Using inpatient hospital discharge data to monitor patient safety events

By Jennifer A. Taylor, PhD, MPH, Ravi S. Pandian, MA, Lu Mao, MSc, and Yvonne L. Michael, ScD, SM

The development of systematic and sustainable surveillance systems is necessary for the creation of patient safety prevention programs and the evaluation of improvement resulting from innovations. To that end, inpatient hospital discharges collected by the Pennsylvania Health Care Cost Containment Council were used to investigate patient safety events (PSEs) in Pennsylvania in 2006. PSEs were identified using external cause of injury codes (E-codes) in combination with the Agency for Healthcare Research and Quality's patient safety indicators (PSIs). Encounters with and without PSEs were compared with regard to patient age, sex, race, length of stay, and cost. Approximately 9% of all Pennsylvania inpatient discharges had a PSE in 2006. Patients with a PSE were on average older, male, and white. The average length of stay for a PSE was 3 days longer and $35 000 more expensive than a non-PSE encounter. It was concluded that E-codes and PSIs were useful tools for the surveillance of PSEs in Pennsylvania, and that administrative data from healthcare organizations provide a consistent source of standardized data related to patient encounters, creating an opportunity to describe PSEs at the population level.

INTRODUCTION

Administrative data exist within any healthcare facility. They are traditionally used to send claims to payors for reimbursement and follow the universal billing form standard. They have also been used for state public health reporting (eg, reporting of child abuse, communicable diseases, or falls to the state health department).

The International Classification of Disease (ICD-9-CM) nomenclature is used to classify hospital encounters into diagnostic categories for diagnoses and reimbursement. For injuries, the "nature of injury" (N-code) and "external cause of injury" (E-code) chapters are of particular importance in documenting both cause and outcome of injury. A diagnostic string containing an E-code within E870–E879 or E930–E949 represents injuries that are specifically related to adverse effects of medical care. Both of the code sets and their attendant validity have been explored for their utility in patient safety surveillance

Published online in Wiley Online Library (wileyonlinelibrary.com) • DOI: 10.1002/jhrm.21107

Table 1: Prior Research Analyzing Administrative Coding to Identify Error-Associated Injury

Date	Authors	Relevant Findings	Validation	Scope
1997	O'Hara and Carson[a]	Analysis of E-codes found 5% of all discharges had adverse events; 4.4% of discharges had misadventures and complications of medical or surgical care; 1% had adverse drug events	E-codes had 65% sensitivity compared to chart review	Victoria (Australia)
2001	Utah Dept of Health[b]	Analysis of administrative coding identified 0.42% of discharges statewide from a five-year period as having error-associated injuries; 6.5% of all discharges had misadventures and complications of medical or surgical care; 2.5% of hospitalizations had adverse drug events	N/A	Utah (USA)
2005	Van Tuinen et al.[c]	N/A	Analysis of administrative coding had a 37% Positive Predictive Value (PPV) for identifying surgery-related adverse events	Missouri (USA)
2005	Layde et al.[d]	Analysis of administrative coding found a medical injury rate of 133.3 per 1000 hospitalizations, of which 56.2 per 1000 were drug events, and 59.5 per 1000 were procedure-related	Administrative data showed 59.9% sensitivity and 97.4% specificity compared to chart review	Wisconsin (USA)

[a]O'Hara DA, Carson NJ. Reporting of adverse events in hospitals in Victoria, 1994–1995. *Med J Aust.* 1997;166(9):460–463
[b]Utah Health Data. *Adverse Events Related to Medical Care, Utah 1995–1999.* Salt Lake City: Utah Department of Health; 2001
[c]Van Tuinen M, Elder S, Link C, et al. *Surveillance of Surgery-Related Adverse Events in Missouri Using ICD-9-CM Codes.* Rockville, MD: Agency for Healthcare Research and Quality; 2005
[d]Layde PM, Meurer LN, Guse C, et al. *Medical Injury Identification Using Hospital Discharge Data.* Rockville, MD: Agency for Healthcare Quality and Research; 2005

(**Table 1**). These findings suggest that E-codes in administrative databases can be used to identify patient safety events with adequate sensitivity and specificity.

