SPECIAL ARTICLE

The Safety of Inpatient Health Care

David W. Bates, M.D., David M. Levine, M.D., M.P.H., Hojjat Salmasian, M.D., Ph.D., M.P.H., Ania Syrowatka, Ph.D., David M. Shahian, M.D., Stuart Lipsitz, Sc.D., Jonathan P. Zebrowski, M.D., M.H.Q.S., Laura C. Myers, M.D., M.P.H., Merranda S. Logan, M.D., M.P.H., Christopher G. Roy, M.D., M.P.H., Christine Iannaccone, M.P.H., Michelle L. Frits, B.A., Lynn A. Volk, M.H.S., Sevan Dulgarian, B.S., B.A., Mary G. Amato, Pharm.D., M.P.H., Heba H. Edrees, Pharm.D., Luke Sato, M.D., Patricia Folcarelli, Ph.D., R.N., Jonathan S. Einbinder, M.D., M.P.H., Mark E. Reynolds, B.A., and Elizabeth Mort, M.D., M.P.H.

ABSTRACT

BACKGROUND

Adverse events during hospitalization are a major cause of patient harm, as documented in the 1991 Harvard Medical Practice Study. Patient safety has changed substantially in the decades since that study was conducted, and a more current assessment of harm during hospitalization is warranted.

METHODS

We conducted a retrospective cohort study to assess the frequency, preventability, and severity of patient harm in a random sample of admissions from 11 Massachusetts hospitals during the 2018 calendar year. The occurrence of adverse events was assessed with the use of a trigger method (identification of information in a medical record that was previously shown to be associated with adverse events) and from review of medical records. Trained nurses reviewed records and identified admissions with possible adverse events that were then adjudicated by physicians, who confirmed the presence and characteristics of the adverse events.

RESULTS

In a random sample of 2809 admissions, we identified at least one adverse event in 23.6%. Among 978 adverse events, 222 (22.7%) were judged to be preventable and 316 (32.3%) had a severity level of serious (i.e., caused harm that resulted in substantial intervention or prolonged recovery) or higher. A preventable adverse event occurred in 191 (6.8%) of all admissions, and a preventable adverse event with a severity level of serious or higher occurred in 29 (1.0%). There were seven deaths, one of which was deemed to be preventable. Adverse drug events were the most common adverse events (accounting for 39.0% of all events), followed by surgical or other procedural events (30.4%), patient-care events (which were defined as events associated with nursing care, including falls and pressure ulcers) (15.0%), and health care–associated infections (11.9%).

CONCLUSIONS

Adverse events were identified in nearly one in four admissions, and approximately one fourth of the events were preventable. These findings underscore the importance of patient safety and the need for continuing improvement. (Funded by the Controlled Risk Insurance Company and the Risk Management Foundation of the Harvard Medical Institutions.)

From the Division of General Internal Medicine and Primary Care, Brigham and Women's Hospital (D.W.B., D.M.L., H.S., A.S., S.L., C.I., M.L.F., S.D., M.G.A., H.H.E., L.S.), the Department of Health Care Policy (E.M.), Harvard Medical School (D.W.B., D.M.L., H.S., A.S., D.M.S., S.L., J.P.Z., M.S.L., H.H.E., L.S., E.M.), the Department of Health Policy and Management, Harvard T.H. Chan School of Public Health (D.W.B.), the Edward P. Lawrence Center for Quality and Safety (D.M.S., J.P.Z., E.M.), the Division of Cardiac Surgery, Department of Surgery (D.M.S.), the Department of Psychiatry (J.P.Z.), the Division of Nephrology (M.S.L.), and the Division of General Internal Medicine (E.M.), Massachusetts General Hospital, and the Controlled Risk Insurance Company and the Risk Management Foundation of the Harvard Medical Institutions (L.S., P.F., J.S.E., M.E.R.) — all in Boston; the Kaiser Permanente Northern California Division of Research, Oakland (L.C.M.); Maine Medical Center, Portland (C.G.R.); and Mass General Brigham, Somerville, MA (L.A.V.). Dr. Bates can be contacted at dbates@bwh.harvard.edu or at the Division of General Internal Medicine and Primary Care, Brigham and Women's Hospital, 1620 Tremont St., Boston, MA 02120.

N Engl J Med 2023;388:142-53.
DOI: 10.1056/NEJMsa2206117

The Harvard Medical Practice Study (HMPS) was conducted in a sample of patients hospitalized in New York State in 1984, and the results were published in 1991.[1,2] Key findings included an adverse event rate of 3.7 events per 100 admissions, of which 28% were judged to have been caused by negligence; 16% led to death or permanent disability. The HMPS was an extensive study that focused on medical injury and litigation. It played a major role in informing the report by the Institute of Medicine (now known as the National Academy of Medicine) titled, "To Err Is Human: Building a Safer Health System,"[3] which brought the problem of patient safety into the public eye. Notable follow-up studies were conducted in the United States, including a study involving hospitals in Utah and Colorado that was conducted by many of the same investigators who were involved in the HMPS and used methods that were similar to those used in the HMPS,[4] as well as a study by Landrigan and colleagues in which data on adverse events were reviewed at 10 hospitals in North Carolina over a 6-year period.[5] Many international studies have also been conducted.[6,7]

Patient safety has changed substantially since the HMPS was performed,[8] as exemplified by the development of effective strategies for preventing specific types of adverse events such as catheter-related bloodstream infections[9] and surgery-related adverse events.[10] In addition, more efficient and reliable approaches have been established for the identification of adverse events, including the use of "triggers," whereby information in a medical record that was previously shown to be associated with adverse events is identified.[11]

