to error were coded as 1, and the rest were coded as 0. Length of time on duty was recorded as a categorical variable: less than 8 hours, 8 to 24 hours, and more than 24 hours.

We hypothesized that the setting of care could affect the factors contributing to error. We therefore tested for differences in reported contributing factors between emergent and nonemergent cases, and between incidents involving an intraoperative error and other incidents, using chi-squared analysis. We also examined whether the most commonly reported factors identified in our descriptive analysis had significant interrelationships using chi-square analysis.

To investigate whether reports occurring fewer than 6 months after the incident or based on M&M review differed in important respects from other incidents, we tested for differences in severity of injury and incidence of common contributing factors using chi-square analysis. Finally, using the subset of incidents that were reported by more than one surgeon, we tested agreement among surgeons' judgments by calculating kappa scores for their reports of leading contributory factors. Analyses were performed using SAS Version 8 (SAS Institute, Cary, NC).

RESULTS

Thirty-eight of the 45 surgeons we approached agreed to participate in the study (participation rate = 84%). In interviews, these surgeons reported one to eight incidents each (mean 4, standard deviation 1.9), providing a total of 146 different incidents for analysis. Fourteen incidents were reported by more than one surgeon.

Fifty-five percent of incidents came from review of interviewees' complications reported at M&M conference during the previous 6 months; of 235 M&M cases reviewed, 34% met the criteria for inclusion. The remaining 45% were independently volunteered.

Table I shows the incidents' clinical characteristics. Two thirds involved events occurring fewer than 6 months before the interview. Sixty percent of incidents occurred in the operating room, 12% in an intensive care unit, and 16% on a nonintensive care hospital floor. Three quarters involved nonemergent care. Most incidents occurred during daytime hours, but 40% occurred after hours. In the 94 incidents (64%) in which information was available on how long the principal clinician had been on duty at the time, 37% had been working (in the hospital or office) for more than 8 hours and 16% for more than 24 hours.

The outcomes were serious: 33% of incidents resulted in permanent disability and 13% resulted in a patient's death. The injuries were diverse: 77% involved injuries directly related to an operation or other procedure, with the three most common subtypes being a visceral injury (eg, a bowel or ureteral laceration), bleeding, and wound infection or dehiscence; 13% involved an unnecessary or inappropriately chosen procedure; 10% involved unnecessary advancement of disease (eg, missed diagnosis of breast cancer due to an incomplete workup).

The reported errors occurred in all phases of surgical care. Most commonly, they occurred during the intraoperative phase of surgical care (66%). In one quarter of the incidents, errors occurred during preoperative management, and in another quarter, during postoperative management. (Percentages sum to more than 100% because, in one in five incidents, surgeons reported a chain of events that spanned more than one phase of care.) Two or more clinicians substantially contributed to error in 70% of the reported incidents, and three or more clinicians contributed in 18%.

Table II shows the systems and cognitive factors that surgeons reported as having contributed to error. Surgeons reported that systems factors contributed to error in 86% of incidents. Two systems factors per incident was the median (range, 0 to 8), with surgeons reporting a median of four systems factors in incidents involving emergency care ($P < .001$ compared with incidents involving nonemergent care). For example, in one illustrative case, an attending surgeon reported an incident that had occurred 1 month earlier involving a trau-

Table I. Clinical characteristics of the incidents (N = 146)

Characteristic	*# of incidents*	*% of incidents*
Time between incident and report		
Less than 1 week	1	1%
Less than 1 month	24	16%
Less than 6 months	101	69%
Less than 2 years	119	82%
Location of the adverse event		
Operating room	87	60%
Hospital room (non-ICU)	23	16%
Intensive care unit	17	12%
Physician's office	4	7%
Emergency room	10	3%
Other	5	2%
Clinical circumstances		
Emergency care	34	23%
Nonemergency care	112	77%
After hours (5 pm to 7 am)	59	40%
During daytime hours (7 am to 5 pm)	87	60%
Number of clinicians contributing to error		
Two or more clinicians	103	70%
Three or more clinicians	26	18%
Phase of care in which error contributed to injury		
Preoperative management	39	27%
Intraoperative management	96	66%
Postoperative management	32	22%
Severity of injury		
Temporary disability	79	54%
Permanent disability	48	33%
Death	19	13%
Type of injury		
Unnecessary advancement of disease	14	10%
Unnecessary/incorrect invasive procedure	18	13%
Injury from appropriate procedure:	114	77%
Bowel, bladder, other visceral injury	31	27%
Bleeding	26	23%
Wound infection/dehiscence	14	12%
Myocardial infarction/CVA	7	8%
Operative failure	4	4%
Pneumonia	4	4%
Musculoskeletal injury	6	2%
Other type of procedural injury	18	15%