The Agency for Healthcare Research and Quality (AHRQ) has demonstrated the usefulness of administrative data for understanding adverse events especially as a vehicle for testing patient safety indicators (PSI).[1] AHRQ's PSI algorithm utilizes variables in hospital administrative data such as primary and secondary diagnoses, procedures, age, gender, admission source, and discharge status to reflect preventable injury status more precisely. The principal difference between the PSIs and traditional injury surveillance metrics (N-codes and E-codes) is that PSIs attempt to tailor numerators and denominators solely to events that are deemed *preventable*, whereas E-codes and N-codes include events for which preventability is difficult to ascertain, or for which prevention strategies are unknown. The AHRQ PSIs have high specificity, moderate sensitivity and moderate predictive positive value, though these characteristics are variable among the PSIs.[2,3]

Administrative data have limitations. Their sensitivity in detecting events when they happen is less than that of chart review, which is considered the gold standard in case finding. Nonetheless, given the simplicity, feasibility, and low expense of using administrative data, the fact that it does not capture all adverse events may not necessarily preclude risk managers from using administrative data as a springboard to problem identification. By using the codes presented herein, risk managers could order the frequency of adverse events that happen in their facilities in descending order. They would then have to explore more deeply which ones are most important for their particular quality goals. Events that happen most frequently (ie, medication

Table 2: Strengths and Limitations of Administrative Data

Strengths	Limitations
Readily available	May be biased by financial incentives ("upcoding", "downcoding")
Inexpensive	Long wait for data
Include large populations and wide geographic	Inconsistent accuracy and completeness
Relatively complete capture of patient contact in hospitals	Lack of clinical detail and relevant information such as "Do Not Resuscitate" orders
Standardized for national and international comparisons	Variability in coding practices across institutions
	Inability to distinguish between adverse events in which no error occurred and true medical errors

Sources: Zhan C, Miller MR. Excess length of stay, charges, and mortality attributable to medical injuries during hospitalization. *JAMA.* 2003;290(14):1868–1874; De Coster C, Quan H, Finlayson A, et al. Identifying priorities in methodological research using ICD-9-CM and ICD-10 administrative data: Report from an international consortium. *BMC Health Serv Res.* 2006;6(1):77.

errors) may not be perceived as immediately actionable when compared to events that happen less frequently but have significant liability repercussions to the facility (ie, wrong-site surgeries). Administrative data do not describe the severity of the adverse event or the long-term outcome—they simply start the case-finding process in a rigorous and reproducible way.

Though they are imperfect, administrative data allow population-scale research because they provide an easily accessible tool for comparisons of rates and trends.[4] Key strengths and limitations of administrative data are summarized in **Table 2**.

STUDY PURPOSE

The purpose of this study was to describe the incidence of patient safety events in Pennsylvania and to identify demographic characteristics associated with increased risk of injury. The ICD-9-CM E-codes and the AHRQ PSIs were used as the case-finding tools. These tools were applied to the 2006 statewide inpatient administrative discharge dataset collected by the Pennsylvania Health Care Cost Containment Council (PHC4). The PHC4 was established in 1986 as an independent state agency responsible for collecting and reporting on hospital utilization with a particular emphasis on analyzing the cost and quality of health care. The PHC4 collects over 4.5 million inpatient hospital discharge and ambulatory/outpatient procedure records each year, and has received critical acclaim for its annual report cards on cardiac surgery and hospital acquired infections.[5]

This analysis adds to prior research identifying healthcare-associated injuries in administrative data by estimating the magnitude and documenting the distribution of events in a large East Coast state.

METHODS

The analyses use Pennsylvania inpatient discharge data purchased from PHC4 for 2006. The data consist of records from all facilities in the state that perform inpatient services, excluding state psychiatric hospitals. For each discharge, a principal diagnosis and up to 10 associated ICD-9-CM diagnostic codes are assigned. In the case of an injury, supplemental E-codes are recorded. The ICD-9-CM E-codes E870–E879 and E930–E949 are used to signify a patient safety event, including adverse events of medical care and adverse events of drugs, respectively.

