Analysis of Ethnic Disparities in Workers' Compensation Claims Using Data Linkage

Lee S. Friedman, PhD, Peter Ruestow, MSc, and Linda Forst, MD, MPH

Objective: The overall goal of this research project was to assess ethnic disparities in monetary compensation among construction workers injured on the job through the linkage of medical records and workers' compensation data. **Methods:** Probabilistic linkage of medical records with workers' compensation claim data. **Results:** In the final multivariable robust regression model, compensation was $5824 higher ($P = 0.030$; 95% confidence interval: 551 to 11,097) for white non-Hispanic workers than for other ethnic groups when controlling for injury severity, affected body region, type of injury, average weekly wage, weeks of temporary total disability, percent permanent partial disability, death, or attorney use. **Conclusions:** The analysis indicates that white non-Hispanic construction workers are awarded higher monetary settlements despite the observation that for specific injuries the mean temporary total disability and permanent partial disability were equivalent to or lower than those in Hispanic and black construction workers.

The construction industry has continuously been one of the most hazardous industries in the United States. Each year several hundred thousand construction workers become ill or are injured as a result of on-the-job hazards. The estimated rates for injuries, illnesses, and fatalities among construction workers are consistently among the highest of any economic sector.[1] The industry is also one of largest sectors in the United States, using approximately 5% of the workforce.[2] This workforce is made up of many low-skilled, low-educated, and immigrant populations that are vulnerable to workplace injury.[3–7]

According to the Bureau of Labor Statistics data, Hispanics make up 29.1% of the construction industry workforce, even though they represent only 14.6% of the total US workforce.[8] Furthermore, Hispanic workers comprise an even larger proportion within specific occupations—many of which are high risk-–including general construction laborers (43.1% of total workforce), roofers, (46.4%), and cement masons (51.5%).[8] Although Hispanics and immigrants are often used in more hazardous occupations involving unskilled labor than native-born non-Hispanics,[9–13] there have been relatively few studies in the peer-reviewed literature investigating ethnic disparities in occupational injuries that include all industry and injury types. The few studies that have been conducted have reported conflicting results. Some studies indicate that work-related injury and lost work time are higher among African Americans and Hispanics,[12,14–16] whereas others have shown no significant differences,[17–19] or that occupational injury is more common among white non-Hispanics.[20] The studies that have been conducted fail to show a consistent picture primarily because of different sampling methods used. None of the studies have specifically investigated differences in severity of injury and many fail to differentiate between non-Hispanics and Hispanics.

From the School of Public Health (Drs Friedman and Forst and Mr Ruestow), Division of Environmental and Occupational Health Sciences, University of Illinois, Chicago; and The Social Policy Research Institute (Dr Friedman), Skokie, IL.

Disclosure: None of the authors have any conflicts of interest that may be relevant to the submitted work.

Address correspondence to: Lee S. Friedman, PhD, The Social Policy Research Institute, 4001 Emerson, Skokie, IL 60076 (lfriedman@tspri.org).

DOI: 10.1097/JOM.0b013e31825a34d1

State-based data repositories can be used to fill in the gaps left by Federal surveillance programs. Currently, the Bureau of Labor Statistics' data are the primary source of occupational surveillance data for Illinois and the United States. Nevertheless, there are several alternative surveillance resources that can be used to help better describe construction injuries in the state. These include the Illinois Workers' Compensation Commission (IWCC) claims data, the Illinois Trauma Registry (ITR), and the Illinois Hospital Discharge database. These databases have been underutilized for occupational surveillance.

One important endpoint in occupational injury surveillance is the level of financial compensation workers receive from workers' compensation after an injury claim. Workers' compensation data provide information about the direct costs paid for claims that are not based on estimates. The workers' compensation system was designed primarily to protect employers from excessive damage awards and to provide a more equitable, reliable, and expeditious system of compensation for injured workers. Workers' compensation data are useful for occupational surveillance because most workers' compensation data sets provide information about the employee, employer, level of impairment after an injury or illness, and the direct costs associated with an injury or illness. Studies evaluating workers' compensation data have reported that industry,[21,22] occupation,[21,23–25] legal counsel,[26] union membership,[24] and health care costs[27] are associated with claim costs. None have looked at ethnic disparities in compensation controlling for wage, injury type, and severity of injury.[7,28,29]

The overall goal of this research project was to assess ethnic disparities in monetary compensation—which comprises medical costs, lost wages, dependent benefits, survivor benefits, settlement payments, attorney fees, and other miscellaneous costs—among construction workers injured on the job through the linkage of medical records and workers' compensation data. This approach provides important and previously unavailable information by combining data on in-hospital treatment (ITR), long-term disability (IWCC), and level of compensation (IWCC) for individuals.

