CSTE

Occupational Health
Sub-State Measures
Technical Guidance and Examples

Council of State and Territorial Epidemiologists
Occupational Health Surveillance Subcommittee
Occupational Health Sub-State Measures Workgroup
June 2016
Table of Contents

Executive Summary................................................................. 4
Introduction .................................................................................. 6
  Purpose ..................................................................................... 6
  Background ................................................................................ 6
Need for Occupational Health and Safety Data at Local and Sub-State Levels .................. 6
Potential Users of Occupational Health Sub-State Measures ........................................ 6
Generating OHIs and Other Measures at Sub-state Levels .......................................... 8
Example 1. Work-related inpatient hospitalizations and/or outpatient hospitalizations (emergency department visits) ......................................................... 9
  Displaying the Data: Example from New York State ............................................. 12
  Displaying the Data: Example from Colorado .................................................... 15
Example 2. Work-related poisonings by county or zip code .................................... 17
  Displaying the Data: Example from Florida (county-level) ................................... 21
  Displaying the Data: Example from Colorado (zip code-level) ......................... 24
Example 3. Elevated blood lead levels among adults ............................................... 28
  Displaying the Data: Example from New York State ........................................... 31
Example 4. Percentage of workers employed in industries at high risk for occupational morbidity by county ................................................................. 32
Example 5. Work-related severe traumatic injury hospitalizations .......................... 36
Example 6. Work-related injuries and illnesses reported to workers’ compensation .... 40
  Displaying the Data: Example from Massachusetts .............................................. 58
Appendix A. Topics to Consider for Sub-State Analyses ........................................... 59
Appendix B. Data Sources at Sub-state Level Relevant to Occupational Safety & Health .... 63
Appendix C. Sample SAS Code for LAUS denominator ........................................... 67
Appendix D. Sample SAS Program to Crosswalk WC employer city/town into the 48 POWPUMAs (v2000 US Census boundaries) for Massachusetts – numerator data ................. 69
Appendix E. Resources for GIS Training and Mapping ............................................ 73
Appendix F. List of Workgroup Members ......................................................... 75
Tables

Table 1. Examples of Sub-state Measures in this Technical Guidance and the Corresponding Occupational Health Indicators (OHIs) ................................................................. 4
Table 2. Measures that Have Potential for Generation at Sub-state Levels .......................... 8

Figures

Figure 1.1. Work-related injury hospitalizations per 10,000 employed persons age 16 years and older by region and counties, New York, 2007-2009 ................................................................. 12
Figure 1.2. Work-related injury hospitalizations per 10,000 employed persons age 16 years and older by region and counties, New York, 2007-2009 ................................................................. 13
Figure 1.3. Work-related injury hospitalizations per 10,000 employed persons age 16 years and older for Albany County, New York, 2006-2008 ................................................................. 14
Figure 1.4. Work-related injury hospitalizations per 100,000 employed persons age 16+ years in Region 1 compared to statewide, Colorado, 2009-2011 ................................................................. 15
Figure 1.5. Work-related emergency department injury rates per 100,000 employed persons age 16+ years in Region 1 compared to statewide, Colorado, 2011 ................................................................. 15
Figure 1.6. Work-related emergency department injury rates per 100,000 employed persons age 16+ years in Region 1 cities, Colorado, 2011 ................................................................. 16
Figure 2.1. Crude rates of occupational exposure calls reported to a poison center, by county of exposure, age 16 and older, Florida, 2003-2013 (N= 15,745) ................................................................. 21
Figure 2.2. Crude rates of occupational exposures to chemicals¥ reported to a poison center, by caller zip code, Age 16 and older, Colorado, 2000-2010 (N=1,459) ................................................................. 24
Figure 2.3. Crude rates of occupational exposures to Fumes/Gases/Vapors¥ reported to a poison center, by caller zip code, Age 16 and older, Colorado, 2000-2010 (N= 956) ................................................................. 25
Figure 2.4. Crude rates of occupational exposures to Pesticides/Fertilizers¥ reported to a poison center, by caller zip code, Age 16 and older, Colorado, 2000-2010 (N=713) ................................................................. 26
Figure 3.1. Elevated blood lead levels (≥ 10 μg/dL) per 100,000 employed persons age 16 years or older in Albany County, New York, 2003-2012 ................................................................. 31
Figure 3.2. Elevated blood lead levels (≥ 10 μg/dL) per 100,000 employed persons age 16 years or older in Albany County, New York, 2003-2012 ................................................................. 31
Figure 6.1. Annual rate of WC claims among local government workers age 16+ by place of work Public Use Microdata Areas (POWPUMA), Massachusetts, 2009-2011 (N= 8,218)* ................................................................. 58
Executive Summary