Table II. Incidents, by contributing factor

Factors cited as contributing to error in an incident	*# of incidents*	*% of incidents**
Systems factors	126	86%
Inexperience/lack of competence	75	53%
Communication breakdown	62	43%
Excessive workload/ inadequate staffing	30	22%
Lack of supervision	29	21%
Fatigue	21	16%
Interruptions/distractions	21	16%
Technology/equipment failure	22	15%
Administrative complexity/ bureaucracy	9	6%
Inappropriate protocol	2	1%
Ergonomics (lighting, space, etc.)	2	1%
Cognitive factors	126	86%
Error in judgment	92	63%
Failure of vigilance	72	49%
Failure of memory	5	3%

*Percentages are calculated for non-missing data. Three incidents were missing data on inexperience, 11 on workload/staffing, 6 on lack of supervision, 14 on interruptions, and 1 each on technology failure, administration, and ergonomics. No incidents were missing data on the cognitive factors, communication breakdown, or use of an inappropriate protocol.

ma patient with a strangulated incisional hernia shortly after open-technique diagnostic peritoneal lavage. Abdominal exploration revealed infarcted bowel and inadequate fascial closure (sutures were not through fascia). A resident had performed the procedure, and the attending cited lack of supervision and inexperience as contributing factors. During review of other potential factors, the attending also reported that interruptions and workload likely contributed (he was called away to another patient), as did fatigue (the resident was on duty more than 24 hours), ergonomics (the bay's overhead lighting was not working), and miscommunication (the attending only learned afterward that the resident had never done the procedure before).

Inexperience or lack of competence with a particular surgical task was the most commonly cited factor, reported in 53% of incidents. Communication breakdowns were reported to contribute in 43%. Excessive workload, fatigue, or both were factors specified in 33%: 22% of incidents involved excessive workload/inadequate staffing, 16% involved fatigue, and 5% involved both. One or more of this group of factors contributed to error in 83% of all incidents. Individual cognitive factors

Table III. Incidents involving emergency versus nonemergency care, by contributing factor

Factors cited as contributing to error in an incident	*Emergency care (n = 34)*	*Nonemergency care (n = 112)*	*P value*
Systems factors			
Inexperience/lack of competence	74% (25/34)	53% (57/108)	.03
Communication breakdown	59% (20/34)	38% (42/112)	.03
Excessive workload/inadequate staffing	29% (9/31)	20% (21/104)	.30
Lack of supervision	42% (14/33)	14% (15/107)	< .001
Fatigue	27% (9/33)	12% (12/101)	.03
Interruptions/distractions	16% (5/31)	16% (16/101)	.97
Technology/equipment failure	24% (8/34)	13% (14/111)	.12
Administrative complexity/bureaucracy	3% (1/34)	6% (8/111)	.37
Inappropriate protocol	18% (6/34)	8% (9/111)	.11
Ergonomics (lighting, space, etc.)	6% (2/33)	0% (0/112)	.009
Cognitive factors			
Error in judgment	41% (14/34)	52% (58/112)	.28
Failure of vigilance	68% (23/34)	62% (69/112)	.52
Failure of memory	6% (2/34)	3% (3/112)	.37

also were commonly cited (86% of incidents), with surgeons citing failures of judgment and vigilance in the majority of incidents.

Several systems factors were significantly more commonly reported in emergency cases, including inexperience, miscommunication, lack of supervision, and fatigue (Table III). Acuity made no difference in the likelihood that judgment, vigilance, or memory failure were reported as contributors.

