AMERICAN JOURNAL OF INDUSTRIAL MEDICINE 56:1149–1156 (2013)

Characterizing the Relationship Between In-Hospital Measures and Workers' Compensation Outcomes Among Severely Injured Construction Workers Using a Data Linkage Strategy

Peter S. Ruestow, MPH[1] and Lee S. Friedman, PhD[1,2]*

Background: *To characterize the relationship between acute measures of severity and three important workers' compensation outcomes associated with a worker's ability to return to work and the cost of a work-related injury.*
Methods: *Probabilistic data linkage of workers' compensation claims made by injured construction workers from 2000 to 2005 with two Illinois medical record registries. Multivariable robust regression models were built to assess the relationship between three in-hospital measures and three outcomes captured in the Workers' Compensation data.*
Results: *In the final multivariable models, a categorical increase in injury severity was associated with an extra $7,830 (95% CI: $4,729–$10,930) of monetary compensation awarded, though not with temporary total disability (TTD) or permanent partial disability (PPD). Our models also predicted that every extra day spent in the hospital results in an increase of 0.51 (95% CI: 0.23–0.80) weeks of TTD and an extra $1,248 (95% CI: $810–$1,686) in monetary compensation. Discharge to an intermediate care facility following the initial hospitalization was associated with an increase of 8.15 (95% CI: 4.03–12.28) weeks of TTD and an increase of $23,440 (95% CI: $17,033–$29,847) in monetary compensation.*
Conclusions: *We were able to link data from the initial hospitalization for an injured worker with the final workers' compensation claims decision or settlement. The in-hospital measures of injury severity were associated with total monetary compensation as captured in the workers' compensation process.* Am. J. Ind. Med. 56:1149–1156, 2013. *© 2013 Wiley Periodicals, Inc.*

KEY WORDS: injury severity; length of stay; disability; worker's compensation; injury; linkage

INTRODUCTION

The dangers of the construction industry continue to pervade occupational surveillance efforts nationwide. According to the Bureau of Labor Statistics (BLS), there were an estimated 244,200 non-fatal occupational injuries and illnesses occurring in US construction workers in 2009. The construction industry as a whole employed an estimated 6,497,870 workers during this time. Nationally, it had the largest number of injury fatalities among all industries (N = 816 deaths). Over the last decade, there has been a downward trend in both fatal and non-fatal construction injuries in the United States [US BLS CFOI, 2009; US BLS SOII, 2009]. However, the latest drop in nonfatal injuries that began in 2007 coincides with the sharp

[1]Division of Environmental and Occupational Health Sciences, School of Public Health, University of Illinois, Chicago, Illinois
[2]The Social Policy Research Institute, Skokie, Illinois
Contract grant sponsor: Center to Protect Workers' Rights through a cooperative agreement with the National Institute for Occupational Safety and Health; Contract grant number: U60 OH009762.
Disclosure Statement: None of the authors have any conflicts of interest that may be relevant to the submitted work.
*Correspondence to: Lee Friedman, PhD, Division of Environmental and Occupational Health Sciences, School of Public Health, University of Illinois at Chicago, 2121W. Taylor Street (Room 504), Chicago, IL 60612.
E-mail: lfried1 @uic.edu

Accepted 8 May 2013
DOI 10.1002/ajim.22212. Published online 3 June 2013 in Wiley Online Library (wileyonlinelibrary.com).

decline in the national housing market and very high unemployment within the industry. There is also growing evidence that the BLS vastly underreports injuries and illnesses [Rosenman et al., 2006; Nestoriak and Pierce, 2009], and that the downward trend may not be as precipitous as BLS data indicates [Friedman and Forst, 2007a].

Though not without limitations, publicly available datasets such as those maintained by the BLS can be useful in describing the magnitude of occupational injuries in the United States. However, alternative data sources should be identified and utilized as occupational surveillance tools in order to obtain clearer and more actionable information regarding injury risk and outcomes at the local level. In Illinois, these include the Illinois Workers' Compensation Commission (IWCC) claims dataset, the Illinois Trauma Registry (ITR), and the Illinois Hospital Discharge database. All are currently underutilized in occupational injury surveillance, despite the fact that each contains distinct demographic, exposure, health, and economic information about injured workers. The ability to match information from two or more of these databases to a single worker would allow interested parties to follow an injury from the point of initial hospitalization through the compensation claims processes.