Occupational health surveillance plays an essential role in maintaining the health of the working population. The standardized Occupational Health Indicators (OHIs) currently recommended by the National Institute for Occupational Safety and Health (NIOSH) and the Centers for Disease Control and Prevention (CDC) measure statewide work-related indices. While state-level occupational health and safety data increase our understanding of the occupational health status of a statewide population, sub-state level data such as at the county or regional levels are needed to inform local decision making and action.

This document is intended to provide guidance for generating occupational health measures at the sub-state geographic level as well as provide examples of reporting and using such measures. These measures can be used to assess regional differences in health and safety risks within states. Such data may be useful to state, regional, and local health departments; federal, state, and local government officials; non-profit/advocacy organizations; health care providers; and academic research institutions.

The Council of State and Territorial Epidemiologists (CSTE) Occupational Health Subcommittee reviewed the existing list of standardized state-level OHIs and selected several measures for which sub-state level data are easily available in many states. The subcommittee adapted state OHI methodology for select OHI measures and created guidance for additional non-OHI measures to enable calculation of sub-state level indicators.

This document details the method of generating these measures and includes the discussion of identifying and accessing numerator and denominator data sources, event inclusion/exclusion criteria, steps for calculating frequencies, calculating rates and other indices, recommendations for use, and some limitations of sub-state analyses. In some cases, sample statistical code is included. Additionally, this document provides examples of ways sub-state measures are being used and reported by state health departments.

Readers are encouraged to adapt these methods to suit their specific needs. The measures presented in this document are examples and not intended to be exhaustive. Rather, the guidance provides a starting point from which states can examine and calculate OHIs and other data at the sub-state levels.

Table 1. Examples of Sub-state Measures in this Technical Guidance and the Corresponding Occupational Health Indicators (OHIs)

<table>
<thead>
<tr>
<th>Sub-state Measure</th>
<th>Corresponding OHI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example 1: Work-related hospitalizations and emergency department visits</td>
<td>OHI 2: Work-Related Hospitalizations</td>
</tr>
<tr>
<td>Example 2: Work-related poisonings by county or zip code</td>
<td>OHI 11: Acute Work-Related Pesticide-Associated Illness and Injury Reported To Poison Control Centers</td>
</tr>
<tr>
<td>Example 3: Elevated blood lead levels among adults</td>
<td>OHI 13: Elevated Blood Lead Levels among Adults</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Example 4: Percentage of workers employed in industries at high risk for occupational morbidity by county</td>
<td>OHI 14: Workers Employed in Industries at High Risk for Occupational Morbidity</td>
</tr>
<tr>
<td>Example 5: Work-related severe traumatic injury hospitalizations</td>
<td>OHI 22: Work-related Severe Traumatic Injury Hospitalizations</td>
</tr>
<tr>
<td>Example 6. Workers’ compensation claims for work-related injuries and illnesses resulting in lost work-days by place of work Public Use Microdata Areas (POWPUMAs)</td>
<td>Not Applicable</td>
</tr>
</tbody>
</table>
Introduction

Purpose
This document is intended to provide guidance and examples of generating and using Occupational Health Indicators and other measures at the sub-state geographical level. Sub-state geographic level can be by city, county or a region. This document reflects the collective efforts of several state occupational epidemiology programs exploring these methods for the first time. Readers are encouraged to adapt these methods as needed and as appropriate for their datasets and purposes.