Twenty categories of PSI are commonly used in patient safety research (http://www.qualityindicators.ahrq.gov/Downloads/Software/SAS/V21R2A/psi_guide_rev2.pdf). The algorithm software developed to identify the PSI in the SAS environment is available free of charge through AHRQ's Web site.[6] Version 4.0 of the software was used on the PHC4 data.

For each patient record, we identified whether a PSE occurred. Our primary definition of a PSE combines the E-code with the PSI algorithm developed by AHRQ.[7] Additionally, in secondary analyses, we evaluated PSEs separately by case finding method.

STATISTICAL ANALYSES

First, we examined the rates for PSEs and described them graphically by age. We conducted this analysis using two denominators. In the primary analysis, the hospital discharges were the denominator. In the secondary analysis, the 2006 census population in Pennsylvania was the denominator. The primary analysis provides the event rate, which is a reflection of risk for hospitalized patients.

The secondary analysis provides a population measure of patient safety and is directly comparable to other injury causes that use a population-based denominator (eg, motor vehicle crashes, fires, etc). Second, we examined the relationship between patient safety errors and age, sex, gender, length of stay and total charges, compared with patients not experiencing a PSE event. Finally, we conducted multivariate logistic regression analyses to examine the associations between PSE events and various patient-level and hospital-level characteristics.

To estimate the association between patient-level and hospital-level characteristics and patient safety events we created an indicator variable called "PSE" that was set to 1 if the patient experienced a PSE event, and 0 if they did not. Traditional linear regression allows one to examine the relationship between a quantitative variable (variables that intuitively measure as a number, and that have ordinal, interval, and ratio properties; examples include age, weight, and distance) and its potential correlates. Logistic regression allows one to examine the relationship between possible correlates and a variable, like PSE, whose values are 0 and 1. This is accomplished via the use of a "link" function, a formula that translates quantitative values into probability values that range from 0 to 1. A logistic regression was executed using the statistical software package Stata and its GLM command, utilizing the "logit" link function, and controlling for hospital-level correlations of PSEs. While PSEs in different hospitals may be assumed to be independent, PSEs within the same hospital may not. The logistic regression model as described would control for such potential intrahospital correlations.

While PSEs in different hospitals may be assumed to be independent, PSEs within the same hospital may not.

Demographic variables of interest include age (categorized into 10-year age groupings), race, and sex. Although significant differences are detected in preliminary analysis between PSEs and non-PSEs groups in patient length of stay (LOS), and in total charges billed for the hospital stay, these differences may be consequences rather than causes of a PSE. In that sense, using them as covariates in a multivariate model would give misleading interpretations on their coefficients and may also distort the effects of other predictors. The final model included age, race, sex, and a clustering variable for each hospital.

Regression coefficients of interest were tested using the Wald test with a significance level of 0.05. However, since the sample size is very large, tests are extremely powerful. Therefore, we also take into consideration the effect size (ie, whether it is of practical meaning) when evaluating the difference between encounters with and without PSE.

RESULTS

Our analysis used data on 1 982 986 inpatient hospital discharges in the Pennsylvania Health Care Cost Containment Council (PHC4) database from 2006. Approximately 9% of all Pennsylvania inpatient discharges had a PSE. Patients with a PSE were on average older, male, and white. The average LOS for a PSE was 3 days longer and $35 000 more expensive than a non-PSE encounter (**Table 3**). These analyses are easily repeatable in any healthcare facility using the methods we described.

Table 4 shows the logistic regression model (allowing for clustering at the hospital level) on patient safety events. There was a significant difference in the risk of patient safety events by age and race. The risk of patient safety events increased with age, but the trend flattens at age 60, where both age groups 60–69 and 70–79 seem to have similar odds of PSEs. The risk of PSE decreased slightly after age 80 years.

Hispanics had a lower risk of patient safety events compared to White, non-Hispanic adults (OR, 0.69; 95% CI, 0.52–0.92). Asian/Pacific Islanders had a higher risk of patient safety events relative to White/non-Hispanics (OR, 1.15; 95% CI, 1.02–1.30).

