Excess Length of Stay, Charges, and Mortality Attributable to Medical Injuries During Hospitalization

Chunliu Zhan, MD, PhD

Marlene R. Miller, MD, MSc

Context Although medical injuries are recognized as a major hazard in the health care system, little is known about their impact.

Objective To assess excess length of stay, charges, and deaths attributable to medical injuries during hospitalization.

Design, Setting, and Patients The Agency for Healthcare Research and Quality (AHRQ) Patient Safety Indicators (PSIs) were used to identify medical injuries in 7.45 million hospital discharge abstracts from 994 acute-care hospitals across 28 states in 2000 in the AHRQ Healthcare Cost and Utilization Project Nationwide Inpatient Sample database.

Main Outcome Measures Length of stay, charges, and mortality that were recorded in hospital discharge abstracts and were attributable to medical injuries according to 18 PSIs.

Results Excess length of stay attributable to medical injuries ranged from 0 days for injury to a neonate to 10.89 days for postoperative sepsis, excess charges ranged from \$0 for obstetric trauma (without vaginal instrumentation) to \$57 727 for postoperative sepsis, and excess mortality ranged from 0% for obstetric trauma to 21.96% for postoperative sepsis ($P<.001$). Following postoperative sepsis, the second most serious event was postoperative wound dehiscence, with 9.42 extra days in the hospital, \$40 323 in excess charges, and 9.63% attributable mortality. Infection due to medical care was associated with 9.58 extra days, \$38 656 in excess charges, and 4.31% attributable mortality.

Conclusion Some injuries incurred during hospitalization pose a significant threat to patients and costs to society, but the impact of such injury is highly variable.

JAMA. 2003;290:1868-1874 www.jama.com

Despite recognition of medical injuries as a leading cause of death and patient safety as a critical area for improvement,[1-4] the overall approach to patient safety (eg, focusing on medical injuries)[5,6] and definitional issues (eg, what is considered preventable)[7,8] remains debated. Medical injuries can happen during all stages of the complicated process of care, vary widely in nature, and are relatively infrequent. The lack of standard taxonomy, in addition to definitional issues, in large part explains why so little is known about the prevalence, adverse outcomes, and effective prevention of medical injuries.[7,9-12]

The limited research on medical injuries has primarily relied on medical record abstraction conducted ad hoc and on a small scale.[2,13-17] Medical records contain rich clinical details that allow identification of various injuries and close calls and analysis of circumstances and causes. However, transforming medical records into useful research data on medical injuries is resource intensive and requires exceptional knowledge and skills in medical context and research methods. Alternative systems for research include mandatory and voluntary reports, drug safety surveillance, nosocomial infection surveillance, and medical malpractice data.[2,18] All of these data systems have limitations, and obtaining access for research purposes may be difficult. For example, approximately 20 US states mandate reporting of serious adverse events,[19] but no published study has ever used these data, most likely because they are strictly guarded from the public and researchers.

Administrative data are a potential source of information on medical injuries. Administrative data are regularly collected and maintained for reimbursement and management purposes; are computer readable, inexpensive to analyze, and longitudinal; and cover large populations. These data have been used to reveal startling small-

For editorial comment see p 1917.

Author Affiliations: Center for Quality Improvement and Patient Safety, Agency for Healthcare Research and Quality, Department of Health and Human Services, Rockville, Md (Dr Zhan); Department of Pediatrics, Johns Hopkins University, Baltimore, Md (Dr Miller).

Corresponding Author and Reprints: Chunliu Zhan, MD, PhD, Center for Quality Improvement and Patient Safety, Agency for Healthcare Research and Quality, 540 Gaither Rd, Rockville, MD 20850 (e-mail: czhan@ahrq.gov).

area variations in health care and clinical practice patterns since the 1970s[20] and to conduct outcomes research since the 1980s.[21] In the early 1990s, Iezzoni et al[22] developed a Complications Screening Program to screen systematically for quality gaps using administrative data. Soon after, researchers at the Agency for Healthcare Research and Quality (AHRQ)[23] developed similar quality indicators. In response to the increased need for patient-safety tools, the AHRQ recently developed and released the Patient Safety Indicators (PSIs), which are specifically designed for screening administrative data for incidences of concern related to patient safety.[24]

This study uses the PSIs and administrative data to assess excess length of stay (LOS), charges, and deaths attributable to medical injuries during hospitalization. Recognizing the limitations of these data, we draw conclusions based on multivariable matching analysis, which compares medical injury cases with controls that share similar characteristics, and regression analysis, which compares all cases and controls and adjusts for confounding factors available in administrative data.