However, documenting the extent to which patient safety has improved has been challenging, despite major efforts such as reports commissioned by the Office of Inspector General of the Department of Health and Human Services that provide national estimates of harm in the Medicare population. Such reports have shown a modest serial decrease in the incidence of health care–associated infections.[12,13] However, in contrast to health care–associated infections, many key safety domains lack metrics that can be easily measured by organizations to routinely track adverse events and assess progress in improving safety.[8] One example is adverse drug events (defined as injuries resulting from drugs that were taken), for which the change in incidence over time remains unclear, given that hospitals do not routinely measure the frequency of such events. They occur much more often than voluntary incident reporting suggests; one study showed a measured incidence that was almost 20 times as high as the incidence identified through voluntary reporting.[14]

Many aspects of health care have changed since the HMPS. For example, electronic health records (EHRs), which were rare when the initial HMPS was conducted, are now in routine use. Furthermore, a substantial proportion of medical care has shifted from the inpatient to the outpatient setting. In the current study, the SafeCare study, we assessed the frequency of adverse events in both inpatients and outpatients; in this report, however, only the former are described.

We report the frequency and types of harm in a cohort of 11 hospitals in Massachusetts. These hospitals all had the same malpractice insurance carrier, which provided support for this study as a component of its mission.

Methods

Hospital Selection

We conducted a retrospective cohort study. The 11 participating hospitals were specifically selected to include a range of both large and smaller hospitals, and the investigators were unaware of the internally measured incidences of adverse events in these hospitals. The hospitals were associated with three health care systems. Two hospitals had fewer than 100 beds, 4 had 100 to 200 beds, 2 had 201 to 500 beds, and 3 had more than 700 beds. The study sample was designed to include hospitals and patient populations that would provide reliable estimates of safety and safety-related metrics among patients 18 years of age or older at each location. All the participating hospitals agreed to undergo review by the Mass General Brigham institutional review board, which approved this study.

Sampling

At each of the participating hospitals, a random sample of admissions records was obtained, with oversampling in the smaller hospitals. The target sample from the participating hospitals in Massachusetts included all inpatient admissions with discharges occurring in 2018, except for the following: admissions for hospice, rehabilitation, or psychiatric care; for addiction treatment; and for

observation only, under the two-midnight rule, which categorizes a hospital stay that does not cross two midnights as an observation-only encounter. A total sample of 2750 admissions (a mean of 250 per hospital) was calculated. Four smaller hospitals were oversampled, which resulted in a final sample of 2836 admissions. Additional details are provided in the Supplementary Appendix, available with the full text of this article at NEJM.org.

RECORD REVIEW

Nine nurses performed reviews of the admissions records to identify possible adverse events. These reviewers followed a detailed manual that outlined the process for chart review and described the specific types of data to be collected. In this study, adverse events were defined as "unintended physical injury resulting from or contributed to by medical care that requires additional monitoring, treatment, or hospitalization, or that results in death."[15] Medical care included the actions of individual hospital staff as well as the broader systems and care processes and included both acts of omission (failure to diagnose or treat) and acts of commission (incorrect diagnosis or treatment, or poor performance).

The reviewers were randomly assigned admissions across the hospitals. If the reviewers identified information on a given chart that warranted further follow-up to identify adverse events related to the index admission, they were permitted to review data that had been recorded in the chart up to 30 days after the patient's discharge. To help the reviewer determine whether harm was related to the index admission, no limit was established for the review of chart information that had been recorded before the index admission. The reviewers followed a protocol detailing the sequence of reviewing an admission in Epic (Epic Systems), the most common EHR system used by the hospitals. Admissions to hospitals that used an EHR system other than Epic were randomly assigned to reviewers who were trained in the use of the other systems, and these reviewers followed a protocol similar to that used for Epic. Eight hospitals used Epic, 2 used Meditech, and 1 used a custom-made EHR system. All data were entered into a data-collection tool (which had been created with the use of Microsoft Access) that performed live data validation.

The reviewers looked for triggers (Fig. S1 in the Supplementary Appendix).[11] For each admission, the reviewers could document up to eight possible adverse events; the maximum number of adverse events was observed in only eight patient charts (0.28%). When the reviewers identified an adverse event, the type of event was classified as a blood-transfusion reaction, a health care–associated infection, an adverse drug event, an event associated with pregnancy or the perinatal period, an event related to a surgical or other procedure, or a patient-care event (which was defined as an event related to nursing care, including falls and pressure ulcers) (Fig. S2). Reviewers also looked for any indication that an error occurred during care, such as an error in diagnosis or in the performance of a procedure (Fig. S4). In addition, the reviewers provided an overall narrative summary of the admission and a separate summary describing each possible adverse event.

Eight physicians reviewed randomly assigned adverse event summaries and either agreed or disagreed with the adverse event type. If these adjudicators disagreed, the event type was changed. When the adjudicators had questions or thought that one adverse event should be counted as several, they could send their questions or comments to the nurse to review again. In addition, the adjudicators ranked the severity of each event with the use of a general severity scale[16] that categorized events as significant, serious, life-threatening, or fatal (see Table S6 for definitions and examples). They also provided assessments of whether the harm was preventable,[17] and they graded their confidence (with the use of a six-point ordinal scale) regarding whether the event was caused by health care management.[18] A confidence score of 4 or higher (which indicated that health care management was slightly more likely than not to have caused the event) indicated that an adverse event had occurred; this confidence threshold aligned with that used in the HMPS.[1] Details are provided in the Supplementary Appendix.

RELIABILITY

Overall, 10% of the possible adverse events that had been identified were randomly selected to be judged by a second physician. Each of these possible events was randomly assigned to an adjudicator who had not previously reviewed the event. Adjudicators were not given any information about the initial review.

