AMERICAN JOURNAL OF INDUSTRIAL MEDICINE 57:1100–1109 (2014)

Injury Classification Agreement in Linked Bureau of Labor Statistics and Workers' Compensation Data

Sara E. Wuellner, MPH* and David K. Bonauto, MD, MPH

Background *Estimates of select occupational injuries and illnesses often differ across data sources. We explored agreement in injury classifications and the impact of differences on case estimates among records reported to multiple data sources.*
Methods *We linked cases reported in the Bureau of Labor Statistics (BLS) annual Survey of Occupational Injuries and Illnesses (SOII) to Washington State workers' compensation (WC) claims and evaluated agreement in injury characteristics coded in each data source according to the same occupational injury and illness classification system.*
Results *Agreement between data sources was greatest for body part and lowest for event or exposure. Agreement on nature of injury varied by condition. WC-assigned injury codes estimated 94% more amputations than SOII-assigned codes while SOII-assigned codes estimated 34% more work-related MSD cases.*
Conclusions *Accounting for classification differences may improve case ascertainment within individual data sources and help align injury and illness estimates derived from different data sources.* Am. J. Ind. Med. 57:1100–1109, 2014.

KEY WORDS: occupational injury; surveillance; Workers' Compensation data; injury classification; injury coding

INTRODUCTION

Occupational health surveillance data are used to monitor workplace exposures and health effects, develop workplace interventions, and guide policy. Accurate data are essential to inform the appropriate allocation of limited research and prevention resources. The United States Bureau of Labor Statistics' (BLS) Survey of Occupational Injuries and Illnesses (SOII), an annual survey of sampled business establishments, is one of the primary sources of work-related injury and illness data in the US, providing both national, and for most states, state level estimates of nonfatal occupational injuries and illnesses based on employer reports of OSHA recordable cases [US Department of Labor, 2012a]. Over the years, SOII has been the focus of many researchers, policy analysts, and others in the occupational health community who question the completeness of the survey data. Several studies report a failure by SOII to capture all eligible injuries [Leigh et al., 2004; Rosenman et al., 2006; Boden and Ozonoff, 2008] and other studies comparing SOII estimates of select conditions to cases identified through other data sources have concluded that SOII underestimates such

Washington State Department of Labor and Industries, Safety and Health Assessment and Research for Prevention (SHARP) Program, Olympia, Washington

Contract grant sponsor: Bureau of Labor Statistics; Contract grant number: OS-20358-11-75-J-53.
Disclosure Statement: The authors report no conflicts of interests.

*Correspondence to: Sara Wuellner, MPH, Washington State Department of Labor and Industries Safety and Health Assessment and Research for Prevention (SHARP) Program, PO Box 44330, Olympia, WA 98504.
E-mail: sara.wuellner@lni.wa.gov

Accepted 25 November 2013
DOI 10.1002/ajim.22289. Published online 17 December 2013 in Wiley Online Library (wileyonlinelibrary.com).

conditions [Morse et al., 2001; Lipscomb et al., 2008a; Kica and Rosenman, 2012; Friedman et al., 2013].

Potential reasons for the observed SOII undercount are many: incomplete employer reports of injuries and illnesses, whether intentional or inadvertent; inadequate understanding of the reporting requirements; and constraints of the survey's methodology which limits the reporting of certain cases, including illnesses with long latencies, injuries that worsen over time, and those that are difficult to attribute to work [Seligman et al., 1988; Pransky et al., 1999; Azaroff et al., 2002; Ruser, 2008; Nestoriak and Pierce, 2009; Dong et al., 2011].

Another possible explanation for apparent low SOII case capture, yet to be explored, is the characterization of reported cases and whether it is consistent across data sources or differs by source. Identical injury events can appear distinct if the characterization of the injury differs by data source. For example, based on variant incident descriptions, a case may be coded as a crushing injury in the SOII but as an amputation in another data source. Depending on the data source used for case surveillance, this incident would be counted toward the total number of one condition instead of the other. Thus, differences in case classification may lead to divergent estimates of specific conditions.

In this study, we assessed injury coding agreement among cases reported to multiple data sources and examined the impact of coding differences on estimates of select occupational injuries and illnesses.

METHODS

Data Sources and Codes

We linked 3 years of BLS SOII case data to Washington WC claims data to assess injury classification agreement among cases reported to both systems.

During the study period, both data sources coded injury and illness characteristics according to the Occupational Injury and Illness Classification manual (OIICS) from 1992, with minor revisions adopted in 2007. OIICS, developed by BLS, provides a classification system for coding four aspects of a work-related injury or illness: the principal physical characteristic or nature of the injury or illness; part of body affected; the source, namely objects, substances or other factors responsible for the injury or illness; and the event or exposure to describe the manner in which the injury or illness occurred [US Department of Labor, 1992, 2007].

OIICS employs a hierarchical structure with up to four digits used to describe each aspect of the case. The first digit designates the division that represents general categories of case characteristics. The second digit designates the major group, and, in certain prescribed instances, a third and sometimes fourth digit are used to designate the group and subgroup, respectively. For example, for the characteristic nature of injury, the division *Traumatic injuries and disorders* (0*) contains ten major groups including *Open wounds* (03*) (an asterisk indicates the inclusion of all codes that begin with the character(s) listed). Nine of the ten groups within *Open wounds*, a partial list of which includes *Animal or insect bites* (032), *Cuts, lacerations* (034), and *Gunshot wounds* (036), allows for no more detail beyond the three digit group level. Among the groups in *Open wounds*, only *Amputations* (031*) includes subgroups: *Amputations, fingertip* (0311) and *Amputations, except fingertip* (0319).