METHODS

Data Sources

We have a data-sharing agreement with the Illinois Department of Public Health and the IWCC. The institutional review board at the University of Illinois, Chicago, has approved this work (no. 2008-0060), as have the research committees at the Illinois Department of Public Health and IWCC.

Illinois Trauma Registry

The ITR is mandated by the state legislature and is managed by the Illinois Department of Public Health. All of the state's level 1 and 2 trauma centers ($N = 62$) are required to report all patients (1) who have sustained traumatic injuries (*International Classification of Diseases, 9th Revision–Clinical Modification* [ICD-9-CM] external injury codes E800-995) and admitted to a trauma center for more than 12 hours, (2) have been transferred to a level 1 or 2 center, or (3) are dead on arrival or die in the emergency department. We provided an assessment of data quality of the ITR in a previous article.[30] The ITR contains data on demographics (age, gender, and

race/ethnicity), exposure (mechanism of injury), health outcomes (diagnoses, measures of injury severity, hospital procedures, and disability status on discharge), and economics (payer source).

IWCC Claims Database

The IWCC operates the administrative court system for workers' compensation cases in Illinois. There are approximately 70,000 claims filed with IWCC for financial compensation each year. Unlike single-carrier states with a well-organized and centralized reporting system (eg, Washington) in Illinois, the IWCC handles only claims in which the employee and the employer are unable to resolve compensation issues for an injury without administrative intervention. Any aspect paid for before initiating a claim through the IWCC that is not disputed by either party is not litigated through IWCC or reported in the data set. An arbitrator initially hears a workers' compensation claim. The arbitrator's decision can subsequently be appealed before a panel of three commissioners.

We obtained a data set of all claims from the IWCC, which included information on both active and closed claims. The data set contains an array of information including employer information (name of company, worksite location), employee demographics (age, gender, marital status, the number of dependents), cause and type of injury or illness, level of temporary and permanent disability, and details of the compensation costs associated with the injury. For this study, we used a data set created in our previous work.[31] The data set contains records of 19,734 workers' compensation claims made by injured construction workers between 2000 and 2005. Compensation costs were not adjusted for inflation. The minimum age in this study group was 16 years because Illinois law prohibits persons younger than 16 years from working in construction. We cleaned and deduplicated the data set, as reported in previous work.[30,31]

Data Linkage

The workers' compensation data set of 19,734 claims[31] was the primary linkage data set. After exclusion of claims without valid dates of birth or dates of injury, the final data set contained 16,794 claims. We used probabilistic linkage methods to link the two data sets. We used a date of birth, date of injury, gender, residential ZIP code, nature of injury, and body part affected to link the data.

Using the linkage variables, cases were matched from the IWCC to the trauma registry. A combination of SQL and SAS code was written to carry out all steps in the linkage, using SAS version 9.1 (SAS Institute Inc, Cary, NC). All linked cases were validated by manually reviewing cause of injury and type of injury codes. If the information on cause and type of injury did not match between the data sets, the linked pair was removed. In the trauma registry, cause of injury and type of injury are coded using ICD-9-CM coding. The workers' compensation data set uses a different coding system for both cause and type of injury.

This study was part of a larger linkage project that involved the two data sets mentioned previously and a hospital discharge data set that includes patients admitted to hospitals without trauma units. Only the trauma registry contained information regarding race and ethnicity. Therefore, this data analysis was restricted to the linked data set containing the trauma registry and workers' compensation data. We expected a linkage rate of approximately 10% to 15% on the basis of this pilot study and analysis of the type of injuries contained in the workers' compensation database; most claims for compensation involve injuries and illnesses that do not result in hospitalization or acute trauma. A total of 1664 workers' compensation records linked to the two hospital registries-–a linkage rate of 9.9%. Between the two medical record data sets, 1039 cases of 16,975 claims in the IWCC data set matched to the Illinois trauma registry (6.1%) whereas 1378 matched to the hospital discharge data set (8.1%). There were 753 cases identified that matched across all three data sets.

A comparison of cases in the final linked data set with claims in the IWCC that did not match found that matched and unmatched cases were very similar in terms of the distribution by gender, average weekly wage, marital status, the number of dependents, age, and population density of employment area. Some of these demographic variables—particularly average weekly wage, marital status, and number of dependents—have a major impact on the amount of compensation received by the worker in the claims process. An approximately equal proportion of matched and unmatched cases had a final decision on their workers' compensation claim (as opposed to dismissed cases or new cases still in progress). Furthermore, the monthly trend in injuries from January 2000 to December 2005 for both groups was very similar.

As expected, unmatched and matched cases differed by injury type. Unmatched cases had a relatively higher occurrence of back, spine, and upper extremity injuries and generally were not treated as inpatients in hospitals or in trauma units. In contrast, matched cases disproportionately involved more severe injuries such as injuries involving multiple body parts (17.43% matched vs 9.29% unmatched) and they were more likely to die from their injuries (1.68% matched vs 0.46% unmatched). Matched cases also had a higher mean percent permanent partial disability (13.9% matched vs 10.4% unmatched) and were more likely to suffer injuries resulting in permanent partial disability of greater than 50% loss of function.

