

Comparing Patient-Reported Hospital Adverse Events with Medical Record Review: Do Patients Know Something That Hospitals Do Not?

Joel S. Weissman, PhD; Eric C. Schneider, MD, MSc; Saul N. Weingart, MD, PhD; Arnold M. Epstein, MD, MA; JoAnn David-Kasdan, RN, MS; Sandra Feibelmann, MPH; Catherine L. Annas, JD; Nancy Ridley, MS; Leslie Kirle, MPH; and Constantine Gatsonis, PhD

Background: Hospitals routinely survey patients about the quality of care they receive, but little is known about whether patient interviews can detect adverse events that medical record reviews do not.

Objective: To compare adverse events reported in postdischarge patient interviews with adverse events detected by medical record review.

Design: Random sample survey.

Setting: Massachusetts, 2003.

Patients: Recently hospitalized adults.

Measurements: By using parallel methods, physicians reviewed postdischarge interviews and medical records to classify hospital adverse events.

Results: Among 998 study patients, 23% had at least 1 adverse event detected by an interview and 11% had at least 1 adverse event identified by record review. The κ statistic showed relatively poor agreement between interviews and medical records for occur-

rence of any type of adverse event ($\kappa = 0.20$ [95% CI, 0.03 to 0.27]) and somewhat better agreement between interviews and medical records for life-threatening or serious events ($\kappa = 0.33$ [CI, 0.20 to 0.45]). Record review identified 11 serious, preventable events (1.1% of patients). Interviews identified an additional 21 serious and preventable events that were not documented in the medical record, including 12 pre-discharge events and 9 postdischarge events, in which symptoms occurred after the patient left the hospital.

Limitations: Patients had to be healthy enough to be interviewed. Delay in reaching patients (6 to 12 months after discharge) may have resulted in poor recall of events during the hospital stay.

Conclusion: Patients report many events that are not documented in the medical record; some are serious and preventable. Hospitals should consider monitoring patient safety by adding questions about adverse events to postdischarge interviews.

Ann Intern Med. 2008;149:100-108.

For author affiliations, see end of text.

www.annals.org

Improvements in patient safety depend in part on the ability to measure and monitor injuries to patients. The Institute of Medicine has called for building a national health information infrastructure that would include data about adverse events (1), but we lack consensus about the best way to measure adverse events.

Common options for reporting adverse events include voluntary and mandatory reporting by providers, automated surveillance of clinical data, administrative data, and record review. Each of these methods has weaknesses (1–4). The medical record is often preferred because of its rich clinical content, but medical record review is expensive, documentation can vary among different clinical settings, and underreporting may occur because of concern about medicolegal liability (5).

In recent years, patients have reported about the quality of care in various settings (6–9). Patients may also be an important source of information about hospital safety.

Many hospitals survey patients after discharge to assess their satisfaction with care, so it may be relatively easy to add questions about safety events or symptoms that may suggest the occurrence of an event. Before embarking on such an effort, it would be useful to know what would be gained from the extra effort required. In other words, can patients identify safety events that are not described in the medical record? If so, what are the characteristics of those events? We assessed the concordance of adverse events detected by patient interviews with those detected by medical record review, and then characterized the adverse events that were detected by interview but not by medical record review.

METHODS

Sample

The sample included patients age 18 years or older who were hospitalized for medical or surgical treatment and were discharged from Massachusetts hospitals from 1 April to 1 October 2003. We devised a 2-stage probability sample. In the first stage, we grouped the 71 acute care hospitals into 4 strata on the basis of size. We selected the 5 largest hospitals with certainty and allocated the remaining 66 to 3 groups of 22 on the basis of size. Five hospitals were selected from each of the 3 remaining strata, for a total of 20 hospitals. Before data collection began, 4 of the

See also:

Print

Editors' Notes 101

Web-Only

Conversion of graphics into slides

original 20 hospitals withdrew because of resource constraints, leaving 16 hospitals. In the second stage, we randomly selected 6003 patients from the hospital discharge lists to achieve a final interview target of approximately 3000 patients after projected exclusions and nonresponse.

Data Collection Protocol

Hospitals supplied patients' names, contact information, age, dates of hospitalization, and discharge status. Eligible patients received a mailing with a toll-free number to opt out of the interview. Staff at the Center for Survey Research at the University of Massachusetts, Boston, made at least 6 calls before classifying someone as a nonrespondent. Because of delays in obtaining patient lists from the hospitals, interviews took place 6 to 12 months after discharge. Patients gave informed consent orally but had to provide a signed release form to allow us to review their medical records.

The study was approved by the institutional review boards at Partners HealthCare System, the Massachusetts Department of Public Health, and each of the study hospitals. We created a firewall to prevent state staff from accessing identifiable data.

Questionnaire

We developed a draft questionnaire after holding focus groups. Initial items were cognitively tested with 10 respondents, followed by a field pretest with 32 persons. In its final form, the interview took about 20 minutes to complete. The first series of questions assessed 4 common categories of hospital treatment or management, including hospital staff administering medicines brought from home, new medicines given in the hospital, diagnostic tests, and surgery. Patients were asked to describe any "negative effects" or "complications" while hospitalized. To ensure completeness, we asked a second series of questions about 11 specified complications and injuries, including heart attack, stroke, uncontrolled bleeding, rash, and others. We recorded open-ended responses to both sets of questions.

As a supplemental inquiry, we sought information about events discovered after the patient left the hospital but were related to care provided during their stay. Complications, such as wound infections, which often manifest after discharge, are potentially of interest to quality improvement efforts. Thus, we asked whether a complication they reported had occurred before, during, or after the hospital stay and whether they "had to go back to any hospital or emergency room specifically because of complications associated with that hospital stay." We also searched the free-text descriptions. We studied these post-discharge adverse events separately from the other interview events and did not use them in our analyses of concordance between interview and medical record review.