BLS SOII data

The BLS administers the SOII annually in partnership with participating states to estimate the incidence of nonfatal OSHA-recordable occupational injuries and illnesses. BLS collects work-related injury and illness data from sampled private sector employers as well as state and local governments. Sampled establishments are asked to submit the number of OSHA recordable cases that occurred within the survey year. For injuries and illnesses resulting in at least 1 day of missed work following the date of injury, employers send detailed worker and incident characteristics including worker name, date of birth or age, sex, date of injury, and a description of the injury or illness as well as the activity immediately preceding the incident. Survey respondents are instructed to report detailed case information from any of the following sources: the OSHA Form 301; a workers' compensation report; an accident report; or an insurance form. Following a protocol established by BLS, participating states code the injury descriptions according to OIICS.

Washington State SOII data for survey years 2006–2008 were obtained through a cooperative agreement between the BLS and the Washington State Department of Labor and Industries (L&I).

Washington WC data

L&I regulates workers' compensation insurance for all non-federal employers operating in Washington State [Washington Revised Code, 1972]. In addition, L&I administers the Washington state fund (SF) workers' compensation insurance program which is the sole workers' compensation insurance provider for all employers in the state except those covered by an alternate workers' compensation system (e.g., Longshore and Harbor Workers' Compensation Act, Federal Employees' Compensation Act), specific employers or occupations exempt from mandatory coverage (e.g., self-employed), or those who are able to self-insure. Approximately 70% of

workers under L&I jurisdiction are covered by SF employers and 30% work for a self-insured employer.

A workers' compensation claim is initiated when a worker and the health care provider complete and submit a report of industrial injury or occupational disease (RIIOD). All filed SF claims, regardless of claim acceptance or award, are coded by trained L&I staff who review the narrative description of the incident and the resulting injury or illness as provided on the RIIOD to assign codes according to OIICS. In contrast, only a portion of self-insured claims awarded indemnity payments are coded for injury characteristics (in Washington, the waiting period for indemnity is 3 calendar days following the day of injury).

All filed workers' compensation claims with an injury date in 2006–2008 were extracted from the L&I workers' compensation database on July 13, 2010. The claims data extracted included claimant name, sex, date of birth, date of injury or illness, employer name and address, and OIICS Nature, Event or Exposure, and Part of Body codes.

Record Linkage

Record linkage procedures were patterned after those developed by researchers for another study linking similar BLS case data to WC claims data [Boden and Ozonoff, 2008]. In preparation for linking, first names and addresses were standardized using the US Census Bureau's data standardization software GDRIVER [US Bureau of the Census, 1999]. Linking was attempted between SOII cases (OSHA recordable injuries and illnesses resulting in 1 or more days of missed work) and all filed WC claims among SF and self-insured employers, and was conducted in two stages. First, we linked cases to claims through identical matches on eight of the nine common variables: worker last name, first initial, sex, date of birth or age at injury, date of injury, employer name, employer address, zip code or city, and federal employer identification number. We then linked the remaining unmatched cases using the probabilistic software program LinkPlus [US Department of Health and Human Services, 2007], which assigns a score to each potential matched pair based on the degree of similarity in the common variables. Two researchers reviewed pairs with lower scores to determine whether to accept the potential match. Disagreements between the two reviewers were settled by a third reviewer. We allowed some difference in injury dates between data sources, although date differences were not allowed to exceed 2 months. Because the injury date in WC is adjudicated to reflect the last injurious exposure, the date reported by the employer may not be identical, especially for non-traumatic conditions with no clear precipitating event. The injury date documented by the employer may be the date the worker received medical treatment or the date of missed or restricted work. Injury dates more than 2 months apart were assumed to reflect separate injury events rather than differences in the characterization of one singular event.

Data Analysis

The analysis of injury and illness classification agreement was limited to linked SF claims because injury classification codes are not systematically assigned to self-insured claims. More commonly assigned OIICS codes were assessed individually while less common codes were aggregated within major groups or divisions.

The BLS assigns each reported case a sample weight that is used to estimate the number and rate of nonfatal occupational injuries and illnesses among the population. Using the SOII sample weights assigned to each case, two population estimates for each selected condition were calculated for comparison: one based on injury and illness characteristics as coded in SOII and a second based on characteristics as coded in WC. We selected for estimation two conditions that state-based surveillance efforts monitor using the SOII data: amputations and musculoskeletal disorders [Council of State and Territorial Epidemiologists, Updated 2012]. The CSTE surveillance definitions of amputations and musculoskeletal disorders (MSD) are based on 1992 OIICS codes. Amputations are defined as cases with an *Amputation* Nature code (031*). Musculoskeletal disorders are defined as cases with an MSD-related Nature code and an MSD-related Event code (OIICS Nature codes: 021 (*Sprains, strains, tears*); 0972 (*Back pain, hurt back*); 0973 (*Soreness, pain, hurt, except back*); 1241 (*Carpal tunnel syndrome*); 153* (*Hernia*); or 17* (*Musculoskeletal system and connective tissue diseases and disorders*) and OIICS Event codes: 211 (*Bending, climbing, crawling, reaching, twisting*); 22* (*Overexertion*); or 23* (*Repetitive motion*)).

