# A Physician-based Voluntary Reporting System for Adverse Events and Medical Errors

*Saul N. Weingart, MD, PhD, Lawrence D. Callanan, Jr., MD, Amy N. Ship, MD, Mark D. Aronson, MD*

***OBJECTIVE:*** **To create a voluntary reporting method for identifying adverse events (AEs) and potential adverse events (PAEs) among medical inpatients.**

***DESIGN:*** **Medical house officers asked their peers about obstacles to care, injuries or extended hospitalizations, and problems with medications that affected their patients. Two independent reviewers coded event narratives for adverse outcomes, responsible parties, preventability, and process problems. We corroborated house officers' reports with hospital incident reports and conducted a retrospective chart review.**

***SETTING:*** **The cardiac step-down, oncology, and medical intensive care units of an urban teaching hospital.**

***INTERVENTION:*** **Structured confidential interviews by postgraduate year-2 and -3 medical residents of interns during work rounds.**

***MEASUREMENTS AND MAIN RESULTS:*** **Respondents reported 88 events over 3 months. AEs occurred among 5 patients (0.5% of admissions) and PAEs among 48 patients (4.9% of admissions). Delayed diagnoses and treatments figured prominently among PAEs (54%). Clinicians were responsible for the greatest number of incidents (55%), followed by workers in the laboratory (11%), radiology (15%), and pharmacy (3%). Respondents identified a variety of problematic processes of care, including problems with diagnosis (16%), therapy (26%), and failure to provide clinical and support services (29%). We corroborated 84% of reported events in the medical record. Participants found voluntary peer reporting of medical errors unobtrusive and agreed that it could be implemented on a regular basis.**

***CONCLUSIONS:*** **A physician-based voluntary reporting system for medical errors is feasible and acceptable to front-line clinicians.**

***KEY WORDS:*** **adverse events; medical error; voluntary reporting.**
**J GEN INTERN MED 2001;16:809–814.**

The authors of the 1999 Institute of Medicine report *To Err is Human* called on health care organizations to create voluntary reporting systems for medical errors.[1] Impressed with NASA's Aviation Safety Reporting System,[2] the authors of the report argued that "voluntary reporting systems play a valuable role in encouraging improvement. . .(and) in enhancing understanding of the factors that contribute to errors." The voluntary system would collect information about near-misses and other errors that result in minor injuries as a complementary approach to a mandatory national error reporting system for egregious acts and incidents that result in substantial injury.

Although researchers elicited voluntary error reports in studies of adverse events, this approach has been expensive, labor intensive, and difficult to sustain.[3–5] Incident reporting systems are in widespread use, but miss many events and usually have poor physician participation.[6] Currently, few physicians use the voluntary error reporting systems in place in health care.[7,8]

The aim of this study was to create and test a nonpunitive reporting method for medical errors that relies on physician participation. It uses structured, confidential peer interviews among house officers.

## METHODS

### Setting

The study site is a 564-bed Boston teaching hospital. We studied 3 inpatient services: a 30-bed cardiac step-down (telemetry) unit, an 18-bed oncology unit, and a 12-bed medical intensive care unit (MICU) between March and June 1999. We chose the cardiac unit for its high patient volume, the oncology unit because of the potential toxicities associated with chemotherapy, and the MICU because of patients' acuity and severity of illness.

### Interview Protocol

We recruited all 10 PGY-2 and -3 medical residents who rotated through the study units during the first 3 days of their 3-week rotations. We sent each resident an e-mail message that described the project, solicited his/her assistance, and offered a $100 honorarium. We followed up by telephone. All agreed to participate. One of us (SNW) met briefly (15 minutes) with each resident to review the interview protocol and data reporting forms and to answer questions. We also met with all 21 interns to describe the project, their roles, confidentiality protections, and risks and benefits of participation. Interns completed 4-week rotations; 3 interns rotated through 2 of the 3 study units on subsequent rotations.