Medical Record Review

We identified adverse events by using a structured implicit review tool adapted from the methodology used by the Harvard Medical Practice Study and others (10–13).

Context

Some hospitals review a patient's hospital record to ascertain adverse events during a recent hospital stay. The accuracy of this procedure is unknown. Asking patients about adverse events is another approach.

Contribution

The authors surveyed randomly chosen patients discharged from all hospitals in Massachusetts. Of 998 patients who returned the surveys and gave permission for medical record review, 21 reported severe, potentially preventable adverse events in addition to the 11 such events recorded in the medical record.

Caution

The surveys took place 6 to 12 months after discharge.

Implication

Postdischarge patient surveys and medical record review are complementary ways to detect adverse events in the hospital.

—The Editors

We defined *adverse events* as "unintended harm to the patient by an act of commission or omission rather than by the underlying disease or condition of the patient" (1). We created a prespecified list of 18 types of events (for example, hospital-acquired infections, adverse drug events) on the basis of a previous study (13) and created explicit evidence-based clinical case definitions to maximize reliability of judgments. The full list of events and the criteria set are available from the authors.

Clinically experienced registered nurses received training by the investigators. The nurses identified adverse events from medical records by using the clinical case definitions and prepared case summaries for physician reviewers to evaluate. Nurses also recorded patients' major diagnostic code; length of stay; and comorbid conditions, which were summarized by a Charlson score (14).

Physician Review

Physician reviewers classified events and scored severity and preventability by using parallel methods for both the medical record reviews and the patient interviews. We analyzed only events confirmed by physician review.

To classify interview events, we flagged patients' affirmative responses to queries about possible complications. Two physicians coded the first 31 cases together to clarify definitions. Next, both reviewers coded 120 reports to assess interrater reliability. Differences were resolved by consensus. One reviewer coded the remaining reports. Reviewers rated their confidence in the validity of the patient report (little or none, modest, strong, or virtually certain). Reviewers chose "little or none" if the patient report had too little information to adequately understand or interpret

it, if it was internally inconsistent, or if the description was not clinically plausible.

To classify medical record events, 2 physician reviewers (not authors) received training similar to that of the nurses. They independently assessed the case summaries and conferred to attempt consensus. By using this process, they always reached consensus without needing a third physician review.

Reviewers classified events according to severity (life-threatening, serious, clinically significant, or trivial or insignificant) and preventability (definitely, probably, probably not, or definitely not) by using written guidelines. An example of a life-threatening event is a nosocomial urinary tract infection that leads to septic shock. Serious incidents generally involve an intervention and prolonged hospitalization (for example, a wound infection that requires debridement). Clinically significant events are more transient (for example, iatrogenic pneumothorax with minimal hemodynamic changes or an adverse drug event causing transient laboratory abnormalities). Trivial or insignificant events include minor injuries, such as pain at a venipuncture site or postural dizziness. Reviewers judged the preventability of the event on the basis of an assessment of the likelihood that it resulted from a clinical intervention that was inconsistent with current clinical practice or was a complication that could have been anticipated, rather than an anticipated manifestation of an underlying disease process.

We used the κ statistic to characterize interrater reliability on the physician scores for the interview-based events (before they conferred about different scores), and weighted κ for ordinal variables. The κ scores were excellent for the presence of adverse events (0.85) and adverse drug events (0.97), type of incident (0.91), and category of hospital care (0.84). Agreement was good for preventability (0.71), fair for severity (0.45), and poor for confidence in the report (0.35). For record review, the intrarater and interrater reliability (between nurses and physicians) was reasonably good (0.61 to 0.96, depending on the measure) and was similar to previous adverse event studies that used physicians to review records (13, 15, 16).

Medical Record Repeated Review

To assess the frequency of missed or unrecorded medical record events, we instituted a post hoc repeated review. All interview events were put into a virtual envelope field in the nurses' computerized medical record review tool. Once the nurses completed the initial review, they were instructed to open the envelope. If the patient had reported an event that the nurse had not recorded, the nurse sought evidence that the event had occurred but did not meet the strict clinical case definition (yes, no, or not sure). If "no" or "not sure," the nurse was instructed to examine the record again to determine whether the event was missed the first time. Nurses received assurance that no one would judge them by their frequency of "missed" events.

Statistical Analysis

We first present descriptive statistics (counts and proportions) of adverse events by using the patient as the unit of analysis and defining a case as someone having 1 or more events. We present the characteristics of patients with adverse events identified by interviews, medical records, and both methods (concordant cases). As in other studies of concordance, we calculated overall agreement, κ scores, and the proportion of cases identified by each data source (sensitivity) by using the other data source as a gold standard (17). Thus, we calculated sensitivity of the interview by using the medical record as the gold standard and the sensitivity of the medical record by using the interview as the gold standard.

Next, we examined individual events. The study nurse determined whether each data source was describing the same event. We constructed 3 lists of events, each was identified by 1 of 3 methods of ascertainment: interview only, medical record only, and both interview and medical record. Because the probability of selection was approximately constant across all strata and preliminary analyses showed little effect from adjustment for differential non-response, the analyses were not weighted.

Role of the Funding Source

The Agency for Healthcare Research and Quality funded the study. The funding source had no role in the design and conduct of the study; the collection, management, analysis, and interpretation of the data; the preparation, review, or approval of the manuscript; or the decision to submit the manuscript for publication.