STATISTICAL ANALYSIS

We used SAS software for all statistical analyses (version 9.1; SAS Institute, Inc). Frequencies of injuries overall and distributions by age, gender, and race/ethnicity were determined. We also described type of injuries and disparities between groups in terms of injury severity, external cause of injury, work location, duration of hospitalization, in-hospital mortality, percent disability, and level of monetary compensation through workers' compensation.

Appropriate parametric (Pearson's chi-square) and nonparametric tests (Wilcoxon ranked sum test) were used to evaluate bivariate relationships, and the *t* test was used to compare mean differences in continuous metrics including injury severity scores (ISSs), percent disability awarded, and level of monetary compensation. The Levenne test was used to test for equal variance between groups to determine whether to use equivariant or nonequivariant statistical measures of significance.

ICD-9-CM N codes were used to assess body region and type of injury on the basis of the Barell classification matrix.[32,33] We calculated ISSs, using Abbreviated Injury Scale scores for six major body regions. The Abbreviated Injury Scale scores are calculated using a standardized computer algorithm on the basis of discharge records rather than admission status to reduce miscoding.[33] Injury severity scores of 16 or greater indicate that the worker suffered a severe injury. We report the number of hospitalization days (length of stay), days of treatment in intensive care units, days on a ventilator, and disability at discharge.

We calculated the cumulative percent permanent partial disability for this study. Permanent partial disability involves partial loss of body function at the point of maximum medical improvement. We used the statutory formula to calculate cumulative percent disability when more than one body part was injured and limited in function. An example of the statutory formula for computing cumulative disability is $A + (1 - A)*B$, where A is the percent disability for a specific injury involving a specific body part and B is the percent disability for a second specific injury involving a specific body part.

For the regression analysis, the dependent variable (total monetary compensation) was not normally distributed (kurtosis = 34.4; skewness = 4.0). In scenarios with extreme or many outliers causing the data to be skewed, ordinary least squares regression produces biased parameter estimates because the parameter estimates will be weighted toward the outliers, which also inflates the

variance. Nevertheless, we did not transform the dependent variable because back transformation of log transformed data leads to biased estimates.[34,35] Although the log transformation makes the data less skewed, it changes the relationship between the dependent and independent variables.[34,35]

Therefore, for the multivariable regression analysis, we used robust M-estimation as implemented in SAS version 9 (PROC ROBUSTREG; SAS Institute, Inc). The parameter estimates derived from robust regression are less influenced by outliers. This is generally achieved by weighting observations whose residuals are large and do not remove them. The outcome variable of interest was total financial compensation of decided claims, excluding claims in progress and dismissed claims. We used a manual stepwise selection method to identify the best model fit for the predictors. Akaike (AIC) criterion and Schwarz information criteria (BIC) were also used for model selection and to identify the best weighting function. In the final model, gender ($P = 0.44$), age at the time of accident ($P = 0.45$), the number of dependents ($P = 0.76$), and interval from the day of accident to the day of filing ($P = 0.58$) were highly insignificant and, therefore, were excluded. Because of the small sample size of the final linked data set, we did not have the statistical power to look at individual ethnic groups. Therefore, in the final multivariable model only, we compared white non-Hispanic construction workers with all the other ethnic groups combined (white non-Hispanic vs all other ethnic/racial groups). The final model included the following covariates (data format, data source): employee's weekly wage (continuous; IWCC), ISS (continuous; ITR), days of initial hospitalization (continuous; ITR), discharged from the hospital to an acute/intermediate care facility or rehabilitation center (dichotomous; ITR), attorney representation (dichotomous; IWCC), total weeks of total temporary disability (continuous; IWCC), and percent permanent partial disability (continuous; IWCC).

To validate our primary multivariable robust regression model, we also ran a linear regression model on a truncated data set in which outliers in monetary compensation were removed. On the basis of an analysis of the distribution of total financial compensation overall and within specific ethnic subgroups, we removed all cases with total financial compensation exceeding $225,000 (greater than 95th centile). The removal of the outliers made the dependent variable near normally distributed. We ran the same multivariable model used in the primary analysis.

RESULTS

In the final linked data set, the ethnic distribution among traumatically injured construction workers was as follows: black, $n = 68$; Hispanic, $n = 168$; white, $n = 724$; and other ethnicities, $n = 79$ (American Indian, Pacific Islander, Asian, other). Table 1 presents the demographic characteristics of the workers included in the final linked data set. Nearly all the injured workers in the data set were men (98.5%). Hispanics were also younger than white non-Hispanics ($P < 0.05$) but not blacks. The average weekly wage for white non-Hispanics was significantly higher than the weekly wage reported for Hispanics ($P < 0.05$; Tukey's comparison of means) but not for the other groups. The proportion of construction workers who used attorney representation for their claim process instead of advocating on one's own behalf was as follows: black, 63 (92.6%); Hispanic, 146 (86.9%); white non-Hispanic, 604 (83.4%); and other ethnicities, 73 (92.4%).