Residents interviewed the 3 interns on their teams about barriers to high-quality care, injuries or hospitalizations that were extended as a result of care, and problems with medications. Interviews took place 3 times weekly as part of morning work rounds. Residents were permitted to self-report additional events. They accepted but did not solicit reports from unit nurses and social workers. Residents recorded the date an event was reported, the

*Received from the Division of General Medicine and Primary Care, Beth Israel Deaconess Medical Center, and Harvard Medical School (SNW, LDC, ANS, MDA), Boston, Mass.*

*Address correspondence and reprint requests to Dr. Weingart: Division of General Medicine and Primary Care, Beth Israel Deaconess Medical Center, 330 Brookline Ave., LY 326A, Boston, MA 02215 (e-mail: sweingar@caregroup.harvard.edu).*

patient's name or hospital identification number, and a brief narrative. They attempted to verify the event in the medical record, with the patient, or with other health care providers. Participation was voluntary and reports, although shared among members of the house staff team and other staff members and students who were present on rounds, were considered confidential. Respondents' names were not recorded. The project was conducted as a peer review activity under the auspices of the Department of Medicine's Quality Improvement Committee.

### Coding and Classification

Residents submitted abbreviated weekly written reports, which we entered verbatim into a spreadsheet. Two internists (SNW and LDC) received forms that included each event narrative and a classification guide. The forms did not include information regarding patient name and medical record number, date of incident, clinical service, treating clinicians, or respondent. Each internist independently coded all incident narratives using a classification scheme that we developed for a previous study.[9] A third internist (ANS), blinded to coder, resolved discrepancies. Coders identified the party responsible for the incident from a list of 30 possibilities, the most important process-of-care deficiency from a list of 29 options, and the most serious adverse outcome the patient experienced from a list of 42 possibilities. They classified as adverse events (AEs) injuries that occurred as a result of medical care rather than the natural course of illness. Potential adverse events (PAEs) were near-misses: errors that could have resulted in injury but did not, due to interception or good fortune. Other quality problems that did not meet the definition of AE or PAE were classified separately. These incidents reflected inefficiencies, inconveniences, or other defects in service quality. For example, one patient missed a meal because the breakfast order was not transcribed. Finally, coders judged if the AE was probably, possibly, or unlikely preventable.

Interrater reliability was substantial for responsible party ($\kappa = 0.87$), adverse event ($\kappa = 0.63$), and process-of-care categories ($\kappa = 0.65$). We do not present data on preventability because agreement was poor (weighted $\kappa = 0.15$) and only moderately improved when we aggregated possibly and probably preventable judgments into a single category ($\kappa = 0.25$).

### Corroboration of Reports

To corroborate reports independently, we reviewed hospital incident reports from the study units that the hospital Risk Management Department maintains in accordance with state and Joint Commission on Accreditation of Healthcare Organization (JCAHO) requirements. We (SNW and LDC) also examined the medical records of patients with house officer–identified incidents, following a protocol approved in advance by the hospital's institutional review board. To maximize the likelihood of corroborating reports, reviewers used the date of the report and event narrative to guide the record review. Reviewers re-examined a 20% sample of charts selected at random. Interrater reliability was substantial ($\kappa = 0.76$). Discrepancies regarding medical record confirmation of house officer–reported events were resolved by a third reviewer (ANS).

### House Officer Survey

We surveyed intern and resident participants at the end of the study by e-mail. We inquired about the time required to participate, obstacles to reporting, and the feasibility of implementing the approach on an ongoing basis. Nonresponders received 2 follow-up reminders.

### Statistical Analyses

Analyses were descriptive. We calculated interrater reliability of coding and chart review using the $\kappa$ statistic to correct for chance variation among categorical variables and weighted $\kappa$ for ordinal variables. We calculated event rates based on admission to the respective units recorded in the hospital administrative database, and compared rates between study units using Fisher's exact test. We planned to collect 3 months of data on each unit, but an unusually small census (fewer than 2 patients per week during 1 month) in the MICU resulted in reassignment of house officers during part of the study. We collected data in the MICU for 6 weeks only. We used STATA software, version 4.0 (STATA Corp., College Station, Tex) for statistical computations.

## RESULTS

### Overview

Of the 987 patients admitted to the 3 study units, 721 were admitted to cardiology, 178 to oncology, and 88 to the MICU. Residents reported 94 incidents. We excluded 6 generic complaints (e.g., "dim light in the call room made it difficult to read sign-outs"). The final sample included 88 incidents (48 in cardiology, 30 in oncology, and 10 in the MICU) affecting 76 patients. Eleven patients experienced 2 or more incidents. Interns reported 48 incidents (62%), followed by PGY-2 and -3 residents' self-reports (29%), and unsolicited reports from nurses (8%) and social workers (1%). There was no pattern by day of the week or week of the study.

