Just in the Wrong Place...?: Geographic Tools for Occupational Injury/Illness Surveillance

Roni A. Neff, PhD, SM,[1]* **Frank C. Curriero, PhD,**[1,2] **and Thomas A. Burke, PhD, MPH**[3]

Background *Geographic analysis is now integral to public health surveillance, but has been underused for occupational injury/illness.*

Methods *Mapping and spatial statistics are used to examine national county-level mean establishment Lost Workday Injury/Illness (LWDII) rates in the Occupational Safety and Health Administration (OSHA) Data Initiative (ODI), 1997–2001. The following questions are explored: Does occupational injury/illness vary geographically at the county level?; Does variation remain after accounting for industry hazard?; Where are rates higher or lower than expected?*

Results *The methods provide evidence of geographic variation in nonfatal occupational injury/illness rates, including after adjusting for industry hazard.*

Conclusions *Geographic analyses can improve intervention targeting, suggest risk factors for investigation, and make the case for targeting resources to prevention in hard-hit areas, as well as improving ongoing surveillance.* Am. J. Ind. Med. 51:680–690, 2008. © 2008 Wiley-Liss, Inc.

KEY WORDS: geography; spatial; surveillance; occupational injury; occupational safety; occupational health; county-level; mapping; geographic information systems; ranking

INTRODUCTION

When a worker's injury or illness is attributed to bad luck, it is often said that he or she was just "in the wrong place." But what if a place is "wrong" for a lot of workers? The more effectively public health programs can characterize the determinants, variation, and extent of health conditions, the better they will be at shaping, targeting and evaluating interventions and at communicating needs and priorities. Spatial statistics and geographic information systems (GIS) have become integral tools used in mainstream public health surveillance [Cromley and McLafferty, 2002; Krieger et al., 2002; Hillemeier et al., 2003; Krieger et al., 2003; Waller and Gotway, 2004]. However, they have been underused in the occupational injury/illness field. In particular, few reports consider variation within states [Neff, 2006]. This article demonstrates the combined use of spatial statistics methods and GIS for examining county-level occupational injury/illness data. Analyses are performed using nonfatal injury/illness rates in the Occupational Safety and Health Administration (OSHA) Data Initiative (ODI), 1997–2001.

The Centers for Disease Control and Prevention (CDC) defines public health surveillance as "the ongoing systematic collection, analysis and interpretation of outcome-specific data for use in the planning, implementation and evaluation

of public health practice'' [Thacker, 2000]. Geographic analysis (using maps and/or spatial statistics to examine distributions) improves surveillance by: increasing the ability to target programs to areas of need and to design and evaluate programs relevant to local risk factors and issues; helping policymakers, program managers, and communities understand and contextualize issues relevant to their jurisdictions; and contributing to hypothesis generation about causative factors. The U.S. has set the Healthy People 2010 objective of ''Increas[ing] the proportion of all major national, state and local health data systems that use geocoding (geographic identifiers that enable mapping) to promote nationwide use of GIS at all levels.'' The target is 90% (Objective 23-3) [U.S. Department of Health and Human Services, 2000].

The terms, ''geographic'' and ''spatial'' are often used interchangeably. We use ''geographic'' herein to denote variation across space, such as in risk factors or health outcomes. ''Area-level'' variation is that between discrete geographic areas such as counties. Area-level characteristics can reflect the aggregated effect of individual characteristics (such as the summed experience of workers in a high hazard industry dominant in an area) or the holistic effect of being in an area with particular conditions (such as a cultural view that occupational injury is normal).

Four key geographic analysis concepts are relevant. *Global spatial dependence* or *autocorrelation* means that across a whole map, nearby areas are more likely to be similar to each other than to those farther away, that is, that the variation in the map can exhibit a clustering pattern rather than being a random array. *Spatial variation in risk* refers to the extent to which risk or odds of an event varies geographically, often after controlling for key risk factors through regression. By contrast with global properties, local properties of a map focus on particular areas. *Cluster detection* involves identifying local ''hotspots'' on a map. Lastly, the viewer's own interaction with the map by comparing findings in areas of interest with those elsewhere—the so-called ''*where's my house?*' phenomenon'' [Waller and Gotway, 2004]—contributes to a map's policy implications.

Why would occupational injury/illness rates vary geographically? Our conceptual framework, developed elsewhere [Neff et al., 2006a], posits four main geographically varying risk factors: the economy; demographics; policy/culture/values; and industry. These affect business and worker incentives and capacity for safety by affecting an area's balance of power between business and workers, among other mechanisms. In turn, workplace-level behavior is affected, and is the proximal cause of injury/illness. ''Behavior'' includes actions directly related to safety and health and those with unintentional influence, such as promoting unhealthy work organization.

To promote geographic analysis, this article demonstrates geographic tools to address the following questions:

(1) Does occupational injury/illness vary geographically at the county level?
(2) What can be learned from comparing observed and expected rates by industry?
(3) Cluster Detection: Where are rates higher/lower than expected?

Finally, we ask whether it is appropriate to use the ODI for surveillance.

BACKGROUND

In the 1970s and 1980s, occupational safety and health was in the vanguard of using geographic analysis for public health surveillance and hypothesis generation [Mason et al., 1975, 1976; Stone et al., 1978; Frazier and Sundin, 1986; Pickle et al., 1987; Sundin and Frazier, 1989; Pickle et al., 1990; Kim, 1998; Devesa et al., 1999; NIOSH, 2006]. The then-new surveillance tools were used widely in public health practice and spawned numerous research studies aimed at hypothesis generation and later, hypothesis testing [Hoover and Fraumeni, 1975; Hoover et al., 1975; Blot and Fraumeni, 1976, 1977; Brinton et al., 1976; Blot et al., 1977; Stone et al., 1978; Frazier et al., 1983; Frazier and Sundin, 1986].

Since that time, geographic analysis has become mainstream throughout public health, but has received relatively little attention in occupational injury/illness research and surveillance; it was not mentioned in NIOSH's 2001 surveillance recommendations, and was not included in the Council of State and Territorial Epidemiologists' (CSTE) Occupational Health Indicators [NIOSH, 2001; CSTE, 2004; Thomsen et al., 2007]. Key barriers to states developing geographic tools include data availability and questions about appropriate aggregation and denominators, because county-level numbers of events can be small.