In approximately half the 75 cases in which inexperience played a role, the inexperience was that of a trainee (Table IV). Surgeons reported a lack of adequate supervision as a contributing factor in 55% of such incidents. In the other half of the incidents involving inexperience, the inexperienced participant was an attending staff member (or equivalent). Such lack of expertise may occur because a procedure is new to medicine or simply new to an individual. It appeared, however, that the latter was the more common situation. In 70% of these nontrainee cases, others with greater expertise at the same task were available to the clinician at the same institution.

Among cases involving communication breakdown, two thirds involved an inadequate handoff of information or a change in the personnel providing a patient's care. Difficulties in ascertaining responsibility also were implicated (Table IV), whether because of a lack of a clear clinician in charge of a particular aspect of care or conflict among clinicians making decisions. Other problems included communication failures between residents and attending surgeons, and between nurses and physicians.

We tested all possible relationships among the most common systems factors (inexperience/lack of competence, communication breakdown, workload, fatigue, lack of supervision), and the two most common cognitive factors (misjudgment, lack of vigilance), and we found several strong and significant associations. Failure of judgment was directly associated with reports of inadequate supervision (odds ratio [OR] = 3.4; 95% confidence interval [CI] = 1.2-9.6), whereas failure of vigilance was inversely associated with inexperience as a reported contributing factor (OR = 0.51; 95% CI = 0.3-1.0). The likelihood that miscommunication was cited doubled when surgeons reported excessive workload to be contributory in a given case (OR = 2.3; 95% CI = 1.6-3.3]). The likelihood that fatigue was a cited factor doubled when surgeons reported excessive workload to be contributory (RR = 2.3, 95% CI = 1.0-5.0]).

We also found inconsistencies in recall of events. Incidents occurring more than 6 months before the interview were more likely to involve permanent injury (42% vs 22%, $P = .008$) and visceral injury (48% vs 19%, $P = .002$) than those occurring within 6 months of the interview. Surgeons also were less likely to report fatigue as a contributing factor (0% vs 35%, $P = .001$). (No other factors were reported significantly more or less often based on time.) There was no significant difference in severity of injury or contributing factors between incidents identified in review of M&M reports and those that were independently volunteered. In the 14 incidents that were reported by more than one surgeon, we also found good agreement among surgeons about the role of leading contributing

factors. Comparing surgeons' judgments about whether inexperience played a contributing role in a given incident, the kappa statistic was 0.7; for judgments about whether communication played such a role, the kappa statistic was 0.8. (Too few of these 14 cases involved fatigue or excessive workload to permit accurate calculation of kappa statistics for these factors.)

Table IV. Incidents involving inexperience or communication breakdown, by subcharacteristic

Characteristic	*Percentage of cases reported*
Inexperience	53% of all incidents (n = 75)
Proportion involving a trainee	55% (n = 41)
Proportion involving a nontrainee	45% (n = 34)
Communication breakdown	43% of all incidents (n = 62)
Handoff or change in personnel cited as contributing to error	66% (n = 41)
Lack of clear clinician in charge cited as contributing to error	15% (n = 9)
Conflict over decision-making cited as contributing to error	15% (n = 9)
Other failure of communication cited as contributing to error	37% (n = 23)

DISCUSSION

We found that confidential interviews with surgeons successfully elicited detailed reports on a large number of surgical adverse events resulting from errors in care. The incidents reported were serious, with one third resulting in permanent disability and 13% in death, and we were able to identify important underlying patterns in the errors.

Contrary to the premises of malpractice law, the vast majority of errors did not appear to be solely the result of individual failure.[33] We found that the vast majority of surgical errors reported involved contributions from more than one clinician and, frequently, a chain of events spanning more than one phase of care. Surgeons reported that cognitive failures (eg, of judgment or vigilance) played a role in more than half the incidents; they also reported that systems factors contributed in 84% of these cases. Furthermore, we found that errors in judgment were strongly associated with reports of inadequate supervision, indicating that systems failures contributed to at least a subset of the cognitive errors.

We identified several leading underlying vulnerabilities. Emergency surgical care, in particular, appeared to pose special risks for patients in our study hospitals. Incidents involving emergency care were associated with significantly more systems failures than others and were more likely to involve problems of inexperience, communication breakdown, supervision, and fatigue. This makes some sense. Emergency care has been associated with increased risk of surgical error in other studies and often increases organizational and team difficulties.[34]

Nonetheless, only one quarter of the incidents involved emergency care. Most involved elective care. However, regardless of the setting, inexperience, communication breakdown, fatigue, and excessive workload were the most common contributing systems factors identified. Factors of organization, planning, and interaction among team members appeared to play a critical and underappreciated role.