Cause of Injury

Black construction workers were disproportionately more likely than other ethnic groups to suffer injuries as a result of assaults and motor vehicle crashes (Table 1), although the cause of injury between Hispanic and white non-Hispanic workers was very similar, with the majority of injuries caused by falls (Table 1).

Type of Injury

The most common injury types across all workers were fractures, internal injuries, and open wounds. Black workers disproportionately suffered internal injuries and open wounds to the head, face, and neck compared with other groups (Table 2). Hispanics were more likely to suffer amputations. Overall the distribution of injuries by body part was nearly identical between white non-Hispanic and Hispanic construction workers.

Measures of Injury Severity

Injured black construction workers stayed in the hospital slightly longer than other workers, and they were more likely to be sent to an intensive care unit and be put on a ventilator, but none of these differences were statistically significant (Table 3). Mean ISSs did not significantly differ between the groups. In addition, both black and white non-Hispanic workers were slightly more likely to be discharged to an intermediate care or rehabilitation facility than Hispanic workers.

Workers' Compensation Claims

The proportion of workers whose workers' compensation claims had been decided or settled by the court system was nearly identical between the various ethnic groups: black, $n = 52$ (76.5%); Hispanic, $n = 132$ (78.6%); white non-Hispanic, $n = 578$ (79.8%); and other, $n = 58$ (73.4%). The median total monetary compensation awarded to the workers differed substantially by ethnicity ($P = 0.01$) (Table 4). Despite a difference in median monetary compensation, total percent permanent partial disability and weeks of total temporary disability did not significantly differ between the ethnic groups (Table 4). The difference in monetary compensation persisted when the data were stratified by injury type. Table 5 presents the median total monetary compensation and mean percent permanent partial disability awarded for the most prevalent specific injuries among the workers in the linked data set. White non-Hispanic construction workers consistently were awarded higher monetary settlements despite the fact that the mean percent permanent partial disability was equivalent to or lower than that in black and Hispanic construction workers, and this was true for amputations, torso injuries, open wounds of the upper extremity, and traumatic brain injury.

In the final multivariable robust regression model, compensation was $5824 higher ($P = 0.030$; 95% confidence interval: 551 to 11,097) for white non-Hispanic workers when controlling for other covariates than for minority workers ($R^2 = 0.25$). In the second validation model in which the extreme payment outliers were removed, compensation was $6844 higher ($P = 0.027$; 95% confidence interval: 781 to 12,907) for white non-Hispanic workers when controlling for other covariates than for minority workers ($R^2 = 0.32$).

DISCUSSION

The analysis indicates that white non-Hispanic construction workers are awarded higher monetary settlements despite the observation that within specific injuries the measures of injury severity were equivalent to or lower than those in Hispanic and black construction workers. Further investigation of these ethnic disparities is necessary. Because the data were restricted to construction workers who were admitted to hospitals following an injury, which generally consist of the most severe injuries, it is possible that the ethnic disparities in compensation are specific to traumatic injuries suffered by these construction workers alone. In a broader data set involving workers of all industries suffering both acute injuries, chronic injuries, and illnesses, the disparities observed in this analysis may not persist. Nevertheless, other studies have shown that African American workers suffering low back injuries received significantly lower settlement awards ($3000 less on average)[36] and experience greater economic struggle after workers' compensation settlements.[29,37]

TABLE 1. Demographic Characteristics of Injured Construction Workers' Illinois Data Linkage Study, 2000–2005