RESULTS

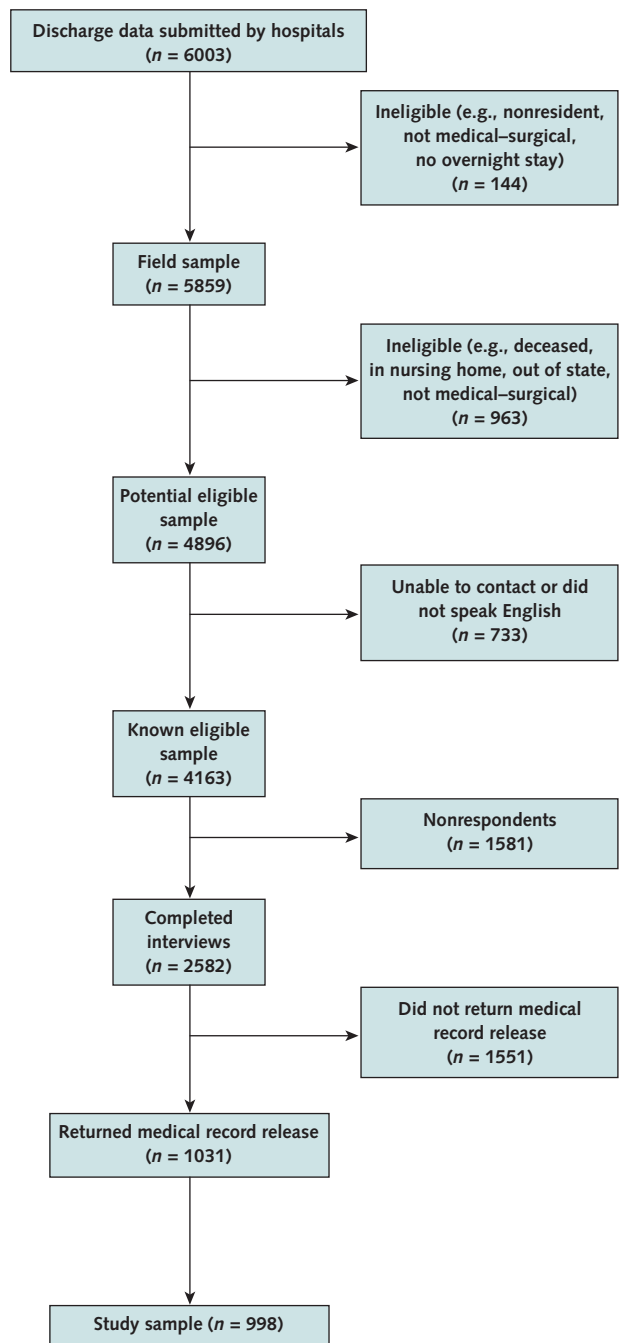
Sample Characteristics

Of 6003 patients identified by the study hospitals, 4163 patients were known to be eligible for the study (Figure). Of these, 2582 completed an interview (62% response rate) and 1031 returned medical record review authorization forms (51% of completed interviews); 998 medical records were available for review (97% of returned forms). These 998 patients constituted the study sample. Respondents who completed an interview were similar to nonrespondents with respect to sex, but they slightly underrepresented those who stayed 6 days or more in the hospital and those age 75 years or older. Patients who authorized medical record review were more likely than those who did not authorize it to be white women older than age 50 years but reported similar numbers and types of adverse events (18). Slightly more than half were women, 54% were older than age 60 years, 95% were white, and most rated their health as good (Table 1).

Characteristics of Patients with Adverse Events

Among interviewees, 229 patients (23% of total) reported 304 adverse events (1.3 events per patient with ≥ 1 event) (Table 2). In contrast, medical record reviewers

Figure. Study flow diagram.



found 105 patients (11%) with 128 events (1.2 events per patient with ≥ 1 event). Of those interviewed, women were more likely than men to report events ($P < 0.001$), as were younger patients and patients with lower Charlson scores (both $P < 0.001$). No subgroup had more adverse events when the medical record method was used. Adverse events were associated with longer lengths of stay for both ascertainment methods ($P < 0.001$).

Patient-Level Concordance of Ascertainment Methods

In the patient-level analysis, 53 patients (5.3% of total) had at least 1 adverse event of any type that was confirmed by both the interview and medical record methods. Overall agreement between the 2 methods was 77% ($\kappa = 0.20$ [95% CI, 0.03 to 0.27]). Sensitivity of the interview was 50.5% when the medical record was used as the gold standard, and sensitivity of the medical record was 23.1% when the interview was used as the gold standard. For events in which the physician reviewers had at least moderately high confidence in their ratings, agreement was 79% ($\kappa = 0.19$ [CI, 0.12 to 0.26]); for serious or life-threatening events, agreement was 94% ($\kappa = 0.33$, [CI, 0.20 to 0.45]).

Analysis of Individual Events

Table 3 summarizes the frequency of events in order to better understand the type of individual events found by 1 ascertainment method but not the other. Interviewees

Table 1. Study Sample Characteristics

Variable	Patients, n (%)
Total	998 (100)
Demographic characteristic*	
Men	473 (47.4)
Women	525 (52.6)
Age†	
18–44 y	123 (12.3)
45–59 y	341 (34.2)
≥ 60 y	534 (53.5)
Race	
White	947 (94.9)
Other	40 (4.0)
Unknown	11 (1.1)
Overall health*	
Excellent, very good, or good	776 (77.8)
Fair or poor	213 (21.3)
Unknown	9 (0.9)
Major diagnostic code‡	
Circulatory	213 (21.3)
Digestive	179 (17.9)
Musculoskeletal or connective	156 (15.6)
Respiratory	82 (8.2)
All others	368 (36.9)
Length of stay‡	
1–2 d	211 (21.1)
3 d	215 (21.5)
4–5 d	271 (27.2)
≥ 6 d	301 (30.2)
Charlson score (age-adjusted)‡§	
Low (0–1)	241 (24.1)
Medium (2)	138 (13.8)
High (3–4)	341 (34.2)
Very high (≥ 5)	278 (27.9)

*Data from interviews.