When we look at the trend among age groups and the rate of PSE as a percentage of all hospital discharges, we notice intriguing patterns as the codes are broken out into more detailed components. For example, when separating out the events detected by external cause of injury codes for "Misadventures to patients during surgical and medical care" (E870–E879) compared to "Drugs, medicinal and biological substances causing adverse effects in therapeutic use" (E930–E949), we observe similar patterns of increasing risk with age (**Figure 1**). However, at the age range of 70–79, we find a steep drop in the rate of adverse medical care events without a corresponding decrease in drug error/adverse drug events (**Figure 1a**). In contrast, PSI did not vary with age. Moreover, when we look at these same error rates using statewide age-specific census population as the denominator, the risk of drug error/adverse drug events sharply increases in this age group and those older (**Figure 1b**).

DISCUSSION

Nine percent of Pennsylvania inpatient discharges in 2006 had a patient safety event. This is similar to seminal studies in New York, Colorado, and Utah that found adverse event rates of 2.9 and 3.7 percent, respectively.[8,9] Compared to white patients, Asian/Pacific Islanders had

Table 3. Characteristics of Inpatient Discharges with and without a Patient Safety Event

Variable		Patient Safety Events[a] (n = 177683)	Non-Patient Safety Events (n = 1805303)	P-value[b]
Age (years)	n	177682	1805234	
	mean	61.9	51.4	<0.01
	stddev	20.5	27.3	
Sex	male	80674 (45.4%)	776246 (43.0%)	<0.01
	female	97008 (54.6%)	1028988 (57.0%)	
Race	white	145525 (81.9%)	1404915 (77.8%)	<0.01
	black	20564 (11.6%)	238451 (13.2%)	
	hispanic	3096 (1.7%)	64195 (3.6%)	
	asian	1090 (0.6%)	12918 (0.7%)	
	native American	95 (0.1%)	2429 (0.1%)	
	other	7312 (4.1%)	82326 (4.6%)	
Length of stay (days)	n	177682	1805231	
	mean	8.1	4.9	<0.01
	stddev	11.2	7.5	
Charges ($)	n	177683	1805303	
	mean	66086.9	30451.5	<0.01
	stddev	136089.3	64096.9	

Note: stddev = standard deviation of the mean.
[a]Patients who were flagged with an AHRQ Patient Safety Indicator, or had an E-code ranging from E870–E879, or had an E-code ranging from E930–E949, or some combination of the three
[b]t-test for continuous variables and Chi-square test for categorical variables

15% higher odds of a PSE and Hispanics had 31% lower odds. This is consistent with prior research suggesting that patients considered to be non-white are less likely to be recommended for high technology treatment, and are therefore more likely to receive standard, lower risk surgical procedures, resulting in lower rates of patient safety events for events like complications of anesthesia and iatrogenic pneumothorax.[10,11]

Risk for an event was greater with increasing age. The sharp drop in the discharge incidence of medical care events after age 70 is a curious finding (**Figure 1**), particularly in light of the fact that nationwide, people age 65 and over have the highest rate of operations, at 4358 per 10 000.[12] When expressed as a population rate, this reduction in medical care adverse events is coupled with a corresponding increase in drug error/adverse drug events for this age group (**Figure 1b**). The divergence in these trends is a subject for future investigation.

The observed greater length of stay and total charges comparing PSE encounters to those without PSE is consistent with prior research finding that complex care, urgent care, and a prolonged hospital stay are associated with more errors.[13] A retrospective medical review of data from 2 acute hospitals in the Greater London area found that each adverse event led to an average of 8.5 additional days in the hospital and additional direct costs of over $400 000.[14] These findings support that reduction of adverse events could reduce healthcare costs significantly, although as noted previously, the LOS and cost may be a cause rather than a consequence of PSEs.