Such a data linkage strategy may also be used to characterize the association between health information obtained during the acute clinical care process and important long-term outcomes captured in workers' compensation data, such as workers' ability to return to work, level of permanent disability, and total monetary compensation awarded. Understanding these relationships would allow key stakeholders to begin planning for work-related outcomes during the acute hospitalization phase. To date, many data linkage projects involving workers' compensation data have been limited to descriptive analyses [Alamgir et al., 2006; Bunn et al., 2007], estimation of underreporting to workers' compensation carriers [Alamgir et al., 2006] or surveys forming national estimates of injuries [O'Leinick et al., 1993; Boden and Ozonoff, 2008], or specific health outcomes and exposures such as asthma [Cherry et al., 2009], diisocyanates [Tarlo et al., 2002], organic solvents [Riise et al., 1995], and motor vehicle crashes [Boufous and Williamson, 2006]. More recently, a research team in the state of Washington, linking workers' compensation claims with the state trauma registry, found that work disability and medical costs increased monotonically with injury severity as measured by the injury severity score (ISS) [Sears et al., 2013].

This data linkage study aims to characterize the relationship between acute measures of severity and three important workers' compensation outcomes associated with a worker's ability to return to work and the cost of a work-related injury—temporary total disability (TTD) (e.g. weeks of lost work), permanent partial disability (PPD) (e.g., percent loss of function), and total monetary cost of workers' compensation claims. We focused our analysis on construction workers, a group at relatively higher risk for both long-term disability and more severe injury [Stover et al., 2007; Dong et al., 2008].

METHODS

Data Sources

We have a data sharing agreement with both the Illinois Department of Public Health and the IWCC. The IRB at UIC 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 was 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 who sustain traumatic injuries (ICD-9-CM external injury codes E800–995) and are admitted to a trauma center for greater than 12 hr, who are transferred to a level I or II center, or who are dead-on-arrival or die in the emergency department. We provide an assessment of data quality in the ITR in a previous paper [Friedman and Forst, 2007b]. 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).

Hospital discharge database

The University of Illinois at Chicago Hospital is a member of the Illinois Hospital Association, the organization that compiles and manages the hospital discharge database. As a member of the Illinois Hospital Association, the PI's research team has ongoing access to the hospital discharge database, which is based on billing records. It includes all patients treated for more than 23 hr in any Illinois hospital (i.e., inpatients only) for any medical reason, and includes data on patient demographics (age and gender), exposure (mechanism of injury), health outcomes (diagnoses, hospital procedures, and discharge status), and economics (hospital charges and payer source). The Illinois Hospital Association compiles, maintains, and conducts quality control of the dataset.

Illinois Workers' Compensation Commission claims database

The IWCC operates the administrative court system for workers' compensation cases in Illinois. There are

approximately 70,000 claims filed with the IWCC for financial compensation each year. Unlike single carrier states with a well organized and centralized reporting system (e.g., Washington), in Illinois the IWCC only handles claims in which the employee and employer are unable to resolve compensation issues for an injury without administrative intervention. Any aspect paid for prior to initiating a claim through the IWCC that is not disputed by either party is not litigated through IWCC or reported in the dataset.

We obtained a dataset of all claims from the IWCC consisting of employer information, employee demographics (age, gender, marital status, and number of dependents), cause and type of injury or illness, level of temporary and permanent disability, and details on the compensation costs associated with the injury. For this study, we used a dataset created in a previous analysis containing records on 19,734 workers' compensation claims made by injured construction workers between 2000 and 2005 [Friedman and Forst, 2009]. The minimum age in this study group was 16 years, as Illinois law prohibits persons under the age of 16 from working in construction.