Kappa statistics, a measure of agreement adjusted for chance [Landis and Koch, 1977], were used to measure agreement in OIICS codes among matched SOII-WC records for three injury characteristics: Nature, Part of Body, and Event. Agreement was assessed for divisions, major groups, groups, and subgroups. When a major group or group was the greatest level of detail available in OIICS, the case was included in the analysis of more detailed classifications. McNemar's test was used to assess differences in proportions of select injury classifications among the matched cases [McNemar, 1947]. Neither data source was regarded as the referent. All analyses were performed using SAS version 9.3.

The Washington State Institutional Review Board approved the study.

RESULTS

Record linkage procedures matched 90% of the 29,862 SOII cases to WC claims, a total of 26,925 linked records.

Links among SF WC claims totaled 15,447 claims (57%) and the remaining cases linked to WC claims among self-insured employers. SOII and WC data documented identical injury dates in 14,049 SF claims (91% of linked SF claims); injury dates were within 7 days in 14,943 linked SF claims (97%).

Among the 15,447 SOII cases linked to SF WC claims, *Sprains, strains, tears* were the most common injury or illness assigned in either SOII or WC (48% and 40%, respectively) and backs were the most frequently affected body part (23% of linked SOII cases, 24% of linked WC claims). WC classified more cases than SOII as non-traumatic (1,519 records compared with 1,099 records, $P < 0.01$) although SOII coded 27 more cases as *Carpal tunnel syndrome*, the most frequently assigned non-traumatic Nature classification in either data source (236 records based on SOII-assigned codes, 209 records based on WC-assigned codes). *Overexertion*, specifically in lifting, was the most frequently assigned exposure, with a similar number of cases classified as *Overexertion* in SOII (4,299 records) as in WC (4,300 records).

One in five linked records (3,205 cases) was classified identically in both data sources for all three injury characteristics: Nature, Part of Body, and Event. For 1,426 linked records (9%), none of the injury characteristics matched across data sources.

Classification agreement varied by injury characteristic and level of coding detail, with the portion of records in agreement decreasing as coding detail increased (Table I). For each injury characteristic (Nature, Body Part, and Event), at least 85% of cases linked to SF claims were coded identically in both data sources at the most general (1-digit) division level. For Nature and Body Part, one in three cases were assigned to groups (3-digit codes) that differed by data source. For Event, over half were assigned to groups (3-digit codes) that differed by data source.

Nature of Injury Classification

Over 90% of linked records were classified in SOII as one of thirteen 3-digit group codes: 10 within traumatic injuries and three within systemic conditions or disorders. Among those thirteen group codes, agreement with WC-assigned Nature was lowest for *Nonspecified injuries and disorders* ($k = 0.19$), and greatest for *Hernia* ($k = 0.89$) and *Heat burns, scalds* ($k = 0.89$). Agreement for the most frequently-assigned code, *Sprains, strains, tears* ($k = 0.68$), was exceeded only by agreement among cases classified as *Hernia* or *Heat burns, scalds*, although one in four cases classified in SOII as *Sprains, strains, tears* was classified differently in WC. Twenty percent of SOII-designated *Sprains, strains, tears* were classified in WC as some other traumatic injury or disorder, including: 7% as *Multiple Traumatic Injuries and Disorders*; 4% as *Dislocations*; and 4% as *Nonspecified injuries and disorders*. Four percent of SOII-designated *Sprains, strains, tears* were classified in WC as *Systemic diseases and disorders*. For four of the thirteen most frequently SOII-assigned groups, cases not coded identically in WC were most often assigned another group within the SOII-assigned major group. When not coded identically in WC, SOII-assigned *Cuts, lacerations* were classified in WC as some other *Open wound* more than any other group code; SOII-assigned *Abrasions, scratches* and *Foreign bodies* were most often classified as some other *Surface wounds and bruises* in WC; and SOII-assigned *Heat burns, scalds* appeared in WC as some other burn. More than 10% of cases within each of the following six SOII-assigned groups had codes assigned in WC that differed at the 1-digit division level: *Rheumatism*; *Peripheral nerve damage*; *Nonspecified injuries and disorders*; *Hernia*; *Foreign bodies*; and *Punctures*. When not classified identically in WC, cases classified in SOII as *Bruises, contusions* or *Fractures* were more often assigned the code for *Multiple Traumatic Injuries and Disorders* than any other classification.

TABLE I. Agreement in Injury Classification Codes* by Level of Coding Detail Among 15,447 SOII Cases Linked to Washington State Fund WC Claims

	Division (1-digit code)			Major Group (2-digit code)			Group (3-digit code)			Subgroup (4 digit-code)		
	n	%	Kappa	n	%	Kappa	n	%	Kappa	n	%	Kappa
Nature of injury or illness	14,525	94	0.62	10,966	71	0.62	10,268	67	0.57	10,041	65	0.55
Part of body	13,376	87	0.82	11,844	77	0.74	10,100	65	0.63	10,091	65	0.63
Event or exposure	13,055	85	0.78	10,672	69	0.64	6,875	45	0.42	6,647	43	0.40

Data presented are number of linked records with matching codes, percent of total linked records, and kappa statistic.