STATISTICAL ANALYSIS

We used a sampling design in which some of the smaller hospitals were oversampled; each patient's admission record that was sampled was assigned a weight for the analyses. The weight was the inverse of the proportion of admission records sampled from that hospital. Using these weights in all the analyses allowed us to obtain estimates of demographic characteristics and outcomes in the population of interest. Along with weighting, all 95% confidence intervals accounted for clustering within a hospital; a generalized estimating equations approach with an exchangeable correlation matrix was used to calculate the marginal probability of an adverse event.[19,20] Confidence intervals were not adjusted for multiplicity; therefore, they should not be used in place of hypothesis testing. We also report the intraclass correlation coefficient that was estimated with generalized estimating equations (equivalent to the exchangeable correlation) as a measure of the variance among hospitals. Patient characteristics associated with admissions are reported as numbers and percentages for categorical variables and as means for continuous variables.

For the assessment of interrater reliability, we used Fleiss' kappa coefficient to determine the degree of agreement between the first and second adjudicators regarding their confidence that an adverse event was caused by health care management. We used percent agreement and Gwet's agreement coefficient with 95% confidence intervals to determine the degree of agreement regarding their confidence as to whether any adverse event had occurred and whether a preventable adverse event had occurred.[21] All the analyses were performed with the use of SAS/STAT software, version 9.4 (SAS Institute).

RESULTS

STUDY SAMPLE

We evaluated 11 hospitals; these included 3 large and 8 smaller hospitals. There were 193,549 admissions to these hospitals during the study period, and 2809 randomly selected admissions (the SafeCare random sample) were included in the analysis (Fig. 1). The estimated intraclass correlation coefficient among hospitals for all adverse events identified was 0.02 (95% confidence interval [CI], 0.00 to 0.04). The weighted random sample was reasonably representative of all inpatient admissions in Massachusetts during the study period (Table 1). The differences between the weighted random sample and statewide admissions with respect to age group, race, and Hispanic ethnic group were modest. However, the percentage of admissions that involved patients who were non-Hispanic was higher in the statewide group than in the weighted random sample (92.1% vs. 80.4%). In addition, the percentages of admissions that involved patients who had Medicare or Medicaid as their primary insurance were higher in the statewide group (50.2% vs. 42.2% for Medicare and 16.5% vs. 9.6% for Medicaid).

ADVERSE EVENTS IN THE WEIGHTED RANDOM SAMPLE

Within the weighted random sample of 2809 admissions, we identified at least one adverse event in 23.6% of the admissions (Table 2). An additional 314 adverse events were present at the time of admission; these events were determined to have occurred before the index admissions and were excluded from the analysis (Fig. 1). We identified 978 adverse events as having occurred during the index admissions, 222 (22.7%) of which were judged to be preventable (Table 3). Among the preventable adverse events, 19.7% were serious (i.e., caused harm that resulted in substantial intervention or prolonged recovery), 3.3% were life-threatening, and 0.5% were fatal. Examples of adverse events, including severity category and preventability assessment, are provided in Table S8.

Among all admissions, 523 (18.6%) involved at least one adverse event that was categorized as significant (i.e., caused unnecessary harm but resulted in rapid recovery), 211 (7.5%) involved a serious adverse event (as defined above), 34 (1.2%) included at least one adverse event that was life-threatening, and 7 (0.2%) involved an adverse event that was fatal. Overall, 191 admissions (6.8%) included at least one adverse event that was deemed to be preventable, and 29 admissions (1.0%) involved at least one adverse event that was assessed as preventable and was categorized as serious, life-threatening, or fatal. Table 2 shows the incidence of adverse events per admission according to demographic characteristics and insurance type. The percentage of admissions that included at least one adverse event was higher among older patients than among younger patients and among men than among women and was lower among Asian

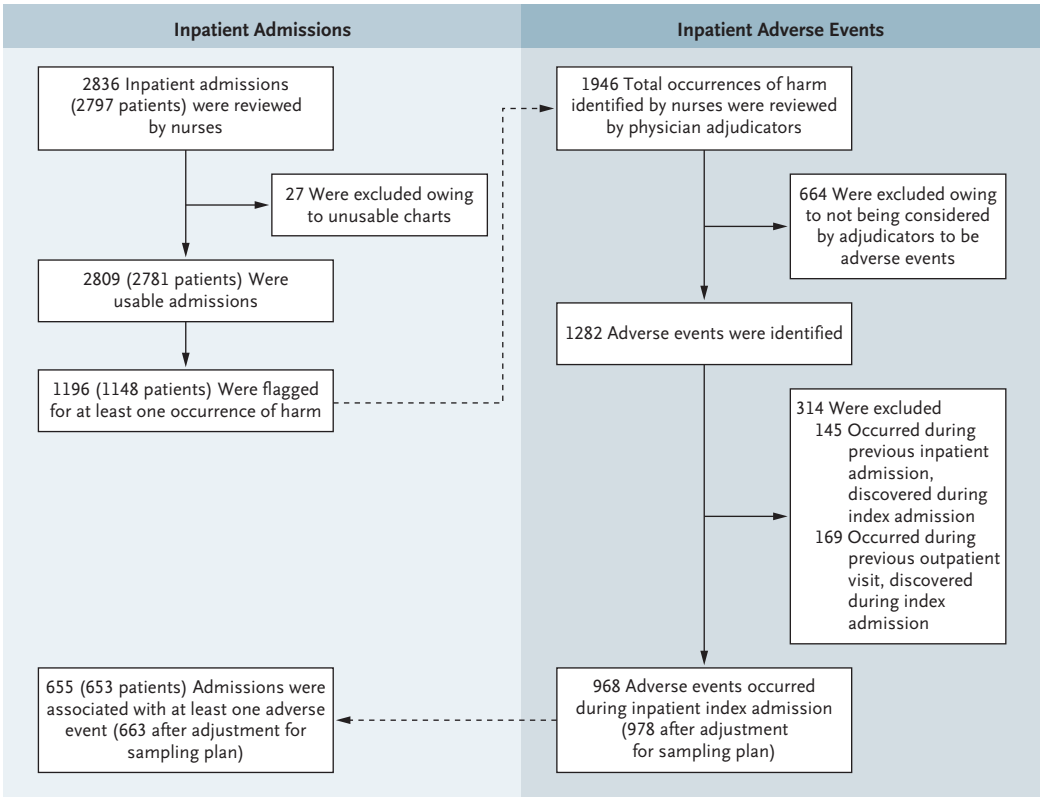

Figure 1. Review and Adjudication Process.