Data Linkage

The IWCC dataset of 19,734 claims [Friedman and Forst, 2009] was the primary linkage dataset. After exclusion of claims without valid birth dates or injury dates, the final dataset contained 16,794 claims.

Table I outlines the probabilistic data linkage procedure used, which was accomplished in multiple steps, or passes. The concept of using multiple passes is to begin with the highest level of precision and then modify the precision with each subsequent step. The initial pass identified records in the IWCC dataset and the two separate medical record datasets that matched exactly on date of birth, date of injury, gender, residential ZIP code, nature of injury, and body part affected. In the second pass, we removed matches from the first pass, and then re-matched the datasets without residential ZIP code (since a worker's address may have changed from the time of injury to the time a claim was filed). In the subsequent passes, we allowed for variation of date of injury and date of birth. For the entire process, 64.6% of cases were matched during the 1st pass, 28.4% in the second pass, 2.5% in the third, 3.8% in the fourth, and 0.7% in the fifth.

TABLE I. Summary of the Data Linkage Steps Used to Merge Three State Datasets: Illinois Trauma Registry, Inpatient Hospital Discharge Data, and Workers' Compensation Claims

	DOB	Date of injury	Gender	ZIP code	Type of injury
Trauma registry to work comp					
1st pass	Exact	Exact	Exact	Exact	Match
2nd pass	Exact	Exact	Exact	Omit	Match
3rd pass	Exact	±1	Exact	Exact	Match
4th pass	Year	Exact	Exact	Exact	Match
Hospital discharge to work comp					
1st pass	Exact	Exact	Exact	Exact	Match
2nd pass	Exact	Exact	Exact	Omit	Match
3rd pass	Exact	±1	Exact	Exact	Match
4th pass	Year	Exact	Exact	Exact	Match
5th pass	Exact	±2	Exact	Exact	Match

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). Cause and type of injury information from each data source were manually reviewed for all linked cases to ensure an accurate match. We expected a linkage rate of approximately 10–15%, based on our pilot study which found that most claims for compensation involve injuries and illnesses that do not result in hospitalization or acute trauma. A total of 1,664 workers' compensation records linked to the two hospital registries—a linkage rate of 9.9%.

A comparison of cases in the final linked dataset with claims in the IWCC that did not match found that matched and unmatched cases were very similar in terms of distribution by gender, average weekly wage, marital status, number of dependents, age, and population density of employment area. Some of these demographic variables, particularly average weekly wage, have a major impact on the amount of compensation received by the worker in the claims process (Table II). Further, the monthly trend in injuries from January 2000 to December 2005 for both groups was very similar.

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

Primary Independent Variables

We focused our analysis on three commonly used measures of acute injury severity—new injury severity score (NISS), length of hospitalization, and discharge from the hospital to an intermediate care facility. Using the latest revision of ICDMAP-90, we mapped ICD-9-CM N-codes for all cases to calculate Abbreviated Injury Scale (AIS) scores for six major body regions. The AIS scores are calculated using a standardized computer algorithm based on discharge

TABLE II. Comparison of Matched and Unmatched Cases, Demographic Characteristics Illinois Workers Compensation Claims Data, 2000–2005

	Unmatched (N = 18,070)		Matched (N = 1,664)	
	N	%	N	%
Gender				
Male	17,224	95.3	1,624	97.6
Female	828	4.6	40	2.4
Unspecified	18	0.1	0	0.0
Marital status				
Single	6,768	37.5	651	39.1
Married	11,034	61.1	995	59.8
Widowed/divorced	40	0.2	1	0.1
Unspecified	228	1.3	17	1.0
Number of dependents				
0	9,454	52.3	864	51.9
1	2,941	16.3	255	15.3
2	3,135	17.3	307	18.4
3	1,639	9.1	151	9.1
4	587	3.2	54	3.2
5 or more	310	1.7	33	2.0
Unspecified	4	0.0	0	0.0
Age (years)				
16–24	1,483	8.2	191	11.5
25– 34	4,713	26.1	474	28.5
35–44	6,284	34.8	528	31.7
45–54	3,828	21.2	333	20.0
55–64	1,372	7.6	120	7.2
65 and older	101	0.6	11	0.7
Unspecified	289	1.6	7	0.4
Population density: place of accident (persons/sq.mi)				
Rural (0–499)	556	3.1	39	2.3
Mid range (500–999)	758	4.2	80	4.8
Urban (≥1,000)	16,466	91.1	1,529	91.9
Unspecified	290	1.6	16	1.0
Attorney representation used	13,922	77.0	1,394	83.8
Claim w/final decision	14,580	80.7	1,318	79.2
	Mean	**SD**	**Mean**	**SD**
Average age	39.6	10.4	38.6	10.8
Average weekly wage	930.3	367.4	882.4	378.1