*Both the SOII and WC cases were coded using the Occupational Injury and Illness Classification System, a hierarchical classification system with general categories containing codes of greater specificity. Most codes available in OIICS are 3- or 4-digits in length, although some are limited to 1- or 2-digits. The most general category, the Division, is designated by the first digit of the code. To assess agreement in assignment of the Division, codes were aggregated to the first digit. Major Group included aggregation to the 2-digit level plus codes with no more detail beyond the 1-digit Division level. Group included aggregation to the 3-digit level plus terminal 1- and 2-digit codes. Subgroup consists of all codes as assigned with no aggregation.

TABLE II. Linked Records by WC and SOII Nature of Injury Coding[†] (n = 15,447)

	SOII-assigned codes									
WC-assigned codes	**Traumatic injuries to bones, nerves, spinal cord (01*)**	**Traumatic injuries to muscles, tendons, ligaments, joints, etc. (02*)**	**Open wounds (03*)**	**Surface wounds and bruises (04*)**	**Burns (05*)**	**Multiple traumatic injuries and disorders (08*)**	**Other traumatic injuries and disorders (09*)**	**Systemic diseases and disorders (1*)**	**All others**	**Total**
Traumatic injuries to bones, nerves, spinal cord (01*)	972 (6)	378 (2)	55 (<1)	41 (<1)	0 (0)	84 (1)	57 (<1)	7 (<1)	3 (<1)	1,597 (10)
Traumatic injuries to muscles, tendons, ligaments, joints, etc. (02*)	62 (<1)	5,760 (37)	15 (<1)	94 (1)	1 (<1)	91 (1)	227 (1)	93 (1)	8 (<1)	6,351 (41)
Open wounds (03*)	28 (<1)	28 (<1)	1,145 (7)	46 (<1)	2 (<1)	39 (<1)	52 (<1)	9 (<1)	4 (<1)	1,353 (9)
Surface wounds and bruises (04*)	39 (<1)	149 (1)	59 (<1)	1,257 (8)	9 (<1)	89 (1)	62 (<1)	31 (<1)	11 (<1)	1,706 (11)
Burns (05*)	0 (0)	2 (<1)	1 (<1)	6 (<1)	204 (1)	1 (<1)	3 (<1)	23 (<1)	0 (0)	240 (2)
Multiple traumatic injuries and disorders (08*)	199 (1)	520 (3)	123 (1)	241 (2)	2 (<1)	561 (4)	72 (<1)	18 (<1)	57 (<1)	1,793 (12)
Other traumatic injuries and disorders (09*)	24 (<1)	362 (2)	18 (<1)	78 (1)	3 (<1)	29 (<1)	213 (1)	29 (<1)	11 (<1)	767 (5)
Systemic diseases and disorders (1*)	6 (<1)	303 (2)	33 (<1)	60 (<1)	17 (<1)	4 (<1)	73 (<1)	794 (5)	11 (<1)	1,301 (8)
All others	14 (<1)	76 (<1)	10 (<1)	41 (<1)	4 (<1)	18 (<1)	36 (<1)	31 (<1)	109 (1)	339 (2)
Total	1,344 (9)	7,578 (49)	1,459 (9)	1,864 (12)	242 (2)	916 (6)	795 (5)	1,035 (7)	214 (1)	15,447 (100)

Data presented are n (%).

*Codes with greater detail aggregated at the level indicated.

[†]Both the SOII and WC cases were coded using the Occupational Injury and Illness Classification System.

Table II presents the distribution of cases by the 2-digit major group codes assigned in each data source characterizing Nature. Even at this more general level of detail, coding disagreements between the two data sources persist. Of the 1,344 cases coded as *Traumatic injuries to bones, nerves, spinal cord* (9% of linked records), 72% were classified the same in WC, 15% had a WC-assigned code for *Multiple traumatic injuries and disorders*, and the remaining 13% of SOII-designated cases of *traumatic injuries to bones, nerves, spinal cord* were classified in WC as 1 of 12 other major groups, including non-traumatic conditions or disorders. For most traumatic major groups, injuries not classified the same in both data sources were often classified as *Multiple traumatic injuries and disorders* in WC.

Agreement was better among traumatic injuries than non-traumatic conditions or disorders. Of the 14,348 linked records classified in SOII as *Traumatic injuries or disorders*, 66% had an identical Nature code assigned in WC, and 29% had a non-identical WC-assigned code within the traumatic injury division, including 8% with a WC-assigned code for multiple traumatic injuries. Among SOII's 1,035 *Systemic diseases and disorders*, 55% were coded identically in WC, 21% were assigned a different code within the *Systemic diseases and disorders* division, and another 21% had a WC-assigned code within the *Traumatic injuries or disorders*. An additional 3% of SOII-designated *Systemic diseases and disorders* were classified in WC within some division other than *Systemic diseases and disorders* or *Traumatic injuries or disorders*.