Unusable charts refers to selected charts that were not eligible for inclusion in the study (e.g., because the admission was canceled or the patient was admitted for research purposes).

patients than among Black or White patients and among Hispanic patients than among non-Hispanic patients; the percentage was also lower among patients who used Medicaid as their primary insurance than among those who used private insurance or Medicare. The percentage of admissions that involved preventable events was higher among older patients, among men, among Black or White patients than among Asian patients, among non-Hispanic patients, and among those who used Medicare. The mean length of stay for admissions with at least one adverse event was more than twice as long as that for admissions without adverse events (9.3 days [95% CI, 7.6 to 11.0] vs. 4.2 days [95% CI, 3.6 to 4.7]). The mean length of stay was 10.8 days (95% CI, 8.5 to 13.1) for admissions with at least one preventable adverse event.

Adverse drug events (Table 3) accounted for 381 (39.0%) of the overall adverse events and were the most common type, followed by events related to a surgical or other procedure (297 events [30.4%]), patient-care events including falls and pressure ulcers (147 events [15.0%]), and health care–associated infections (116 events [11.9%]). Events related to a surgical or other procedure were most likely to be rated as life-threatening, and health care–associated infections were most likely to be fatal. Patient-care events (57 of 147 events [38.8%]) and adverse drug events (102 of 381 events [26.8%]) were more likely to be preventable than other event types. Only 10 diagnostic errors (e.g., a delayed diagnosis of sepsis or renal failure or an incorrect diagnosis of seizure) that resulted in an adverse event were identified; this number was only a small fraction of all harms identified.

Across the 11 hospitals, adverse event rates ranged from 15.1 to 47.0 events per 100 admissions (Table S4). Larger hospitals had higher event rates

Table 1. Demographic Characteristics, Insurance Type, and Length of Stay in a Weighted Random Sample, in the Hospitals Included in the Study, and According to Admissions in Massachusetts.*

Variable	Study Weighted Random Sample of Admissions† (N=2809)	Admissions in Cohort of 11 Massachusetts Hospitals (N=193,549)	Massachusetts Inpatient Admissions‡ (N=581,234)
Age			
Mean (95% CI) — yr	59.9 (57.9–61.8)	59.9 (59.8–60.0)	59.9 (59.8–59.9)
Distribution — no. (% [95% CI])			
18 to 44 yr	680 (24.2 [19.9–28.4])	48,311 (25.0 [24.8–25.2])	146,824 (25.3 [25.1–25.4])
45 to 64 yr	831 (29.6 [25.6–33.6])	57,100 (29.5 [29.3–29.7])	167,246 (28.8 [28.7–28.9])
65 to 84 yr	1,056 (37.6 [34.6–40.6])	69,305 (35.8 [35.6–36.0])	200,292 (34.5 [34.3–34.6])
≥85 yr	243 (8.7 [5.7–11.6])	18,833 (9.7 [9.6–9.9])	66,872 (11.5 [11.4–11.6])
Sex — no. (% [95% CI])			
Female	1,561 (55.6 [52.3–58.9])	109,645 (56.6 [56.4–56.9])	334,751 (57.6 [57.5–57.7])
Male	1,225 (43.6 [40.3–47.0])	83,866 (43.3 [43.1–43.6])	246,474 (42.4 [42.3–42.5])
Unknown	22 (0.8 [0.0–1.8])	38 (0.02 [0.01–0.03])	9 (0.002 [0.001–0.003])
Race — no. (% [95% CI])§			
Asian	96 (3.4 [2.5–4.4])	6,970 (3.6 [3.5–3.7])	15,658 (2.7 [2.7–2.7])
Black	288 (10.3 [6.0–14.5])	17,416 (9.0 [8.9–9.1])	43,027 (7.4 [7.3–7.5])
White	2,117 (75.4 [68.2–82.5])	151,258 (78.1 [78.0–78.3])	477,241 (82.1 [82.0–82.2])
Other	174 (6.2 [3.8–8.6])	10,693 (5.5 [5.4–5.6])	29,055 (5.0 [4.9–5.1])
Unknown	134 (4.8 [2.1–7.4])	7,212 (3.7 [3.6–3.8])	16,253 (2.8 [2.8–2.8])
Ethnic group — no. (% [95% CI])§			
Hispanic	159 (5.7 [2.0–9.3])	14,854 (7.7 [7.6–7.8])	46,099 (7.9 [7.9–8.0])
Non-Hispanic	2,259 (80.4 [70.4–90.4])	156,534 (80.9 [80.7–81.1])	535,077 (92.1 [92.0–92.1])
Unknown	392 (14.0 [3.1–24.8])	22,161 (11.4 [11.3–11.6])	58 (0.01 [0.01–0.01])
Type of insurance — no. (% [95% CI])			
Private	1,305 (46.5 [38.5–54.4])	91,956 (47.5 [47.3–47.7])	170,530 (29.3 [29.2–29.5])
Medicare	1,185 (42.2 [36.9–47.5])	80,332 (41.5 [41.3–41.7])	291,884 (50.2 [50.1–50.3])
Medicaid	271 (9.6 [6.2–13.1])	18,122 (9.4 [9.2–9.5])	95,865 (16.5 [16.4–16.6])
Uninsured	28 (1.0 [0.5–1.5])	1,657 (0.9 [0.8–0.9])	4,067 (0.7 [0.7–0.7])
Unknown or other	19 (0.7 [0.0–1.4])	1,482 (0.8 [0.7–0.8])	18,888 (3.2 [3.2–3.3])
Mean length of hospital stay (95% CI) — days	5.4 (4.5–6.2)	5.4 (5.3–5.4)	4.4 (4.4–4.5)