records rather than admission status in order to reduce miscoding [Baker et al., 1974]. The AIS scores were used to calculate the NISS, which differs from the traditional formulation in that it uses the three highest abbreviated injury scores regardless of body region [Baker et al., 1974; Osler et al., 1997]. NISS scores are based on all N-codes for injuries identified during the course of hospitalization and have been shown to be a better predictor of adverse outcomes than the traditional formulation [Osler et al., 1997]. To make comparisons and calculate measures of association with workers' compensation outcomes in this analysis, the NISS was categorized in accordance with previous findings about its distribution—<9, 9–15, 16–24, and ≥25 [Stevenson et al., 2001; Kilgo et al., 2004]. Discharge to an intermediate care facility included skilled nursing facilities, rehabilitation facilities, acute care facilities, or residential facilities.

Primary Outcome Variables

TTD refers to the time period an injured worker is temporarily unable to do work, or is cleared for light duty, as indicated by a physician. TTD is defined in terms of weeks.

PPD involves partial loss of body function at the point of maximum medical improvement, and is measured as a percentage. Factors the IWCC considers when determining PPD include the nature of the accident, nature of the initial injury, objective physician physical findings, lost time, ability to return to work, job description, weekly wage, and current subjective complaints [Rusin, 2003]. 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 $A + (1 - A) \times B$, 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.

Our analysis also considered total financial compensation of decided and settled claims, excluding claims in progress and dismissed claims. Total financial compensation included medical costs, dependent benefits, survivor benefits, settlement payments, attorney fees, and other miscellaneous costs. Total compensation dollars were adjusted for inflation using the Consumer Price Index for all urban consumers.

Statistical Analysis

We used SAS software for all statistical analyses (v.9.1). We excluded deaths (N = 28) from the analysis because these workers would be missing TTD and PPD data. We also only included claims for which a final decision or settlement was reached. After exclusions, there were 1,295 cases in our dataset.

All three outcome variables were skewed to the right, therefore non-parametric statistical methods were employed (e.g., Kruskal–Wallis test). Subsequently, we characterized the relationship between disability ratings and trauma center severity measures using regression modelling to adjust for confounding and to identify interactions. Because none of the outcome variables were normally distributed, we used robust M-estimation as implemented in SAS Version 9 (PROC ROBUSTREG; SAS Institute, Inc.) [Huber, 1973, 1981; Hampel et al., 1986]. A manual stepwise selection method

was employed to identify the best model fit for the predictors, including covariates such as demographic variables, wage, injury type, and attorney representation. Akaike criterion and Schwarz information criteria were also used for model selection and to identify the best weighting function. No evidence of multicolinearity among the independent variables was indicated. A two-sided *P*-value less than 0.05 was considered statistically significant.

RESULTS

Demographics for cases that linked across at least one of the datasets (N = 1,664) are shown in Table II. The majority of cases (91%) were males under the age of 55. About 60% of all workers in the sample were married at the time of injury, with slightly less than 50% having at least one dependent. The median weekly wage was $882.00. Table III displays the mean and median values for TTD, PPD, and total compensation for cases within categories of NISS, length of stay, and discharge status. Beta estimates and associated *P*-values for crude and multivariable regression models are displayed in Table IV.

