# CLASSIC PAPER

# Incidence of adverse events and negligence in hospitalized patients: results of the Harvard Medical Practice Study I*

T A Brennan, L L Leape, N M Laird, L Hebert, A R Localio, A G Lawthers, J P Newhouse, P C Weiler, H H Hiatt

.................................................................................................................

**Background:** As part of an interdisciplinary study of medical injury and malpractice litigation, we estimated the incidence of adverse events, defined as injuries caused by medical management, and of the subgroup of such injuries that resulted from negligent or substandard care.
**Methods:** We reviewed 30 121 randomly selected records from 51 randomly selected acute care, non-psychiatric hospitals in New York State in 1984. We then developed population estimates of injuries and computed rates according to the age and sex of the patients as well as the specialties of the physicians.
**Results:** Adverse events occurred in 3.7% of the hospitalizations (95% confidence interval 3.2 to 4.2), and 27.6% of the adverse events were due to negligence (95% confidence interval 22.5 to 32.6). Although 70.5% of the adverse events gave rise to disability lasting less than 6 months, 2.6% caused permanently disabling injuries and 13.6% led to death. The percentage of adverse events attributable to negligence increased in the categories of more severe injuries (Wald test $\chi^2 = 21.04$, $p<0.0001$). Using weighted totals we estimated that among the 2 671 863 patients discharged from New York hospitals in 1984 there were 98 609 adverse events and 27 179 adverse events involving negligence. Rates of adverse events rose with age ($p<0.0001$). The percentage of adverse events due to negligence was markedly higher among the elderly ($p<0.01$). There were significant differences in rates of adverse events among categories of clinical specialties ($p<0.0001$), but no differences in the percentage due to negligence.
**Conclusions:** There is a substantial amount of injury to patients from medical management, and many injuries are the result of substandard care.

See end of article for
authors' affiliations
.......................

Over the past decade there has been a steady increase in the number of malpractice claims brought against healthcare providers[1 2] and in the monetary damages awarded to plaintiffs.[3–5] This increase has precipitated numerous state programs designed to moderate the number of claims and encourage providers to develop quality of care initiatives.[6 7] Advocates of tort reform argue that the existing system of malpractice litigation is inefficient in compensating patients injured by medical practice and in deterring the performance of poor quality care that is sometimes responsible for the injuries.[8] Others defend the role of tort litigation.[9] These debates will probably continue even as claims rates begin to decrease.[10]

Controversy over the virtues of common law malpractice litigation occurs without much empirical information regarding the epidemiology of poor quality care and iatrogenic injury. The most widely quoted estimates of the incidence of iatrogenic injury and substandard care were developed over 10 years ago.[11] Other reviews by physicians to identify poor quality care or adverse events have been restricted to non-random samples of much smaller numbers of records.[12 13]

To address the need for empirical information we undertook the Harvard Medical Practice Study. A primary goal was to develop more current and more reliable estimates of the incidence of adverse events and negligence in hospitalized patients. We defined an adverse event as an injury that was caused by medical management (rather than the underlying disease) and that prolonged the hospitalization, produced a disability at the time of discharge, or both. We defined negligence as care that fell below the standard expected of physicians in their community. To estimate the incidence of these critical events, we reviewed a random sample of more

than 31 000 hospital records using techniques we have previously described.[14–16]

## METHODS
### Sample selection and record review
We have presented our methods of record review and our sampling strategy in detail elsewhere.[16] We used a two-stage sampling process to create a weighted sample of 31 429 records of hospitalized patients from a population of 2 671 863 non-psychiatric patients discharged from non-federal acute care hospitals in New York in 1984. Initially, the records were screened by trained nurses and medical records analysts; if a record was screened as positive, two physicians independently reviewed it. The physicians, almost all of whom were board certified internists or surgeons, were trained by us to assess the medical records for evidence of adverse events and negligence (Appendix I) and to grade their confidence that an adverse event had occurred on a scale of 0 to 6 (the causation score).

Because we were interested in estimating the statewide incidence of adverse events, the physician reviewers recorded not only adverse events that occurred and were discovered during the index hospitalization, but also those caused by medical management before the index hospitalization and first discovered during it. In calculating incidence rates we counted only events discovered during the sampled 1984 hospitalizations. By including adverse events that occurred earlier but were first discovered during the index hospitalization, we compensated for adverse events caused during the

.........................................................

index hospitalization but discovered only after discharge. In order to avoid overstating incidence we excluded events that were caused during the 1984 index hospitalization but were discovered during a subsequent hospitalization in 1984.

If the reviewers' confidence in the occurrence of an adverse event was greater than 1 on a six point scale, they assessed the disability it caused. Next, they judged whether there was evidence of negligence and indicated their level of confidence in that judgment. Throughout the process they could consult New York specialists recruited for the purpose. Discrepancies between the two physician reviewers in the identification of adverse events were noted by a medical records analysis supervisor overseeing the screening process and were resolved in an independent review by a supervising physician (one of six physicians from Boston who directed the record review in one region in New York).

### Testing reliability and validity

To test the validity of the process of screening by medical records analysts, 1% of all records were reviewed again by a medical records analysis supervisor using a blank screening form. The validity of the initial review was tested by considering the supervisor's review a gold standard.