* Data according to the patient's primary language are provided in the Supplementary Appendix. Confidence intervals were not adjusted for multiplicity and should not be used in place of hypothesis testing. Percentages may not total 100 because of rounding.
† The use of weighting of admission records allowed for adjustment for oversampling of the smaller hospitals. The numbers of admissions may not sum to 2809 because of weighting and rounding.
‡ Included are all Massachusetts admissions with discharges in 2016 from the State Inpatient Databases, excluding Major Diagnostic Categories 19 (mental diseases and disorders) and 20 (alcohol or drug use or induced mental disorders); patients younger than 18 years of age; admissions with a diagnosis of palliative care (Z51.5), as defined by the criteria of the International Classification of Diseases, version 10, at the time of admission; admissions for observation; and admissions at a rehabilitation or long-term acute care hospital.[22]
§ Race and ethnic group were reported by the patient.

than smaller hospitals. The rates of preventable adverse events ranged from 0.9 to 10.9 events per 100 admissions.

Of all possible adverse events adjudicated by physicians, 194 were included in the calculations of reliability. The evaluation of agreement between the adjudicators regarding their confidence that the harm was caused by health care manage-

Table 2. Weighted Incidence, Severity, and Preventability of Adverse Events per Admission, Overall and According to the Demographic Characteristics and Insurance Type Associated with the Admissions.*

Variable	Admissions with ≥1 Adverse Event	Admissions with ≥1 Preventable Adverse Event†	Admissions with ≥1 Significant Adverse Event‡	Admissions with ≥1 Serious Adverse Event‡	Admissions with ≥1 Life-Threatening Adverse Event	Admissions with ≥1 Fatal Adverse Event	Admissions with ≥1 Preventable Serious, Life-Threatening, or Fatal Adverse Event†
Overall — no. of admissions/total no. (% [95% CI])	663/2809 (23.6 [19.9–27.3])	191/2809 (6.8 [5.4–8.2])	523/2809 (18.6 [15.6–21.6])	211/2809 (7.5 [5.3–9.7])	34/2809 (1.2 [0.5–1.9])	7/2809 (0.2 [0.0–0.5])	29/2809 (1.0 [0.4–1.7])
Age group — no. of admissions/total no. (%)							
18 to 44 yr	106/680 (15.6)	24/680 (3.5)	81/680 (11.9)	36/680 (5.3)	4/680 (0.6)	0	4/680 (0.6)
45 to 64 yr	190/831 (22.9)	60/831 (7.2)	157/831 (18.9)	61/831 (7.3)	8/831 (1.0)	2/831 (0.2)	9/831 (1.1)
65 to 84 yr	296/1056 (28.0)	81/1056 (7.7)	231/1056 (21.9)	99/1056 (9.4)	19/1056 (1.8)	4/1056 (0.4)	11/1056 (1.0)
≥85 yr	71/243 (29.2)	25/243 (10.3)	55/243 (22.6)	15/243 (6.2)	3/243 (1.2)	1/243 (0.4)	5/243 (2.1)
Sex — no. of admissions/total no. (%)							
Female	326/1561 (20.9)	92/1561 (5.9)	256/1561 (16.4)	107/1561 (6.9)	14/1561 (0.9)	1/1561 (0.1)	13/1561 (0.8)
Male	335/1225 (27.3)	98/1225 (8.0)	265/1225 (21.6)	104/1225 (8.5)	20/1225 (1.6)	6/1225 (0.5)	16/1225 (1.3)
Unknown	2/22 (9.1)	1/22 (4.5)	2/22 (9.1)	0	0	0	0
Race — no. of admissions/total no. (%)							
Asian	14/96 (14.6)	3/96 (3.1)	11/96 (11.5)	6/96 (6.2)	1/96 (1.0)	0	1/96 (1.0)
Black	69/288 (24.0)	19/288 (6.6)	57/288 (19.8)	22/288 (7.6)	4/288 (1.4)	1/288 (0.3)	5/288 (1.7)
White	520/2117 (24.6)	150/2117 (7.1)	409/2117 (19.3)	162/2117 (7.7)	27/2117 (1.3)	5/2117 (0.2)	19/2117 (0.9)
Other	26/174 (14.9)	7/174 (4.0)	17/174 (9.8)	10/174 (5.7)	0	0	2/174 (1.1)
Unknown	34/134 (25.4)	12/134 (9.0)	28/134 (20.9)	11/134 (8.2)	2/134 (1.5)	1/134 (0.7)	2/134 (1.5)
Ethnic group — no. of admissions/total no. (%)							
Hispanic	28/159 (17.6)	9/159 (5.7)	19/159 (11.9)	10/159 (6.3)	2/159 (1.3)	1/159 (0.6)	3/159 (1.9)
Non-Hispanic	553/2259 (24.5)	159/2259 (7.0)	432/2259 (19.1)	177/2259 (7.8)	30/2259 (1.3)	5/2259 (0.2)	22/2259 (1.0)
Unknown	83/392 (21.2)	23/392 (5.9)	72/392 (18.4)	24/392 (6.1)	2/392 (0.5)	1/392 (0.3)	4/392 (1.0)