Some basic state-level geographic tools, particularly counts and rates of occupational injuries and fatalities, do continue to be used to inform policy and practice [NIOSH, 1991; NIOSH, 1993; Windau et al., 1998; Adekoya and Pratt, 2001; U.S. Bureau of Labor Statistics, 2003; CSTE, 2004; AFL-CIO, 2008a]. An example of the policy impact of these reports comes from NIOSH's 1993 chartbook on National Traumatic Occupational Fatality data. Its maps showing Alaska's fatality rate to be far above the U.S. rate led to a multiagency collaborative on the issue—today considered a major success story for the field [NIOSH, 1993; Smith, 2001]. Some county-level surveillance tools also remain current; advances in Internet technology have enabled the ability to allow users to generate their own county-level maps from the Cancer Atlas and respiratory mortality [NIOSH, 2006; National Cancer Institute, 2007]. At least four states have also presented county-level maps in recent years [Brooks and Davis, 1996; Stanbury et al., 2004; New Jersey

Department of Health and Senior Services, 2005; Bonauto, 2007] and New York has developed a subset of the CSTE indicators at the county level [NY State Department of Health, January 2007]. The U.S. House of Representatives Committee on Education and Labor hosts a national interactive map showing locations of about 10% of 2007 U.S. workplace fatalities [House Committee on Education and Labor, 2007].

MATERIALS AND METHODS

Data

Injury/illness data

Surveillance tools are demonstrated using the OSHA Data Initiative (ODI) as the database [OSHA Directorate of Compliance Programs, 1998, 1999, 2000, 2001, 2002]. This database, not previously described in the peer reviewed literature to our knowledge, is unique in providing national establishment-level injury/illness data. The ODI is conducted to improve OSHA's enforcement targeting, and focuses on high injury/illness rate industries. While the ODI has significant limitations, its existence and public availability following Freedom of Information Act Requests [OSHA/The Memory Hole, 2005; AFL-CIO, 2008b] present an opportunity for analysis. Variables include establishment-level Lost Workday Injury/Illness (LWDII) rates, reflecting the number of cases requiring time away from work or reassigned duties per 100 full time equivalent workers.

From 1997 to 2001, OSHA annually surveyed roughly 80,000 establishments with over 40–60 employees in industries with the highest injury/illness rates in the Bureau of Labor Statistics' Survey of Occupational Injuries and Illnesses, plus all manufacturing industries. The sample excluded construction and industries not regulated by OSHA such as mining and most government workers. Table I describes the ODI database for 1997–2001. After

2001, OSHA replaced LWDII with "Days Away from work, Restricted work or job Transfer injury and illness" (DART), so later years are not included here.

Several exclusions were made: establishments with under 60 employees; six states with low numbers: Hawaii, Minnesota, Oregon, South Carolina, Washington, and Wyoming; and the 0.1% of establishments reporting LWDII above 50 per 100 employees (due to the likelihood of error, based on the recommendations of Joseph DuBois, PhD, OSHA Office of Statistics and author judgment). We combined 5 years of data for most analyses to increase sample size. Drawbacks to combining years include inconsistent sampling, unavoidably counting some establishments more than once due to lack of unique identifiers in the ODI database, and the inability to examine temporal variation. Sensitivity analyses supported the decisions, but suggested potential further benefit from segmenting years.

Establishment locations were geocoded (linked to maps) and assigned to counties based on zip code centroids. Establishments that could not be geocoded were omitted. The outcome variable, "mean LWDII," was calculated to reflect the mean reported LWDII of all responding establishments in a county.

While some studies have shown stronger effects considering smaller geographic units [Krieger et al., 2002; Ringquist, 2005], we used county-level analysis for the following reasons. The area immediately surrounding a workplace, such as an industrial district, may have little to do with the area level risk factors affecting workers. Workers tend to come in from elsewhere, although they do tend to live within the same counties (67.4% of workers ages 16 and over worked in the county where they lived in 2000 [Census, 2000]). Counties also have a certain policy relevance, in that states often divide funds and administrative initiatives at the county level, and certain social and policy factors do vary at the county level. In the CSTE Occupational Health Indicators for states, counties were chosen as the meaningful level of analysis in those instances where sub-state level indicators are used [CSTE, 2004].

TABLE I. OSHA Data Initiative Database, 1997–2001: Establishment-Level Lost Workday Injury/Illness (LWDII) Rates in Most High-Rate Industries

Year	N (after exclusions[a])	LWDII mean (SD)	# not geocoded (%)	Response rate	National average LWDII[b]
1997	44,199	7.41 (6.68)	544 (1.2)	NA	3.3
1998	36,043	8.14 (6.77)	332 (0.9)	91%	3.1
1999	42,270	7.66 (6.45)	369 (0.9)	95%	3.0
2000	45,026	7.30 (6.21)	232 (0.5)	96%	3.0
2001	49,308	5.93 (5.65)	434 (0.9)	94%	2.8
Total	216,846	7.22 (6.38)	1,941 (0.9)	NA	

[a]Exclusions: HI, MN, OR, SC, WA, WY; establishments with <60 employees; establishments reporting LWDII > 50.
[b]National average from Bureau of Labor Statistics, covering all private industry.

Index of Area Industry Hazard

An Index of Area Industry Hazard was created to reflect the rates that would be expected based on county industry mix alone. The Index was calculated by assigning each establishment its industry mean from the Bureau of Labor Statistics (BLS), Survey of Occupational Injury and Illness (SOII) in the corresponding year. (For industry mean, 4-digit Standard Industrial Classification code was used where available, otherwise 3- or 2-digit SIC.) County average "expected" mean was then calculated.

Spatial Statistical Analyses

The statistical methods used to examine the following questions were selected primarily based on the appropriateness of the intended analysis. Since a general goal of the manuscript is to promote the use of geographic information and corresponding analysis tools, other reasons included the widespread use of these methods in other applications and the available computing resources at an accessible technical level.

Does occupational injury/illness vary geographically at the county level?

County mean occupational injury/illness rates were mapped. To examine whether there is a pattern to the observed variation or whether rates are just randomly distributed, testing for global geographic autocorrelation is performed at the county level. Moran's I, comparable to a correlation coefficient, is a widely used method [Waller and Gotway, 2004] reflecting the correlation between values in one location with those in other locations.