	Black ($n = 68$)	Hispanic ($n = 168$)	Other ($n = 79$)	White ($n = 724$)
Gender				
Male	66 (97.1%)	167 (99.4%)	79 (100.0%)	711 (98.2%)
Female	2 (2.9%)	1 (0.6%)	0 (0.0%)	13 (1.8%)
Mean age (SD), yr	38.7 (9.7)	35.1 (9.6)	35 (10.3)	39.3 (11.1)
16–24	4 (5.9%)	24 (14.3%)	13 (16.5%)	80 (11.0%)
25–34	25 (36.8%)	68 (40.5%)	36 (45.6%)	186 (25.7%)
35–44	22 (32.4%)	48 (28.6%)	12 (15.2%)	240 (33.1%)
45–54	16 (23.5%)	23 (13.7%)	14 (17.7%)	148 (20.4%)
55–64	0 (0.0%)	3 (1.8%)	3 (3.8%)	62 (8.6%)
65 and older	1 (1.5%)	1 (0.6%)	0 (0.0%)	6 (0.8%)
Marital status				
Single	39 (57.4%)	43 (25.6%)	37 (46.8%)	293 (40.5%)
Married	28 (41.2%)	124 (73.8%)	40 (50.6%)	423 (58.4%)
Widowed/divorced	0 (0.0%)	0 (0.0%)	1 (1.3%)	0 (0.0%)
Unspecified	1 (1.5%)	1 (0.6%)	1 (1.3%)	8 (1.1%)
Mean number of dependents (SD)	1.2 (1.8)	1.6 (1.6)	1.3 (1.6)	0.9 (1.2)
Average weekly wage (SD)	887 (332)	814 (357)	775 (378)	909 (376)
Cause of injury				
Assault	7 (10.3%)	3 (1.8%)	2 (2.5%)	6 (0.8%)
Motor vehicle crash	11 (16.2%)	17 (10.1%)	6 (7.6%)	77 (10.6%)
Cutting or piercing instrument	3 (4.4%)	11 (6.5%)	5 (6.3%)	44 (6.1%)
Electrocution	0 (0.0%)	1 (0.6%)	2 (2.5%)	15 (2.1%)
Falls	27 (39.7%)	89 (53.0%)	45 (57.0%)	394 (54.4%)
Machinery	4 (5.9%)	15 (8.9%)	6 (7.6%)	64 (8.8%)
Struck by or against an object	9 (13.2%)	20 (11.9%)	5 (6.3%)	68 (9.4%)
Caught between objects	1 (1.5%)	6 (3.6%)	2 (2.5%)	20 (2.8%)
Other	6 (8.8%)	6 (3.6%)	6 (7.6%)	36 (5.0%)

TABLE 2. Injury Type by Body Part Affected Among Construction Workers Filing Workers' Compensation Claims by Ethnicity

	Black ($n = 68$)	Hispanic ($n = 168$)	Other ($n = 79$)	White ($n = 724$)
Fracture				
Head, face, and neck	4 (5.9%)	17 (10.1%)	12 (15.2%)	78 (10.8%)
Back and spine	6 (8.8%)	25 (14.9%)	7 (8.9%)	131 (18.1%)
Torso	5 (7.4%)	17 (10.1%)	9 (11.4%)	102 (14.1%)
Lower extremities	17 (25.0%)	44 (26.2%)	19 (24.1%)	210 (29.0%)
Upper extremities	10 (14.7%)	33 (19.6%)	23 (29.1%)	175 (24.2%)
Open wound				
Head, face, and neck	19 (27.9%)	24 (14.3%)	15 (19.0%)	109 (15.1%)
Back and spine	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Torso	2 (2.9%)	1 (0.6%)	1 (1.3%)	14 (1.9%)
Lower extremities	2 (2.9%)	2 (1.2%)	0 (0.0%)	6 (0.8%)
Upper extremities	5 (7.4%)	21 (12.5%)	13 (16.5%)	71 (9.8%)
Amputation				
Lower extremities	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.1%)
Upper extremities	1 (1.5%)	11 (6.5%)	0 (0.0%)	24 (3.3%)
Internal injury				
Head, face, and neck	21 (30.9%)	22 (13.1%)	14 (17.7%)	93 (12.8%)
Back and spine	0 (0.0%)	2 (1.2%)	0 (0.0%)	3 (0.4%)
Torso	5 (7.4%)	13 (7.7%)	8 (10.1%)	80 (11.0%)

*Totals in each column may exceed column totals because individual workers may have suffered more than one injury.

TABLE 3. Measures of Injury Severity and Discharge Status Among Construction Workers Filing Workers' Compensation Claims by Ethnicity

	Black ($n = 68$)	Hispanic ($n = 168$)	Other ($n = 79$)	White ($n = 724$)
Hospital treatment				
Mean days in hospital (SD)	4.7 (6.8)	4.0 (7.5)	7.1 (25.5)	4.2 (6.7)
Sent to intensive care unit (ICU)	19 (27.9%)	27 (16.1%)	15 (19.0%)	132 (18.2%)
Mean days in ICU* (SD)	3.4 (5.9)	4.3 (5.6)	4.3 (9.1)	3.3 (6.6)
Required mechanical ventilation	6 (8.8%)	9 (5.4%)	3 (3.8%)	35 (4.8%)
Mean days on ventilator* (SD)	2.4 (4.6)	6.3 (6.4)	4.1 (4.6)	2.3 (6.1)
Underwent operation	27 (39.7%)	90 (53.6%)	38 (48.1%)	389 (53.7%)
Severity of injury				
Penetrating injuries ($N =$)	8 (11.8%)	16 (9.5%)	8 (10.1%)	65 (9.0%)
Mean injury severity score (SD)	8.0 (6.4)	8.1 (7.4)	7.8 (6.8)	8.3 (8.4)
ISS 16 and higher (severe injuries)	9 (13.2%)	23 (13.7%)	11 (13.9%)	96 (13.3%)
Discharge status				
Discharged home	57 (83.8%)	144 (85.7%)	62 (78.5%)	587 (81.1%)
Acute care/rehabilitation/skilled nursing	9 (13.2%)	16 (9.5%)	16 (20.3%)	113 (15.6%)
Left against medical advice	1 (1.5%)	4 (2.4%)	0 (0%)	4 (0.6%)
Morgue	1 (1.5%)	4 (2.4%)	1 (1.3%)	20 (2.8%)

*Mean days in intensive care unit and on ventilator include only patients sent to an intensive care unit or required mechanical ventilation.