The reliability of judgments of adverse events (causation) and substandard care (negligence) was tested by a team consisting of a medical records analysis supervisor, several physician reviewers, and a physician supervisor, which completed a second review of all records initially screened as positive at two hospitals. The results of this review were compared with those of the original review with use of the kappa statistic.

### Follow up of missing records and adjustments

Several months after the initial review of records, we asked all the hospitals to attempt to identify the current status of any records that they had not located earlier. We reviewed all the records found in this follow up search using our regular review process. This enabled us to estimate the rates of adverse events and negligence in missing records. We also adjusted for possible differential selection of missing records according to hospital and case type, and we used imputation to fill in the missing items of data, conditional on a reviewer's response, to other items.[17]

### Definition of variables

To establish that an adverse event or negligence had occurred, we used as a criterion an average confidence score of 4 or higher (on a six point scale). For patient disability scores we used the ratings given by both reviewers and assigned half the weight for each case to each of the two reviewers. Data concerning age, sex, and primary discharge diagnosis were obtained from the data base of the New York Statewide Planning and Research Cooperative System (SPARCS).[18] Specialties were determined on the basis of diagnosis related groups (DRGs) (Appendix II).

### Statistical analysis

We report our results as the percentage of discharges with adverse events, the percentage of adverse events due to negligence, and population estimates of the numbers of adverse events and adverse events due to negligence according to disability category. We calculated all percentages and population projections using the selection weights, adjusted as described above. We used the SESUDAAN software package to calculate standard errors.[19] The significance of differences in rates was tested with the Wald statistic.

For five age groups we computed the crude rate of adverse events and a rate directly standardized to control for the inherent risk that a particular diagnosis would give rise to an adverse event. We standardized the rate using four risk categories obtained as follows. Three physician supervisors individually rated all 470 DRGs on a scale of l to 6, reflecting their belief that the DRG was most (6) or least (1) likely, on clinical grounds, to be associated with an adverse event. We averaged the three ratings to define four risk categories of DRG (Appendix II). We did not standardize the percentage of negligence according to DRG risk. Since the denominator of the percentage of negligence was the number of adverse events, this acted as an implicit control for the complexity of care.

To compare rates of adverse events and negligence according to sex, we used directly standardized rates controlling for five categories of patient age and four categories of risk that a particular diagnosis would give rise to an adverse event. Only two age categories (<65 and ≥65 years) were used to standardize the percentage of negligence.

## RESULTS

We completed the initial review of 30 195 of the 31 429 records (96.1%) in the original random sample. Among these, the medical records analysts found 7817 positive according to the screening criteria. Physicians reviewed 7743 of them at the second level review. The results reported here are thus based on 30 121 records, including 22 378 with negative screens and 7743 reviewed by physicians. Using the incidence categories described above, the physicians identified 1278 adverse events and 306 adverse events due to negligence (fig 1). The incidence rates presented here are based on the 1133 adverse events and 280 negligent ones discovered during 1984 admissions (categories l, 4, and 5; table 1).

We estimated the statewide incidence rate of adverse events to have been 3.7% (95% confidence interval 3.2 to 4.2) and the rate of adverse events due to negligence to have been 1.0% (95% confidence interval 0.8 to 1.2). The percentage of adverse events due to negligence was 27.6% (95% confidence interval 22.5 to 32.6). Using the weighting procedure we calculated that, of the 2 671 863 patients discharged from acute care hospitals in New York State in 1984, there were 98 609 adverse events and 27 179 adverse events due to negligence.

Most adverse events (mean (SE) 56.8 (1.6)%) resulted in minor impairment with complete recovery in one month. Another 13.7 (1.1)% led to disabilities that lasted more than one but less than six months. However, 2.6 (0.4)% of the adverse events gave rise to permanent total disability and 13.6 (1.7)% caused death. Extrapolating to the state of New York in 1984, we estimated that 2550 patients suffered permanent total disability and that 13 451 died at least in part as a result of adverse events (table 2).

Negligence was more frequent in patients who had more severe adverse events. Of the adverse events that led to temporary disability lasting less than 1 month, 22.2 (2.8)% were caused by negligence. On the other hand, of those that caused permanent total disability, 34.4 (8.1)% were caused by negligence. In addition, 51.3 (6.9)% of the deaths from adverse events were caused by negligence. These differences in the percentage of negligence according to category were significant (Wald test $\chi^2 = 21.04$, p<0.0001).