Our findings suggest new directions for remedies and research. Gaining understanding of how the organization of emergency surgical care contributes to error and how it could be improved appears to be a critical next step in error research. Regarding errors in which inexperience played a role, we found that inadequate supervision was a frequently cited factor when trainees were involved. We also found that the majority of attending staff who erred while performing tasks that were new to them, or that they were inexpert with, reported having colleagues on staff with greater expertise. Thus, as has been recognized in other settings, improving supervision and formalizing skills training could be important directions for intervention and future study.[25]

Breakdowns in the accurate transfer of information, in particular during "handoffs" between personnel, were the second most common factor reported to contribute to error. This is similar to findings in internal medicine,[35] anesthesiology,[36] and other specialties.[37] Standardization of handoffs has been shown to reduce these types of errors and deserves strong consideration for wider use in surgery.[38]

Surgeons also reported heavy workloads and fatigue as frequent (and related) contributors to error. This was consistent with our findings that 37% of the 94 incidents in which surgeons' schedules were available to us involved a surgeon who had been on duty for more than 8 hours at the time of error (16%, more than 24 hours). The correlation we identified between workload and miscommunication raises the possibility that heavy workloads may actually produce miscommunication in the generation of errors. We must also con-

sider, however, the reverse effect. Strategies that are now being introduced to reduce workload and fatigue commonly involve increasing the number of personnel providing care, which in turn increases the number of handoffs. Unless measures minimize errors from handoffs, the net result could be a paradoxical increase in adverse events.[35]

Several cautions in interpreting our results are warranted. First, it is unclear how representative the incidents reported by interviewees are of all the incidents that occur. Incident reporting is voluntary, and comes from the perspective of only one participant. Although 6 months of M&M cases were reviewed and discussed, many relevant incidents are not reported to M&M conferences. Inevitably, only a subset of complications resulting from error are identified. Errors resulting in visceral injuries, for example, appeared to be more readily recalled in older cases, and errors of omission appear to be less easily remembered or recognized than errors of commission. Of note, the characteristics of the incidents identified in our study were similar in several important respects to those identified through chart reviews of 15,000 hospital admissions in a previous study of preventable surgical adverse events.[3] In both studies, visceral injury, bleeding, and wound complications were the three most common types of injuries reported (accounting for 62% of injuries in the present study and 50% in the previous study). Also, the setting for the majority of incidents in both studies was the operating room (60% versus 66%). The broad similarity between the types of incidents that we identified and those incidents found using different methods suggests that the incidents reported here are reasonably similar, at least in general typology, to those detected in other, larger scale epidemiologic investigations.

A second concern is that incident reports cannot be expected to gather perfectly accurate information about events, depending as they do on a clinician's recall. Medicine has no universal flight data recorder as yet. Interviews did allow us to gather finer, more intimate detail about the causes of events than other methods of error analysis have to date, but memory for events is known to be fallible, and we did find inconsistencies between reports on newer versus older events.[39] People appear to be particularly prone to underestimating the influence of factors such as fatigue and interruptions. Third, because we did not have multiple interviewers, the possibility of bias in data gathering must also be considered.

Finally, findings from three teaching hospitals may not be generalizable to other hospitals, particularly nonteaching ones. Previous studies have found that nonteaching hospitals do not have significantly different rates of preventable adverse events.[40] However, the causative nature of events can undoubtedly vary by setting.

Nonetheless, incident reports appear to be a useful and important source of information regarding the nature of surgical errors, and interviews appear to be an efficient, effective way of eliciting these reports. Previous analyses of incident reports gathered through formal, confidential interview methods like those we used have proved remarkably valuable in other fields, both inside and outside of medicine. Most prominently, Cooper et al's 1978 interview study of anesthesia mishaps identified several leading causal factors, including poor anesthesia equipment design, inadequate monitoring, and handoffs.[25] This work provoked a series of targeted interventions that greatly reduced mortality from general anesthesia.[41,42] Use of these methods to examine other aspects of surgical care could well achieve similar results.