TABLE 4. Measures of Injury Severity and Discharge Status Among Construction Workers Filing Workers' Compensation Claims by Ethnicity

	Black ($n = 68$)	White, Hispanic ($n = 168$)	Other ($n = 79$)	White Non-Hispanic ($n = 724$)
Total workers' compensation, US$				
Median	26,385.00	27,121.00	23,768.00	33,056.00
Mean	47,935.00	48,519.00	43,048.00	60,431.00
Mean temporary total disability, wk				
Median	12.4	12.1	8.9	12.8
Mean	21.8	26.7	28.3	29.5
Mean percent permanent partial disability, %				
Median	15.0	18.9	11.3	20.0
Mean	22.7	24.8	21.8	25.7

This study does not explain why white non-Hispanic construction workers would receive higher compensation. All that can be stated is that the observed higher compensation among white non-Hispanic construction workers is not diminished when controlling for injury severity, affected body region, type of injury, or common factors associated with higher compensation-–average weekly wage, weeks of temporary total disability, percent permanent partial disability, injuries resulting in death, or the use of an attorney. Nevertheless, the multivariable models do not explain all the variance, indicating that there are other factors that are important in determining final claim compensation. It is possible that ethnicity is acting as a proxy measure for these other determinants.

The value of a data linkage strategy rests on the supposition that more can be learned from multiple information sources together than from any one alone. By focusing on construction workers, we have demonstrated a few of the ways health information from medical records can enhance information collected during the legal process of workers' compensation. Nevertheless, to achieve the statistical power necessary to draw confident conclusions using this information, the public databases must offer ways to more efficiently link data across them. For worker's compensation data, linkage would be more successful and informative if the worker's compensation data were linked to data sets that better represented the type of injuries and illnesses that result in claims, such as outpatient records, rehabilitation and physical therapy records, or claim management services used by private and self-insurers. Because the workers' compensation data set is missing information about ethnicity and the medical records data sets do not capture employer information, only through data linkage was this analysis on ethnic disparities of workers' compensation claims possible in Illinois. Furthermore, the linkage to the medical records provided greater detail on the cause and nature of injury suffered by these construction workers.

LIMITATIONS

Selection bias represents the greatest potential limitation to this study. Only 10% of the original sample of construction worker claims linked with the medical records, although this was the linkage rate expected on the basis of our previous pilot study. The demographic distributions among linked cases were very similar to those IWCC cases that did not link to either of the other databases. That the types of injuries differed between the linked and nonlinked cases is not surprising. Not every worker who files a claim requires treatment in a trauma center or inpatient setting. But linkage with inpatient medical records allows for an analysis of serious acute

TABLE 5. Median Monetary Compensation and Average Permanent Partial Disability Among Construction Workers Filing Workers' Compensation Claims by Ethnicity

	Black ($n = 68$)	White, Hispanic ($n = 168$)	Other ($n = 79$)	White Non-Hispanic ($n = 724$)
Fracture of back and spine				
Number of cases	6 (8.8%)	25 (14.9%)	7 (8.9%)	131 (18.1%)
Median monetary compensation, US$	16,843.00	24,872.00	60,250.00	34,852.00
Average temporary total disability, wk	67.80	43.50	61.80	39.30
Average permanent partial disability, %	24.4	17.1	8.9	21.6
Open wound upper extremities				
Number of cases	5 (7.4%)	21 (12.5%)	13 (16.5%)	71 (9.8%)
Median monetary compensation, US$	30,856.87	6,407.19	6,102.59	17,235.51
Average temporary total disability, wk	7.80	29.30	27.50	26.90
Average permanent partial disability, %	25.5	37.7	26.9	29.0
Amputation upper extremities				
Number of cases	1 (1.5%)	11 (6.5%)	0 (0.0%)	24 (3.3%)
Median monetary compensation, US$	. . .	$9,571.00	. . .	$14,517.00
Average temporary total disability, wk	. . .	8.20	. . .	34.10
Average permanent partial disability, %	. . .	49.3	. . .	42.0
Internal injury torso				
Number of cases	5 (7.4%)	13 (7.7%)	8 (10.1%)	80 (11.0%)
Median monetary compensation, US$	14,498.00	20,386.63	122.12	27,034.42
Average temporary total disability, wk	36.00	61.00	74.70	38.20
Average permanent partial disability, %	28.8	26.6	10.4	21.7
Type 1 and 2 traumatic brain injuries				
Number of cases	23 (33.8%)	35 (20.8%)	20 (25.3%)	129 (17.8%)
Median monetary compensation, US$	31,324.02	28,687.59	38,359.18	45,600.88
Average temporary total disability, wk	20.60	27.70	19.10	26.40
Average permanent partial disability, %	16.0	16.5	18.0	18.8

injuries. The conclusions from this analysis cannot be generalized to injuries and illnesses that do not result in hospitalization (eg, carpal tunnel syndrome or back injuries).