We also analyzed the distribution of adverse events among different patient populations. Rates of adverse events increased strongly with increasing age (p<0.0001). Persons 65 or older had more than double the risk of persons 16–44 years of age (table 3). Unlike the rates of adverse events, the percentage of adverse events due to negligence did not increase monotonically with age, but the rate of negligence among those older than 64 was higher than that of any other
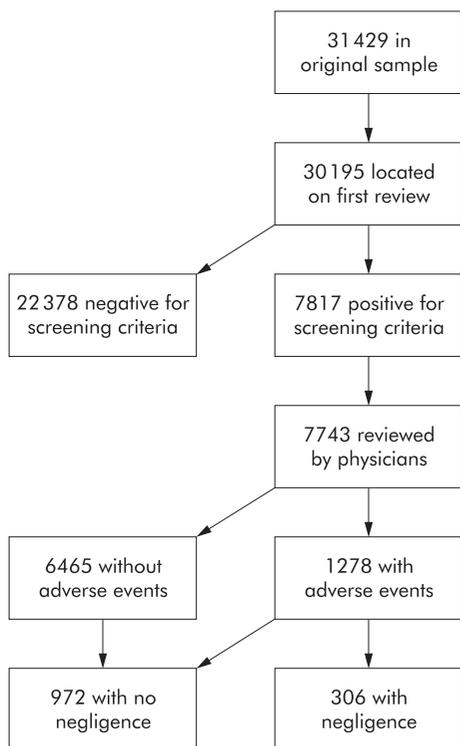
**Figure 1** The record review process. Numbers of medical records are shown.

age group, a difference that remained after standardizing for DRG risk category.

After standardizing for age and DRG risk category, we found no significant differences between sexes in rates of adverse events (male, 3.8 (0.4)%; female, 3.7 (0.4)%) or in the percentage of adverse events due to negligence (male, 27.4 (2.8)%; female, 25.0 (2.8)%).

Table 4 shows the rates of adverse events and negligence for groups of clinical specialties based on DRG groupings, as well as population estimates for each specialty. Rates of adverse events varied significantly, ranging from a low of 0.6 (0.1)% for neonatal DRGs to a high of 16.1 (3.0)% for vascular surgery DRGs, a more than 25-fold difference. Rates of negligence did not vary significantly.

We checked the accuracy of our results in several ways. First, we found 154 of the 326 missing records (47.2%) in

follow up visits to the six hospitals. The rates of adverse events (2.5%) and negligence (0.7%) among the missing records were lower than among the records originally reviewed. Second, a test by the medical records analysis supervisors of the validity of the screening criteria revealed a sensitivity of 89%. Third, the reliability of the judgments by the physicians was comparable to that in our pilot studies.[14] The agreement on the presence of an adverse event was 89% (kappa = 0.61). With regard to negligence, the agreement was 93%, but the kappa statistic was much lower (0.24, table 5).

## DISCUSSION
As part of a comprehensive empirical assessment of medical injury and medical malpractice,[16] we estimated the rates of adverse events and the subgroup of those adverse events caused by negligent care in hospitalized patients in New York State in 1984. Our results should be understood in the context of both medical malpractice litigation and quality assessment. The concepts of adverse event and negligence are derived explicitly from the theory of tort law, of which medical malpractice is a part. Malpractice litigation is intended in part to promote better quality care by fixing economic sanctions on those who provide substandard care that leads to injuries. Thus, malpractice litigation should in theory be linked to quality assurance. We left aside the aspects of compensation and corrective justice in tort litigation in this analysis.[20]

Adverse events do not, of course, necessarily signal poor quality care; nor does their absence necessarily indicate good quality care. For example, a drug reaction that occurs in a patient who has been appropriately prescribed the drug for the first time is an adverse event, but one that is unavoidable given today's technology. If, on the other hand, the drug reaction occurs in a patient who is given the drug despite a known sensitivity to it, the adverse event is properly judged to be due to negligence. Such care, which may reasonably lead to successful tort litigation, should be a target of quality assurance programs.

Using our methods, we estimated that 3.7% of the patients hospitalized in 1984 suffered adverse events, whereas the rate of adverse events due to negligence was 1.0%. These results may be compared with those of the only other large scale effort to estimate the incidence of iatrogenic injury and substandard care, the California Medical Association's Medical Insurance Feasibility Study.[11] Investigators there found 870 potentially compensable events (a category comparable to our adverse events) in a convenience sample

**Table 1** Categories of incidence of adverse events and negligence

| Category | Timing of adverse event | Adverse events* | Adverse events due to negligence* |
|---|---|---|---|
| 1 | Occurred and discovered during index hospitalization | 647 (50.6) 55046 (49.4) | 156 (51.0) 15257 (51.2) |
| 2 | Occurred during index hospitalization, discovered during subsequent outpatient care | 78 (6.1) 6327 (5.7) | 7 (2.3) 776 (2.6) |
| 3 | Occurred during index hospitalization, discovered during subsequent hospitalization | 67 (5.2) 6526 (5.9) | 19 (6.2) 1857 (6.2) |
| 4 | Occurred during outpatient care before index hospitalization but discovered during index hospitalization | 167 (13.1) 16142 (14.5) | 59 (19.3) 6019 (20.2) |
| 5 | Occurred during earlier hospitalization but discovered during index hospitalization | 319 (25.0) 27420 (24.6) | 65 (21.2) 5903 (19.8) |

*For each category the first row of values indicates the sample count, and the second row the weighted population total. Data are presented as number (%).