METHODS

Data and Variables

The primary source of data for this study was the 2000 Healthcare Cost and Utilization Project Nationwide Inpatient Sample (NIS) developed by the AHRQ. The 2000 NIS file contained 7.45 million uniform hospital discharge abstracts for all inpatient stays from 994 acute-care general hospitals across 28 states, approximating a 20% stratified sample of nonfederal acute-care hospitals in the United States.[25] The NIS includes variables on sources and types of admissions, diagnosis related groups (DRGs), 15 *International Classification of Diseases, 9th Edition, Clinical Modification (ICD-9-CM)* diagnosis codes, 15 procedure codes, procedure days from admission, discharge status, LOS, total charges, patient demographic characteristics, insurance coverage, and a few hospital characteristic variables.

To account for severity of illness at admission, we used the method previously described by Elixhauser et al[26] to define 30 comorbid diseases and 3M Corporation's (St Paul, Minn) All Patient Refined (APR)–DRG system[27] to define 4 disease stages pertaining to risk of death and expected resource consumption. The comorbid diseases were summarized into indices in a 2-step process. We first estimated the regression coefficients of the 30 comorbidities on in-hospital mortality, LOS, and total charges, respectively, using all discharges in the 2000 NIS and controlling for age, sex, race/ethnicity, insurance status, and DRG fixed effects. The logistic regression of mortality produced a *c* statistic of 0.926. In the second step, the significant positive coefficients, which indicated independent effects of the presence of comorbid diseases on LOS, charges, and probability of death, were used as relative weights to summarize the comorbidities into indices. Preliminary analysis showed that the resulting 3 comorbidity indices were highly correlated ($R>0.96$) and were also correlated with disease stages ($R>0.50$). We recognized that the presence of a comorbid disease could shorten LOS and reduce charges by leading to an early death, resulting in negative coefficients in LOS and charge regressions contrary to the objective of the measures. The APR-DRG, however, may inappropriately incorporate preventable complications in defining disease staging.[27] Based on these preliminary analyses and considerations, we used the methods of Charlson et al,[28] Romano et al,[29,30] and Ghali et al[31] in choosing the index based on excess probability of death associated with comorbidities for our analysis.

Identifying Medical Injuries During Hospitalizations

We used the AHRQ PSIs to identify medical injuries during hospitalization. The University of California–Stanford Evidence-Based Practice Center, with funding and collaboration from the AHRQ, developed the PSIs through a 5-step process.[32,33] First, the literature was reviewed to develop a list of candidate indicators and to collect information about their validity and reliability. Second, several clinician panels were formed to evaluate the clinical sensibility and to suggest revisions of the candidate indicators. Third, *ICD-9-CM* coding experts were consulted to ensure that the definition of each indicator reflected the intended clinical situation. Fourth, an empirical analysis of the promising indicators was conducted using Healthcare Cost and Utilization Project data. Lastly, PSI software and documentation were produced for public release by the AHRQ.

The AHRQ PSIs include 20 indicators with reasonable validity, specificity, and potential for fostering quality improvement. Details on variables and *ICD-9-CM* codes used to define these PSIs are available elsewhere.[24,34] Our analysis excluded 2 indicators, death in low-mortality DRGs and failure to rescue, because cases in both of these indicators died during hospitalization, and analyses of excess LOS and charges would not be meaningful.

Analytical Approach

Adequate control of confounding is the major challenge to attributing adverse outcomes correctly to medical injuries in administrative data–based analyses. The LOS, charges, and mortality for hospital stays differ by principal reasons for admission, severity of illness, comorbid diseases, and other patient demographic and socioeconomic characteristics.[35] Patient outcomes also differ across hospitals.[13] Of special concern is the variation across hospitals in numbers of diagnosis codes used and coding variation,[36] which inevitably affect the medical injuries flagged by AHRQ PSIs and comorbidities defined by the approach described by Elixhauser et al.[26]

Identification of a relatively homogeneous risk pool for each PSI removes some confounding variation.[22,23] The standard approach to risk adjustment is regression analysis with outcomes regressed against an indicator and confounding factors. How-

Table 1. Hospital Characteristics (N = 994)

Hospital Characteristics	No. (%) of Hospitals
Hospital control	
Public	234 (23.5)
Nonprofit	620 (62.4)
For profit	140 (14.1)
Location/teaching status	
Rural	447 (45.0)
Urban nonteaching	381 (38.3)
Urban teaching	166 (16.7)
No. of beds*	
Small	344 (34.6)
Medium	315 (31.7)
Large	335 (33.7)
Region	
Northeast	139 (14.0)
Midwest	286 (28.8)
South	381 (38.3)
West	188 (18.9)

*The Healthcare Cost and Utilization Project hospital size designations of small, medium, and large based on number of beds varied by region and rural/urban locality.