The probabilistic linkage strategy results in a level of uncertainty that cannot be avoided. Nevertheless, we used the trauma registry to estimate the real mismatch probability using the linkage variables (details to be published in another article). The trauma registry, unlike the other two data sets, includes patient name, which provides us with an additional level of information for confirming matches. Across all the patients in the trauma registry between 2000 and 2009 ($N = 452{,}491$), we identified duplicates using the same linkage variables. Identifying duplicates within data sets or between data sets involves similar procedures. Based on this real-world exercise, the probability of mismatches using the main linkage variables (date of birth, date of injury, gender, residential ZIP code, nature of injury, and body part affected) is very low (0.3%, or 3 per 1000 records).

In the data set associated with trauma units, the standard for determining race/ethnicity is based on observation and record review[38] rather than on self-identification. Hispanic workers may be "black" or "white," but "black" Hispanics are frequently misclassified as African American because ethnicity is determined by observation and not self-report. This bias likely results in an undercounting of Hispanics. Sample size of the ethnic/racial subgroups was small in the final linked data set, which prevented us from statistically assessing differences between specific subgroups. Furthermore, the sample size contributed to the high level of uncertainty of the parameter estimates as indicated by the wide confidence intervals reported in our multivariable analysis.

CONCLUSION

Data linkage methodologies show promise in the field of occupational epidemiology. By linking various existing data sources, interested parties can greatly increase the amount of information available for injury cases. For instance, by linking the medical records with workers' compensation data as done in this study, one would be able to follow a subset of workers through both the health care system and the workers' compensation claims structure. The different information contained within each database might help us to better characterize risk factors and design more effective interventions. With the expansion of electronic health records, the value of data linkage will grow, but for it to succeed using workers' compensation records there needs to be an improvement in access to outpatient records and coding systems within the workers' compensation data systems that can be more easily reconciled with other medical record systems.

ACKNOWLEDGMENTS

This study was supported by a small study grant from the Center to Protect Workers' Rights through a cooperative agreement with the National Institute for Occupational Safety and Health (U60 OH009762).

REFERENCES

1. Bureau of Labor Statistics. *Workplace Injuries and Illnesses—2010.* Washington, DC: US Department of Labor; 2011. Publication No. USDL-11-1502.