AHRQ PSIs have a substantially lower rate of PSE detection than E-codes. This was expected based on the focus in the development of the PSIs on *preventable* injury, whereas E-codes capture both preventable events and those currently not considered preventable (eg, adverse drug events). However, this analysis suggests that it is prudent to track all PSEs, regardless of preventability, so that if novel prevention methods arise in the future, their impact can be measured through data that have been collected both prior to and after implementing these interventions.[15] E-codes and PSIs should be used together to comprehensively illuminate the spectrum of potential PSEs in any healthcare facility. In some cases, there may be overlap between the events detected by both case-finding methods, but we have shown in previous work that

Table 4: PSE Logistic Regression Model, Clustering for Hospital

Parameter		Odds Ratio	95% Confidence Limits		P-value
Age	0–9*	1.00	1.00	1.00	–
	10–19	2.60	1.82	3.72	<0.01
	20–29	2.41	1.24	4.66	<0.01
	30–39	3.13	1.57	6.23	<0.01
	40–49	4.42	2.20	8.86	<0.01
	50–59	5.83	2.91	11.69	<0.01
	60–69	6.91	3.45	13.84	<0.01
	70–79	6.92	3.46	13.83	<0.01
	80 +	5.62	2.81	11.23	<0.01
Race	Other	1.05	0.87	1.27	0.61
	Native American	0.48	0.17	1.38	0.18
	Asian/Pacific Island	1.15	1.02	1.30	0.03
	Hispanic	0.69	0.52	0.92	0.01
	Black	0.97	0.85	1.08	0.49
	White*	1.00	1.00	1.00	–
Gender	Female	0.94	0.91	0.93	<0.01
	Male*	1.00	1.00	1.00	–

*Indicates the reference group for the interpretation of that parameter's Odds Ratios. For example, the odds of a patient 40–49 years old having a PSE is 5.83 times the odds of a patient 0–9 years old. Similarly, the odds of a female patient having a PSE is 0.94 times the odds of a male patient.

there is very little overlap between them, so they are best used in an additive capacity.[15]

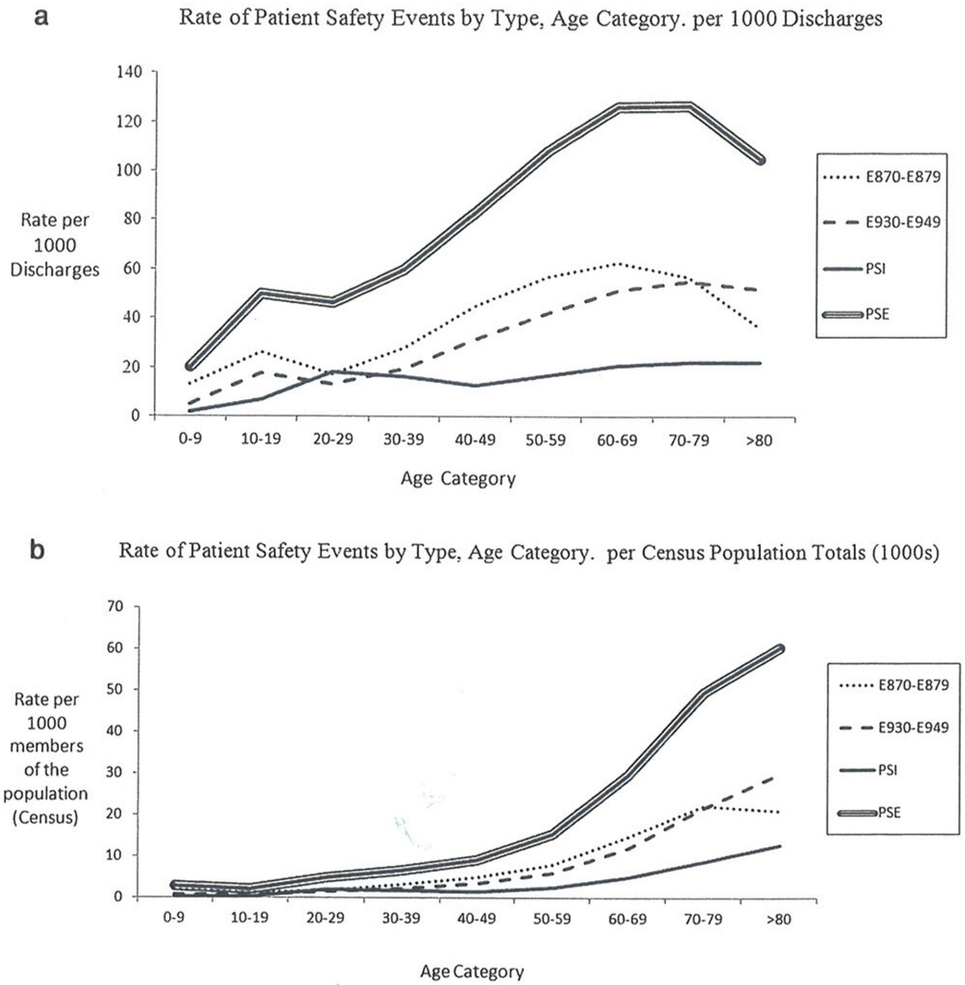
We have shown in this article methods to detect PSEs from easily accessible administrative data in Pennsylvania hospitals. The methods that we described can be used in any healthcare facility, in any locality or state, and—because of the use of the ICD—in any country using the ICD. We encourage risk managers to reproduce our methods in their facilities and to gather administrative data from their geographic areas to make comparisons similar to ours. More important, it would be interesting to hear from organizations using these methods as to how they determined which events they made actionable in their quality improvement programs. We have provided the road map; risk managers will provide the solution.