Table 2. Patient Characteristics (N = 7 448 149)

Patient Characteristics	Patients, %
Age, y	
≤17	17.5
18-39	22.4
40-64	25.1
≥65	34.9
Sex	
Male	40.7
Female	59.3
Race/ethnicity	
White	53.4
Black	10.0
Hispanic	8.7
Other/unknown	27.9
Annual income by ZIP code area, $	
1-24 999	8.6
25 000-34 999	29.3
35 000-44 999	25.9
≥45 000	34.4
Primary insurance	
Medicare	36.0
Medicaid	16.5
Private	39.1
Uninsured	4.5
Other	3.8
Comorbidity	
Without comorbidity	51.7
Comorbidity index*	3.77 (2.39)

*Mean (SD) comorbidity index score (possible range of scores, 0 = 33.81, with higher scores indicative of greater comorbidity).

ever, these regressions easily produce small, statistically significant, and often biased findings because of limited clinical details, extreme imbalance of few cases vs large number of controls, substantial misclassification of medical injuries in terms of false-negative and false-positive results, and difficulties in model specification.[37]

Our primary approach to confounding was multivariable matching, which is a relatively new method in assessing treatment effects in observational studies in which assignment to case or control group is not at random.[38-42] This method matches cases and controls on known characteristics to produce a control group that mimics the case group. When the cases are rare and controls are plentiful, as is the case in administrative data–based medical injury analysis, each case can be matched to 1 or to many controls.[42,43] Our approach expanded on several earlier studies of medical injuries that used matching. Classen et al[44] and Bates et al[45] matched cases of adverse drug reactions (ADRs) with controls on DRG, severity, comorbidity, and patient sociodemographic characteristics to estimate excess LOS, charges, and deaths attributable to ADRs. In another study, Bates et al[46] matched patients with ADRs with patients from the same hospital unit and similar pre-event LOS to study risk factors for ADRs. Jensen et al[47] matched cases of hospital-acquired *Staphylococcus aureus* infections with patients with the same principal diagnosis at admissions to identify risk factors among unmatched factors, such as age and anemia.

We matched each identified medical injury case with up to 4 controls from the same hospitals and with the same DRG, sex, white or nonwhite race, and age within 10 years. We further matched cases without any comorbidity with controls without any comorbidity and matched cases and controls with comorbidities within a 1% difference in risk of death due to comorbidities. The matching algorithm first selects controls that meet the matching criteria and then randomly selects 4 controls if more than 4 eligible controls are found. Matching more controls with each case reduces the impact of random variation in eligible controls. However, even armed with a large database, we were not able to find 4 controls for most cases. Our initial analysis suggests that a 1-to-4 match is a good compromise.

For any given PSI, excess LOS was calculated as the difference between LOS for a case and mean LOS for available controls for each matched set, and the mean excess LOS and its SE were then calculated for all matched sets. Unmatched cases and controls were discarded. Similar processes were used to calculate excess charges and excess mortality.

We also computed linear and logistic regressions to estimate excess outcomes attributable to medical injuries to provide comparisons with matching analyses. Regression analyses used SAS PROC GLM, PROC MIXED, and an SAS macro for the mixed-effect model (GLIMMIX)[48] to adjust for hospital and DRG fixed effects and patient variables. For computation convenience, we eventually used PROC GLM to estimate the excess LOS and charges where DRG and hospital fixed effects were adjusted but not calculated. PROC GLM was also used to estimate excess mortality to ensure direct comparability with the estimates from matching analysis. In contrast with matching analysis, where only matched cases and controls were used, regression analysis included all cases and all controls from associated risk pools. The regression analyses provide a cross-validation of our matching estimates in light of the limitations of administrative data. $P<.05$ was considered statistically significant for all analyses.

RESULTS

TABLE 1 and TABLE 2 describe the characteristics of the 994 hospitals and the 7 448 149 patient discharges in the 2000 NIS data. Although the 2000 NIS was constructed from 28 states that voluntarily contributed data, characteristics of the included patients and hospitals closely mirror the 2000 National Hospital Discharge Survey in patient age and sex[49] and the 2000 American Hospital Association Annual Survey in the 4 hospital variables (data available from the authors).

TABLE 3 lists the 18 PSIs along with the number of PSI events, the risk pool, and the event rate for each indicator. The rates of the PSIs ranged from 0.004 (transfusion reactions) to 224.21 (obstetric trauma in vaginal delivery with instrumentation) per 1000 discharges. The rates cannot be directly compared across PSIs because the denominator for each PSI is determined by specific inclusion and exclusion criteria. For example, the risk pool for postoperative sepsis includes patients admitted for selective surgery without a primary diagnosis of infection and without any diagnosis indicating immunocompromised state or cancer, whereas the risk pool for trauma to a neonate includes only live births.