2. Bureau of Labor Statistics. Table 2.1 Employment by major industry sector, 1998, 2008, and projected 2018. Available at: http://www.bls.gov/emp/ep_table_201.pdf. Last accessed November 25, 2011.
3. Dembe A. Social inequalities in occupational health and health care for work-related injuries and illnesses. *Int J Law Psychiatry*. 1999;22:567–579.
4. Scherzer T, Rugulies R, Krause N. Work-related pain and injury and barriers to workers' compensation among Las Vegas hotel room cleaners. *Am J Public Health.* 2005;95:483–488.
5. Pransky G, Moshenberg D, Benjamin K, et al. Occupational risks and injuries in non-agricultural immigrant Latino workers. *Am J Ind Med.* 2002;42:117–123.
6. Dembe AE, Savageau JA, Amick BC, et al. Racial and ethnic variations in office-based medical care for work-related injuries and illnesses. *J Nat Med Assoc*. 2005;97:498–507.
7. >Dong X, Ringen K, Men Y, et al. Medical costs and sources of payment for work-related injuries among Hispanic construction workers. *J Occup Environ Med.* 2007;49:1367–1375.
8. Bureau of Labor Statistics. Table 11. Employed persons by detailed occupation, sex, race, and Hispanic or Latino ethnicity. Available at: http://www.bls.gov/cps/cpsaat11.pdf. Last accessed November 25, 2011.
9. Gabor M, Houlder D, Carpio M. *US Department of Labor (USDOL)*. 5th ed. Report on the American Workforce. Washington, DC: US Department of Labor; 2001.
10. Greenberg E, Macı´as RF, Rhodes D, et al. *National Center for Education Statistics*. *English Literacy and Language Minorities in the United States*. Washington, DC: US Department of Education; 2001. NCES 2001-464.
11. Murray LR. Sick and tired of being sick and tired: scientific evidence, methods, and research implications for racial and ethnic disparities in occupational health. *Am J Public Health.* 2003;93:221–226.
12. Strong LL, Zimmerman FJ. Occupational injury and absence from work among African American, Hispanic, and non-Hispanic white workers in the National Longitudinal Survey of Youth. *Am J Public Health.* 2005;95:1226–1232.
13. Dong X, Platner JW. Occupational fatalities of Hispanic construction workers from 1992 to 2000. *Am J Ind Med.* 2004;45:45–54.
14. Robinson JC. Exposure to occupational hazards among Hispanics, blacks, and non-Hispanic whites in California. *Am J Public Health*. 1989;79:629–630.
15. Weddle MG, Bissell RA, Shesser R. Working women at risk. Results from a survey of Hispanic injury patients. *J Occup Med.* 1993;35:712–715.
16. Friedman LS, Forst L. Ethnic disparities in traumatic occupational injury. *J Occup Environ Med.* 2008;50:350–358.
17. Wagener DK, Winn DW. Injuries in working populations: black–white differences. *Am J Public Health.* 1991;81:1408–1414.
18. Mcaig LF, Burt CW, Stussman BJ. A comparison of work-related injury visits and other injury visits to emergency departments in the United States, 1995–1996. *J Occup Environ Med.* 1998;40:870–875.
19. Oh JH, Shin EH. Inequalities in nonfatal work injury: the significance of race, human capital, and occupations. *Soc Sci Med.* 2003;57:2173–2182.
20. Smith GS, Wellman HM, Sorock GS, et al. Injuries at work in the US adult population: contributions to the total injury burden. *Am J Public Health.* 2005;95:1213–1219.
21. Waehrer G, Dong XS, Miller TR, et al. Costs of occupational injuries in construction in the United States. *Accid Anal Prev.* 2007;39:1258–1266.
22. Leigh JP, Waehrer GM, Miller TR, et al. Costs of occupational injury and illness across industries. *Scan J Work Environ Health*. 2004;30:199–205.
23. Shah SM, Bonauto D, Silverstein B, et al. Injuries and illnesses from wood framing in residential construction, Washington State, 1993–1999. *J Occup Environ Med*. 2003;45:1171–1182.
24. Lipscomb HJ, Dement JM, Behlman R. Direct costs and patterns of injuries among residential carpenters, 1995–2000. *J Occup Environ Med.* 2003;45:875–880.
25. Horwitz IB, McCall BP. Disabling and fatal occupational claim rates, risks, and costs in the Oregon construction industry 1990–1997. *J Occup Environ Hyg.* 2004;1:688–698.
26. Bernacki EJ, Tao XG. The relationship between attorney involvement, claim duration, and workers' compensation costs. *J Occup Environ Med.* 2008;50:1013–1018.
27. Appel DA. Health care costs in workers' compensation. *Benefits Q 4th Q*. 1993;9:6–8.
28. Bonauto DK, Smith CK, Adams DA, et al. Language preference and non-traumatic low back disorders in Washington State workers' compensation. *Am J Ind Med*. 2010;53:204–15.
29. Chibnall JT, Tait RC, Andresen EM, et al. Race and socioeconomic differences in post-settlement outcomes for African American and Caucasian workers' compensation claimants with low back injuries. *Pain.* 2005;114:462–472.
30. Friedman L, Forst L. Occupational injury surveillance of traumatic injuries in Illinois using the Illinois Trauma Registry, 1995–2003. *J Occup Environ Med*. 2007;49:401–410.
31. Friedman LS, Forst L. Workers' compensation costs among construction workers: a robust regression analysis. *J Occup Environ Med.* 2009;51:1306–1313.
32. Barell V, Aharonson-Daniel L, Fingerhut LA, et al. An introduction to the Barell body region by nature of injury diagnosis matrix. *Inj Prev.* 2002;8:91–96.
33. Baker SP, O'Neill B, Haddon W, et al. The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care. *J Trauma*. 1974;14:187–196.
34. Parkhurst DF. Arithmetic versus geometric means for environmental concentration data. *Environ Sci Technol*. 1998;32:92A–98A.
35. Huybrechts T, Thas O, Dewulf J, et al. How to estimate moments and quantiles of environmental data sets with non-detected observations? A case study on volatile organic compounds in marine water samples. 2002;975:123–133.
36. Tait RC, Chibnall JT, Andresen EM, Hadler NM. Management of occupational back injuries: differences among African Americans and Caucasians. *Pain*. 2004;112:389–396.
37. Tait RC, Chibnall JT. Legal sequelae of occupational back injuries a longitudinal analysis of Missouri judicial records. *Spine*. 2011;36:1402–1409.
38. Friedman DJ, Cohen BB, Averbach AR, et al. Race/ethnicity and OMB directive 15: implications for state public health practice. *Am J Public Health.* 2000;90:1714–1719.