Over 75% of linked cases were categorized in SOII using 1 of 12 body part codes. Among those 12, high agreement with WC coding was found for five body parts (33% of SOII-designated cases): *Eye*, *Knee*, *Finger*, *Ankle*, and *Shoulder* ($k > 0.80$), and lower agreement was found for *Wrist(s)*, *Foot*, *Lumbar region*, *Hand*, and *External neck injuries* (kappa ranged from 0.76 to 0.60). Among body parts with less agreement, often codes assigned in WC described proximate body parts. Of the SOII-designated wrist injuries, 71% not coded identically in WC were coded as *Finger*, *Hand*, *Arm*, or *Multiple upper extremities*. Among SOII-classified injuries involving an unspecified part of the foot, 81% of cases those without a matching code in WC were classified in WC as *Toe(s)*, *Ankle(s)*, or some other part of the foot. Three out of four cases coded in SOII as involving the lumbar region and a different body part assigned in WC were classified as some other region of the back in WC, often *Multiple back regions* or *Back, including spine, spinal cord, unspecified*. Among SOII hand injuries, 85% of those not classified as such in WC were classified as *Finger(s)*, *Wrist(s)*, *Arm(s)*, or *Multiple upper extremities*.

Body part classifications not in agreement at the division level were often coded as injuries to *Multiple Body Parts* in one of the two data sources (Table III). In both data sources, 11% of records were classified as injuries to *Multiple Body*

TABLE III. Linked Records by WC and SOII Body Part Coding† ($n = 15,447$)

	SOII-assigned codes									
WC-assigned codes	**Head (0*)**	**Neck, including throat (1*)**	**Back, including spine, spinal cord (23*)**	**Trunk, excluding back (2* other than 23)**	**Finger(s), fingernail(s) (34)**	**Upper extremities, excluding fingers (3* other than 34)**	**Lower Extremities (4*)**	**Multiple Body Parts (8)**	**All others**	**Total**
Head (0*)	914 (6)	7 (<1)	4 (<1)	3 (<1)	7 (<1)	11 (<1)	7 (<1)	37 (<1)	1 (<1)	991 (6)
Neck, including throat (1*)	7 (<1)	163 (1)	5 (<1)	12 (<1)	0 (0)	5 (<1)	2 (<1)	50 (<1)	1 (<1)	245 (2)
Back, including spine, spinal cord (23*)	5 (<1)	43 (<1)	3,228 (21)	114 (1)	4 (<1)	8 (<1)	22 (<1)	345 (2)	2 (<1)	3,771 (24)
Trunk, excluding back (2* other than 23)	7 (<1)	8 (<1)	132 (1)	1,640 (11)	3 (<1)	38 (<1)	20 (<1)	167 (1)	19 (<1)	2,034 (13)
Finger(s), fingernail(s) (34)	6 (<1)	0 (0)	9 (<1)	3 (<1)	1,229 (8)	108 (1)	19 (<1)	6 (<1)	4 (<1)	1,384 (9)
Upper extremities, excluding fingers (3* other than 34)	5 (<1)	6 (<1)	10 (<1)	53 (<1)	155 (1)	1,766 (11)	35 (<1)	79 (1)	12 (<1)	2,121 (14)
Lower Extremities (4*)	2 (<1)	1 (<1)	24 (<1)	28 (<1)	7 (<1)	20 (<1)	2,945 (19)	64 (<1)	5 (<1)	3,096 (20)
Multiple Body Parts (8)	109 (1)	67 (<1)	149 (1)	200 (1)	12 (<1)	90 (1)	121 (1)	904 (6)	15 (<1)	1,667 (11)
All others	10 (<1)	1 (<1)	5 (<1)	15 (<1)	1 (<1)	8 (<1)	11 (<1)	8 (<1)	79 (1)	138 (1)
Total	1,065 (7)	296 (2)	3,566 (23)	2,068 (13)	1,418 (9)	2,054 (13)	3,182 (21)	1,660 (11)	138 (1)	15,447 (100)

Data presented are n (%).
*Codes with greater detail aggregated to the level indicated.
†Both the SOII and WC cases were coded using the Occupational Injury and Illness Classification System.

Parts ($P = 0.86$), however, within-case agreement was modest ($k = 0.49$).

Over ninety percent of cases were classified in SOII within three Event or Exposure divisions: *Bodily reaction and exertion* (44%); *Contact with objects and equipment* (27%); and *Falls* (19%). Within *Bodily reaction and exertion*, *Overexertion* (28%, $k = 0.77$) was most commonly assigned, followed by *Bodily reaction* (12%, $k = 0.55$), and *Repetitive motion* (4%, $k = 0.59$) (Table IV). When not assigned an identical code in WC, *Bodily reaction* cases were generally classified as *Overexertion* and SOII-designated *Overexertion* cases were often coded in WC as *Bodily reaction*.

Within the division *Contact with objects and equipment*, agreement was greatest for *Rubbed or abraded by fiction or pressure* ($k = 0.74$) and lowest for *Struck against object or equipment* (0.43). Like *Bodily reaction and exertion*, differences in coding were usually limited to the major group level, and agreement was high for the division ($k = 0.81$). Among *Falls*, agreement was greater for *Falls to lower level* ($k = 0.71$) than *Falls on same level* ($k = 0.65$). More than one in three cases classified in SOII as *Falls on same level* were classified as some other event or exposure in WC including: *Falls to lower level*; *Bodily reaction*; and *Struck against object or equipment*.