† Range, 19–97 y.

‡ Data from medical records.

§ The Charlson score assesses the burden of comorbid illness by assigning a score to hospitalized patients predictive of 1-year survival (14).

Table 2. Patients with 1 or More Adverse Events*

Variable	All Patients, n (%)	Patients with ≥1 Adverse Event		
		Interview, n (%)	Medical Record, n (%)	Interview and Medical Record, n (%)
Total†	998	229 (23)†	105 (11)†	53 (5.3)
Demographic characteristic‡				
Men	473	89 (18.8)	47 (9.9)	17 (3.6)
Women	525	140 (26.7)	58 (11.0)	36 (6.9)
<i>P</i> value§	–	0.003	0.57	0.022
Age‡				
18–44 y	123	52 (42.3)	12 (9.8)	10 (8.1)
45–59 y	341	93 (27.3)	43 (12.6)	27 (7.9)
≥60 y	534	84 (15.7)	50 (9.4)	16 (3.0)
<i>P</i> value§	–	<0.001	0.30	0.002
Race‡				
White	947	217 (22.9)	101 (10.7)	51 (5.4)
Other	40	10 (25.0)	3 (7.5)	2 (5.0)
Unknown	11	2 (18.2)	1 (9.1)	0 (0.0)
<i>P</i> value (excludes unknown)§	–	0.76	0.79	1.00
Overall health‡				
Excellent, very good, or good	776	178 (22.9)	83 (10.7)	41 (5.3)
Fair or poor	213	50 (23.5)	21 (9.9)	12 (5.6)
Unknown	9	1 (11.1)	1 (11.1)	0 (0.0)
<i>P</i> value (excludes unknown)§	–	0.87	0.72	0.84
Major diagnostic code 				
Circulatory	213	35 (16.4)	21 (9.9)	9 (4.2)
Digestive	179	49 (27.4)	26 (14.5)	17 (9.5)
Musculoskeletal or connective	156	50 (32.1)	15 (9.6)	5 (3.2)
Respiratory	82	16 (19.5)	6 (7.3)	4 (4.9)
All others	368	79 (21.5)	37 (10.1)	18 (4.9)
<i>P</i> value§	–	0.004	0.37	0.082
Length of stay‡				
1–2 d	211	27 (12.8)	7 (3.3)	2 (0.9)
3 d	215	42 (19.5)	5 (2.3)	4 (1.9)
4–5 d	271	69 (25.5)	20 (7.4)	12 (4.4)
≥6 d	301	91 (30.2)	73 (24.3)	35 (11.6)
<i>P</i> value§	–	<0.001	<0.001	<0.001
Charlson score (age-adjusted) 				
Low (0–1)	241	79 (32.7)	20 (8.3)	14 (5.8)
Medium (2)	138	40 (29.0)	17 (12.3)	10 (7.2)
High (3–4)	341	64 (18.8)	30 (8.8)	16 (4.7)
Very high (≥5)	278	46 (16.5)	38 (13.7)	13 (4.7)
<i>P</i> value§	–	<0.001	0.13	0.65

* Based on interview reports and medical record reviews.

† 229 patients reported 304 events from the interview; 105 patients had 128 events identified by means of medical record review. The percent agreement of these data was 77.2% = (53 + 717)/998. By using the medical record as the gold standard, sensitivity was 50.5% (95% CI, 40.5%–60.4%) (53 of 105 patients); specificity was 80.3% (CI, 77.5%–82.8%) (717 of 893 patients); positive predictive value was 23.1% (CI, 17.8%–29.1%) (53 of 229 patients); and negative predictive value was 93.2% (CI, 91.2%–94.9%) (717 of 769 patients). By using the patient as the gold standard, sensitivity was 23.1% (CI, 17.8%–29.1%) (53 of 229 patients); specificity was 93.2% (CI, 91.2%–94.9%) (717 of 769 patients); positive predictive value was 50.5% (CI, 40.5%–60.4%) (53 of 105 patients); and negative predictive value was 80.3% (CI, 77.5%–82.8%) (717 of 893 patients).

‡ Data from interviews.

§ *P* values for the hypothesis of no association between each patient characteristic and the presence of ≥1 adverse event (assessed by the indicated method) were based on chi-square tests, when appropriate, and Fisher exact tests in all other cases.

|| Data from medical records.

reported 253 events that were not found in the medical records. Medical record review identified 82 events that were not reported in the interviews. Forty-six events occurred in both sources.

Of the 253 events identified only by interview, 135

(53%) were adverse drug events. The next most frequent category was “other,” which included difficulty breathing, decreased blood pressure, excessive pain, and non-specific rashes. Of the 253 events discovered only by interview, 34 (13%) were serious and 2 (1%) were life-

threatening; 73 (29%) and 2 (1%) were considered to be probably preventable and definitely preventable, respectively. Events that were both serious and preventable represented 4.7% of events identified only on interview.