What can be learned from comparing observed and expected rates by industry?

Industry hazard is one of the strongest predictors of occupational injury and illness [Waehrer et al., 2004], so taking it out of the equation may help identify areas with other important factors contributing to area rates. The Index was mapped and areas where rates were greater or lesser than expected based on industry hazard were discussed.

Cluster detection: where are rates higher/lower than expected?

Local Indicators of Spatial Autocorrelation (LISAs) were calculated and mapped, assigning values to each county based on similarity with nearby counties. A related tool of interest for surveillance is ranking areas based on their injury/illness rates to aid in prioritization and to motivate intervention. We present ranks based on five methods:

(1,2)   **Direct ranks for states and for counties:** Because many counties have small numbers included in the ODI, the highest rates for direct ranks will often be based on small numbers, reflecting outliers. Three methods address this concern.

(3)   **Counties with the highest means after Bayesian shrinkage:** Louis and Shen recommend using a Bayesian statistical method to "shrink" ranks towards the mean, with those counties having the largest variance in their means getting shrunk the most [Louis and Shen, 1999]. This Bayesian rank essentially reflects risk balanced with variance. Bayesian ranking was performed separately by year; 1998 data are reported.

(4)   **Highest rates among large counties:** This method limits the ranking to counties with at least 30 establishments in the sample. While excluding almost 90% of counties, the method does focus on counties with significant industry, where intervention may be especially valuable.

(5)   **Prevention index:** The prevention index has been used in occupational safety and health to help prioritize industries for intervention [Silverstein et al., 2002; Bonauto et al., 2006]. In this method, a county's ranks for the number of establishments and for mean LWDII are averaged to create a new prevention index ranking. This index provides a way to balance population and injury/illness rates, but does not address variance, as the Bayesian method does.

Software

Mapping and some of the geographic analyses were performed using ArcGIS version 9.0 [ESRI, 2005]. Stata Intercooled version 9.1 [StataCorp, 2005] and Microsoft Excel 2003 were used for calculations. The freeware program, Geoda, was used for calculation of Moran's I and LISAs [Anselin, 2003]. The R statistical package was used to generate the Bayesian ranks, Index of Area Industry Hazard, and related analyses [R Development Core Team, 2005]. For code and detailed methods, see Neff et al. [2006b].

RESULTS

The 216,846 establishments in the 1997–2001 ODI sample were located in 2,657 of the 3,141 U.S. counties. Counties had up to 5,685 (Los Angeles) establishments, although the median was 25 and the 90th percentile was 178.

Does Occupational Injury/Illness Vary Geographically at the County Level?

Map 1 depicts the national distribution of mean LWDII for included counties. County mean Lost Workday Injury/Illness rate (LWDII) ranged from 0 to 25.2, with a mean of 7.2 (SD 2.9), excluding outlier establishments reporting LWDII over 50. Sixteen counties had mean LWDIIs of zero, all based on small numbers. There was substantially more variability between counties than within counties; the mean squared error of the LWDII between counties was 253.8, while within counties it was 38.0 ($F = 6.68$, $P = 0.00$). The Moran's I for mean county LWDII was 0.25 ($P < 0.001$), showing positive spatial autocorrelation across the whole map, and confirming the presence of spatial variation. This finding provides the foundation for further explorations of the nature and extent of the variation, and lays the groundwork for the possibility of geographically targeting interventions.

What Can be Learned From Comparing Observed and Expected Rates by Industry?

The Index of Area Industry Hazard (Map 2) showed a very different distribution of mean county "expected" rates from the ob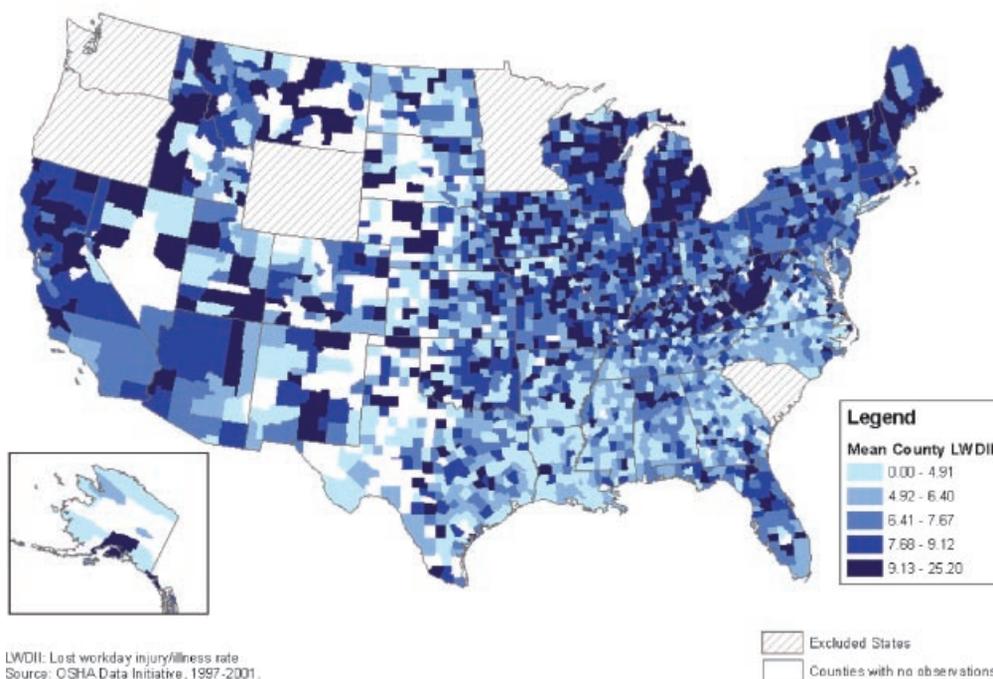served. The Index ranged from 0.7 to 12.18, with a mean of 5.97, substantially lower than the ODI mean of 7.22 for these years. As can be seen from the map, the highest expected rates were near the center of the country. The mean Index for counties in Southern states differed by only 0.3 from that for counties in non-Southern states; by contrast, in the observed data, Southern states had LWDII's 2.1 points less than in non-Southern states. Similarly, using the Index, New England, California, and Michigan-Wisconsin do not appear to have elevated rates as they do in the observed data.

Cluster Detection? Where Are Rates Higher/Lower Than Expected?