Table 3 also reports the percentage of cases with at least 1 matching control. The matching rates varied from 33% for postoperative sepsis to 99% for obstetric trauma following cesarean delivery. As expected, matching rates were higher in homogeneous situations such as birth-related discharges than in more complicated situations such as surgery-related discharges. Unmatched cases, which were deleted from the matching analysis, differed from matched cases in patient demographics and severity and the degree and direction of difference varied across PSIs. For example, unmatched cases with a foreign body left during a procedure had more comorbid diseases than did matched cases (25% vs 74%), but unmatched cases of iatrogenic pneumothorax had fewer comorbid diseases than did matched cases (57% vs 50%). The number of matches obtained also varied among PSIs. Using foreign body left during a procedure as an example, 69% of cases had at least 1 match, 57% had at least 2 matches, 47% had at least 3 matches, and 40% had 4 matches.

TABLE 4 presents the excess LOS, charges, and mortality attributable to the PSI events based on the matching analysis. Excess LOS attributable to PSI events ranged from 0 days for injury to a neonate to 10.89 days for postoperative sepsis; excess charges ranged from \$0 for obstetric trauma (without vaginal instrumentation) to \$57 727 for postoperative sepsis; and excess mortality ranged from 0% for obstetric trauma to 21.92% for postoperative sepsis ($P<.001$). The most serious event appeared to be postoperative sepsis, followed by postoperative wound dehiscence, with 9.42 extra days, \$40 323 in excess charges, and 9.63% attributable mortality. Infection due to medical care was associated with 9.58 extra days, \$38 656 in excess charges, and 4.31% attributable mortality. Birth trauma and anesthesia complications appeared to have the lowest impact on these outcome measures, although the patients with these events undoubtedly experienced other significant outcomes that were not captured by the PSIs.

Regression analysis based on all cases and controls produced similar results (data available from the authors). The largest differences between the 2 sets of estimates were 8% for LOS estimates and 18% for charge estimates. Although the differences between excess mortality estimates were small, the matching estimates doubled the regression estimates for foreign body left during a procedure (1.13% vs 2.14%) and postoperative wound dehiscence (4.32% vs 9.63%).

COMMENT

This study identified a substantial number of cases that were likely to be medical injuries that resulted from failure in the process of care in hospitals. The PSI rates may represent only the tip of the iceberg and are in fact substantially lower than those reported from other sources of data for individual types of events.[50] For example, previous studies reported that foreign bodies left during procedures were approximately 1 in 1000 to 1500 intra-abdominal operations[51,52] and reported decubitus ulcer at 7.5% of discharges,[53] but our analysis suggests approximately 1 in 100 000 and 2.2%, respectively.

Although this study adds little to the scant knowledge about the national prevalence of various types of medical injuries, it provides significant insights into the adverse effects of selected medical injuries on patients and health care resources. Previously, a number of studies used medical records from single institutions to examine excess LOS,[44]

Table 3. Patient Safety Events and Multivariable Matching Rates*

Patient Safety Indicators	No. of Events	Risk Pool	Rate per 1000 Discharges at Risk	Match Rate, %
Accidental puncture or laceration	11 810	5 628 112	3.32	75
Birth trauma, injury to neonate	4740	720 021	6.53	96
Complications of anesthesia	1369	1 933 085	0.71	74
Decubitus ulcer	41 440	1 932 676	21.51	56
Foreign body left during procedure	536	6 572 845	0.09	69
Iatrogenic pneumothorax	3919	5 861 689	0.67	66
Obstetric trauma, cesarean birth	1138	191 227	6.97	99
Obstetric trauma, vaginal birth with instrumentation	12 518	51 225	224.21	95
Obstetric trauma, vaginal birth without instrumentation	51 223	591 752	86.61	99
Postoperative hemorrhage or hematoma	3494	1 695 495	2.06	69
Postoperative hip fracture	1068	1 397 898	0.77	51
Postoperative physiologic and metabolic derangement	799	801 702	1.00	44
Postoperative pulmonary embolism or deep vein thrombosis	15 704	1 689 662	9.34	61
Postoperative respiratory failure	2275	633 855	3.58	37
Postoperative sepsis	2592	229 853	11.25	33
Postoperative wound dehiscence	843	411 099	2.05	55
Selected infection due to medical care	11 449	5 752 102	1.99	63
Transfusion reaction	30	6 572 845	0.004	80

*Number of events and denominators are defined by the Agency for Healthcare Research and Quality Patient Safety Indicator (PSI) algorithm. Four matching controls were sought for each PSI case following the predefined matching rules: same hospital, diagnosis related group, and sex; white vs nonwhite race; within 10 years of difference in age; no comorbidity; and risk of death due to comorbidities less than 1% among those with comorbidities. Matching rates are the percentages of PSI cases that were matched with at least 1 control.

cost,[44,45] and deaths[44] attributable to adverse drug events in hospitalized patients. The Harvard Medical Practice Study used data from medical records from small samples of hospitals to estimate deaths and total health care costs attributable to medical injuries in New York State[54,55] and Colorado and Utah.[16] Based on such studies of limited scale, the landmark report by the Institute of Medicine[2] in 1999 concluded that medical injuries result in 44 000 to 98 000 deaths and $17 billion direct health care costs annually. The study by Kalish et al[56] is an exception. They used hospital administrative data from 404 California acute-care hospitals to show that patients undergoing major surgery who had 1 or more of the 26 types of complications identified had longer stays (13.5 vs 5.4 days) and higher total charges ($30 896 vs $9239). However, like all previous work, this analysis grouped all medical injuries together and failed to account for the clinical heterogeneity of medical injuries and their different consequences. Our study for the first time provides specific estimates for excess LOS, charges, and mortality due to each of the 18 specific types of medical injuries.