Events Discovered through Repeated Medical Record Review and after Supplemental Analysis

We repeated the medical record review to determine whether the lack of agreement with the interviews was because of the strict case definitions used in record review or

mistakes by reviewers. During this process, we identified an additional 64 patients as having 1 or more medical record events (62 had not met the original case definitions and 2 were missed). Agreement between interview and medical record review increased to 84% ($\kappa = 0.49$ [CI, 0.42 to 0.55]). The sensitivity of the interview (when the record review was used as the gold standard) increased to 69%, whereas the sensitivity of the medical record (when the interview was used as gold standard) increased to 51%. Although the number of individual

Table 3. Types of Individual Adverse Events, by Method of Ascertainment

Adverse Event	Interview, n (%)	Medical Record, n (%)	Interview and Medical Record, n (%)
Total*	253 (100)	82 (100)	46 (100)
Type			
Wound infection	4 (1.6)	9 (11.0)	4 (8.7)
Hospital-acquired urinary tract infection	3 (1.2)	4 (4.9)	1 (2.2)
Hospital-acquired pneumonia	1 (0.4)	3 (3.7)	1 (2.2)
Hospital-acquired bacteremia	2 (0.8)	1 (1.2)	0 (0)
Hospital-acquired sepsis	0 (0)	1 (1.2)	0 (0)
Operative nerve injury	22 (8.7)	1 (1.2)	0 (0)
Operative organ injury	1 (0.4)	3 (3.7)	3 (6.5)
Operative vessel injury	9 (3.6)	8 (9.8)	5 (10.9)
Postoperative acute myocardial infarction	2 (0.8)	1 (1.2)	1 (2.2)
Postoperative stroke	0 (0)	0 (0)	0 (0)
Postoperative shock	1 (0.4)	0 (0)	0 (0)
Postoperative respiratory distress	1 (0.4)	2 (2.4)	0 (0)
Iatrogenic pneumothorax	0 (0)	3 (3.7)	2 (4.3)
Hospital-acquired pulmonary embolism	0 (0)	0 (0)	0 (0)
Hospital-acquired deep venous thrombosis	3 (1.2)	0 (0)	3 (6.5)
Fall	4 (1.6)	6 (7.3)	3 (6.5)
Pressure sore	4 (1.6)	5 (6.1)	1 (2.2)
Adverse drug event	135 (53.4)	32 (39.0)	21 (45.7)
Other	61 (24.1)	3 (3.7)	1 (2.2)
Severity			
Life-threatening	2 (0.8)	1 (1.2)	3 (6.5)
Serious	34 (13.4)	32 (39.0)	15 (32.6)
Clinically significant	158 (62.5)	41 (50.0)	26 (56.5)
Insignificant, trivial	59 (23.3)	8 (9.8)	2 (4.3)
Preventability			
Definitely	2 (0.8)	2 (2.4)	3 (6.5)
Probably	73 (28.9)	29 (35.4)	13 (28.3)
Probably not	171 (67.6)	23 (28.0)	17 (37.0)
Definitely not	7 (2.8)	0 (0)	0 (0)
Unable to determine	NA	28 (34.1)	13 (28.3)
Serious and preventable†	12 (4.7)	11 (13.4)	9 (19.6)
Confidence of raters			
Modest	44 (17.4)	22 (26.8)	13 (28.3)
Strong	209 (82.6)	60 (73.2)	33 (71.7)
Category of hospital care			
Own medications	16 (6.3)	NA	NA
Hospital medications	115 (45.5)	NA	NA
Surgical complications	80 (31.6)	NA	NA
Procedures or tests	18 (7.1)	NA	NA
Miscellaneous experience	24 (9.5)	NA	NA

NA = not applicable.

* Percentages may not total to 100% because of rounding. Six events identified by interview were determined to be duplicates (the same event) and were excluded.

† Includes events classified as "serious" or "life-threatening" and "probably" or "definitely" preventable.

Table 4. Serious and Preventable Adverse Events Reported in Interview Only*

Timing and Type	Description
During hospital stay (predischARGE) (n = 12)	
Hospital-acquired deep venous thrombosis	After open heart surgery, the patient developed pain in his leg, which was later discovered to be deep venous thrombosis.
Adverse drug event	After placement of a left ventricular assist device, the patient received large amounts of diuretics. He developed electrolyte abnormalities, which caused delirium.
Operative vessel injury	After placement of an implantable cardiac defibrillator, the patient developed bleeding that led to a prolonged hospitalization.
Adverse drug event	After receiving new medications in the hospital, the patient became disoriented and confused for 24 hours.
Hospital-acquired pneumonia	After a Whipple procedure, the patient developed postoperative pneumonia.
Adverse drug event	After receiving sleeping medication, the patient became disoriented, climbed out of bed, and acted disruptively.
Wound infection	The patient developed a wound infection after a knee replacement.
Hospital-acquired deep venous thrombosis	After a knee replacement, the patient developed deep venous thrombosis in the lower leg.
Other	After splenectomy, it was discovered that the patient had a bone marrow disorder that could have been treated with medication alone.
Postoperative respiratory distress	After surgery for repair of an ankle fracture, the patient developed difficulty breathing.
Operative vessel injury	The patient bled after laparoscopic gallbladder surgery, which necessitated additional surgery to stop the bleeding.
Operative nerve injury	After lung resection for cancer, the patient noted arm numbness and weakness of the hand.
Postdischarge (n = 9)	
Other	After repair of a ventral hernia, the patient returned to the hospital with wound dehiscence.
Wound infection	After bilateral surgery on her legs, the patient returned to the hospital with infected surgical wounds.
Wound infection	After back surgery, the patient returned to the hospital with an infection requiring reopening of the wound.
Wound infection	After abdominal surgery to relieve bowel obstruction, the patient returned to the hospital for additional surgery to address an infection involving the screen used to repair a hernia.
Wound infection	The patient had surgery on the arteries of his legs. He developed a wound infection requiring a return to the hospital for additional surgery and treatment.
Hospital-acquired deep venous thrombosis	The patient had an admission for leg surgery. After discharge, he developed deep venous thrombosis necessitating readmission to the hospital for treatment.
Wound infection	The patient had implants placed for radiation therapy. One of the implants became infected, requiring a return to the hospital for additional treatment.
Wound infection	After surgery for a broken leg, the patient developed a <i>Staphylococcus aureus</i> infection and was readmitted for additional surgery and treatment.
Operative organ injury	The patient developed a bile leak after laparoscopic gallbladder removal and returned to the hospital for operative repair and pain control.