Map 3 depicts the statistically significant clusters identified in the LISA analysis. "High-high" counties (black), which have high rates and are surrounded by counties with high rates, are in Michigan-Wisconsin, Kentucky-West Virginia, Northern New England, Northern California and scattered elsewhere. Clusters of "low–low" counties (medium-gray) are seen throughout the South, particularly in Louisiana.

Tables II and III present ranking comparisons using the above-described ranking schemes. The top and bottom 10 states based on mean LWDII reflect the above-described geographic distribution, and ranged from Vermont (LWDII = 9.77) to Louisiana (LWDII = 4.98).

Mean County LWDII, Quintiles

Legend
Mean County LWDII
0.00 - 4.91
4.92 - 6.40
6.41 - 7.67
7.68 - 9.12
9.13 - 25.20

Excluded States
Counties with no observations

LWDII: Lost workday injury/illness rate
Source: OSHA Data Initiative, 1997-2001.
Excludes HI, MN, OR, SC, WA, WY; establishments with <60 employees or rates >50.

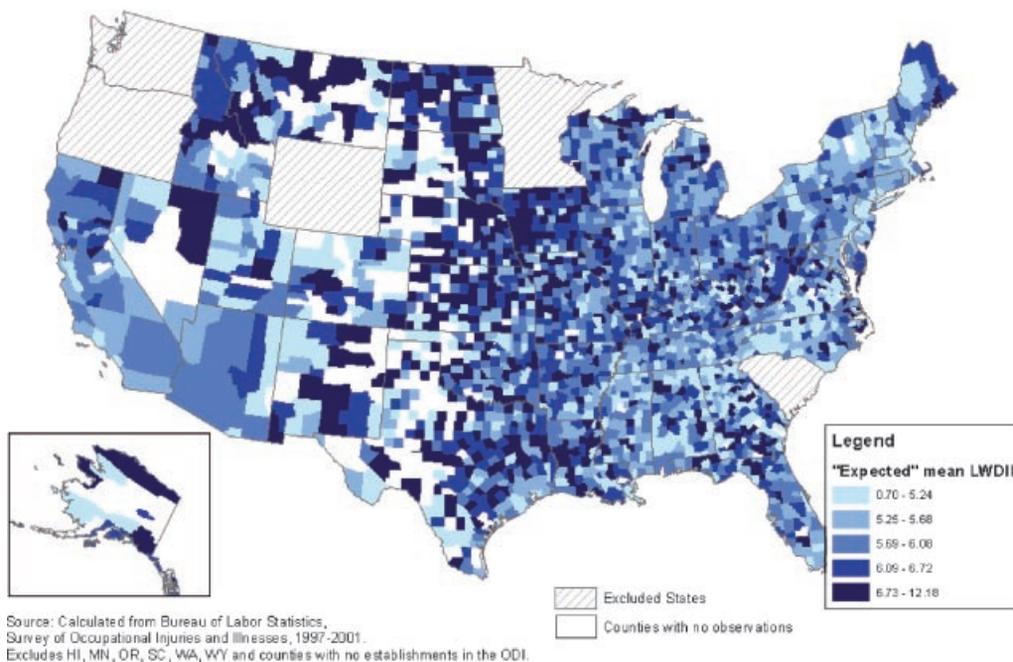**MAP 1.**

Index of Area Industry Hazard based on BLS Data

"Expected" mean county LWDII if every establishment had its industry mean

Legend

"Expected" mean LWDII

- 0.70 - 5.24
- 5.25 - 5.68
- 5.69 - 6.08
- 6.09 - 6.72
- 6.73 - 12.18

Excluded States

Counties with no observations

Source: Calculated from Bureau of Labor Statistics,
Survey of Occupational Injuries and Illnesses, 1997-2001.
Excludes HI, MN, OR, SC, WA, WY and counties with no establishments in the ODI.

MAP 2.

Clusters of High and Low Rate Counties
(Local Moran's I)

Legend

- No cluster
- High rate cluster
- Low rate cluster

Source: LWDII from OSHA Data Initiative 1997-2001.
Excludes HI, MN, OR, SC, WA, WY, counties with no establishments in the ODI, and establishments
with <60 employees or LWDII>=50. Calculated in GeoDa.
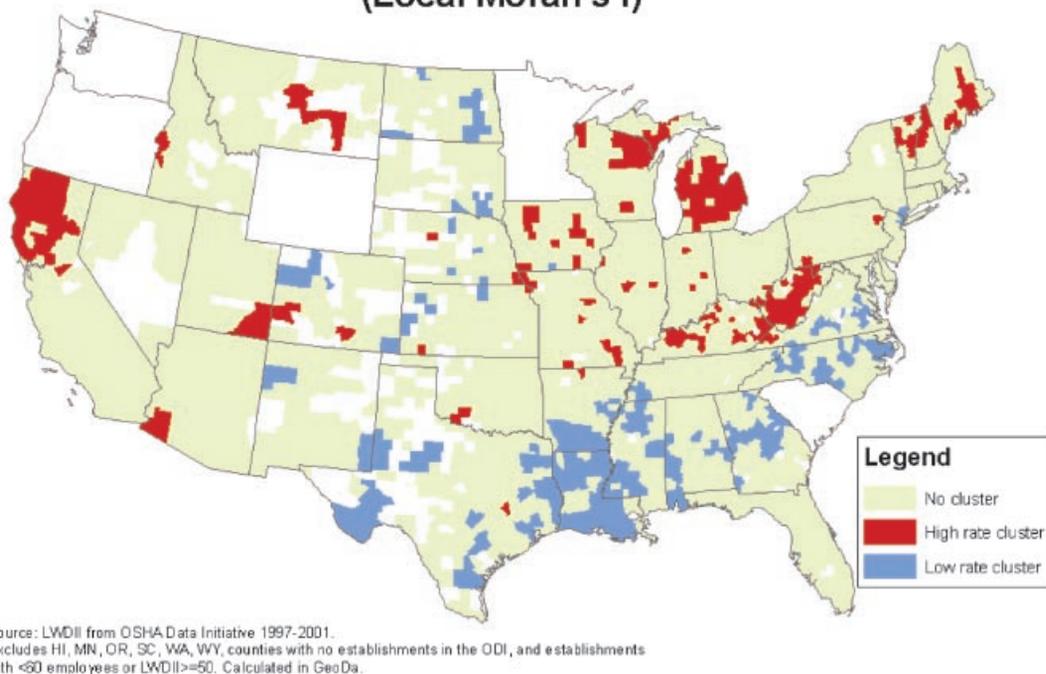
MAP 3.