**Table 2** Population distribution of adverse events according to category of disability

| Category of disability | Adverse events* | Adverse events due to negligence* | Percentage due to negligence |
|---|---|---|---|
| Minimal impairment, recovery 1 month | 56042 (56.8 (1.6)) | 12428 (45.7 (3.7)) | 22.2 (2.8) |
| Moderate impairment, recovery >1–6 months | 13521 (13.7 (1.1)) | 3302 (12.1 (2.2)) | 24.4 (4.8) |
| Moderate impairment, recovery >6 months | 2762 (2.8 (0.5)) | 817 (3.0 (1.0)) | 29.6 (8.6) |
| Permanent impairment, ⩽50% disability | 3807 (3.9 (0.6)) | 869 (3.2 (1.1)) | 22.8 (6.8) |
| Permanent impairment, >50% disability | 2550 (2.6 (0.4)) | 877 (3.2 (0.8)) | 34.4 (8.1) |
| Death | 13451 (13.6 (1.7)) | 6895 (25.4 (4.2)) | 51.3 (6.9) |
| Could not reasonably judge disability | 6477 (6.6 (0.7)) | 1989 (7.3 (1.3)) | 30.7 (5.9) |
| Total† | 98610 | 27177 | 27.6 (2.4) |

*Values are numbers with mean (SE) %.
†Values differ from the sums of those reported above because of rounding.

**Table 3** Rates of adverse events and negligence according to age

| Age of patient | No of cases reviewed | Mean (SE) rate of adverse events (%) | | Rate of negligence |
|---|---|---|---|---|
| | | Crude | Standardized* | |
| Newborn | 3595 | 0.6 (0.1) | 1.4 (0.3) | 20.8 (7.1) |
| ⩽15 years | 3066 | 2.1 (0.4) | 2.7 (0.6) | 21.9 (5.0) |
| 16–44 years | 11101 | 2.6 (0.2) | 2.6 (0.2) | 26.7 (2.8) |
| 45–64 years | 7379 | 4.7 (0.4) | 4.4 (0.4) | 20.6 (2.4) |
| ⩾65 years | 4980 | 5.9 (0.5) | 5.7 (0.6) | 33.1 (4.2) |
| p value† | | | <0.0001 | <0.01 |

*According to DRG class.
†For the distribution of rates of events.

of 20 864 records, for an overall rate of 4.6%. This rate was 26% higher than our estimate of 3.7%. The California study revealed a negligence rate of 0.8%, 20% lower than the result of our review.

Because our sample of hospital records was random, we could provide for the first time population estimates of adverse events and adverse events due to negligence. Among the 2 671 863 discharges from New York hospitals in 1984, we estimate that there were 98 609 adverse events. Although 56 042 of them (56.8%) led to minimal disability with complete recovery in one month and another 13 521 (13.7%) to moderate disability with complete recovery in 6 months, 2550 (2.6%) produced permanent total disability and 13 451 (13.6%) led to death. The burden of iatrogenic injury was thus large.

Even more disturbing was the number of adverse events caused by negligence. We estimated that 27 179 injuries, including 6895 deaths and 877 cases of permanent and total disability, resulted from negligent care in New York in 1984. Under the tort system, all of these could have led to successful litigation. We could not measure all negligent

acts, and made no attempt to, but measured only those that led to injury. Medical records are probably a poor source of information on negligence that does not cause injury. Thus, our figures reflect not the amount of negligence, but only its consequences.

The analyses of rates of adverse events and the percentage of adverse events due to negligence according to characteristics of the patient are of special interest. Identifying risk factors for adverse events, whether negligent or not, constitutes a crucial first step toward their prevention, an important goal of quality assurance. In this study we focused on patient age and sex and on clinical specialty groups.

To increase the precision of our analyses of risk factors, we standardized the data according to our estimates of the risk of a particular DRG's giving rise to an adverse event. This risk categorization was found to correlate well with the observed rates of adverse events, but not with rates of negligence (Appendix II). The absence of an effect of DRG risk category on negligence was expected, for our physicians' judgments regarding the standard of care reflected the inherent riskiness of a procedure or disease state.

**Table 4** Rates of adverse events and negligence among clinical specialty groups*

| Specialty | Rate of adverse events | | Rate of negligence | |
|---|---|---|---|---|
| | Mean (SE) % | Population estimate | Mean (SE) % | Population estimate |
| Orthopedics | 4.1 (0.6) | 6746 | 22.4 (4.7) | 1514 |
| Urology | 4.9 (0.8) | 4819 | 19.4 (6.5) | 933 |
| Neurusurgery | 9.9 (2.1) | 2987 | 35.6 (8.6) | 1063 |
| Thoracic and cardiac surgery | 10.8 (2.4) | 3588 | 23.0 (9.3) | 826 |
| Vascular surgery | 16.1 (3.0) | 3187 | 18.0 (8.1) | 575 |
| Obstetrics | 1.5 (0.2) | 5013 | 38.3 (7.0) | 1920 |
| Neonatology | 0.6 (0.1) | 1713 | 25.8 (6.9) | 442 |
| General surgery | 7.0 (0.5) | 22324 | 28.0 (3.4) | 6247 |
| General medicine | 3.6 (0.3) | 37135 | 30.9 (4.4) | 11475 |
| Other | 3.0 (0.4) | 11097 | 19.7 (4.9) | 2183 |
| p value† | <0.0001 | | 0.64 | |

*Values differ from the sums of those reported above because of rounding.
†For the distribution of rates of events.