Impact of Discordant Injury Characterization on Case Estimates of Select Conditions

BLS weights each case reported in SOII to estimate the incidence of cases among the population. To examine the impact of coding differences on occupational injury and illness case estimates, we applied the SOII sample weights to reported cases of amputations and musculoskeletal disorders, identified through SOII-assigned codes or WC-assigned codes.

Among the 15,447 matched cases, 98 amputations were identified from the SOII-assigned codes and 119 from the WC-assigned codes ($k = 0.65$). Among the 119 cases classified in WC as *Amputations*, 60% were classified in SOII as *Amputations*, 18% as *Cuts, lacerations*, 8% as *Avulsions*, 6% as *Fractures*, 5% as *Crushing injuries*, and the remaining 3% as some other traumatic injury. Among the 98 SOII-identified *Amputations*, 72% appeared in WC as *Amputations*, 11% as *Multiple traumatic injuries and disorders*, 11% as some other *Open wound*, and 5% as some other traumatic injury. Applying the SOII sample weights, there were an estimated 449 amputations based on the SOII nature of injury classifications and an estimated 871 amputations based on the WC classifications, an increase of 94% over the SOII estimate. Based on classification codes assigned in either SOII or WC, 146 cases were identified as amputations, representing an estimated 985 amputations.

TABLE IV. Linked Records by WC and SOII Event or Exposure Coding† ($n = 15{,}447$)

	SOII-assigned codes								
WC-assigned codes	**Contact with objects and equipment (0*)**	**Falls (1*)**	**Bodily reaction (21*)**	**Overexertion (22*)**	**Repetitive motion (23*)**	**Exposure to harmful substances or environments (3*)**	**Transportation accidents (4*)**	**All others**	**Total**
Contact with objects and equipment (0*)	3,646 (24)	204 (1)	155 (1)	106 (1)	15 (<1)	42 (<1)	94 (1)	26 (<1)	4,288 (28)
Falls (1*)	145 (1)	2,367 (15)	190 (1)	41 (<1)	0 (0)	1 (<1)	14 (<1)	11 (<1)	2,769 (18)
Bodily reaction (21*)	76 (<1)	217 (1)	1,087 (7)	265 (2)	83 (1)	0 (0)	2 (<1)	1 (<1)	1,731 (11)
Overexertion (22*)	152 (1)	48 (<1)	290 (2)	3,601 (23)	175 (1)	5 (<1)	3 (<1)	26 (<1)	4,300 (28)
Repetitive motion (23*)	7 (<1)	5 (<1)	18 (<1)	59 (<1)	333 (2)	0 (0)	0 (0)	3 (<1)	425 (3)
Exposure to harmful substances or environments (3*)	34 (<1)	2 (<1)	2 (<1)	3 (<1)	5 (<1)	467 (3)	0 (0)	7 (<1)	520 (3)
Transportation accidents (4*)	37 (<1)	14 (<1)	5 (<1)	5 (<1)	0 (0)	1 (<1)	453 (3)	3 (<1)	518 (3)
All others	132 (1)	94 (1)	114 (1)	219 (1)	61 (<1)	51 (<1)	14 (<1)	211 (1)	896 (6)
Total	4,229 (27)	2,951 (19)	1,861 (12)	4,299 (28)	672 (4)	567 (4)	580 (4)	288 (2)	15,447 (100)

Data presented are n (%).
*Codes with greater detail aggregated to the level indicated.
†Both the SOII and WC cases were coded using the Occupational Injury and Illness Classification System.

SOII-designated coding identified 5,922 MSD cases, 1,299 more cases than the 4,623 MSD cases based on WC classifications ($k = 0.71$). Among the 2,071 cases identified as an MSD in one data source but not the other, 57% were assigned an MSD related nature of injury code but lacked an MSD related event or exposure code, 29% had an MSD related event or exposure code but lacked an MSD related nature of injury code, and 14% lacked both the event or exposure codes and the nature of injury codes used to identify MSD. After applying the SOII sample weights, the estimated number of MSD cases was 34% higher based on the SOII-assigned classification codes compared with the WC-assigned codes (32,172 estimated cases compared with 24,066 estimated cases). Based on classification codes assigned in either SOII or WC, 6,278 cases were identified as MSD, representing an estimated 34,216 MSD cases.

DISCUSSION

Linking occupational injury and illness records from multiple data sources allowed us to compare injury classification codes assigned by two different systems to the same work-related incident. This is the first study we know of to measure agreement in injury coding between SOII cases and WC claims and to assess the impact of coding differences on case estimates.

There are several possible explanations for the differences in the injury classification codes assigned in SOII compared with WC. First, the forms used to collect the descriptions of the events differ. Compare the questions posed by each system to illicit the injury narratives that are then coded.

SOII form:

- *What happened? Tell us how the injury or illness occurred. Examples: "When ladder slipped on wet floor, worker fell 20 feet," "worker was sprayed with chlorine when gasket broke during replacement:" "Worker developed soreness in wrist over time,"* and
- *What was the injury or illness? Tell us the part of the body that was affected and how it was affected; be more specific than "hurt," "pain," or "sore." Examples: "strained back;" chemical burn, hand;" "carpal tunnel syndrome."*

WC accident report:

- *Describe in detail how your injury or exposure occurred. Include tools, machinery, chemicals, or fumes that may have been involve*.
- *Part of body injured or Exposed*.