Our matching methods drew conclusions based primarily on matched cases and controls, which greatly reduces confounding but also limits our ability to use weights and make national extrapolation. However, given that our data include discharge abstracts from an approximate 20% sample of US hospitals and, therefore, the national estimates could be 5 times higher, one can infer that the 18 types of medical injuries may add to a total of 2.4 million extra days of hospitalization, $9.3 billion excess charges, and 32 591 attributable deaths in the United States annually. Assuming an average cost-charge ratio of approximately 0.5, as previously suggested,[57] the total national health care costs for the 18 types of medical injuries identified by the PSIs could be $4.6 billion. It is not surprising that our estimates are lower than the Institute of Medicine estimates because our estimates cover only selected types of medical injuries that were discovered during hospitalization and were recorded as *ICD-9-CM* codes. Nevertheless, our estimates clearly support the Institute of Medicine's contention that medical injuries are a serious epidemic confronting our health care system.

Several limitations should be recognized in interpreting these results. First, the reliability and validity of the AHRQ PSIs depend on the accuracy and completeness of *ICD-9-CM* coding in the administrative data. There may be various coding errors. Some relevant diagnoses may not be coded. Number of codes recorded and coding practices may vary across hospitals.[36] Financial incentives and fear of retribution could also affect the accuracy of the codes.[58,59] Second, the *ICD-9-CM* system was not designed to identify medical injuries and, therefore, is not clinically precise for this purpose.[60-62] Some PSIs, such as a foreign body left during a procedure, are unequivocally medical injuries in the process of care, but others, such as postoperative hemorrhage, may be in part due to patient conditions and in part due to failure in care. Administrative data–based tools, although generally high in specificity (ie, low in false-positive results),[62] are low in sensitivity (ie, high in false-negative results) in identifying medical injuries.[63-66] Because of these issues, coupled with substantial underreporting, the AHRQ PSIs are not definitive measures of medical injuries and have limited validity as measures for

Table 4. Excess Length of Stay, Charges, and Mortality Attributable to Patient Safety Events*

Patient Safety Indicators	Excess LOS, d	*P* Value	Excess Charge, $	*P* Value	Excess Mortality, %	*P* Value
Accidental puncture or laceration	1.34 (0.08)	<.001	8271 (344)	<.001	2.16 (0.20)	<.001
Birth trauma, injury to neonate	−0.09 (0.08)	.27	298 (295)	.32	−0.08 (0.07)	.27
Complications of anesthesia	0.17 (0.90)	.26	1598 (660)	.02	0.24 (0.36)	.51
Decubitus ulcer	3.98 (0.10)	<.001	10 845 (368)	<.001	7.23 (0.23)	<.001
Foreign body left during procedure	2.08 (0.68)	.002	13 315 (3329)	<.001	2.14 (1.06)	.04
Iatrogenic pneumothorax	4.38 (0.24)	<.001	17 312 (1091)	<.001	6.99 (0.73)	<.001
Obstetric trauma, cesarean birth	0.43 (0.14)	.003	2718 (551)	<.001	−0.02 (0.02)	.32
Obstetric trauma, vaginal birth with instrumentation	0.07 (0.02)	<.001	220 (104)	.03	0.00	.32
Obstetric trauma, vaginal birth without instrumentation	0.05 (0.01)	<.001	−93 (66)	.16	0.00	>.99
Postoperative hemorrhage or hematoma	3.94 (0.27)	<.001	21 431 (1257)	<.001	3.01 (0.46)	<.001
Postoperative hip fracture	5.24 (0.69)	<.001	13 441 (1945)	<.001	4.52 (1.34)	<.001
Postoperative physiologic and metabolic derangement	8.89 (0.75)	<.001	54 818 (5099)	<.001	19.81 (2.27)	<.001
Postoperative pulmonary embolism or deep vein thrombosis	5.36 (0.15)	<.001	21 709 (747)	<.001	6.56 (0.33)	<.001
Postoperative respiratory failure	9.08 (0.57)	<.001	53 502 (3121)	<.001	21.84 (1.46)	<.001
Postoperative sepsis	10.89 (0.90)	<.001	57 727 (3077)	<.001	21.92 (1.47)	<.001
Postoperative wound dehiscence	9.42 (0.72)	<.001	40 323 (3467)	<.001	9.63 (1.55)	<.001
Selected infection due to medical care	9.58 (0.23)	<.001	38 656 (1026)	<.001	4.31 (0.35)	<.001
Transfusion reaction	3.44 (1.94)	.09	18 929 (10 068)	.07	−1.04 (1.04)	.33