**Table 5** Results of duplicate review process

| Review process A | Review process B | | | | | |
|---|---|---|---|---|---|---|
| | Adverse events | | | Negligence | | |
| | Absent | Present | Total | Absent | Present | Total |
| Absent | 249 | 21 | 270 | 293 | 12 | 305 |
| Present | 13 | 35 | 48 | 9 | 4 | 13 |
| Total | 262 | 56 | 318 | 302 | 16 | 318 |
| Kappa statistic | | 0.61 | | | 0.24 | |

We found that both crude and standardized rates of adverse events increased with age. This suggests that elderly people are at higher risk of an adverse event, and it may reflect in part the fact that older people are likely to have more complicated illnesses and often require more complicated intervention. It may also be ascribable in part to their greater fragility. Such differences highlight the importance of controlling for age when comparing population groups. People over the age of 64 were at higher risk of an adverse event associated with negligence, a finding not readily explained by differences in the severity of illness. Presumably, this means that care for the elderly less frequently meets the standard expected of reasonable medical practitioners. Sex did not appear to represent a risk factor for adverse events or negligence.

There is great variation among specialties with regard to the riskiness of the procedures employed and the severity of illness in the patients for whom care is provided. The finding that patients in certain specialty groups, as defined by DRGs, were at higher risk of adverse events was therefore not surprising. The percentage of adverse events due to negligence did not, however, vary according to specialty. The momentary lapse on the part of an internist who forgets to ask about sensitivity to an antibiotic until the end of an interview (but before writing a prescription) may have far different consequences than the neurosurgeon's momentary lapse during an operation on the brain or spinal cord. One goal of our study was to examine such issues, for the nature of medical injury and of medical injury due to negligence will help guide investigators who seek to reduce the occurrence of such injuries.

The observations concerning rates of adverse events and negligence among specialties have implications relevant to today's system of malpractice insurance. Practitioners of certain specialties are sued more frequently and thus pay much higher premiums than others.[3] We found that these specialties (neurosurgery, cardiac and thoracic surgery, and vascular surgery) had higher rates of adverse events, but not higher rates of negligence. Our data suggest that variations among specialties in rates of litigation do not reflect differing levels of competence, but rather differences in the kinds of patients and diseases for which the specialist cares.

There were a number of potential sources of error in our estimates. One was missing records, but we were reassured by the fact that the rates of adverse events and negligence in the follow up study were lower overall than in the initial survey. Another possible source of error was our use of hospital records for information on adverse events and negligence. We had, however, previously demonstrated the integrity of hospital records in this capacity.[15] Of course, our findings cast little light on practice in physicians' offices.

Error may also have been introduced by our review methods. We realize that judgments regarding the causes of adverse events and negligent care are difficult and sometimes inaccurate. In previous studies we addressed the reliability and validity of our process.[14 15] We repeated some of these tests in our record review in New York. We found that the screening process had a higher level of validity than our previous estimates had suggested, with a sensitivity of 89% compared with 85% in our pilot study. The reliability of physicians' judgments about the presence of adverse events was good (kappa = 0.61)

However, the more difficult judgments regarding negligence had a lower degree of reliability (kappa = 0.24), although the overall agreement on judgments of negligence was excellent (93%). The low kappa statistic indicates that, in the records with evidence of negligence, physicians disagreed frequently about the extent of substandard care. If we use the presence of any evidence of negligence (rather than a combined confidence of more than 50–50) as a threshold to test reliability, the statistic increases considerably (kappa = 0.49). Moreover, using the confidence-in-negligence score as an ordinal measure produces an intraclass correlation coefficient of 0.41 for negligence. It is also important to note that, because of budgetary and time constraints, this test of reliability involved only two teams of physicians. Our pilot test, which showed a higher degree of reliability on judgments of negligence, involved numerous sets of physicians and perhaps better reflected the variation from physician to physician.[14]

Nonetheless, all of this underlines the fact that physicians find it difficult to judge whether a standard of care has been met—hardly a surprising fact in view of the complexity of clinical decision making. The relatively low level of reliability tends to bias estimates toward the null. The differences that emerged in the group comparisons are therefore that much more likely to be true. In addition, as table 5 demonstrates, the rates from both review processes were quite similar, suggesting that our overall estimates are accurate, even given some unreliability of judgments.

Physicians' estimates of disability were another potential source of error. The physicians based their decisions on evidence in the medical records, which sometimes described hospitalizations subsequent to the index admission. Without complete follow up information on the patient, however, absolutely accurate estimates of disability were not, of course, possible.

The judgments of physicians that an adverse event led to death also require a note of caution. Many patients who died after an adverse event had very serious underlying disease, and several surely had shortened life expectancies independent of their iatrogenic injury. Physicians could not, and were not asked to, estimate the number of days of life lost as a result of the adverse event. This is a critical issue, particularly in the case of a terminally ill person. For instance, a pneumothorax injury sustained during the insertion of a central venous catheter may have been the immediate cause of death in a comatose patient with metastatic lung cancer who was undergoing mechanical ventilation because of respiratory failure. Although this patient might have lived only a few more hours or days had the adverse event not occurred, the death was judged to have resulted from the

**Table 6** Population estimates of rates of adverse events and negligence according to DRG category

| DRG category | Mean (SE) % adverse events | Mean (SE) % adverse events due to negligence |
|---|---|---|
| 1 | 1.82 (0.32) | 24.4 (6.6) |
| 2 | 3.34 (0.21) | 30.0 (3.4) |
| 3 | 4.26 (0.36) | 27.0 (2.9) |
| 4 | 7.13 (0.91) | 23.0 (4.6) |
| p value* | <0.0001 | 0.47 |

*By the Wald statistic.

medical injury. In addition, some patients may have requested and received limited care, even though the fact was not documented in the medical record. Although we trained physician reviewers to be alert to this issue, it may still have led to some error in our estimates. None of this is to say that deaths of sick, elderly patients due to adverse events are excusable, only that the number of deaths we report here is not directly comparable in economic terms to the number of deaths from automobile accidents, for example, in which the victims are generally younger and healthier.