### Adverse Events and Potential Adverse Events

House officers reported 5 AEs. Adverse events affected 0.3%, 1.1%, and 1.1% respectively of cardiac step-down, oncology, and MICU admissions (Table 1). Adverse events included acute renal failure following aggressive diuresis without adequate monitoring; a 4-hour preoperative delay that resulted in progressively worsening hemodynamic

**Table 1. Rate of Adverse Events and Errors per 100 Admissions by Study Unit**

| | Cardiology, *n* = 721 | | Oncology, *n* = 178 | | MICU, *n* = 88 | | Total, *n* = 987 | |
|---|---|---|---|---|---|---|---|---|
| | *n* | Rate, % | *n* | Rate, % | *n* | Rate, % | *n* | Rate, % |
| Adverse events (AEs) | 2 | 0.3 | 2 | 1.1 | 1 | 1.1 | 5 | 0.5 |
| Potential AEs | 24 | 3.3 | 19 | 10.7 | 5 | 5.7 | 48 | 4.9 |
| Other quality problems | 22 | 3.1 | 9 | 5.1 | 4 | 4.5 | 35 | 3.5 |
| Total incidents | 48 | 6.7 | 30 | 16.9 | 10 | 11.4 | 88 | 8.9 |

*MICU, medical intensive care unit.*

instability in a patient with an expanding pseudoaneurysm following cardiac catheterization; extension of a superior vena cava thrombus with arm swelling and dyspnea when a patient's heparin was discontinued; excessive sedation following a 4-fold overdose of fentanyl for mucositis pain; and hypotension in a mechanically ventilated patient who inadvertently received a double strength infusion of fentanyl.

Potential adverse events occurred more frequently than AEs. Potential adverse events or "near-miss" errors represented 3.3%, 10.7%, and 5.7%, respectively of cardiac stepdown, oncology, and MICU admissions (Table 1). Delayed diagnoses accounted for 33% of PAEs (Table 2). For example, neurosurgical consultation was inexplicably delayed for 5 hours in a patient with a subdural hematoma and an acute change in mental status. Delayed treatment accounted for 21% of PAEs. For example, caregivers delayed treatment for chest pain 1.5 hours in a patient awaiting sestamibi injection. In paired comparisons of AE and PAE rates by study unit, the only statistically significant difference was between PAE rates in the cardiology and oncology units (3.3% vs 10.7%, respectively; $P < .001$).

## Responsible Parties

Table 2 shows the distribution of incidents by responsible party. Clinicians, including attending physicians, nurses, house officers, emergency room staff, and subspecialty consultants, accounted for 55% of reports. Workers in the laboratory, radiology, and pharmacy together accounted for 30% of incidents. House officer respondents identified themselves as the responsible party 8% of the time.

## Problematic Processes

Table 3 shows the problematic processes of care associated with reported incidents. Therapeutic errors accounted for 26%. The most common therapeutic errors involved medications (17%): wrong dose, route, or time; failure to order a drug; and failure to recognize a contraindication. Process problems involving the delivery of