Type of insurance — no. of admissions/ total no. (%)							
Private	296/1305 (22.7)	71/1305 (5.4)	238/1305 (18.2)	97/1305 (7.4)	14/1305 (1.1)	3/1305 (0.2)	9/1305 (0.7)
Medicare	314/1185 (26.5)	102/1185 (8.6)	245/1185 (20.7)	95/1185 (8.0)	18/1185 (1.5)	4/1185 (0.3)	16/1185 (1.4)
Medicaid	47/271 (17.3)	15/271 (5.5)	34/271 (12.5)	16/271 (5.9)	2/271 (0.7)	0	4/271 (1.5)
Uninsured	4/28 (14.3)	2/28 (7.1)	4/28 (14.3)	2/28 (7.1)	0	0	0
Unknown or other	1/19 (5.3)	1/19 (5.3)	1/19 (5.3)	1/19 (5.3)	0	0	0

* Multiple adverse events with different severity levels could have occurred during a single admission. Results according to the patient's primary language are provided in the Supplementary Appendix. Percentages may not total 100 because of rounding, and the numbers of admissions may not sum to 2809 because of weighting and rounding.
† This category includes adverse events that were assessed as preventable or probably preventable.
‡ A significant adverse event was defined as an event that caused unnecessary harm but resulted in rapid recovery. A serious adverse event was defined as an event that caused harm that resulted in substantial intervention or prolonged recovery.

ment (as assessed with the use of the full, six-point ordinal confidence scale) resulted in a kappa coefficient of 0.70 (95% CI, 0.59 to 0.81). The assessment of agreement between the adjudicators regarding their confidence that an adverse event had occurred (as indicated by a confidence score of ≥4) resulted in a Gwet's agreement coefficient of 0.54 (95% CI, 0.41 to 0.66) and a percent agreement of 73.7%. The assessment of agreement between the adjudicators in their confidence that the adverse event was preventable resulted in a Gwet's agreement coefficient of 0.64 (95% CI, 0.54 to 0.75) and a percent agreement of 75.3%. Examples of harm that were not considered to be adverse events caused by health care management are provided in Table S9.

DISCUSSION

We evaluated the frequency and types of health care–associated adverse events approximately three decades after the original HMPS and found that adverse events remain common and are preventable nearly one fourth of the time. Preventable adverse events were identified in approximately 7% of all admissions, and preventable adverse events categorized as serious, life-threatening, or fatal were identified in approximately 1%. Adverse drug events were the most common type, followed by adverse events related to a surgical or other procedure, patient-care events such as falls and pressure ulcers, and health care–associated infections. Patient-care events and adverse drug events were the most likely events to be preventable.

Direct comparison of adverse event rates with those of other studies is challenging and warrants consideration of several caveats. In part because of their focus on malpractice, both the HMPS and the study of hospitals in Utah and Colorado included in their evaluations adverse events that were present on admission.[1,2,4] In the study by Landrigan and colleagues, adverse event data were evaluated at 10 hospitals in North Carolina over a 6-year period with the use of the Institute for Healthcare Improvement Global Trigger Tool for Measuring Adverse Events; their evaluation also included events present on admission.[5] Although we tracked such events, we did not count them in the calculations of our primary rates because we were assessing the incidence of events during hospitalization. Our approach may have resulted in conservative esti-

Table 3. Weighted Incidence, Severity, and Preventability of Adverse Events, Overall and According to Type of Event.*

Type of Adverse Event	No. of Events	Percentage of Overall Events	Events/100 Admissions (95% CI)	Significant Adverse Events	Serious Adverse Events	Life-Threatening Adverse Events	Fatal Adverse Events	Preventable Adverse Events†
Overall events identified — no. (% [95% CI])	978	100.0	34.8 (29.2–40.5)	661 (67.6 [64.3–70.9])	265 (27.1 [23.4–30.8])	44 (4.5 [2.1–7.0])	7 (0.7 [0.1–1.3])	222 (22.7 [19.4–26.0])
Adverse drug events — no. of events (%)‡								
All events	381	39.0	13.6 (3.8–23.4)	291 (76.4)	83 (21.8)	5 (1.3)	2 (0.5)	102 (26.8)
Hypotension	64	6.5	2.3 (0.6–4.0)	43 (67.2)	21 (32.8)	0	0	21 (32.8)
Mental status change§	43	4.4	1.5 (0.5–2.6)	38 (88.4)	3 (7.0)	1 (2.3)	0	14 (32.6)
Acute kidney injury	42	4.3	1.5 (0.4–2.6)	37 (88.1)	5 (11.9)	0	0	8 (19.0)
All other	232	23.7	8.3 (1.9–14.6)	172 (74.1)	53 (22.8)	4 (1.7)	2 (0.9)	59 (25.4)
Events related to a surgical or other procedure — no. of events (%)								
All events	297	30.4	10.6 (0.8–20.3)	155 (52.2)	113 (38.0)	27 (9.1)	2 (0.7)	40 (13.5)
Hypotension	39	4.0	1.4 (0.0–2.9)	18 (46.2)	17 (43.6)	2 (5.1)	1 (2.6)	5 (12.8)
Hemorrhage	33	3.4	1.2 (0.2–2.2)	3 (9.1)	25 (75.8)	5 (15.2)	0	5 (15.2)
Urinary retention	27	2.8	1.0 (0.3–1.7)	26 (96.3)	1 (3.7)	0	0	2 (7.4)
All other	197	20.1	7.0 (0.2–13.8)	107 (54.3)	69 (35.0)	20 (10.2)	1 (0.5)	28 (14.2)
Patient-care events — no. of events (%)¶								
All events	147	15.0	5.2 (1.9–8.5)	122 (83.0)	21 (14.3)	4 (2.7)	0	57 (38.8)
Intravenous infiltrate	30	3.1	1.1 (0.3–1.9)	30 (100.0)	0	0	0	10 (33.3)
Pressure ulcer	30	3.1	1.1 (0.1–2.0)	26 (86.7)	4 (13.3)	0	0	15 (50.0)
Anemia	7	0.7	0.2 (0.0–0.6)	2 (28.6)	5 (71.4)	0	0	0
All other	80	8.2	2.8 (0.8–4.9)	64 (80.0)	12 (15.0)	4 (5.0)	0	31 (38.8)