Model Outcome: Temporary Total Disability

Hospital length of stay and discharge to a rehabilitation or intermediate care facility, but not the NISS, remained significant after adjustment for age, gender, and attorney representation (Table IV). According to this multivariable model, each additional day spent in the hospital resulted in an extra half week of awarded TTD (0.51, 95%CI: 0.23–0.80). Patients requiring continued care at another facility after discharge from the hospital were awarded 8 more weeks of TTD than those discharged home (8.15, 95%CI: 4.03–12.28).

Model Outcome: Permanent Partial Disability

The relationship between PPD and each of the in-hospital measures of severity considered—the NISS, length of hospitalization, and discharge to a rehabilitation/intermediate care facility—became statistically insignificant after controlling for age, gender, and attorney representation (Table IV).

Model Outcome: Total Compensation

All in-hospital measures of severity considered were found to be strongly associated with total monetary compensation after controlling for age, gender, attorney representation, marital status, and number of body parts injured (Table IV). The adjusted model indicated that each categorical increase in the NISS increases the amount of total compensation awarded by almost $8,000 ($7,830, 95%CI: $4,729–$10,930). Each day a patient remains in the hospital

TABLE III. Distributions of Outcome Variables by Measures of In-Hospital Severity for Injured Construction Workers (N = 1,295)

	Temporary total disability (weeks)			Permanent partial disability (%)			Total compensation (USD$)		
	n	Mean	Median	n	Mean (%)	Median (%)	n	Mean ($)	Median ($)
NISS*,**,***									
0–8	864	25.56	10.36	863	23	17	865	46,052	24,001
9–15	264	33.61	17.36	264	27	27	264	65,693	38,119
16–24	124	32.53	14.50	125	25	15	125	78,924	50,287
≥25	38	47.19	16.57	38	20	10	38	99,356	44,597
LOS (days)*,***									
1	450	23.20	9.00	450	24	18	451	41,735	21,778
2–3	416	27.52	13.00	415	23	19	416	51,892	30,098
4–7	261	29.61	16.00	261	26	25	261	57,792	35,978
>7	164	44.49	20.00	165	24	15	165	93,198	62,956
Discharge to rehab/intermed care (Y/N)[a],*,**,***									
Yes	1,057	24.50	11.86	1,056	24	20	1,058	46,484	27,123
No	236	46.73	19.00	237	24	12	237	91,684	53,011

Total compensation adjusted for inflation using the U.S. Consumer Price Index.

[a]Discharge to a rehab/intermediate care facility included skilled nursing facilities, rehabilitation facilities, acute care facilities, or residential facilities. N sizes may not sum to the total (N = 1,295) due to missing data in one or both variables within each cross-tabulation.

*Kruskal–Wallis test for temporary total disability significant ($P < 0.05$).

**Kruskal–Wallis test for permanent partial disability significant ($P < 0.05$).

***Kruskal–Wallis test for total compensation significant ($P < 0.05$).

TABLE IV. Association Between Disability Ratings and In-Hospital Indicators of Severity Among Construction Workers Filing Workers' Compensation Claims

Variable	Crude model estimate	Standard error	*P*-value	Multivariable model estimate	Standard error	*P*-value
Model outcome: temporary total disability (weeks)[a]						
New injury severity score (NISS)*	3.68	0.96	<0.001	1.56	1.02	0.129
Length of stay (LOS)	0.83	0.13	<0.001	0.51	0.15	<0.001
Discharge to rehabilitation or intermediate care facility[b]	11.87	1.97	<0.001	8.15	2.10	<0.001
Model outcome: permanent partial disability (%)[a]						
New injury severity score (NISS)[c]	0.84	0.89	0.346	1.10	0.95	0.245
Length of stay (LOS)	−0.06	0.12	0.642	−0.06	0.13	0.649
Discharge to rehabilitation or intermediate care facility[b]	−0.02	0.02	0.263	−1.77	1.94	0.362
Model outcome: total monetary compensation (USD$)[d]						
New injury severity score (NISS)[c]	14,195	1,528	<0.001	7,830	1,582	<0.001
Length of stay (LOS)	2,212	202	<0.001	1,248	223	<0.001
Discharge to rehabilitation or intermediate care facility[b]	35,761	3,128	<0.001	23,440	3,269	<0.001