**Table 2. Proportion of Adverse Events by Type of Incident and Responsible Party**

| | Adverse Events | | Potential Adverse Events | | Other Quality Problems | | Total | |
|---|---|---|---|---|---|---|---|---|
| | *n* | % | *n* | % | *n* | % | *n* | % |
| Adverse outcome | | | | | | | | |
| Injury | 5 | 100.0 | 0 | 0.0 | 0 | 0.0 | 5 | 5.7 |
| Delayed diagnosis | 0 | 0.0 | 16 | 33.3 | 6 | 17.1 | 22 | 25.0 |
| Delayed treatment | 0 | 0.0 | 10 | 20.8 | 5 | 14.3 | 15 | 17.0 |
| Difficult discharge | 0 | 0.0 | 0 | 0.0 | 7 | 20.0 | 7 | 8.0 |
| None | 0 | 0.0 | 22 | 45.8 | 17 | 48.6 | 39 | 44.3 |
| Responsible party | | | | | | | | |
| House officer | 1 | 20.0 | 6 | 12.5 | 0 | 0.0 | 7 | 8.0 |
| Attending MD | 0 | 0.0 | 0 | 0.0 | 5 | 14.3 | 5 | 5.7 |
| Nurse | 2 | 40.0 | 6 | 12.5 | 2 | 5.7 | 10 | 11.4 |
| Emergency room staff | 0 | 0.0 | 8 | 16.7 | 6 | 17.1 | 14 | 15.9 |
| Subspecialty consultants | 1 | 20.0 | 3 | 6.3 | 8 | 22.9 | 12 | 13.6 |
| Laboratory | 0 | 0.0 | 8 | 16.7 | 2 | 5.7 | 10 | 11.4 |
| Pharmacy | 0 | 0.0 | 2 | 4.2 | 1 | 2.9 | 3 | 3.4 |
| Radiology | 0 | 0.0 | 8 | 16.7 | 5 | 14.3 | 13 | 14.8 |
| Support services | 0 | 0.0 | 3 | 6.3 | 4 | 11.4 | 7 | 8.0 |
| Other | 1 | 20.0 | 4 | 8.3 | 2 | 5.7 | 7 | 8.0 |

**Table 3. Proportion of Adverse Events by Process of Care Problems**

| | Adverse Events | | Potential Adverse Events | | Other Quality Problems | | Total | |
|---|---|---|---|---|---|---|---|---|
| Process of Care | *n* | % | *n* | % | *n* | % | *n* | % |
| Diagnosis | | | | | | | | |
| Inadequate evaluation | 0 | 0.0 | 5 | 10.4 | 3 | 8.6 | 8 | 9.1 |
| Diagnostic error | 0 | 0.0 | 2 | 4.2 | 0 | 0.0 | 2 | 2.3 |
| Delayed consultation | 0 | 0.0 | 2 | 4.2 | 2 | 5.7 | 4 | 4.5 |
| Therapy | | | | | | | | |
| Medication-related | | | | | | | | |
| Delayed administration | 0 | 0.0 | 3 | 6.3 | 0 | 0.0 | 3 | 3.4 |
| Wrong dose or route | 2 | 40.0 | 6 | 12.5 | 0 | 0.0 | 8 | 9.1 |
| Failure to order drug | 0 | 0.0 | 3 | 6.3 | 0 | 0.0 | 3 | 3.4 |
| Failure to recognize contraindication | 0 | 0.0 | 1 | 2.1 | 0 | 0.0 | 1 | 1.1 |
| Delayed procedure | 1 | 20.0 | 1 | 2.1 | 6 | 17.1 | 8 | 9.1 |
| Prevention | | | | | | | | |
| Failure to monitor or follow up | 0 | 0.0 | 1 | 2.1 | 0 | 0.0 | 1 | 1.1 |
| Inadequate supervision | 0 | 0.0 | 1 | 2.1 | 0 | 0.0 | 1 | 1.1 |
| Clinical services | | | | | | | | |
| Failure or delay in performing a test | 0 | 0.0 | 9 | 18.8 | 3 | 8.6 | 12 | 13.6 |
| Failure to report test results | 0 | 0.0 | 2 | 4.2 | 0 | 0.0 | 2 | 2.3 |
| Lost specimen | 0 | 0.0 | 3 | 6.3 | 0 | 0.0 | 3 | 3.4 |
| Other laboratory error | 0 | 0.0 | 1 | 2.1 | 0 | 0.0 | 1 | 1.1 |
| Support services | | | | | | | | |
| Failure to transport | 0 | 0.0 | 2 | 4.2 | 1 | 2.9 | 3 | 3.4 |
| Failure to draw blood | 0 | 0.0 | 0 | 0.0 | 1 | 2.9 | 1 | 1.1 |
| Failure to provide for patient comfort | 0 | 0.0 | 0 | 0.0 | 3 | 8.6 | 3 | 3.4 |
| Other support services | 0 | 0.0 | 1 | 2.1 | 0 | 0.0 | 1 | 1.1 |
| Problematic discharge | 1 | 20.0 | 0 | 0.0 | 3 | 8.6 | 4 | 4.5 |
| Other process problems | | | | | | | | |
| Poor communication | 0 | 0.0 | 1 | 2.1 | 4 | 11.4 | 5 | 5.7 |
| Miscellaneous | 1 | 20.0 | 4 | 8.3 | 9 | 25.7 | 14 | 15.9 |
| Total | 5 | 100.0 | 48 | 100.0 | 35 | 100.0 | 88 | 100.0 |

clinical and support services accounted for 30% of incidents. The majority were due to failure or delay in performing a test (14%). For example, a thoracentesis specimen disappeared, only to reappear "miraculously" (according to the respondent) at the laboratory 2 days later. Diagnostic problems accounted for 16% of reports. A radiologist, for example, misread a hip radiograph as normal rather than fractured in a patient with metastatic prostate cancer and hip pain.