* Reported in interview but not documented in the medical record. Includes events classified as “serious” or “life-threatening” and “probably” or “definitely” preventable. Each event listed in this table represents a unique patient, and each patient had only 1 of these events.

events identified by both sources increased from 46 on the initial review to 129 when reviewed again, 170 events identified by interview but not the medical record remained.

The interviews identified 27 postdischarge events, 21 of which were the result of hospital events that did not manifest until the patient was at home. The remaining 6 events did not represent new hospital-related adverse events.

Description of Serious and Preventable Events Discovered by Interview

Hospitals may be most interested in serious and preventable events identified by means of interview but not documented in medical records. Table 4 shows the 12 serious and preventable pre-discharge events (occurring in 1.2% of the 998 patients in the sample) and 9 postdischarge events (0.9% of patients) identified by the interview that were not documented in the medical record. Together, they represented 2.1% of patients. For example, 1 patient developed deep venous thrombosis in the lower part of his leg after a knee replacement. Another patient

developed a postdischarge wound infection after bilateral surgery on her legs.

DISCUSSION

We examined postdischarge interviews as a new method to identify patients who had adverse events while hospitalized by comparing their findings to a review of medical records. By using strict case definitions of adverse events, we found that the medical record identified approximately one fourth of the events reported by patients. The 2 detection methods have only moderate concordance, which suggests that neither one alone can serve as a gold standard. Several methods may be required to obtain a full account of patient injury.

Regulatory agencies increasingly require hospitals to track adverse events as a measure of performance and to monitor patient safety. In a review of a series of rigorous medical record review studies in the United States, Canada, Europe, and Australia, the reported incidence of ad-

verse events among hospitalized patients was between 2.9% and 16.7% (10, 19–24). The Institute for Healthcare Improvement recently called for hospitals to review patient records to identify adverse events (25). As hospitals allocate resources to respond to these requirements, they will have to decide whether conducting routine patient surveys to identify adverse events is a productive use of resources.

Although postdischarge interviews identified only a small number of additional serious and preventable events—only 1% to 2% of discharges—they represented a substantial increase in yield compared with the number of events identified from medical record review. Nevertheless, with such a low incidence of serious preventable events, it seems unlikely that hospitals will undertake routine medical record reviews or single-purpose surveys for monitoring unless they are required to do so. The cost of surveying patients could be minimized, however, if questions about adverse events were added to extant postdischarge surveys, which is now commonly done.

This study adds to the literature about the value of using patients to report on their own medical treatment and quality of care. Most previous studies about concordance between patients and medical records have focused on reporting diseases (26–28), the use of services (29, 30), or ambulatory adverse drug events (31). To our knowledge, only 1 study examined inpatient events and reported that 55% of patient-reported events had corroborating information in the medical record; however, that study involved only 228 patients on a single unit at 1 teaching hospital (32). We had no independent method to verify the occurrence of events not found in the medical record. However, the agreement between the medical record and the interviews was much higher for serious events (94%) than for clinically significant and trivial events (77% overall), which suggests that patient reports are generally valid.

Getting accurate patient reports of adverse events is difficult because of problems with recall, clinical knowledge, and social desirability bias (33–35). Patients may be reluctant to identify events because they fear alienating their caregivers. Some may be too sick to report, although concordance in our study was higher among patients with more severe adverse events. Thus, we suggest that the best use of postdischarge questions about adverse events may be to screen cases for targeted review of selected medical records or additional investigation.

Our study has several limitations that may affect its generalizability. First, patients must have survived their hospitalizations and have been healthy enough to participate. Patients who died during their stay may have been at higher risk for an adverse event. Second, some patients had more than 1 hospital admission, and they may have attributed the events inaccurately. Third, we could not interview patients until 6 to 12 months after discharge, which may mean that the concordance rates are lower than if the interview took place soon after hospital discharge. Fourth, only 51% of patients completing interviews authorized

record review (not uncommon for complex consent forms requesting access to sensitive information) (18). Nevertheless, persons authorizing access may have been more likely to have had a serious adverse event. Fifth, most respondents were white and non-Hispanic. Other racial or ethnic communities may report complications differently. Sixth, our physician reviewers judged the preventability of an event on the basis of evidence in the medical record or from the patient; a root cause analysis might have provided a more accurate assessment. Seventh, older patients reported lower rates of events, which could be because of poorer recall, failure to recognize events, a tendency to give caregivers the benefit of the doubt, or greater debility. Finally, we did not report on the marginal cost of obtaining data about adverse events by means of patient interviews.

Hospitals could use surveys to detect adverse events to identify patients for a targeted record review, in place of medical record reviews, or as a complementary search strategy. To reduce costs, hospitals could add items to the routine postdischarge surveys that many already conduct.