Total compensation adjusted for inflation using the U.S. Consumer Price Index.
[a]Model adjusted for attorney representation, injury to extremities, age, and gender.
[b]Discharge to a rehab/intermediate care facility included skilled nursing facilities, rehabilitation facilities, acute care facilities, or residential facilities.
[c]NISS is categorized (0–8, 9–15, 16–24, ≥25) and treated as ordinal.
[d]Model adjusted for attorney representation, age, gender, marital status, and number of body parts injured.

increases this award by about $1,250 ($1,248, 95%CI: $810–$1,686), according to the model, and patients requiring continued care after hospital discharge are awarded about $23,000 ($23,440, 95%CI: $17,033–$29,847) of additional total compensation.

DISCUSSION

The NISS, length of hospitalization, and discharge status were strongly associated with total monetary compensation as awarded through workers' compensation. Length of hospitalization and discharge status were also reasonable predictors of TTD awarded. The NISS did not seem to be prognostic of TTD, and none of the measures of severity considered were associated with PPD in this sample of Illinois construction workers. It is surprising that discharge status was not associated with permanent disability, unless this reflects the benefits of access to rehabilitation services on outcomes [Wood et al., 1999; Pransky et al., 2005; Sullivan et al., 2005]. However, discharge to an intermediate care facility was associated with substantially higher total monetary compensation.

TTD, which indicates the number of weeks of missed work or restricted duty following an injury, increased by slightly over a week for each 2 days an injured worker remained in the hospital. Workers discharged to a rehabilitation or other intermediate care facility missed approximately 8 weeks more than patients able to go directly home after hospitalization. This information could be useful to employees, employers, and insurance providers. Information on expected time away from work would allow an injured worker to plan finances accordingly, which may help to alleviate psychosocial as well as pecuniary issues [Mason et al., 2002]. For an employer, this information would help with planning for temporary employment contracting and insurance costs. However, this analysis was restricted to injured construction workers. It is unclear if injured workers in other industries would show similar associations between acute injury measures and long-term outcomes post-injury.

The general rehabilitation literature demonstrates a relationship between in-hospital measures of injury severity (ISS and length of stay) and longer term outcomes, such as a worker's ability to return to work and functional disability [Bull, 1985; Brenneman et al., 1997; Vles et al., 2005; Holtslag et al., 2007a, b]. However, data on the direct relationship between these acute measures and outcomes determined during the workers' compensation process is lacking.

The results of this study highlight the value of data linkage strategies in injury surveillance, but also expose the difficulties inherent in merging disparate state databases in their existing format. Despite these obstacles, we achieved a greater understanding of how hospital health information relates to the workers' compensation claims process. The value of a data linkage strategy rests on the supposition that more can be learned from multiple information sources together than from any single dataset. By focusing on construction workers, we feel that we demonstrated a few of the ways health information may relate to or differ from the

information collected during the legal process in a high-risk group. However, 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.

Limitations

Though the overall linkage rate was slightly lower than expected (9.9%), the demographic distributions among linked cases were very similar to those IWCC cases that did not link to either of the medical record databases. The fact that proportions by type of injury differed between the linked and non-linked cases is not surprising. Not every worker who filed a claim necessarily sought care at a trauma center or inpatient setting.

The low sample size of the final linked dataset may affect the level of precision of the statistical analysis. This may have been a result of the probabilistic linkage strategy that had to be employed in the absence of a shared case identifier among the datasets. There was also a level of uncertainty about "true matches" inherent to this methodology. However, we feel that the linkage process resulted in highly accurate matches among the cases that did link across the databases. This was manually verified by comparing ITR and IHD diagnosis codes with IWCC nature of injury and body part affected fields. We also used the trauma registry to estimate the real mismatch probability using the linkage variables. The trauma registry, unlike the other two datasets, 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 datasets or between datasets 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 1,000 records). However, we have no information on cases involving the same person across datasets that did not link because of data entry errors.