TABLE II. Top- and Bottom-Ranked States

State	Mean LWDII	# Estabs in ODI
High rate states: top 10 (direct rank)		
Vermont	9.77	773
West Virginia	9.76	1,481
Michigan	9.67	10,193
Maine	9.54	1,456
Kentucky	8.99	4,453
Idaho	8.92	1,062
Iowa	8.70	3,909
Alaska	8.59	271
Wisconsin	8.54	8,818
Montana	8.47	624
Low rate states: bottom 10 (direct rank)		
Louisiana	4.98	2,836
D.C.	5.16	139
North Carolina	5.45	9,290
Delaware	5.67	639
Georgia	5.77	7,012
New Jersey	5.81	6,642
Mississippi	5.98	2,708
Arkansas	6.26	3,075
New Mexico	6.34	354
Virginia	6.40	4,463

TABLE III. County Level Ranking Comparisons Using Multiple Methods

State	County	Mean LWDII	# Estabs in ODI
Top 10 direct rank counties			
Kentucky	Carter	25.20	5
Kentucky	Edmonson	23.87	5
Wisconsin	Menominee	23.76	5
Montana	Phillips	22.89	5
Kentucky	Martin	22.57	5
West Virginia	Lincoln	20.49	1
Kentucky	McLean	20.27	5
Texas	Robertson	20.18	9
Colorado	San Miguel	20.04	1
Michigan	Gladwin	20.02	14
Top 10 Bayesian ranked counties (1998)			
Indiana	Fountain	17.53	2
New Mexico	Chaves	17.10	2
Iowa	Carroll	14.29	2
California	Plumas	13.06	3
Utah	Iron	14.37	2
Georgia	Pickens	9.83	2
Texas	Uvalde	10.51	2
Illinois	Warren	17.52	3
Iowa	Shelby	13.44	2
Michigan	Iron	18.84	2
Top 10 large counties			
Michigan	Genesee	12.58	40
Wisconsin	Sheboygan	12.53	50
Michigan	St. Clair	12.02	37
Connecticut	New London	11.96	40
Michigan	Wayne	11.78	323
Wisconsin	Rock	11.46	37
Indiana	St. Joseph	11.31	64
Wisconsin	Racine	11.31	46
Michigan	Muskegon	11.22	42
Michigan	Livingston	11.16	31
Top 10 prevention index counties (in order)			
Michigan	Wayne	10.9	1627
Georgia	Ware	10.3	40
Iowa	Black Hawk	13.2	152
Michigan	Muskegon	11.3	209
Michigan	Barry	15.8	50
Kentucky	Carter	25.2	5
Michigan	Gladwin	20.0	14
Kentucky	Edmonson	23.9	5
Wisconsin	Menominee	23.8	5
Montana	Phillips	22.9	5

Top *counties* based on mean LWDII had rates from 25.2 to 20.02 per 100 workers. Direct county means can be contrasted with 1998 Bayesian ranks. Among the counties with top 10 *Bayesian* ranks, *direct* ranks start at 38, and the distribution of included states is quite different from that in other lists. These counties have fewer ODI establishments than in the above lists because only 1 year is used and because even with low numbers, their means still had low variance.

To address these problems, another ranking was performed with the sample limited to counties with at least 30 establishments in the 1998 sample (n = 283 of 2,537). The identified counties were heavily dominated by the ''rust belt'' areas including Michigan and Wisconsin. The prevention index selected counties from similar states.

DISCUSSION

Key Findings

These analyses demonstrate the utility of studying geographic variation in occupational injury and illness. The maps and Moran's I show that reported nonfatal occupational injury/illness rates do vary by geography in the ODI. Many of the analyses found key areas of high rates in West Virginia-Kentucky, Michigan-Wisconsin, Northern New England, and Northern California. The South had low reported rates. These areas were consistently highlighted in the ''observed versus expected'' maps. Finally, the maps using the Index of Area Industry Hazard to compare observed and expected rates by

industry suggested that these high rate areas were not simply those with the highest industry-related risk factors. Other area level risk factors that might explain the elevated rates in some areas after ''taking industry out of the equation'' include social and economic factors, and geographically biased underreporting. This analysis underlines the benefit of mapping surveillance data to help generate questions for follow-up analysis.

Interpretation

The geographic distributions observed in the ODI were comparable to those seen in other statewide analyses of nonfatal occupational injury/illness, such as the distribution of LWDII reported to the SOII, as mapped in NIOSH's *Chartbook* [NIOSH, 2004]. However, both differ from the state distribution of occupational fatalities seen in the National Traumatic Occupational Fatalities Database [NIOSH, 2004], which is relatively comparable to the distribution of the Index of Area Industry Hazard—although clearly, there are different risk factors for nonfatal versus fatal injury rates. Further, as explored through mapping and multilevel modeling in Neff et al. [2006a], nonfatal events also seem to be associated with area factors including low poverty and white race. A key benefit of mapping surveillance data is to identify patterns and trends, supporting generation of hypotheses that can be explored in further research.

It is a significant possibility that geographically differential underreporting may contribute to the differences between the observed geographic distribution of events and both the expectation based on industry hazard and the distribution of fatal events. Many studies have documented widespread underreporting of nonfatal occupational injury and illness in surveillance systems, and based on the work of Leigh et al. and Rosenman et al., nationally about 2/3 of occupational injuries may be missed in surveillance systems [Pollack and Keimig, 1987; Conway and Svenson, 1998; Pransky et al., 1999; Azaroff et al., 2002, 2004; Leigh et al., 2004; Smith et al., 2005; Rosenman et al., 2006]. Azaroff et al. [2004] documented powerful economic, demographic, political, and social factors in occupational injury/illness underreporting, which may contribute to geographically differential underreporting. The ODI's enforcement usage also provides an incentive for underreporting, and concerns have been raised by the U.S. Government Accountability Office (GAO) [2005] and others.

Discussion of Methods

If selecting one measure to report for surveillance purposes, the basic map of outcomes is recommended, preferably at the county level or other sub-state aggregation. Moran's I and LISA cluster analyses help with interpretation.