The Institute of Medicine and others view several sources of adverse event reporting as complementary (1, 4). We have tested this concept by asking unselected patients to report adverse events and by reviewing their medical records for further evidence of adverse events. We found that even many months after discharge from the hospital, 23% of patients identified an adverse event that affected their care, compared with a rate of 11% using medical record review. The interviews identified 21 serious and preventable adverse events that the medical record review did not detect. Each hospital will decide whether this yield justifies the effort and expense of a survey. A hospital's decision will depend on its patient safety culture and on the marginal value it places on identifying additional adverse events. Policymakers will have their own set of considerations when they decide whether to ask hospitals to survey for adverse events. Hopefully, both will find ways to use patients' experiences to inform their efforts to serve the public.

From Institute for Health Policy, Massachusetts General Hospital, Harvard School of Public Health, Dana-Farber Cancer Institute, and Massachusetts Department of Public Health, Boston, Massachusetts; Massachusetts Hospital Association, Burlington, Massachusetts; and Brown University, Providence, Rhode Island.

Acknowledgment: The authors thank Jack Fowler, PhD; Brian Claridge, PhD; and Dragana Bolcic-Jankovic, MA, for help and support with questionnaire design and administration, and Nancy Wong for assistance with data analysis and programming.

Grant Support: By a cooperative agreement from the Agency for Healthcare Research and Quality to the Massachusetts Department of Public Health (U18 HS11928).

Potential Financial Conflicts of Interest: *Grants received:* J.S. Weissman (Agency for Healthcare Research and Quality), E.C. Schneider (Agency for Healthcare Research and Quality), S.N. Weingart (Agency

for Healthcare Research and Quality), A.M. Epstein (Agency for Healthcare Research and Quality), J. David-Kasdan (Agency for Healthcare Research and Quality), S. Feibelman (Agency for Healthcare Research and Quality), C.L. Annas (Agency for Healthcare Research and Quality), N. Ridley (Agency for Healthcare Research and Quality), L. Kirle (Agency for Healthcare Research and Quality).

Reproducible Research Statement: *Study protocol and data set:* Available from Dr. Weissman (e-mail, joel.weissman@state.ma.us). *Statistical code:* Not available.

Requests for Single Reprints: Joel S. Weissman, PhD, Massachusetts Executive Office of Health and Human Services, One Asburton Place, Room 1109, Boston, MA 02108; e-mail, joel.weissman@state.ma.us.

Current author addresses and author contributions are available at www.annals.org.