A selection bias may have occurred because the cases in the IWCC dataset only include injuries with disputed claims. However, based on BLS data on number of injuries in Illinois in the construction sector, it appears that the majority of claims involving TTD or PPD end up disputed in Illinois. It is also worth noting, again, that only settled cases were included in this analysis. Cases with pending settlements had higher ISSs and longer lengths of stay. Though we cannot predict how these cases would impact our results, it is possible that the effect would be significant.

Two overarching themes emerged from this analysis. First, the in-hospital measures of injury severity are associated with outcomes captured in the workers' compensation process, in particular TTD and total monetary compensation. This suggests that workers' compensation may be predicted or estimated using information obtained during the acute phase. Second, data linkage methodologies are a valuable approach to enhancing existing occupational injury data. In this analysis, we were able to follow an injured worker from the initial hospitalization through the final workers' compensation claims decision or settlement. Combining information contained within each database helped to better characterize outcomes from occupational injuries, and may prove useful in designing more effective interventions for getting construction workers back to work, and may help employers and insurance carriers to more effectively plan for costs associated with injuries. This research should be expanded to look at other industries and to integrate more comprehensive workers' compensation carrier data into the linkage process.

REFERENCES

Alamgir H, Koehoorn M, Ostry A, Tompa E, Demers PA. 2006. How many work-related injuries requiring hospitalization in British Columbia are claimed for workers' compensation? Am J Ind Med 49(6):443–451.

Baker SP, O'Neill B, Haddon W, Long WB. 1974. The injury severity score: A method for describing patients with multiple injuries and evaluating emergency care. J Trauma 14(3):187–196.

Boden LI, Ozonoff A. 2008. Capture-recapture estimates of nonfatal workplace injuries and illnesses. Ann Epidemiol 18(6):500–506.

Boufous S, Williamson A. 2006. Work-related traffic crashes: A record linkage study. Accid Anal Prev 38(1):14–21.

Brenneman FD, Redelmeier DA, Boulanger BR, McLellan BA, Culhane JP. 1997. Long-term outcomes in blunt trauma: Who goes back to work? J Trauma 42(5):778–781.

Bull JP. 1985. Disabilities caused by road traffic accidents and their relation to severity scores. Accid Anal Prev 17:387–397.

Bunn T, Slavova S, Bathke A. 2007. Data linkage of inpatient hospitalization and workers' claims data sets to characterize occupational falls. J Ky Med Assoc 105(7):313–320.

Cherry N, Beach J, Burstyn I, Fan X, Guo N, Kapur N. 2009. Data linkage to estimate the extent and distribution of occupational disease: New onset adult asthma in Alberta, Canada. Am J Ind Med 52(11): 831–840.

Dong X, Men Y, Fujimoto A, Welch L, McCann M, Gittleman J, Trahan C. 2008. The construction chart book, the U.S. construction industry and its workers, 4th edition. Silver Spring, MD: CPWR—The Center for Construction Research and Training.

Friedman L, Forst L. 2007a. The impact of OSHA recordkeeping regulation changes on occupational injury and illness trends in the U.S.: A time-series analysis. Occup Environ Med 64:454–460.

Friedman L, Forst L. 2007b. Occupational injury surveillance of traumatic injuries in Illinois using the Illinois Trauma Registry, 1995–2003. J Occup Environ Med 49(4):401–410.

Friedman L, Forst L. 2009. Workers' compensation costs among construction workers: A robust regression analysis. J Occup Environ Med 51(11):1306–1313.

Hampel FR, Ronchett EM, Rousseeuw PJ, Stahel WA. 1986. Linear models: Robust estimation, robust statistics: The approach based on influence functions. New York: John Wiley & Sons. p 315–328.

Holtslag HR, van Beeck EF, Lindeman E, Leenen LP. 2007a. Determinants of long-term functional consequences after major trauma. J Trauma 62(4):919–927.

Holtslag HR, Post MW, van der Werken C, Lindeman E. 2007b. Return to work after major trauma. Clin Rehabil 21(4):373–383.