The ranking analyses demonstrate that different methods produce very different ranks. These differences are to be expected based on the different criteria used in the methods, and choices among these ranking tools can be made based on conceptual priorities as well as ease of use. Direct ranks are simplest and most comprehensible but should not be used if it is expected that the top counties have high variance in their outcomes. In those cases, the Bayesian method will be more appropriate. The prevention index may be especially helpful for intervention targeting as it incorporates both industry density and risk. The analysis limited to larger counties may also be useful for targeting, where the aim is to reach as many at risk workers as possible. The Bayesian method is a useful way of evaluating risk and variance together, although it may be more technically challenging to calculate and present.

Should the OSHA Data Initiative be Used for Surveillance?

This article may provide the first broad descriptive examination of the OSHA Data Initiative in the literature. We did not have all the information needed to fully evaluate the ODI in terms of the CDC criteria for surveillance databases, however, based on what is known, we expect that ODI limitations would include the flexibility to modify it based on surveillance needs, data quality, sensitivity (based on the likelihood of significant underreporting), and representativeness (given the industries and employers not included) [CDC, 2001]. Also, the changes in sampling strategy by year make the ODI somewhat inconsistent across time. The GAO has questioned whether data limitations, particularly underreporting, make the ODI appropriate for OSHA's own purposes, never mind for surveillance [GAO, 2005].

At the same time, the ODI is the only publicly available database providing national establishment-level occupational injury/illness data. Further, the state-level distribution of its findings did not differ dramatically from those in the more scientifically valid SOII, even though the ODI focused on high injury/illness rate industries. As noted above, underreporting is not unique to the ODI. In terms of CDC criteria for surveillance systems, ODI strengths may include simplicity of use, predictive value positive, timeliness in obtaining the data, and stability [CDC, 2001]. OSHA determined that the database was adequate for a longitudinal analysis [ERG, 2004].

We suggest that ongoing analyses of data in the ODI can be used to supplement other surveillance databases, so long as data are presented with appropriate caveats. At least as importantly, county-level analyses of ODI data can help *OSHA* target areas where reported rates are higher and lower than expected, to stimulate increased intervention and/or audits.

Beyond the ODI, it is useful to consider what other tools might facilitate increased use of mapping and county level

surveillance by state occupational injury/illness surveillance programs. In discussions with surveillance experts, three key barriers were identified: data sources, aggregation, and denominator [personal communications, 2007: Letitia Davis, ScD; David Bonauto, MD, MPH; Henry Anderson, MD; John Sestito, JD; Kitty Gelberg, PhD]. A listing of potential data sources for county level analysis can be found in [Neff, RA Unpublished Dissertation http://jhir.library.jhu.edu/handle/1774.2/32401, pp. 41–44]. Given that the county level may yield small numbers for some outcomes, aggregation may be beneficial. This may be done by year or by selecting meaningful regional groupings of counties.

As for the third concern, the ideal denominator for a count of events in a county would be the number at risk in that county, including by industry or occupation. County employment by industry and time for most areas may be obtained from the Census Bureau's Local Employment Dynamics database [U.S. Bureau of the Census, 2007]. General county employment data can be found from the Census' Current Population Survey, although this survey's relatively small sample size can make estimates unstable. Richardson et al. [2004] found that the Decennial Census of Population provided a suitable substitute in many cases for calculating fatality rates. When the population at risk includes nonworkers, Census population figures, extrapolated to the relevant year, may be appropriate.

Study Limitations

Making data visually appealing and interpretable can be a drawback when appearances are both deceptive and convincing. Problems include the visual impact of large Western counties versus smaller Eastern ones and possible underreporting. The ranking computations demonstrate that results are dependent on the chosen method. Reasons for exploring county-level analysis are outlined in the methods. Results are restricted to be interpreted at the county level so as not commit errors of ecological fallacy [Waller and Gotway, 2004]. Drawbacks to using the ODI as a data source are elaborated elsewhere in the text.

Study Strengths

This study is important in that it suggests methods for bringing geographic analysis to the task of occupational injury/illness surveillance and ultimately to prevention activities. The methods identified clear geographic variation in consistent patterns across analyses. While the county level has drawbacks, counties are meaningful to the extent their policy, politics, and history have shaped the space across time, and they incorporate many commuting workers. Most importantly, they provide data relevant for policy. The presented models of geographic analysis should be readily replicable using other databases. This article may contribute

the first data description of the ODI to the peer reviewed literature and suggests its utility for surveillance and intervention, as well as highlighting underreporting concerns for follow-up.

CONCLUSION

In the U.S., there were 4.1 million reported nonfatal occupational injuries/illnesses in 2006 and 55,000 estimated occupational deaths [Steenland et al., 2003; BLS, 2007]. ''The wrong place'' is a real place where too many workers spend their days (and/or nights.)

Budgets for occupational injury/illness prevention are small and face political challenges; the AFL-CIO reports that in 2006 there were only 2,112 federal and state OSHA inspectors for 8 million workplaces nationwide [AFL-CIO, 2008a]. There is great need to identify efficient targeting schemes. Based on the findings in this article, pilot projects should explore geographic surveillance for states and OSHA should consider a pilot of intervention targeting by area. Such a program could be a time and cost efficient way to reach many high-rate establishments at once, and could change area safety norms and expectations about the certainty of enforcement. Geographic targeting also offers the opportunity for local strategies with cross-industry impact, such as worker training in workplace rights and safety, management training related to safety and health programs, and media training. Finally, by helping to visualize the occupational injury/illness distribution, geographic tools can help in advocacy to increase the overall level of resources devoted to occupational injury/illness prevention.

The occupational injury/illness field has taken little recent advantage of the explosion of new geographic tools, methods, and databases. To support geographic projects, further surveillance investments should be targeted to: expanding the georeferencing of occupational injury and illness databases as recommended in *Healthy People 2010*; developing geographic indicators to incorporated into the CSTE OHI program; and addressing underreporting.

The better we can see the lay of the land, the more effectively we can turn ''wrong'' places into just ''places.''

ACKNOWLEDGMENTS

REFERENCES

Adekoya N, Pratt SG. 2001. Fatal unintentional farm injuries among persons less than 20 years of age in the United States: Geographic profiles. DHHS (NIOSH) Publication No. 2001-131.