References

- Institute of Medicine. Patient Safety: Achieving a New Standard for Care. Washington, DC: National Academy Press; 2003.
- Beckmann U, Bohringer C, Carless R, Gillies DM, Runciman WB, Wu AW, et al. Evaluation of two methods for quality improvement in intensive care: facilitated incident monitoring and retrospective medical chart review. *Crit Care Med.* 2003;31:1006-11. [PMID: 12682464]
- Marchev M, Rosenthal J, Booth M. How states report medical errors to the public: issues and barriers. Portland, ME: National Academy for State Health Policy; 2003.
- Thomas EJ, Petersen LA. Measuring errors and adverse events in health care. *J Gen Intern Med.* 2003;18:61-7. [PMID: 12534766]
- Thomas EJ, Studdert DM, Runciman WB, Webb RK, Sexton EJ, Wilson RM, et al. A comparison of iatrogenic injury studies in Australia and the USA. I: context, methods, casemix, population, patient and hospital characteristics. *Int J Qual Health Care.* 2000;12:371-8. [PMID: 11079216]
- Goldstein E, Farquhar M, Crofton C, Darby C, Garfinkel S. Measuring hospital care from the patients' perspective: an overview of the CAHPS Hospital Survey development process. *Health Serv Res.* 2005;40:1977-95. [PMID: 16316434]
- Safran DG, Karp M, Coltin K, Chang H, Li A, Ogren J, et al. Measuring patients' experiences with individual primary care physicians. Results of a statewide demonstration project. *J Gen Intern Med.* 2006;21:13-21. [PMID: 16423118]
- Castle NG, Brown J, Hepner KA, Hays RD. Review of the literature on survey instruments used to collect data on hospital patients' perceptions of care. *Health Serv Res.* 2005;40:1996-2017. [PMID: 16316435]
- Sofaer S, Crofton C, Goldstein E, Hoy E, Crabb J. What do consumers want to know about the quality of care in hospitals? *Health Serv Res.* 2005;40:2018-36. [PMID: 16316436]
- 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. [PMID: 1987460]
- 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. [PMID: 1824793]
- Thomas EJ, Burstin HR, Orav EJ, Zeena T, Williams EJ, Howard KM, et al. Incidence of and risk factors for adverse events and negligent care in Colorado and Utah in 1992. *J Gen Intern Med.* 1997;12(Suppl):81.
- Weissman JS, Rothschild JM, Bendavid E, Sprivilis P, Cook EF, Evans RS, et al. Hospital workload and adverse events. *Med Care.* 2007;45:448-55. [PMID: 17446831]
- Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis.* 1987;40:373-83. [PMID: 3558716]
- Forster AJ, Murff HJ, Peterson JF, Gandhi TK, Bates DW. Adverse drug events occurring following hospital discharge. *J Gen Intern Med.* 2005;20:317-23. [PMID: 15857487]
- Forster AJ, Murff HJ, Peterson JF, Gandhi TK, Bates DW. The incidence and severity of adverse events affecting patients after discharge from the hospital. *Ann Intern Med.* 2003;138:161-7. [PMID: 12558354]
- Tisnado DM, Adams JL, Liu H, Damberg CL, Chen WP, Hu FA, et al. What is the concordance between the medical record and patient self-report as data sources for ambulatory care? *Med Care.* 2006;44:132-40. [PMID: 16434912]
- Bolcic-Jankovic D, Clarridge BR, Fowler FJ Jr, Weissman JS. Do characteristics of HIPAA consent forms affect the response rate? *Med Care.* 2007;45:100-3. [PMID: 17279027]
- 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. [PMID: 7476634]
- Michel P, Quenon JL, Djihoud A, Tricaud-Vialle S, de Sarasqueta AM. French national survey of inpatient adverse events prospectively assessed with ward staff. *Qual Saf Health Care.* 2007;16:369-77. [PMID: 17913779]
- Vincent C, Neale G, Woloshynowych M. Adverse events in British hospitals: preliminary retrospective record review. *BMJ.* 2001;322:517-9. [PMID: 11230064]
- World Health Organization. National Study on Hospitalisation-Related Adverse Events. Geneva: World Health Organization; 2006.
- Baker GR, Norton PG, Flintoft V, Blais R, Brown A, Cox J, et al. The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. *CMAJ.* 2004;170:1678-86. [PMID: 15159366]
- 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;38:261-71. [PMID: 10718351]
- Griffin FA, Resar RK. IHI Global Trigger Tool for Measuring Adverse Events. Cambridge, MA: Institute for Healthcare Improvement; 2007.
- Fowles JB, Fowler EJ, Craft C. Validation of claims diagnoses and self-reported conditions compared with medical records for selected chronic diseases. *J Ambul Care Manage.* 1998;21:24-34. [PMID: 10181337]
- Fowles JB, Rosheim K, Fowler EJ, Craft C, Arrichiello L. The validity of self-reported diabetes quality of care measures. *Int J Qual Health Care.* 1999;11:407-12. [PMID: 10561032]
- Kwon A, Bungay KM, Pei Y, Rogers WH, Wilson IB, Zhou Q, et al. Antidepressant use: concordance between self-report and claims records. *Med Care.* 2003;41:368-74. [PMID: 12618640]
- Rozario PA, Morrow-Howell N, Proctor E. Comparing the congruency of self-report and provider records of depressed elders' service use by provider type. *Med Care.* 2004;42:952-9. [PMID: 15377927]
- Weissman JS, Levin K, Chasan-Taber S, Massagli MP, Seage GR 3rd, Scampini L. The validity of self-reported health-care utilization by AIDS patients. *AIDS.* 1996;10:775-83. [PMID: 8805870]
- Gandhi TK, Weingart SN, Borus J, Seger AC, Peterson J, Burdick E, et al. Adverse drug events in ambulatory care. *N Engl J Med.* 2003;348:1556-64. [PMID: 12700376]
- Weingart SN, Pagovich O, Sands DZ, Li JM, Aronson MD, Davis RB, et al. What can hospitalized patients tell us about adverse events? Learning from patient-reported incidents. *J Gen Intern Med.* 2005;20:830-6. [PMID: 16117751]
- Ritter PL, Stewart AL, Kaymaz H, Sobel DS, Block DA, Lorig KR. Self-reports of health care utilization compared to provider records. *J Clin Epidemiol.* 2001;54:136-41. [PMID: 11166528]
- Wallihan DB, Stump TE, Callahan CM. Accuracy of self-reported health services use and patterns of care among urban older adults. *Med Care.* 1999;37:662-70. [PMID: 10424637]
- Groves R, Fower F, Couper M, Lepkowski J, Singer E, Tourangeau R. *Survey Methodology.* Hoboken, NJ: Wiley; 2004.

Current Author Addresses: Dr. Weissman: Massachusetts Executive Office of Health and Human Services, One Ashburton Place, Room 1109, Boston, MA 02108.

Drs. Schneider and Epstein: Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115.

Dr. Weingart: Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115.

Ms. David-Kasdan and Ms. Feibelman: Institute for Health Policy, Massachusetts General Hospital, 50 Staniford Street, 9th Floor, Boston, MA 02114.

Ms. Annas: Eastern Massachusetts Healthcare Initiative, 14 Story Street, 2nd Floor, Cambridge, MA 02138.

Ms. Ridley: Massachusetts Department of Public Health, 250 Washington Street, Boston, MA 02108.

Ms. Kirle: Tufts Medical Center, 800 Washington Street, Boston, MA 02111.

Dr. Gatsonis: Brown University, Center for Statistical Sciences, Box GS-121, Providence, RI 02912.

Author Contributions: Conception and design: J.S. Weissman, E.C. Schneider, S.N. Weingart, A.M. Epstein, C.L. Annas, N. Ridley, L. Kirle, C. Gatsonis.

Analysis and interpretation of the data: J.S. Weissman, E.C. Schneider, S.N. Weingart, A.M. Epstein, J. David-Kasdan, L. Kirle.

Drafting of the article: J.S. Weissman, J. David-Kasdan, S. Feibelman, N. Ridley, L. Kirle.

Critical revision of the article for important intellectual content: J.S. Weissman, E.C. Schneider, S.N. Weingart, A.M. Epstein, J. David-Kasdan, S. Feibelman, C.L. Annas, N. Ridley.

Final approval of the article: J.S. Weissman, E.C. Schneider, S.N. Weingart, A.M. Epstein, J. David-Kasdan, S. Feibelman.

Statistical expertise: C. Gatsonis.

Obtaining of funding: J.S. Weissman, A.M. Epstein, N. Ridley, L. Kirle, C. Gatsonis.

Administrative, technical, or logistic support: J.S. Weissman, S.N. Weingart, J. David-Kasdan, S. Feibelman.

Collection and assembly of data: J.S. Weissman, J. David-Kasdan.