NEXT STEPS AND POLICY IMPLICATIONS

The Medical Care Availability and Reduction of Error Act became Pennsylvania law in 2002 and created the Patient Safety Authority (PSA). The PSA made Pennsylvania the first state to mandate the collection of confidential, anonymous reports of patient safety events through the Pennsylvania Patient Safety Reporting System (PA-PSRS). Since 2009, the system has received over 1.5 million reports,[16] which are estimated to comprise 30% of reportable incidents.[5] While data from PA-PSRS are crucial for understanding the burden of PSEs in Pennsylvania, an accurate numerator and denominator are necessary to determine the risk of patient safety events. PA-PSRS collects "serious events" (patient harm) and "incidents" (no patient harm, but potentially harmful).[16] Both event types represent what would comprise the numerator in a risk ratio. However, the system was not designed to collect data on the population at risk, which would comprise the dominator part of the risk ratio. Additionally, the data from PA-PSRS only include age and gender information, excluding important demographic elements like race and patient ZIP Code (a proxy for socioeconomic status). As a result, the data cannot be used for epidemiologic

Figure 1:

Patient Safety Events by Type, Age Category, as Rate of Total Discharges and Census Population (1000s)



E870–E879: *Misadventure during or due to medical care;* **E930–E949:** *Adverse effects of therapeutic drugs, biological;* **PSI:** *AHRQ Patient Safety Indicator;* **PSE:** *The patient experienced an E870–E879 OR an E930–E949 OR a PSI*

analyses such as those presented in this study. Further, while serious events resulting in harm *are* reportable to PA-PSRS, 96% of the reports describe incidents in which harm did not reach the patient, making it an excellent source of near-miss events but not of patient injuries.[16] Therefore, the combined analysis of PA-PSRS with injuries found in administrative data (using the methods described herein) will allow a more complete understanding of Pennsylvania's patient safety burden and potential directions for improvement. Future policy initiatives should seek to integrate these systems. Other states using similar adverse event reporting systems could do likewise.

CONCLUSION

E-codes and PSIs are a useful way to deepen the surveillance of patient safety events in Pennsylvania, and indeed in any of the 50 states. Administrative data are appealing because they are routinely collected, easily acquired, and are also used in other countries—creating the possibility of national and international comparisons. Despite the limitations of administrative data, they are enormously useful as a starting point for hospital-based patient safety surveillance. Our results from Pennsylvania confirm prior research findings that patients experiencing a PSE were older and a patient safety error was associated with longer

 DOI: 10.1002/jhrm

hospital stay and greater cost. Improved and consistently administered surveillance is needed to develop and evaluate hospital-based programs to reduce patient safety errors. The process developed and explained in this article provides a useful toolkit for all states and healthcare facilities seeking to enhance surveillance systems toward improved care, outcomes, and reduced costs.