AFL-CIO. 2008a. Death on the job: The toll of neglect. Available at: http://www.aflcio.org/issues/safety/memorial/doj_2007.cfm. Accessed: September 3, 2007.

AFL-CIO. 2008b. Job Tracker: Health and Safety Sources. Available at: www.workingamerica.org/jobtracker/sources.cfm. Accessed: June 15, 2008.

Anselin L. 2003. GeoDa 0.9 user's guide. Spatial Analysis Laboratory, University of Illinois, Urbana-Champaign, IL. Version 0.9.5.

Azaroff LS, Lax MB, Levenstein C, Wegman DH. 2004. Wounding the messenger: The new economy makes occupational health indicators too good to be true. Int J Health Serv 34(2):271–303.

Azaroff LS, Levenstein C, Wegman DH. 2002. Occupational injury and illness surveillance: Conceptual filters explain underreporting. Am J Public Health 92(9):1421–1429.

Blot WJ, Brinton LA, Fraumeni JF, Stone BJ. 1977. Cancer mortality in U.S. counties with petroleum industries. Science 198:51–53.

Blot WJ, Fraumeni JF. 1976. Geographic patterns of lung cancer: Industrial correlations. Am J Epidemiol 103(6):539–550.

Blot WJ, Fraumeni JF. 1977. Geographic patterns of oral cancer in the United States: Etiologic implications. J Chron Dis 30(745): 757.

Bonauto D, Silverstein B, Adams D, Foley M. 2006. Prioritizing industries for occupational injury and illness prevention and research, Washington state workers' compensation claims, 1999–2003. J Occup Environ Med 48(8):840–851.

Bonauto DK. 2007. Washington State Department of Labor and Industries, Safety and Health Assessment and Research for Prevention (SHARP) Program. Nonfatal occupational injury, 2001–2003. Available at: http://www.doh.wa.gov/HWS/.

Brinton LA, Stone BJ, Blot WJ, Fraumeni JF. 1976. Nasal cancer in U.S. furniture industry counties (letter). Lancet 2:628.

Brooks DR, Davis LK. 1996. Work-related injuries to Massachusetts teens, 1987–1990. Am J Ind Med 29:153–160.

Bureau of Labor Statistics. 2007. News Release: Workplace injuries and illnesses in 2006. Available at: http://www.bls.gov/iif/oshwc/osh/os/osnr0028.txt. Accessed April 7, 2008.

CDC. 2001. Updated Guidelines for Evaluating Public Health Surveillance Systems. MMWR 50(RR13):1–35.

Committee on Education & Labor. U.S. House of Representatives. 2007. Partial reporting of worker deaths in the U.S. in 2007. 2007. Available at: http://edlabor.house.gov/issues/workerdeaths.shtml?. Accessed September 1, 2007.

Conway H, Svenson J. 1998. Occupational injury and illness rates, 1992-96: Why they fell. Mon Labor Rev 121:36–58.

Council of State and Territorial Epidemiologists in Collaboration With NIOSH. 2004. Occupational health indicators: A guide for tracking occupational health conditions and their determinants.

Cromley EK, McLafferty SL. 2002. GIS and public health. New York: The Guilford Press.

Devesa SS, Grauman DJ, Blot WJ, Pennello GA, Hoover RN, Fraumeni JF. 1999. Atlas of cancer mortality in the United States: 1950–1994. NIH Publication Number 99-4564.

ERG. 2004. Evaluation of OSHA's impact on workplace injuries and illnesses in manufacturing using establishment-specific targeting of interventions: Programmed inspections and high hazard notification letters.

ESRI. 2005. ArcGIS version 9.0. Redlands, CA.

Frazier TM, Lalich NR, Pedersen DH. 1983. Uses of computer-generated maps in occupational hazard and mortality surveillance. Scand J Work Environ Health 9: (2 Spec No):148–154.

Frazier TM, Sundin DS. 1986. Industrial demographics and population at risk for silica exposures. In: Goldsmith DF, Winn DM, Shy CM, editors. Silica, silicosis and cancer. Philadelphia: Praeger Scientific. pp 3–9.

Hillemeier MM, Lynch J, Harper S, Casper M. 2003. Measurement issues in social determinants: Measuring contextual characteristics for community health. Health Serv Res 8(6p2):1645–1717.

Hoover R, Fraumeni JF. 1975. Cancer mortality in U.S. counties with chemical industries. Environ Res 9:196–207.

Hoover R, Mason TJ, McKay FW, Fraumeni JF. 1975. Cancer by county: New resource for etiologic clues. Science 189:1005–1007.

Kim JH. 1998. Atlas of respiratory disease mortality, United States: 1982–1993. DHHS (NIOSH) Number 98-157.

Krieger N, Chen JT, Waterman PD, Soobader MJ, Subramanian SV, Carson R. 2002. Geocoding and monitoring of US socioeconomic inequalities in mortality and cancer incidence: Does the choice of area-based measure and geographic level matter? Am J Epidemiol 156:471–482.

Krieger N, Chen JT, Waterman PD, Soobader MJ, Subramanian SV, Carson R. 2003. Choosing area-based socioeconomic measures to monitor social inequalities in low birth weight and childhood lead poisoning: The public health disparities geocoding project (US). J Epidemiol Commun Health 57:186–199.

Leigh JP, Marcin JP, Miller TR. 2004. An estimate of the US government's undercount of nonfatal occupational injuries. J Occup Environ Med 46(1):10–18.

Louis TA, Shen W. 1999. Innovations in Bayes and Empirical Bayes methods: Estimating parameters, populations and ranks. Statist Med 18:2493–2505.

Mason TJ, McKay FW, Hoover R, Blot WJ, Fraumeni JF, Jr. 1975. Atlas of cancer mortality for U.S. counties: 1950–1969. DHEW Publ. No. (NIH) 75-780.

Mason TJ, McKay FW, Hoover R, Blot WJ, Fraumeni JF, Jr. 1976. Atlas of cancer mortality among U.S. non-whites: 1950–1969. DHEW Publ. No. (NIH) 76-1204.

National Cancer Institute. 2007. Customizable mortality maps. Available at: http://cancercontrolplanet.cancer.gov/atlas/index.jsp. Accessed June 30, 2007.