Health care–associated infections — no. of events (%)								
All events	116	11.9	4.1 (0.4–7.9)	70 (60.3)	36 (31.0)	7 (6.0)	3 (2.6)	20 (17.2)
Urinary tract infection	33	3.4	1.2 (0.1–2.2)	31 (93.9)	2 (6.1)	0	0	7 (21.2)
Pneumonia	21	2.1	0.7 (0.0–1.6)	5 (23.8)	11 (52.4)	5 (23.8)	0	2 (9.5)
Colitis‖	14	1.4	0.5 (0.0–1.0)	9 (64.3)	4 (28.6)	1 (7.1)	0	0
Surgical site infection	14	1.4	0.5 (0.1–0.9)	8 (57.1)	5 (35.7)	1 (7.1)	0	2 (14.3)
All other	34	3.5	1.2 (0.0–2.4)	17 (50.0)	14 (41.2)	0	3 (8.8)	9 (26.5)

* As a result of weighting, counts and percentages of all events may not sum to the total number of adverse events in each row or to the total number of events within an adverse event type. Data for events associated with pregnancy or the perinatal period and blood-transfusion reactions are provided in the Supplementary Appendix.
† This category includes adverse events that were assessed as preventable or probably preventable.
‡ Adverse drug events were defined as injuries resulting from drugs that were taken.
§ This adverse event term includes oversedation, delirium, and confusion.
¶ A patient-care event was defined as an event related to nursing care, including falls and pressure ulcers.
‖ This adverse event term includes *Clostridioides difficile* colitis.

mates of event rates. In addition, in the HMPS and the study in Utah and Colorado, an adverse event was defined as an event that resulted in prolonged hospitalization, disability at the time of discharge, or death; in contrast, similar to Landrigan and colleagues, we used a more inclusive definition that counted temporary patient harm as an adverse event.

Our ability to detect certain types of adverse events, such as health care–associated infections, has improved in the interval since those studies were performed. In addition, several new adverse events, such as failure to treat patients with decompensating events (which was not tracked carefully in the original HMPS), are now included. However, our ability to measure many important types of adverse events in an efficient, reliable, and continuous manner remains limited, and our results underscore the need to develop practical measurement tools. For example, in our study, we identified only 10 errors in diagnosis that led to adverse events; the trigger method is not well suited to finding these types of errors, and different approaches, including those involving machine learning, may be more effective.[23]

Internationally, in a systematic review published in 2008, de Vries and colleagues found that adverse events occurred in approximately 1 in 10 admissions across multiple countries, and almost half the events were considered to be preventable.[6] A more recent international meta-analysis published in 2019 by Panagioti and colleagues supported these findings.[7] In most studies, events related to a surgical or other procedure and adverse drug events were the most common. In the review by de Vries and colleagues, across six studies, the median percentage of events that were related to a surgical or other procedure was 39.6%, and the median percentage of adverse drug events was 15.1%.[6] In our study, however, adverse drug events were more common than events related to a surgical or other procedure (39.0% vs. 30.4%).

A recent study in which temporal trends in adverse event rates were evaluated showed that rates have declined substantially over the course of the past decade overall and specifically for health care–associated infections, adverse drug events, and patient-care events including falls and pressure ulcers, although this study tracked only a fraction of adverse drug events.[24] Moreover, the challenges and strains on the health care system created by the coronavirus disease 2019

pandemic appear to have reversed these trends, with substantial increases in health care–associated infections and patient-care events documented in 2020.[25]

Even today, many U.S. hospitals rely solely on voluntary reporting of adverse events, which results in substantial undercounting and, in some cases, misleading reports of zero harm. Identification of adverse events in EHRs in the future will probably be performed by means of computerization of triggers and also through leveraging of artificial intelligence.[26,27] Commercial tools that can identify some types of harm in hospitalized patients, including adverse drug events and health care–associated infections, are already available and widely used, although a broadening of the harms assessed by these tools is warranted. In addition, the Centers for Medicare and Medicaid Services is currently mandating that certain metrics are collected for hospitalized patients.[28,29]

This study has several limitations. First, the hospitals that were selected may not be representative of hospitals at large, although they were selected to include hospitals of varying sizes. Second, our study population included more patients who had private insurance, and fewer patients who had Medicare or Medicaid as their primary insurance, than the overall inpatient population in Massachusetts. Third, our approach almost certainly missed some adverse events, and fourth, agreement between the pairs of adjudicators was only fair.

Three decades after the HMPS drew attention to the issue of health care–associated patient harm, in-hospital adverse events continue to be common, and although only approximately one fourth of the adverse events identified in this study were deemed to be preventable, all adverse events negatively affect medical care and outcomes. Over the course of this 30-year interval, care has become more complex, and diagnostic and therapeutic options to treat disease and alleviate human suffering have advanced. The health care delivery system itself has changed dramatically with the advent of EHRs and the movement of complex care to ambulatory sites, which has resulted in the most severely ill patients being treated in acute care hospitals. Despite stunning advances in medical science, we still have important gaps in patient safety.