## Corroboration of Reports

During the study period, clinicians filed 58 hospital incident reports for 35 patients in the study units: 31 in the cardiac step-down unit, 25 in oncology, and 2 in the MICU. The Risk Management Department classified 45 (78%) as level 1(no significant injury) and the remaining 13 (22%) as level 2 (minor injury). Two of the level 2 events were also reported by house officers: the fentanyl overdose for mucositis pain described above (an AE), and a patient who received an overdose of heparin but had no adverse outcome (a PAE). The remaining 11 level 2 incident reports included 2 falls, 4 adverse drug reactions, 2 missed or late doses, 1 case of phlebitis at the site of an intravenous catheter, a groin hematoma with pseudoaneurysm after cardiac catheterization, and 1 chemical spill with cleaning products.

We obtained and reviewed the medical records for 75 (85.2%) of the 88 house officer–reported events. In the 13 missing cases, we had insufficient information to identify patients uniquely by name or medical record number. We found substantial evidence in the medical record to confirm 63 (84.0%) of the 75 events. In 2 cases, the report was inaccurate. In the remaining unsubstantiated cases, the record provided insufficient information to confirm the event.

## House Officer Survey

To assess the acceptability of voluntary peer reporting of medical errors, we surveyed the house officers who participated in the study. All 10 PGY-2 and -3 residents (100%) and 12 of 21 interns (57%) completed surveys. Residents spent 15–120 minutes (mean = 68, SD = 40) per week on the project, depending on the volume of reports. Interns spent 3–25 minutes (mean = 11, SD = 6) per week. Three residents found it occasionally awkward to elicit reports that might reflect poorly on nurse and physician colleagues. No intern indicated that reporting disrupted the workday or interfered with patient care. All but 1 intern

were confident that responses were handled confidentially; one was "somewhat confident." Two interns expressed some discomfort talking about quality problems. Six reported their own mistakes; the rest were aware of no personal errors. Respondents agreed unanimously that structured confidential interviews could be performed on a regular basis. One resident said that the project helped "...the interns to think critically about the system that they unconditionally accept."

## DISCUSSION

In this study of patients admitted to the cardiac step-down unit, oncology unit, and MICU of a Boston teaching hospital, medical house officers readily identified AEs, PAEs, and other quality problems. Potential adverse events outnumbered AEs (4.9% compared to 0.5% of admissions), suggesting that this approach may be particularly well suited to identify near-misses. Delayed diagnoses and treatments accounted for the greatest number of incidents (42%). Respondents provided sufficiently detailed event narratives to identify responsible parties and processes of care that contributed to the events. Participants found the project unobtrusive and were favorably disposed toward establishing the system on an ongoing basis.

These results are consistent with findings of several earlier studies demonstrating the capacity of front-line clinicians to identify quality problems and adverse events. When Welsh et al. asked medical house officers at a Colorado VA hospital to identify adverse events during morning report, respondents identified more events and more serious events than the usual incident reporting system.[3] Using daily e-mail reminders to elicit adverse event reports from medical house officers at a Boston teaching hospital, O'Neill et al. found adverse event rates comparable to chart review (2.8% vs 2.7%).[4] Using chart review and daily queries of nurses, pharmacists, and clerical staff, Bates et al. identified adverse drug events (ADEs) and potential ADEs in 6.5% and 5.5%, respectively, of admissions to two Boston teaching hospitals.[5] A serious limitation of earlier studies was the heavy investment of investigators' time required for data collection.