Neff RA. 2006. All over the map: A review of literature on geography and occupational injury and illness. Dissertation Chapter.

Neff RA, Burke TA, Curriero FC. 2006a. Social predictors of county occupational injury/illness rates. Dissertation Chapter.

Neff RA, Burke TA, Curriero FC. 2006b. "In the wrong place…?: Spatial tools for occupational injury/illness surveillance." Dissertation Chapter.

New Jersey Department of Health and Senior Services. 2005. Fatal occupational injuries in New Jersey: Ten year report 1993–2002. 2006. Available at: http://www.state.nj.us/health/eoh/survweb. Accessed April 17, 2006.

NIOSH. 1991. Work-related lung disease surveillance report. DHHS (NIOSH) Pub. No. 91-113.

NIOSH. 1993. Fatal injuries to workers in the United States, 1980–1989: A decade of surveillance. Pub. No. 93-108.

NIOSH. 2001. Tracking occupational injuries, illnesses, and hazards: The NIOSH surveillance strategic plan.

NIOSH. 2004. Worker health chartbook 2004. Pub. No. 2004-146.

NIOSH. 2006. Division of respiratory disease studies, surveillance branch. National occupational respiratory mortality surveillance system (NORMS). Available at: http://webappa.cdc.gov/ords/norms.html. Accessed: April 23, 2006.

NY State Department of Health. January 2007. Technical notes about occupational health data. Unpublished. Available at: http://www.health.state.ny.us/statistics/chac/chai/about/about_occupationalhealth.htm. Accessed June 30, 2007.

OSHA Directorate of Compliance Programs. 1998. Interim Plan for Inspection Targeting, OSHA Directive Number 98-1 (CPL 2).

OSHA Directorate of Compliance Programs. 1999. Site Specific Targeting. OSHA Directive Number 99-3 (CPL 2). Appendix A.

OSHA Directorate of Compliance Programs. 2000. Site Specific Targeting 2000 (SST-MM), OSHA Directive Number 00-05 (CPL 2). Appendix A.

OSHA Directorate of Compliance Programs. 2001. Site Specific Targeting 2001 (SST-01). OSHA Directive Number 01-01 (CPL 2). Appendix A.

OSHA Directorate of Compliance Programs. 2002. Site Specific Targeting 2002 (SST-02), OSHA Directive Number 02-02 (CPL 2). Appendix A.

OSHA/The Memory Hole. 2005. November. OSHA's lost workday injury and illness database. Available at: http://www.thememoryhole.org/osha/lwdii.htm. Accessed March 28, 2008.

Pickle LW, Mason TJ, Howard N, Hoover R, Fraumeni JF, Jr. 1987. Atlas of U.S. cancer mortality among whites: 1950–1980. DHHS Publ. No. (NIH) 87-2900.

Pickle LW, Mason TJ, Howard N, Hoover R, Fraumeni JF, Jr. 1990. Atlas of U.S. cancer mortality among nonwhites: 1950–1980. DHHS Publ. No. (NIH) 90-1582.

Pollack ES, Keimig DG. 1987. Counting injuries and illnesses in the workplace: Proposals for a better system. Washington, DC: National Academies Press.

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

R Development Core Team. 2005. R: A language and environment for statistical computing. Version 2.1.1.

Richardson D, Loomis D, Bailer AJ, Bena J. 2004. The effect of rate denominator source on U.S. fatal occupational injury rate estimates. Am J Ind Med 46(3):261–270.

Ringquist EJ. 2005. Assessing evidence of environmental inequities: A meta-analysis. J Policy Anal Manage 24(2):223–247.

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.

Silverstein B, Viikari-Juntura E, Kalat J. 2002. Use of a prevention index to identify industries at high risk for work-related musculoskeletal disorders of the neck, back, and upper extremity in Washington state, 1990–1998. Am J Ind Med 41(3):149–169.

Smith GS. 2001. Public health approaches to occupational injury prevention: Do they work? Inj Prev 7 (Suppl I):i3–i10.

Smith GS, Wellman HM, Sorock GS, Courtney TK, Pransky GS. 2005. Injuries at work in the US adult population: Contributions to the total injury burden and opportunities for prevention. Am J Public Health 95(7):1213–1219.

Stanbury M, Largo TW, Granger J, Cameron L, Rosenman K. 2004. Profiles of occupational injuries and diseases in Michigan. Available at: http://www.michigan.gov/documents/mdch_ProfilesReport_94691_7.pdf. Accessed September 28, 2005.

StataCorp. 2005. Stata statistical software, version 9. College Station, TX.

Steenland K, Burnett C, Lalich N, Ward E, Hurrell J. 2003. Dying for work: The magnitude of US mortality from selected causes of death associated with occupation. Am J Ind Med 43:461–482.

Stone BJ, Blot WJ, Fraumeni JF. 1978. Geographic patterns of industry in the United States: An aid to the study of occupational disease. J Occup Med 20(7):472–477.

Sundin DS, Frazier TM. 1989. Hazard surveillance at NIOSH. Am J Public Health 79: (Suppl):32–37.

Thacker SB. 2000. Historical development. In: Teutsch SM, Churchill RE, editors. Principles and practice of public health surveillance, 2nd edition. New York: Oxford University Press.

Thomsen C, McClain J, Rosenman K, Davis L. 2007. Indicators for occupational health surveillance. MMWR 56: (RR01):1–7.

U.S. Bureau of the Census. 2000. Summary file 3.

U.S. Bureau of the Census. 2007. QWI online [NAICS]. (September 1).

U.S. Department of Health and Human Services. 2000. Healthy people 2010.

U.S. Department of Labor, Bureau of Labor Statistics. 2003. Fatal occupational injuries in the United States, 1995–1999: A chartbook. BLS Report 965.

U.S. Government Accountability Office. 2005. Workplace safety and health: Safety in the meat and poultry industry, while improving, could be further strengthened. GAO-05-96.

Waehrer G, Leigh JP, Cassady D, Miller TR. 2004. Costs of occupational injury and illness across states. J Occup Environ Med 46:1084–1095.

Waller LA, Gotway CA. 2004. Applied spatial statistics for public health data. Hoboken, NJ: John Wiley and Sons, Inc.

Windau JA, Jack TA, Toscano GA. Summer 1998. State and industry fatal occupational injuries, 1992-96. Compensation and Working Conditions Online.