*Data are expressed as mean (SE). Excess length of stay (LOS) is the difference in LOS for a case and a matching control or mean LOS for controls if multiple matching controls were found. The paired *t* test was used to test the hypothesis of whether mean excess LOS is significantly different from 0. Excess mortality and charges were calculated similarly.

quality of care. Third, lack of clinical details limits the potential of administrative data in risk adjustment, which could result in biases in both matching and regression analysis. Fourth, the sheer size of administrative data can often give the illusion of great precision and power,[37] which, coupled with missing confounding variables and difficulties in statistical model selection, often results in bias in administrative data–based analysis. Lastly, a PSI is defined by a group of *ICD-9-CM* codes that may not necessarily reflect clinically homogeneous problems. In the case of anesthesia complications, for example, the problems could be anesthetic overdose, reaction, or endotracheal tube misplacement, each affecting patients in a different way.

The AHRQ PSIs incorporated the latest understanding of the potentials and limitations of administrative data and drew on broad consultations of clinical and coding experts to ensure their clinical validity and reliability.[32,33] The multivariable matching method took advantage of the large number of records in administrative data and allowed us to derive estimates based on matched cases and controls that share similar characteristics in treatment settings, patient conditions, and care processes. Because the PSIs flag few false-positive results (controls misidentified as cases) and a relatively large number of false-negative results (unidentified medical injury cases), as discussed earlier, it was more likely that matches were done between flagged cases and false-negative results than between false-positive results and false-negative results, resulting in conservative estimates. Matching, in combination with our comorbidity index, reduced the possibility of bias associated with the lack of clinical details on patient severity and acuity and avoided the statistical dilemma associated with large sample size for risk adjustment. In addition, similar results from matching and regression models offered extra confidence in our results.

In conclusion, our results clearly show that medical injuries in hospitals pose a significant threat to patients and incur substantial costs to society. This study also points to the need for more research to assess patient outcomes beyond death, charges, and LOS, to understand circumstances and risk factors associated with medical injuries, and to develop strategies to prevent medical injuries.

Author Contributions: *Study concept and design:* Zhan, Miller.
Acquisition of data: Zhan.
Analysis and interpretation of data: Zhan, Miller.
Drafting of the manuscript: Zhan.
Critical revision of the manuscript for important intellectual content: Zhan, Miller.
Statistical expertise: Zhan, Miller.
Obtained funding: Zhan, Miller.
Administrative, technical, or material support: Zhan, Miller.
Study supervision: Zhan.
Disclaimer: The authors of this article are responsible for its contents. No statement in this article should be construed as an official position of the Agency for Healthcare Research and Quality or the US Department of Health and Human Services.
Acknowledgment: We acknowledge comments and suggestions made by Maureen Smith, MD, PhD, Population Health Sciences, University of Wisconsin–Madison Medical School, and Daniel Stryer, MD, James Battles, PhD, Judy Sangl, ScD, and William Baine, MD, AHRQ. We also acknowledge the excellent computing support from Susan Tian, PhD, Social and Scientific Systems Inc, Silver Spring, Md.