The examples provided in the SOII questions may lead the survey respondent's description of the incident and injury or illness to conform to the examples provided. Indeed, among matched cases, a greater portion was coded in SOII as "strained backed" and "carpal tunnel syndrome" while WC codes reflected a greater variety of injury types.

Another possible explanation for differences in injury assignments may relate to the individuals involved in the documentation of the incidents. An injury or illness record keeper from the sampled establishment provides the narrative description of the incident for the SOII. There may be a desire to downplay the severity of the injury (more injuries characterized in SOII as *Surface Wounds or Bruises*, injuries classified in WC as *Amputations* that were classified in SOII as *Cuts, Lacerations*, *Avulsions*, or some other injury) or, in characterizing the event, shifting blame on the worker to mitigate an employer's sense of culpability.

WC claims are coded from injured workers' narratives of the injury and event on the initial accident report. The report is completed by the worker and the attending physician, each of whom are responsible for separate sections of the form. The worker's interaction with the health care provider likely influences the worker's description of the injury. The worker's account may be more medically technical and detailed after discussing the condition with the health care provider. This may explain the higher portions of systemic diseases and multiple injuries among WC codes compared with SOII codes. Additionally, when coders are unable to classify the injury based on the accident report, they review the medical records related to the WC claim to gather the necessary information. Employer injury and illness records are unlikely to include the detail contained in medical documents, and can be expected to describe the injuries using commonplace terms rather than medical terminology.

The timing of the injury reports may offer yet another explanation for the differences in codes. Employers are required to record injuries on their OSHA logs (which provide the source data for SOII) within 7 days of recordability. In Washington, workers, with their health care providers, are entitled to file a WC claim within 1 year of injury and within 2 years of written notification of an occupational disease and the ability to file a claim. Descriptions of injuries may be modified over time as the injury is evaluated and diagnoses are refined so that the employer's description of the injury or illness at the time the case is recorded on the OSHA log may differ from the description provided by the worker at the time of claim filing.

Compared to SOII estimates, studies utilizing other sources of occupational injury and illness data including medical, hospital, WC data, or data combined from multiple sources consistently identify more cases of select conditions including amputations [Anderson et al., 2010; Friedman et al., 2013], burns [Kica and Rosenman, 2012], and musculoskeletal disorders [Silverstein et al., 1998; Lipscomb et al., 2008b]. While other factors may contribute to differences in case capture (including different case definitions, inclusion of non-surveyed populations, and underreporting to SOII), some of the discrepancy between the SOII case estimate and the case ascertainment achieved using other data sources may be

explained by discordant injury characterization. The degree of discordant injury classification can be determined only by matching complete data sources, prior to applying exclusion or selection criteria. The implication of not doing so is the possible erroneous conclusion of incomplete case capture within individual data systems.

In this study, estimates of amputations and MSD varied based on the source of the injury and illness classification. WC classifications resulted in an estimated number of amputations that was nearly twice the number of cases estimated from SOII classifications, while approximately one-third more MSD cases were estimated from SOII injury classifications compared to WC injury classifications. The greater difference in the amputation estimate may be due to the restrictiveness of the surveillance definition that is limited to a single group code. The MSD definition employed by CSTE is broader and encompasses multiple possible codes. For example, a case classified as *Sprains, strains, tear*s in one data source and as a *Back pain* in another source could be included in the estimate of MSD cases. Also, as a rare occurrence, each amputation reported constitutes a greater portion of the total number of amputation cases compared to the contribution of a single MSD case to the total number of MSD cases. Classification differences among rare events may have a substantial impact on case estimates compared to estimates of more common injuries and conditions.

There are several limitations to this study. Narrative descriptions of the injuries were not reviewed to assess the accuracy of the assigned codes. Thus, we are unable to know whether the assigned codes appropriately characterized the incident. Additionally, we were unable to assess the training provided to either group of coders to determine whether it might account for differences between SOII-assigned codes and WC-assigned codes. Another limitation is that the WC data for the study was limited to SF claims data from Washington State. Data from WC systems with different coding procedures, claim filing processes, and injury documentation may result in alternate findings. Coding agreement will likely be greater in systems with higher source dependence between SOII data and WC records, and less in systems with independent data sources.

Our record linking procedures may not have identified all true matches between SOII cases and WC claims, failing to link some true matches and, conversely, linking some false matches. False linkages would likely have dissimilar injury characteristics, artificially lowering coding agreement. True matches left unlinked because of dissimilar record linkage variables may or may not have dissimilar injury characteristics. It is unknown whether omitting these true matches improves coding agreement.

Further studies comparing the description of the injury provided in SOII to the description reported on the WC incident report, as well as any available incident or medical documentation would be able to assess whether classification differences are attributable to inter-rater coding choices, injury development that occurred between employer recording and WC claim filing, or the individual perspectives of those providing the narratives.

BLS adopted a new version of the injury and illness classification system, OIICS 2.0 followed shortly by version 2.01, beginning with 2011 SOII data [US Department of Labor, 2012b]. One of the objectives of the major revision was to increase uniformity by clarifying coding rules [Northwood et al., 2012]. This newer version of the classification system may ease the selection of codes among conditions more difficult to characterize using the old version, potentially improving the accuracy of SOII estimates and reducing some of the observed differences in injury classification by data source. Additional coder training focused on the appropriate use of non-specific and multiple injury codes may further improve injury data.