The present study extends an approach to voluntary peer reporting of medical error that we developed in a pilot study on a general medicine unit.[9] PGY-2 and -3 medical residents reported AE and PAE rates of 2.6% and 4.7%, respectively. Using a similar approach on a surgical unit, nurse respondents reported AE and PAE rates of 4.7% and 1.9%.[10] Most reports on the surgical unit involved mishaps related to the surgery or postoperative pain management, while medical unit reports featured a variety of diagnostic errors, therapeutic mishaps, and delays. The initial studies relied on a single interviewer in only 2 clinical settings.

This study has several potential limitations. First, reviewers were not blinded to patient outcomes, so their judgments about AEs may have been biased by the presence of serious injuries. Since we identified few adverse events, the magnitude of bias is likely small. Second, interrater reliability was poor regarding preventable events. This is a common problem among studies of heath care quality, perhaps due to the lack of consensus among clinicians about the definition of "preventable" events.[5,11] Third, underreporting is likely present. House officers may not recognize or report their errors or those of coworkers, perhaps because they fear punishment or the disapprobation of their peers and supervisors. They may be unaware of incidents that occurred when they were off duty or that were not recorded in the medical record. In addition, the skill and enthusiasm of interviewers varied by resident physician. Fourth, we did not undertake a comprehensive root cause analysis of reported incidents. Although we recognize its importance, such an undertaking was beyond the scope of the project.

Finally, the results may not be generalizable beyond our teaching hospital. The hospital's quality improvement program is led by a respected teacher-clinician. The hospital has a secure quality improvement mailbox for confidential reporting of adverse events and an electronic incident reporting system. The approach may be less successful in organizations without the visible commitment of senior clinicians and a nonpunitive approach to error reporting. Its use in nonteaching hospitals may also be limited by attending physicians' perceived vulnerability to legal consequences of reporting.

Nevertheless, voluntary peer reporting by physicians is a promising approach for detecting adverse events and near-miss errors. Our Department of Medicine adopted the approach as a key component of its quality improvement portfolio; an effort to implement physician-based voluntary reporting by the hospitalist service began in late 2000. The approach complements the hospital incident report system. It is inexpensive, easy to administer, and acceptable to clinician participants. It facilitates discussions among providers about error, increases awareness and, combined with root cause analysis, may help to drive hospital-based improvement efforts. Voluntary adverse event reporting merits further study.

## REFERENCES

1. Kohn LT, Corrigan JM, Donaldson MS, eds. To Err is Human: Building a Safer Health System. Washington, DC: National Academy Press; 1999.
2. Billings CE, Reynard WD. Human factors in aircraft incidents: Results of a 7-year study. Avn Space Environ Med. 1984;55:960–5.
3. Welsh CH, Pedot RP, Anderson RJ. Use of morning report to enhance adverse event detection. J Gen Intern Med. 1996;11: 454–60.
4. O'Neil AC, Petersen LA, Cook F, Bates DW, Lee TH, Brennan TA. Physician reporting compared with medical-record review to identify adverse medical events. Ann Intern Med. 1993;119: 370–6.
5. Bates DW, Cullen DJ, Laird N, et al. Incidence of adverse drug events and potential adverse drug events. JAMA. 1995;274:29–34.

6. Cullen CJ, Bates DW, Small SD, Cooper JB, Nemeskal AR, Leape LL. The incident reporting system does not detect adverse drug events: a problem for quality improvement. Jt Comm J Qual Improv. 1995;12:541–52.
7. Summary of 1999 Information Submitted to MedMARx$^{SM}$: A National Database for Hospital Medication Error Reporting. Rockland, Md: US Pharmacopeia, 2000. Also available at http://www.usp.org/medmarx.
8. http.www.fda.gov/medwatch.
9. Weingart SN, Ship AN, Aronson MD. Confidential clinician-reported surveillance of adverse events among medical inpatients. J Gen Intern Med. 2000;15:470–7.
10. Weingart SN. Identifying "barriers to care": confidential clinician-reported surveillance of adverse events among medical and surgical inpatients. In: Proceedings of Enhancing Patient Safety and Reducing Errors in Health Care; 8–10 November 1998. Rancho Mirage, California: National Patient Safety Foundation; 1998:292–6.
11. Wilson RM, Runciman WB, Gibberd RW, Harrison BT, Newby L, Hamilton JD. The Quality in Australian Health Care Study. Med J Aust. 1995;163:458–71.