In summary, we reviewed a random sample of 30 121 medical records from New York State in 1984, analyzing them for the presence of adverse events and substandard care. We believe that our findings indicate that there are certain risk factors, many definable, for the occurrence of adverse events and negligence.

. . . . . . . . . . . . . . . . . . . . .

**Authors' affiliations**
**T A Brennan, H H Hiatt,** Division of General Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Mass, USA
**L L Leape, L Hebert, A R Localio, A G Lawthers, J P Newhouse, H H Hiatt,** Department of Health Policy and Management, Harvard School of Public Health, Boston, Mass, USA
**N M Laird,** Department of Biostatistics, Harvard School of Public Health, Boston, Mass, USA
**J P Newhouse,** Department of Health Care Policy, Harvard Medical School, Boston, Mass, USA
**J P Newhouse,** Kennedy School of Government, Harvard University, Cambridge, Mass, USA

## APPENDIX I: EXAMPLES OF ADVERSE EVENTS AND NEGLIGENCE

Case 1: During angiography to evaluate coronary artery disease, a patient had an embolic cerebrovascular accident. The angiography was indicated and was performed in standard fashion, and the patient was not at high risk for a stroke. Although there was no substandard care, the stroke was probably the result of medical management. The event was considered adverse but not due to negligence.

Case 2: A patient with peripheral vascular disease required angiography. After the procedure, which was performed in standard fashion, the patient's renal function deteriorated as a result of exposure to angiographic dye. The hospital course was stormy because of kidney failure, but the patient's renal function slowly returned to normal. The adverse event caused the prolonged hospital stay, but there was no negligence. The event was considered adverse but not due to negligence.

Case 3: During a therapeutic abortion after 13 weeks of pregnancy, the physicians unknowingly perforated the patient's uterine wall with a suction device and lacerated the colon. The patient reported severe pain, but was discharged without evaluation. She returned one hour later to a hospital emergency room with even greater pain and evidence of internal bleeding. She required a two-stage

surgical repair over the ensuing four months. The event was considered adverse and due to negligence.

Case 4: A middle aged man had rectal bleeding. The patient's physician completed only a limited sigmoidoscopy, which was negative. The patient had continued rectal bleeding but was reassured by the physician. Twenty two months later, after a 14 kg (30 lb) weight loss, he was admitted to a hospital for evaluation. He was found to have colon cancer with metastases to the liver. The physicians who reviewed his medical record judged that proper diagnostic management might have discovered the cancer when it was still curable. They attributed the advanced disease to substandard medical care. The event was considered adverse and due to negligence.

## APPENDIX II: CLASSIFYING PATIENTS ACCORDING TO SPECIALTY GROUP AND RISK OF ADVERSE EVENTS

In order to classify hospitalizations according to clinical specialty, we used the principal discharge diagnosis. Beginning with the Fetter classification of 24 specialties based on diagnosis related groups (Fetter RB: Preliminary research document: assignment of diagnosis related groups using ICD-9-CM codes to clinical subspecialties, School of Organization and Management, Yale University, 1980), we made four alterations to reduce the number of specialty groups to 10. First, the following specialties were combined with general medicine: cardiology, nephrology, dermatology, neurology, endocrinology, pulmonology, gastroenterology, rheurnatology, and hematology. Second, the following specialties were combined in a residual group: dentistry, gynecology, ophthalmology, and otolaryngology. Third, medical back problems (DRG 243) was moved from the orthopedics specialty to the general medicine specialty. Fourth, psychiatric discharges were not included in this study.

The principal discharge diagnosis was used to measure the risk of adverse events associated with severity of disease. To obtain DRG risk groups, three senior physicians were asked to rate on a scale of 1 (low) to 6 (high) the likelihood that a patient in each of the 470 DRGs would have an adverse event. All DRGs received at least one rating of the likelihood of adverse events. By selecting natural breakpoints in the distribution, we grouped the scores into four risk categories.

The risk groups formed by the physicians' judgments were validated first by comparing the rates of adverse events among these groups with use of the data from the Harvard Medical Practice Study pilot project. They were validated again with the 30 121 observations of this study. Both sets of data exhibited monotonic increases in the rates of adverse events with DRG level. The rate of adverse events according to DRG level in this study is shown in table 6.