REFERENCES

1. President's Advisory Commission on Consumer Protection and Quality in the Health Care Industry. *Quality First: Better Health Care for All Americans–Final Report to the President of the United States*. Washington, DC: US Government Printing Office; 1998.
2. Kohn LT, Corrigan JM, Donaldson M, et al. *To Err Is Human: Building A Safer Health System.* Washington, DC: Institute of Medicine; 1999.
3. Department of Health and Human Services. *The Challenge and Potential for Assuring Quality of Health Care for the 21st Century.* Washington, DC: Department of Health and Human Services; 2000.
4. Hurtado M, Swift E, Corrigan JM, eds. *Envisioning National Health Care Quality Report*. Washington, DC: National Academy of Sciences; 2001.
5. Layde P, Maas L, Teret S, et al. Patient safety efforts should focus on medical injuries. *JAMA.* 2002;287:1993-1997.
6. McNutt R, Abrams R, Aron D. Patient safety efforts should focus on medical errors. *JAMA.* 2002;287:1997-2001.
7. Hayward RA, Hofer TP. Estimating hospital deaths due to medical errors: preventability is in the eyes of the reviewer. *JAMA.* 2001;286:415-420.
8. Leape LL. Lucian Leape on the causes and prevention of errors and adverse events in health care [interview by Peter I. Buerhaus]. *Image J Nurs Sch.* 1999;31:281-286.
9. McDonald CJ, Weiner M, Hui SL. Deaths due to medical errors are exaggerated in Institute of Medicine report. *JAMA.* 2000;284:93-95.
10. Leape LL. Institute of Medicine medical error figures are not exaggerated. *JAMA.* 2000;284:95-97.
11. Shojania KG, Duncan BW, McDonald KM, Wachter RM. Safe but sound: patient safety meets evidence-based medicine. *JAMA.* 2002;288:508-513.
12. Leape LL, Berwick DM, Bates DW. What practices will most improve safety: evidence-based medicine meets patient safety. *JAMA.* 2002;288:501-507.
13. Brennan TA, Hebert LE, Laird NM, et al. Hospital characteristics associated with adverse events and substandard care. *JAMA.* 1991;265:3265-3269.
14. Classen DC, Pestonik SL, Evans RS, Burke JP. Computer surveillance of adverse drug events in hospital patients. *JAMA.* 1991;266:2847-2851.
15. Bates DW, Cullen DJ, Laird N, et al. Incidence of adverse drug events and potential adverse drug events. *JAMA.* 1995;274:29-34.
16. Thomas EJ, Studdert DM, Newhouse JP, et al. Costs of medical injuries in Utah and Colorado. *Inquiry.* 1999;36:255-264.
17. Thomas EJ, Studdert DM, Burstin HR. Incidence and types of adverse events and negligent care in Utah and Colorado. *Med Care.* 2000;38:261-271.
18. Kravitz R, Rolph J, McGuigan K. Malpractice claims data as a quality improvement tool: epidemiology of errors in four specialties. *JAMA.* 1991;266:2087-2092.
19. Leape LL. Reporting of adverse events. *N Engl J Med.* 2002;347:1633-1638.
20. Wennberg J, Gittlesohn A. Small area variations in health care delivery. *Science.* 1973;182:1102-1108.
21. Mitchell J, Bubolz T, Paul J, et al. Using Medicare claims for outcomes research. *Med Care.* 1994;32:JS38-JS51.
22. Iezzoni LI, Foley SM, Heeren T, et al. A method for screening the quality of hospital care using administrative data: preliminary validation results. *QRB Qual Rev Bull.* 1992;18:361-371.
23. Johantgen M, Elixhauser A, Bali JK, Goldfarb M, Harris DR. Quality indicators using hospital discharge data: state and national applications. *Jt Comm J Qual Improv.* 1998;24:88-105.
24. Agency for Healthcare Research and Quality. *AHRQ Quality Indicators–Patient Safety Indicators: Software Documentation.* Rockville, Md: Agency for Healthcare Research and Quality; 2003.
25. Nationwide Inpatient Sample (NIS). Available at: http://www.ahrq.gov/data/hcup/hcupnis.htm. Accessed April 18, 2003.
26. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care.* 1998;36:8-27.
27. Romano PS, Chan BJ. Risk-adjusting acute myocardial infarction mortality: are APR-DRGs the right tool? *Health Serv Res.* 2000;34:1469-1489.
28. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis.* 1987;40:373-384.
29. Romano PS, Roos LL, Jollis JG. Adapting a clinical comorbidity index for use with *ICD-9-CM* administrative data: differing perspectives. *J Clin Epidemiol.* 1993;46:1075-1079.
30. Romano PS, Roos LL, Jollis JG. Further evidence concerning the use of a clinical comorbidity index with *ICD-9-CM* administrative data. *J Clin Epidemiol.* 1993;46:1085-1090.
31. Ghali WA, Hall RE, Rosen AK, Ash AS, Moskovitz MA. Searching for an improved clinical comorbidity index for use with *ICD-9-CM* administrative data. *J Clin Epidemiol.* 1996;49:273-278.
32. University of California, San Francisco–Stanford Evidence-Based Practice Center. *Evidence Report for Measures of Patient Safety Based on Hospital Administrative Data—the Patient Safety Indicators.* Rockville, Md: Agency for Healthcare Research and Quality; 2002.
33. Romano PS, Geppert J, Davies S, et al. A national profile of patient safety in US hospitals. *Health Aff (Millwood).* 2003;22:154-166.
34. Measures of patient safety based on hospital administrative data: the patient safety indicators. Available at: http://www.qualityindicators.ahrq.gov/. Accessed April 18, 2003.