Critical elements in the success of interviews are their confidentiality, the limited time commitment required for interviewees, their interactive nature, and a belief that participating will result in information that is both valuable and nonpunitive. Interviews need not be the only effective method of gathering useful incident reports. Reporting that relies on Internet technology, e-mail, or other interactive, structured forms of information gathering may be able to reproduce our success on a larger scale. Based on our experiences, however, we believe that making such reporting mandatory would likely undermine the level of candor and detail about the nature of errors that clinicians provide.

Chart review studies have found that 50% to 67% of surgical adverse events are preventable. Similarly, in this study, surgeons acknowledged that more than one third of their M&M cases resulted from error. Research and innovation on error reduction must be a central component of efforts to improve surgical outcomes. Carefully elicited, voluntary incident reports were found to be a simple, unique, and rich source of specific information about how errors occur and how to reduce them.

The authors have no known financial or other conflicts of interest in conducting or reporting this work.

REFERENCES

1. Leape LL, Brennan TA, Laird N, Lawthers AG, Localio AR, Barnes BA, et al. The nature of adverse events in hospitalized patients: results of the Harvard Medical Practice Study II. N Engl J Med 1991;324:377-84.

2. Thomas EJ, Studdert DM, Burstin HR, Orav EJ, Zeena T, Williams EJ, et al. Incidence and types of adverse events and negligent care in Utah and Colorado. Med Care 2000;38:247-9.
3. Gawande AA, Thomas EJ, Zinner MJ. The incidence and nature of surgical adverse events in Colorado and Utah in 1992. Surgery 1999;126:66-75.
4. Thomas EJ, Studdert DM, Newhouse JP, et al. Costs of medical injuries in Utah and Colorado. Inquiry 1999;36:255-64.
5. Couch NP, Tilney NL, Rayner AA, Moore FD. The high cost of low-frequency events: the anatomy and economics of surgical mishaps. N Engl J Med 1981;304:634-7.
6. Hannan EL, Siu AL, Kumar D, Chassin MR, et al. The decline of coronary artery bypass graft surgery mortality in New York State. The role of surgeon volume. JAMA 1995;273:209-13.
7. Sosa JA, Bowman HM, Tielsch JM, Powe NR, Gordon TA, Udelsman R, et al. The importance of surgeon experience for clinical and economic outcomes from thyroidectomy. Ann Surg 1998;228:320-30.
8. Ruby ST, Robinson D, Lynch JT, Mark H. Outcome analysis of carotid endarterectomy in Connecticut: the impact of volume and specialty. Ann Vasc Surg 1996;10:22-6.
9. Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. N Engl J Med 1979;301:1364-9.
10. Hannah EL, O'Donnell JF, Kilburn H, et al. Investigation of the relationship between volume and mortality for surgical procedures performed in New York state hospitals. JAMA 1989; 262:503-510.
11. Sosa JA, Bowman HM, Gordon TA, Bass EB, Yeo CJ, Lillemoe KD, et al. Importance of hospital volume in the overall management of pancreatic cancer. Ann Surg 1998;228:428-38.
12. Sexton JB, Thomas EJ, Helmreich RL. Error, stress, and teamwork in medicine and aviation: cross sectional surveys. BMJ 2000;320:745-9.
13. Taffinder NJ, McManus IC, Gul Y, Russell RC, Darzi A. Effect of sleep deprivation on surgeons' dexterity on laparoscopy simulator. Lancet 1998;352:1191.
14. Lunn JN. The National Confidential Enquiry into Perioperative Deaths. J Clin Monitoring 1994;10:426-8.
15. Keyes C, Hammond J. Every defect is a treasure: supervision of junior staff. Intl J Qual Health Care 1997;9:391-2.
16. Leape LL. Error in medicine. JAMA 1994;272:1851-7.
17. Young GJ, Charns MP, Daley J, Forbes MG, Henderson W, Khuri SF, et al. Best practices for managing surgical services: the role of coordination. Health Care Manage Rev 1997;22:72-81.