## REFERENCES
1 **General Accounting Office**. *Medical malpractice: six states case studies.* GAO publication No HRD–87–21. Washington, D.C.: Government Printing Office, 1986.
2 **General Accounting Office**. *Medical malpractice: insurance costs.* GAO publication No HRD–86–112. Washington, DC: Government Printing Office, 1986.
3 **General Accounting Office**. *Medical malpractice: characteristics of claims closed in 1984.* GAO publication No HRD–87–55. Washington, DC: Government Printing Office, 1986.
4 **Florida Academic Task Force for the Review of the Insurance and Tort System**. *Preliminary fact-finding report on medical malpractice.* Gainesville, Fla: State of Florida, 1987.
5 **Department of Health and Human Services**. *Report of the task force on medical liability and malpractice.* DHHS publication No 190–412170133. Washington, DC: Department of Health and Human Services, 1987.
6 **Siegel DR**. An overview of medical malpractice legislation. *NY State J Med* 1986;**86**:377–91.
7 **Robinson G**. The medical malpractice crisis of the 1970's: a retrospective. *Law Contemp Prob* 1986;**49**:5–36.
8 **Priest G**. Modern tort law and its reform. *Valparaiso Univ Law Rev* 1987;**22**:1–38.

9  **Nader R**. The corporate drive to restrict their victims' rights. *Gonzaga Law Rev* 1987;**22**:15–30.
10 **Pear T**. Insurers reducing malpractice fees for doctors in U.S. *New York Times* September 1990;**23**:1.
11 **California Medical Association**. *Report of the Medical Insurance Feasibility Study*. San Francisco: California Medical Association, 1977.
12 **Steel K**, Gertman PM, Crescenzi C, *et al*. Iatrogenic illness on a general medical service at a university hospital. *N Engl J Med* 1981;**304**:638–42.
13 **Couch NP**, Tilney NL, Rayner AA, *et al*. The high cost of low-frequency events: the anatomy and economics of surgical mishaps. *N Engl J Med* 1981;**304**:634–7.
14 **Brennan TA**, Localio AR, Laird NM. Reliability and validity of judgments concerning adverse events and negligence. *Med Care* 1989;**27**:1148–58.
15 **Brennan TA**, Localio AR, Leape LL, *et al*. Identification of adverse events suffered by hospitalized patients: a cross-sectional study of litigation, quality assurance, and medical records at two teaching hospitals. *Ann Intern Med* 1990;**112**:221–6.
16 **Hiatt HH**, Barnes BA, Brennan TA, *et al*. A study of medical injury and medical malpractice: an overview. *N Engl J Med* 1989;**321**:480–4.
17 **Harvard Medical Practice Study**. *Patients, doctors and lawyers: studies of medical injury, malpractice litigation, and patient compensation in New York*. Boston: Harvard Medical Practice Study, 1990. Technical appendix 5.V. 1.
18 **New York State Department of Health**. *Statewide planning and research cooperative system: annual report series, 1984*. Albany, NY: Department of Health, 1985.
19 **Shah BV**. *SESUDAAN: standard errors program for computing standardized rates from sample survey data*. Research Triangle Park, NC: Research Triangle Institute, 1981.
20 **Brennan TA**. *Just doctoring: medical ethics in the liberal state*. Berkeley: University of California Press (in press).

·············· **COMMENTARY** ··············

## HARVARD MEDICAL PRACTICE STUDY

The Harvard Medical Practice Study (HMPS)[1 2] was not the first study to examine adverse events in healthcare organizations, but it established the standard by which adverse events are measured and laid the groundwork for policy discussions on patient safety in several countries. This commentary examines the impact of the study on research and policy in the US and elsewhere.

The methods used in the HMPS were based on the 1977 California medical insurance feasibility study.[3] The refining and rigorous application of these methods to a random sample of patients and hospitals offered one of the first large sample estimates of adverse events in the health services research literature.

Today the HMPS is best known for the methods developed to identify adverse events and estimate their incidence. Yet this was only one of the investigators' goals. Defining the incidence of adverse events was necessary for evaluating whether the tort system was effective in rewarding those who are injured as a result of their care in hospitals and assessing the economic consequences of such injuries. The dramatic finding that adverse events were a common component of hospital care has largely overshadowed the attention given to the evaluation of the tort system and assessment of costs.

The HMPS method for identifying adverse events is based on a two stage chart review. The first stage is carried out by nurses to screen patient records that are likely to include an adverse event. Selected charts are then reviewed in more detail by physicians to confirm the presence of adverse events and to assess the extent to which these events indicate substandard care. This review process has become the benchmark method for research on adverse events in hospitals. However, it should be noted that the methods have drawn criticism for several reasons. Firstly, the documentation in patient records may be incomplete allowing some adverse events to escape notice and, secondly, it is often difficult to untangle the contribution of medical intervention from the underlying disease processes. Thus, even with the carefully structured review process created in the Harvard study, there is substantial variation in the judgments of physician reviewers in that study and others who have used this method. Reliability estimates on the assessment of adverse events are only moderate; those relating to negligence and the degree of impairment attributable to the adverse event are even lower. Other methods, including direct observation and stimulated recall, yield higher numbers of adverse events. Detection using administrative data systems, computer screens, and error reporting systems are less sensitive, but also less costly.[4] Regardless of this, chart review—perhaps because it relies upon the written history of patients' experiences and provides a longitudinal view not available through any other method (except for computerized records)—is often considered the best method for identifying adverse events.