Measuring adverse events in a reliable and efficient way and developing standard approaches to the identification of and focus on preventable adverse events are critical to supporting persons charged with improving safety. Some types of adverse events, such as health care–associated infections, can be identified much more effectively than others, which suggests a need to improve routine tracking, especially for events such as adverse drug events. There is considerable variability among hospitals in adverse event rates, with larger sites having rates of approximately 40% or higher; this finding suggests that if hospitals had data that were more reliable and more routinely collected, it is possible that monitoring could be improved, adverse event rates could be reduced, and improvement strategies could be shared through careful study of interventions. Other key organizational elements such as safety culture and strong leadership with respect to safety and quality are also needed to advance performance. Our findings are an urgent reminder to all health care professionals of the need for continuing improvement in the safety of the care we deliver.

Supported by a grant from the Controlled Risk Insurance Company and the Risk Management Foundation of the Harvard Medical Institutions.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank the leadership and staff from the participating study sites and the many chart reviewers and adjudicators, without whom this work could not have been performed; and Ruth Akindele, Clark Davis, Juliette Randazza, and Sonam Shah for their early contributions to the project.

REFERENCES

1. Brennan TA, Leape LL, Laird NM, 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.
2. Leape LL, Brennan TA, Laird N, 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.
3. Kohn LT, Corrigan JM, Donaldson MS, eds. To err is human: building a safer health system. Washington, DC: National Academies Press, 2000.
4. Thomas EJ, Studdert DM, Burstin HR, et al. Incidence and types of adverse events and negligent care in Utah and Colorado. Med Care 2000;38:261-71.
5. Landrigan CP, Parry GJ, Bones CB, Hackbarth AD, Goldmann DA, Sharek PJ. Temporal trends in rates of patient harm resulting from medical care. N Engl J Med 2010;363:2124-34.
6. de Vries EN, Ramrattan MA, Smorenburg SM, Gouma DJ, Boermeester MA. The incidence and nature of in-hospital adverse events: a systematic review. Qual Saf Health Care 2008;17:216-23.
7. Panagioti M, Khan K, Keers RN, et al. Prevalence, severity, and nature of pre-

ventable patient harm across medical care settings: systematic review and meta-analysis. BMJ 2019;366:l4185.
8. Bates DW, Singh H. Two decades since to err is human: an assessment of progress and emerging priorities in patient safety. Health Aff (Millwood) 2018;37:1736-43.
9. Pronovost P, Needham D, Berenholtz S, et al. An intervention to decrease catheter-related bloodstream infections in the ICU. N Engl J Med 2006;355:2725-32.
10. Haynes AB, Weiser TG, Berry WR, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. N Engl J Med 2009;360:491-9.
11. Classen DC, Resar R, Griffin F, et al. 'Global trigger tool' shows that adverse events in hospitals may be ten times greater than previously measured. Health Aff (Millwood) 2011;30:581-9.
12. Levinson DR. Adverse events in hospitals: national incidence among Medicare beneficiaries. Rockville, MD: Department of Health and Human Services, 2010 (https://oig.hhs.gov/oei/reports/oei-06-09-00090.pdf).
13. Grimm CA. Adverse events in hospitals: a quarter of Medicare patients experienced harm in October 2018. Rockville, MD: Department of Health and Human Services, 2022 (https://oig.hhs.gov/oei/reports/OEI-06-18-00400.pdf).
14. Cullen DJ, 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;21:541-8.
15. Patient Safety Network. Adverse events, near misses, and errors. Rockville, MD: Agency for Healthcare Research and Quality, 2019 (https://psnet.ahrq.gov/primer/adverse-events-near-misses-and-errors).
16. Folli HL, Poole RL, Benitz WE, Russo JC. Medication error prevention by clinical pharmacists in two children's hospitals. Pediatrics 1987;79:718-22.
17. Dubois RW, Brook RH. Preventable deaths: who, how often, and why? Ann Intern Med 1988;109:582-9.
18. Hiatt HH, Barnes BA, Brennan TA, et al. A study of medical injury and medical malpractice. N Engl J Med 1989;321:480-4.
19. Liang K-Y, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika 1986;73:13-22.
20. Lipsitz SR, Fitzmaurice GM, Orav EJ, Laird NM. Performance of generalized estimating equations in practical situations. Biometrics 1994;50:270-8.
21. Wongpakaran N, Wongpakaran T, Wedding D, Gwet KL. A comparison of Cohen's kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. BMC Med Res Methodol 2013;13:61.
22. Healthcare Cost and Utilization Project: State Inpatient Databases (SID). Rockville, MD: Agency for Healthcare Research and Quality, 2016 (https://www.hcup-us.ahrq.gov/sidoverview.jsp).
23. Griffin JA, Carr K, Bersani K, et al. Analyzing diagnostic errors in the acute setting: a process-driven approach. Diagnosis (Berl) 2021;9:77-88.
24. Eldridge N, Wang Y, Metersky M, et al. Trends in adverse event rates in hospitalized patients, 2010–2019. JAMA 2022;328:173-83.
25. Fleisher LA, Schreiber M, Cardo D, Srinivasan A. Health care safety during the pandemic and beyond — building a system that ensures resilience. N Engl J Med 2022;386:609-11.
26. Sammer C, Miller S, Jones C, et al. Developing and evaluating an automated all-cause harm trigger system. Jt Comm J Qual Patient Saf 2017;43:155-65.
27. Bates DW, Levine D, Syrowatka A, et al. The potential of artificial intelligence to improve patient safety: a scoping review. NPJ Digit Med 2021;4:54.
28. Electronic clinical quality measures basics. Baltimore, MD: Centers for Medicare and Medicaid Services, 2022 (https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/ClinicalQualityMeasures).
29. Classen DC, Munier W, Verzier N, et al. Measuring patient safety: the Medicare patient safety monitoring system (past, present, and future). J Patient Saf 2021;17(3):e234-e240.