18. Pearse RM, Dana EC, Lanigan CJ, Pook JAR. Organisational failures in urgent and emergency surgery: a potential peri-operative risk factor. Anaesthesia 2001;56:670-89.
19. Couch NP, Tilney NL, Rayner AA, Moore FD. The high cost of low-frequency events: the anatomy and economics of surgical mishaps. N Engl J Med 1981;304:634-7.
20. Andrews LB, Stocking C, Krizek T, Gottlieb L, Krizek C, Vargish T, et al. An alternative strategy for studying adverse events in medical care. Lancet 1997;349:309-13.
21. Weingart SN, Wilson RM, Gibberd RW, Harrison B. Epidemiology of medical error. BMJ 2000;320:774-7.
22. Kohn LT, Corrigan JM, Donaldson MS, eds. To err is human: building a safer health system. Washington, DC: National Academy Press; 2000.
23. Joint Commission on Accreditation of Health Care Organizations. Root cause analysis in health care: tools and techniques. Oakbrook Terrace, IL: Joint Commission on Accreditation of Health Care Organizations; 2000.
24. Klein GA, Calderwood R, MacGregor D. Critical decision method for eliciting knowledge. IEEE Transactions on Systems, Man, and Cybernetics 1989;19:462-72.
25. Cooper JB, Newbower RS, Long CD, McPeek B. Preventable anesthesia mishaps: a study of human factors. Anesthesiology 1978;49:399-406.
26. Brennan TA, Leape LL, Laird NM, Hebert L, Localio AR, Lawthers AG, et al. Incidence of adverse events and negligence in hospitalized patients: results of the Harvard Medical Practice Study I. N Engl J Med 1991;324:370-6.
27. Reason JT, Human error. Cambridge, MA: Cambridge University Press; 1990.
28. Feldman L, Barkun J, Barkun A, Sampalis J, Rosenberg L. Measuring postoperative complications in general surgery patients using an outcomes-based strategy: comparison with complications presented at morbidity and mortality rounds. Surgery 1997;122:711-9.
29. Brennan TA, Localio AR, Leape LL, Laird NM, Peterson L, Hiatt HH, et al. Identification of adverse events occurring during hospitalization: a cross-sectional study of litigation, quality assurance, and medical records at two teaching hospitals. Ann Intern Med 1990;112:221-6.
30. Flanagan JC. The critical incident technique. Psychological Bulletin 1954;51:327-58.
31. Crandall B, Getchell-Reiter K. Critical decision method: a technique for eliciting concrete assessment indicators from the "intuition" of NICU nurses. ANS Adv Nurs Sci 1993:16:42-51.
32. Vincent C, Taylor-Adams S, Stanhope N. Framework for analysing risk and safety in clinical medicine. BMJ 1998;316:1154-7.
33. Studdert DM, Brennan TA. No-fault compensation for medical injuries: the prospect for error prevention. JAMA 2001;286:217-23.
34. Gawande AA, Studdert DM, Orav EJ, Brennan TA, Zinner MJ. Risk factors for retained instruments and sponges after surgery. New Engl J Med 2003;348:229-35.
35. Petersen LA, Brennan TA, O'Neil AC, Cook EF, Lee TH. Does housestaff discontinuity of care increase the risk for preventable adverse events? Ann Intern Med 1994;121:866-72.
36. Cooper JB, Long CD, Newbower RS, Philips JH. Critical incidents associated with intraoperative exchanges of anesthesia personnel. Anesthesiology 1982;56:456-61.
37. Neale G, Woloshynowych M, Vincent C. Exploring the causes of adverse events in NHS hospital practice. J R Soc Med 2001;94:322-30.
38. Petersen LA, Orav EJ, Teich JM, O'Neil AC, Brennan TA. Using a computerized sign-out program to improve continuity of inpatient care and prevent adverse events. Jt Comm J Qual Improv 1998;24:77-87.
39. Loftus EF. Eyewitness testimony. 2nd Edition. Cambridge, MA: Harvard University Press; 1996.
40. Thomas EJ, Orav EJ, Brennan TA. Hospital ownership and preventable adverse events. J Gen Intern Med 2000;15:211-9.
41. Pierce EC. The 34th Rovenstine lecture. Anesthesiology 1996;84:965-75.
42. Eichhorn JH, Cooper JB, Cullen DJ, Gessner JS, Holzman RS, Maier WR, et al. Anesthesia practice standards at Harvard: a review. J Clin Anesth 1988;55:64-5.