Huber PJ. 1973. Robust regression: Asymptotics, conjectures and the Monte Carlo. Ann Stat 1:799–821.

Huber PJ. 1981. The basic types of estimates, robust statistics. New York: John Wiley & Sons. p 43–55.

Kilgo PD, Meredith JW, Hensberry R, Osler TM. 2004. A note on the disjointed nature of the injury severity score. J Trauma 57(3):479–485.

Mason S, Wardrope J, Turpin G, Rowlands A. 2002. The psychological burden of injury: An eighteen month prospective cohort study. J Trauma 53(1):98–103.

Nestoriak N, Pierce B. 2009. Comparing workers' compensation claims with establishments' responses to the SOII. Mon Lab Rev 132 (5):57–64.

O'Leinick A, Guire KE, Hawthorne VM, Schork MA, Gluck JV, Lee B, La S. 1993. Current methods of estimating severity for occupational injuries and illnesses: Data from the 1986 Michigan Comprehensive Compensable Injury and Illness Database. Am J Ind Med 23(2): 231–252.

Osler T, Baker SP, Long W. 1997. A modification of the injury severity score that both improves accuracy and simplifies scoring. J Trauma 43 (6):922–925.

Pransky G, Gatchel R, Linton SJ, Loisel P. 2005. Improving return to work research. J Occup Rehabil 15(4):453–457.

Riise T, Kyvik KR, Moen B. 1995. A cohort study of disability pensioning among Norwegian painters, construction workers, and workers in food processing. Epidemiology 6(2):132–136.

Rosenman KD, Kalush A, Reilly MJ, Gardiner JC, Reeves M, Luo Z. 2006. How much work-related injury and illness is missed by the current national surveillance system? J Occup Environ Med 48(4):357–365.

Rusin ME. 2003. Permanent partial disability in Illinois: Unraveling the mystery. Rusin Maciorowski & Friedman, Ltd. Chicago, IL.

Sears JM, Blanar L, Bowman SM, Adams D, Silverstein BA. 2013. Predicting work-related disability and medical cost outcomes: Estimating injury severity scores from workers' compensation data. J Occup Rehabil 23(1):19–31.

Stevenson M, Segui-Gomez M, Lescohier I, Di Scala C, McDonald-Smith G. 2001. An overview of the injury severity score and the new injury severity score. Inj Prev 7:10–13.

Stover B, Wickizer TM, Zimmerman F, Fulton-Kehoe D, Franklin G. 2007. Prognostic factors of long-term disability in a workers' compensation system. J Occup Environ Med 49(1):31–40.

Sullivan MJ, Feuerstein M, Gatchel R, Linton SJ, Pransky G. 2005. Integrating psychosocial and behavioral interventions to achieve optimal rehabilitation outcomes. J Occup Rehabil 15(4):475–489.

Tarlo SM, Liss GM, Yeung KS. 2002. Changes in rates and severity of compensation claims for asthma due to diisocyanates: A possible effect of medical surveillance measures. Occup Environ Med 59(1):58–62.

U.S. Department of Labor, Bureau of Labor Statistics. 2009. Census of fatal occupational injuries 2009 [database online]. Washington, DC: U. S. Department of Labor, Bureau of Labor Statistics. Retrieved on February 10, 2011. http://www.bls.gov/iif/oshcfoi1.htm.

U.S. Department of Labor, Bureau of Labor Statistics. 2009. Survey of Occupational Injuries and Illnesses 2009 [database online]. Washington, DC: U.S. Department of Labor, Bureau of Labor Statistics. Retrieved February 10, 2011. http://www.bls.gov/iif/oshcfoi1.htm.

Vles WJ, Steyerberg EW, Essink-Bot ML, van Beeck EF, Meeuwis JD, Leenen LP. 2005. Prevalence and determinants of disabilities and return to work after major trauma. J Trauma 58(1):126–135.

Wood RL, McCrea JD, Wood LM, Merriman RN. 1999. Clinical and cost effectiveness of post-acute neurobehavioural rehabilitation. Brain Inj 13(2):69–88.