Given the differences in injury and illness classifications, surveillance efforts that compare estimates of select conditions across data sources without matching cases undoubtedly will conclude that case numbers or estimates differ by data source. Injuries reported to a system, such as SOII or WC, but classified in a way that excludes them from meeting a particular surveillance definition can make a valuable contribution to surveillance data. A multi-facetted approach that incorporates various aspects of the incident may improve case ascertainment; however it may do so at the expense of specificity.

REFERENCES

Anderson NJ, Bonauto DK, Adams D. 2010. Work-related amputations in Washington state, 1997–2005. Am J Indus Med 53:693–705.

Azaroff LS, Levenstein C, Wegman DH. 2002. Occupational injury and illness surveillance: Conceptual filters explain underreporting. Am J Public Health 92:1421–1429.

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

Council of State and Territorial Epidemiologists Updated. 2012. Occupational health indicators: A guide for tracking occupational health conditions and their peterminants. Accessed at http://www.cste.org/dnn/ProgramsandActivities/OccupationalHealth/Occupational-HealthIndicators/tabid/85/Default.aspx on November 28, 2012.

Dong XS, Fujimoto A, Ringen K, Stafford E, Platner JW, Gittleman JL, Wang X. 2011. Injury underreporting among small establishments in the construction industry. Am J Ind Med 54:339–349.

Friedman L, Krupczak C, Brandt-Rauf S, Forst L. 2013. Occupational amputations in Illinois 2000–2007: BLS vs. data linkage of trauma registry, hospital discharge, workers compensation databases and OSHA citations. Injury 44:667–673.

Kica J, Rosenman KD. 2012. Multisource surveillance system for work-related burns. J Occup Environ Med 54:642–647.

Landis JR, Koch GG. 1977. The measurement of observer agreement for categorical data. Biometrics 33:159–174.

Leigh JP, Marcin JP, Miller TR. 2004. An estimate of the U.S. Government's undercount of nonfatal occupational injuries. J Occup Environ Med 46:10–18.

Lipscomb HJ, Cameron W, Silverstein B. 2008a. Back injuries among union carpenters in Washington State, 1989–2003. Am J Ind Med 51:463–474.

Lipscomb HJ, Cameron W, Silverstein B. 2008b. Incident and recurrent back injuries among union carpenters. Occup Environ Med 65:827–834.

McNemar Q. 1947. Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika 12:153–157.

Morse T, Dillon C, Warren N, Hall C, Hovey D. 2001. Capture–recapture estimation of unreported work-related musculoskeletal disorders in Connecticut. Am J Ind Med 39:636–642.

Nestoriak N, Pierce B. 2009. Comparing workers' compensation claims with establishments' responses to the SOII. Monthly Labor Rev 132:8.

Northwood JM, Sygnatur EF, Windau JA. 2012. Updated BLS occupational injury and illness classification system. Monthly Labor Rev 135:19–28.

Pransky G, Snyder T, Dembe A, Himmelstein J. 1999. Under-reporting of work-related disorders in the workplace: A case study and review of the literature. Ergonomics 42:171–182.

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: 357–365.

Ruser J. 2008. Examining evidence on whether BLS undercounts workplace injuries and illnesses. Monthly Labor Rev 131:13.

Seligman PJ, Sieber WK, Pedersen DH, Sundin DS, Frazier TM. 1988. Compliance with OSHA record-keeping requirements. Am J Public Health 78:1218–1219.

Silverstein B, Welp E, Nelson N, Kalat J. 1998. Claims incidence of work-related disorders of the upper extremities: Washington state, 1987 through 1995. Am J Public Health 88:1827–1833.

US Bureau of the Census. 1999. Record Linkage Software User Documentation. Accessed at http://nedinfo.nih.gov/docs/US%20Census%20Bureau%20Record%20Linkage%20SW%20User%20Documentation.pdf on August 13, 2013.

US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion. 2007. Link Plus Accessed at. http://www.cdc.gov/cancer/npcr/ on September 28, 2012.

US Department of Labor, Bureau of Labor Statistics. 1992. Occupational Injury and Illness Classification Manual. Accessed at http://wwwn.cdc.gov/wisards/oiics/Doc/OIICS%20Manual%201992%20v1.pdf on August 30, 2013.

US Department of Labor, Bureau of Labor Statistics. 2007. Occupational Injury and Illness Classification Manual. Accessed at http://www.bls.gov/iif/oiics_manual_2007.pdf on March 20, 2013.

US Department of Labor, Bureau of Labor Statistics. 2012a. BLS Handbook of Methods, Chapter 9. Occupational Safety and Health Statistics. Accessed at http://www.bls.gov/opub/hom/pdf/homch9.pdf on March 20, 2013.

US Department of Labor, Bureau of Labor Statistics. 2012b. Occupational Injury and Illness Classification Manual Version 2.01. Accessed at http://www.bls.gov/iif/oiics_manual_2010.pdf on March 20, 2013.

Washington Revised Code. 1972. § 51.12.010. Accessed at http://apps.leg.wa.gov/rcw/default.aspx?cite=51.12.010 on March 20, 2013.