Despite the considerable weight of its findings, the full impact of the HMPS was not felt until the release of the Institute of Medicine (IOM) report, *To Err is Human*,[5] in late 1999. The authors of this report developed population estimates of the numbers of Americans who die in hospitals as a result of preventable adverse events based on extrapolations from the HMPS and the more recent Utah-Colorado study.[6] Before publication of the IOM report patient safety was a hidden issue in American health care, but following its publication patient safety became a focal point for reform. The HMPS study contributed important evidence to the ensuing policy debates on the steps needed to assess patient safety and reduce the injury burden.

Another important impact of the HMPS is the use of these methods by researchers in other countries. The HMPS results stimulated interest among Australian researchers and policy makers who replicated the study in a sample of 28 hospitals in 1995. The Australians were more interested in the quality of hospital care than in the performance of the malpractice system, so they reoriented the chart review assessment from judgments of negligence (was the care substandard?) to assessments of potential improvement (could the adverse event be prevented?) This orientation, together with some alterations in the methods, yielded substantially higher results. While the Harvard study found 3.7% of hospital patients in New York State had experienced an adverse event, the Australian study reported that 16.6% of hospital admissions were associated with an adverse event.[7] Later analyses comparing the Australian methods with those of the Utah-Colorado study reduced the magnitude of these differences.[8 9]

In addition to the Australian study, the HMPS methods have been replicated in the UK,[10] Denmark, and New Zealand.[11 12] A recent study in France compared the Harvard methods with other approaches,[13] while a Canadian study of adverse events will be published shortly. The existence of benchmarks in other jurisdictions heightens the appeal of the HMPS methods as a means of assessing the status of patient safety in hospitals around the world.

The HMPS identified adverse drug events as the second most common type of event. This result helped to stimulate research on the epidemiology of adverse drug events[14] and on methods to reduce them.[15] The Harvard study and the more recent study in Utah and Colorado have also contributed to policy discussions about tort reform and the effectiveness of the current medical malpractice system in the US.[16]

The next steps for improving adverse event reporting and investigation will require flexible and efficient tools that can accurately identify patients at risk and target areas for improvement. Beyond the issues of reliability, chart review methods are limited by the retrospective nature of such reviews and the expense involved in clinical assessment of patient records. However, these limitations could be reduced if the screens used in the first stage review were computerized, or other methods were developed that identified

patients with a high likelihood of adverse events. Chart review inevitably points toward individual activities rather than system problems that underlie preventable adverse events. Identifying an adverse event or a pattern of events can therefore only be the first step in creating more effective care systems. Nevertheless, the information gleaned from such reviews may help to stimulate improvement. The next generation of tools needs to be applicable at a reasonable cost and linked to ongoing reviews of patient care. Computerization of the chart review tools would extend the use of these methods from research to quality improvement.

**G R Baker**

Department of Health Policy, Management and Evaluation, University of Toronto, Toronto, Canada M5S 1A8; ross.baker@utoronto.ca

## 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**:370–7.
2 **Leape L**, Brennan T, 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 **California Medical Association**. *Report on the medical insurance feasibility study*. San Francisco: Sutter, 1977.
4 **Thomas EJ**, Petersen LA. Measuring errors and adverse events in health care. *J Gen Intern Med* 2003;**18**:61–7.
5 **Kohn LT**, Corrigan JM, Donaldson MS, eds. *To err is human: building a safer health system*. Washington, DC: National Academy Press, 1999.
6 **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.
7 **Wilson RM**, Runciman WB, Gibberd RW, *et al.* The Quality in Australia Health Care Study. *Med J Aust* 1995;**163**:458–76.
8 **Thomas EJ**, Studdert DM, Runciman WB, *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.
9 **Runciman WB**, Webb RK, Helps SC, *et al.* A comparison of iatrogenic injury studies in Australia and the USA. II: Reviewer behaviour and quality of care. *Int J Qual Health Care* 2000;**12**:379–88.
10 **Vincent C**, Neale G, Woloshynowych M. Adverse events in British hospitals: preliminary retrospective record review. *BMJ* 2001;**322**:517–9.
11 **Davis P**, Lay-Yee R, Schug S, *et al.* Adverse events regional feasibility study: indicative findings. *NZ Med J* 2001;**114**:203–5.
12 **Davis P**, Lay-Yee R, Schug S, *et al.* Adverse events regional feasibility study: methodological results. *NZ Med J* 2001;**114**:200–2.
13 **Michel P**, Quenon JL, de Sarasqueta AM, *et al.* Comparison of three methods for estimating rates of adverse events and rates of preventable adverse events in acute care hospitals. *BMJ* 2004;**328**:199–200.
14 **Bates DW**, Leape LL, Cullen DJ, *et al.* Effect of computerized physician order entry and a team intervention on prevention of serious medication errors. *JAMA* 1998;**280**:1311–6.
15 **Bates DW**, Cullen DJ, Laird N, *et al.* Incidence of adverse drug events and potential adverse drug events—implications for prevention. *JAMA* 1995;**274**:29–34.
16 **Studdert DM**, Brennan TA. No-fault compensation for medical injuries: the prospect for error prevention. *JAMA* 2001;**286**:217–23.