35. Iezzoni LI. Dimensions of risk. In: Iezzoni LI, ed. *Risk Adjustment for Measuring Health Outcomes*. 2nd ed. Chicago, Ill: Health Administration Press; 1997:43-168.
36. Iezzoni LI. Data sources and implications: administrative database. In: Iezzoni LI, ed. *Risk Adjustment for Measuring Health Outcomes*. 2nd ed. Chicago, Ill: Health Administration Press; 1997:169-242.
37. Mark DH. Race and the limit of administrative data. *JAMA.* 2001;285:337-338.
38. Rubin D. Matching to remove bias in observational studies. *Biometrics.* 1973;29:159-183.
39. Rubin DB. Using multivariate matched sampling and regression adjustment to control bias in observational studies. *J Am Stat Assoc.* 1979;74:318-328.
40. Rosenbaum P, Rubin D. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *Am Stat.* 1985;39:33-38.
41. Rosenbaum PR. Optimal matching in observational studies. *J Am Stat Assoc.* 1989;84:1024-1032.
42. Smith HL. Matching with multiple controls to estimate treatment effects in observational studies. *Sociol Methodol.* 1997;27:325-353.
43. Ming K, Rosenbaum PR. Substantial gains in bias reduction from matching with a variable number of controls. *Biometrics.* 2000;56:118-124.
44. Classen DC, Pestotnik SL, Evans RS, Lloyd JF, Burke JP. Adverse drug events in hospitalized patients: excess length of stay, extra costs, and attributable mortality. *JAMA.* 1997;277:301-306.
45. Bates DW, Spell N, Cullen DJ, et al. The costs of adverse drug events in hospitalized patients. *JAMA.* 1997;277:307-311.
46. Bates DW, Miller EB, Cullen DJ, et al. Patient risk factors for adverse drug events in hospitalized patients. *Arch Intern Med.* 1999;159:2553-2560.
47. Jensen AG, Wachmann GH, Poulsen KB, Espersen F, Scheibel JS, Frimodt-Moller N. Risk factors for hospital-acquired *Staphylococcus aureus* bacteremia. *Arch Intern Med.* 1999;159:1437-1444.
48. Littel R, Milliken G, Stroup W, Wolfinger R. *SAS System for Mixed Models*. Cary, NC: SAS Institute Inc; 1996.
49. Kozak L, Hall M, Owings M. National Hospital Discharge Survey: 2000 annual summary with detailed diagnosis and procedure data. *Vital Health Stat 13.* 2002;153:1-194.
50. Fleming S. Complications, adverse events, and iatrogenesis: classifications and quality of care measurement issues. *Clin Perform Qual Health Care.* 1996;4:137-147.
51. Jason R, Chisolm A, Lubetsky H. Retained surgical sponge simulating a pancreatic mass. *J Natl Med Assoc.* 1979;71:501-503.
52. Hyslop J, Maull K. Natural history of the retained surgical sponge. *South Med J.* 1982;75:657-660.
53. Whitefield M, Kaltenthaler E, Rakehurst R, Walters S, Paisley S. How effective are prevention strategies in reducing the prevalence of pressure ulcer? *J Wound Care.* 2000;9:261-266.
54. 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;324:370-376.
55. Johnson W, Brennan T, Newhouse JP, et al. The economic consequences of medical injuries. *JAMA.* 1992;267:2487-2492.
56. Kalish RL, Daley J, Duncan CC, Davis RB, Coffman GA, Iezzoni LI. Costs of potential complications of care for major surgery patients. *Am J Med Qual.* 1995;10:48-54.
57. Friedman B, La Mare J, Andrews R, McKenzie D. Practical options for estimating cost of hospital inpatient stays. *J Health Care Finance.* 2002;29(1):1-13.
58. Simborg DW. DRG creep: a new hospital-acquired disease. *N Engl J Med.* 1981;304:1602-1604.
59. Hsia D, Ahern C, Ritchie B, Moscoe L, Krushat W. Accuracy of diagnostic coding for Medicare patients under the Prospective Payment System. *N Engl J Med.* 1988;11:303-311.
60. Zarling E, Piontek F, Kohli R. The utility of hospital administrative data for generating a screening program to predict adverse events. *Am J Med Qual.* 1999;14:242-247.
61. Geraci J, Ashton C, Kuykendall D, Johnson M, Wu L. International Classification of Diseases, 9th Revision, Clinical Modification codes in discharge abstract are poor measures of complication occurrence in medical inpatients. *Med Care.* 1997;35:589-602.
62. Romano PS, Chan BK, Schembri M, Rainwater J. Can administrative data be used to compare postoperative complication rates across hospitals? *Med Care.* 2002;40:856-867.
63. Lawthers AG, McArtney RG, Davis RB. Identification of in-hospital complications from claims data: is it valid? *Med Care.* 2000;38:785-795.
64. Bates DW, O'Neil AC, Petersen LA, Lee L, Brennan TA. Evaluation of screening criteria for adverse events in medical patients. *Med Care.* 1995;33:452-462.
65. Weingart S, Iezzoni L, Davis R, et al. Use of administrative data to find substandard care: validation of the complications screening program. *Med Care.* 2000;38:796-808.
66. McCarthy E, Iezzoni L, Davis R, et al. Does clinical evidence support *ICD-9-CM* diagnosis coding of complications? *Med Care.* 2000;38:868-876.

Wisdom entereth not into a malicious mind, and science without conscience is but the ruin of the soul.
—François Rabelais (c 1494-1553)