REFERENCES

1. Romano PS, Geppert JJ, Davies S, Miller MR, Elixhauser A, McDonald KM. A national profile of patient safety in US hospitals. [Research Support, US Gov't, P.H.S.]. *Health Aff (Millwood).* 2003;22(2):154–166.
2. Gallagher B, Cen L, Hannan EL, et al. Validation of AHRQ's patient safety indicator for accidental puncture or laceration. *Adv Patient Saf.* 2005;2:27–38.
3. Romano PS. Lessons from the AHRQ PSI validation pilot project. Slide presentation from the AHRQ 2008 Annual Conference (Text Version). http://www.ahrq.gov/about/annualmtg08/091008slides/Romano.htm. Published 2009. Accessed May 2012.
4. Sedman A, Harris JM 2nd, Schulz K, et al. Relevance of the Agency for Healthcare Research and Quality patient safety indicators for children's hospitals. *Pediatrics.* 2005;115(1):135–145.
5. Pennsylvania Health Care Cost Containment. http://www.phc4.org. Published 2010. Accessed May 2012.
6. Agency for Healthcare Research and Quality. PSI SAS software. http://www.qualityindicators.ahrq.gov/software/SAS.aspx. Published 2012. Accessed May 2012.
7. Agency for Healthcare Research and Quality. Patient safety indicators. http://www.qualityindicators.ahrq.gov/modules/psi_overview.aspx. Published 2010. Accessed May 2012.
8. Brennan TA, Leape LL. Adverse events, negligence in hospitalized patients: Results from the Harvard Medical Practice Study. *Perspect Healthc Risk Manage.* 1991;11(2):2–8.
9. Gawande AA, Thomas EJ, Zinner MJ, Brennan TA. The incidence and nature of surgical adverse events in Colorado and Utah in 1992. *Surgery.* 1999;126(1):66–75.
10. Schulman, KA, Berlin JA, Harless W, et al. The effect of race and sex on physicians' recommendations for cardiac catheterization. *N Engl J Med.* 1999;340(8):618–626.
11. Coffey RM., Andrews RM, Moy E. Racial, ethnic, and socioeconomic disparities in estimates of AHRQ patient safety indicators. *Med Care.* 2005;43(3 Suppl):I48–I57.
12. DeFrances CJ, Lucas CA, Buie VC, Golosinskiy A. 2006 national hospital discharge survey. *Natl Health Stat Report.* 2008;5:1–20.
13. Weingart SN, Wilson RM, Gibberd RW, Harrison B. Epidemiology of medical error. *West J Med.* 2000;172(6):390–393.
14. Vincent C, Neale G, Woloshynowych M. Adverse events in British hospitals: Preliminary retrospective record review. *BMJ.* 2001;322(7285):517–519.
15. Taylor JA, Gerwin D, Morlock L, Miller MR. Triangulating case-finding tools for patient safety surveillance: A cross-sectional case study of puncture/laceration. *Inj Prev.* 2011;17(6):388–393.
16. Patient Safety Authority. Patient safety. http://patientsafetyauthority.org/PatientSafetyAuthority/Documents/FINAL%202011%20Annual%20Report.pdf. Published 2011. Accessed December 2012.

ABOUT THE AUTHORS

Jennifer A. Taylor, PhD, MPH, *is an assistant professor at the Drexel University School of Public Health, Department of Environmental and Occupational Health.* ***Ravi S. Pandian, MA,*** *is a research analyst and programmer at the Drexel University School of Public Health, Department of Environmental & Occupational Health.* ***Lu Mao, MSc,*** *is a doctoral candidate in the Department of Biostatistics, University of North Carolina at Chapel Hill.* ***Yvonne L. Michael, ScD, SM,*** *is an associate professor at the Drexel University School of Public Health, Department of Epidemiology and Biostatistics.*