Background
Each year, millions of workers in the United States are injured on the job or become ill as a result of exposure to health hazards at work. The burden of work-related injuries and illnesses impacts workers, employers, and the entire communities in which they live. Occupational health surveillance plays an essential role in preventing these injuries and illnesses through the collection, analysis, interpretation, and dissemination of occupational health data.

In 1998, the Council of State and Territorial Epidemiologists (CSTE) and the National Institute for Occupational Safety and Health (NIOSH) formed the NIOSH-States Occupational Health Surveillance Work Group to develop standardized indicators of occupational health. These Occupational Health Indicators (OHIs) measure statewide work-related diseases, injuries, or factors associated with health (i.e. workplace exposures, hazards, or interventions) and allow states to compare their health or risk statuses with that of other states and to evaluate trends over time.

Need for Occupational Health and Safety Data at Local and Sub-State Levels
OHIs are generally reported at the state level. However, within each state, there are regional differences in economy, industry and occupation representation, and workforce demographics which may impact health and safety risks. Local and regional health departments, non-profit organizations, and health care providers play a key role in assuring the occupational health and safety of a community. While state-level occupational health and safety data increase our understanding of the occupational health status of the state’s population as a whole, sub-state level data at the county or regional levels enable us to identify potential variations across communities within a state and inform local decision making and action.

Potential Users of Occupational Health Sub-State Measures
Occupational health surveillance programs: Occupational health measures calculated at the sub-state geographic level can help programs identify regional disparities and localized trends in work-related injuries and illnesses. It is well-known that industry distribution varies across
the state. Identifying work-related injury and illness rates by regions with high-risk industries can help prioritize areas for outreach activities.

**Local Health Agencies:** Public health agencies that represent counties and/or regions hold strong connections with their communities and can be better equipped to address community-specific health problems. These health agencies depend on localized health data to plan and prioritize occupational health activities. Sub-state occupational health measures can facilitate the development of tailored prevention and intervention efforts.

**Hospitals and health care providers:** Hospitals and health care providers are not only charged with improving the health of the patients they treat but also the communities they serve. Sub-state occupational health measures can help hospitals and health care providers better understand the occupational risks of the working patient population as well as better characterize the health and financial burdens of work-related injuries and illnesses on their communities. Localized occupational health data can also be useful for hospitals conducting community health needs assessments, a requirement for many hospitals under the Affordable Care Act.

**Town, city, and county governments:** Local government officials often hold decision-making authority in community development plans and allocation of resources for community services. A better understanding of the burden of work-related injury and illness and the occupational hazards faced by persons in their communities can help to inform decisions about business development, health protection and promotion services, built environment infrastructures, and other community planning.

**Academic researchers:** Academic programs are often at the forefront of research involving community impact and health assessment, often comparing one community to another or describing the effects of a major community or industrial development change. Data at the community level are crucial for studying and quantifying the health and health risks of community residents.

**Federal or state Occupational Safety and Health (OSHA) programs:** OSHA programs are responsible for development and enforcement of health and safety regulations as well as targeted training and outreach to high-risk employers and worker populations. Assessing health and safety outcomes data at a sub-state level can help OSHA to focus its limited resources even more efficiently on high-risk industries that experience a disproportionate burden of occupational injuries, illnesses, and/or hazardous exposures.

**Others:** Sub-state occupational health measures can also benefit a variety of other groups and individuals who plan local training and service programs or serve as a common voice for community or worker health and safety issues. This includes worker advocacy groups, unions, safety organizations, and rural health organizations.
Generating OHIs and Other Measures at Sub-state Levels

The following table lists OHIs and other measures that may be feasibly generated at sub-state geographic levels such as county or region, depending on data availability, data quality, and resources available in each state.

Table 2. Measures that Have Potential for Generation at Sub-state Levels

<table>
<thead>
<tr>
<th>Occupational Health Indicators</th>
<th>Example Guidance is provided?</th>
</tr>
</thead>
<tbody>
<tr>
<td>OHI 2: Work-Related Hospitalizations (and Emergency Department Visits)</td>
<td>Yes – Ex. 1</td>
</tr>
<tr>
<td>OHI 5: Workers’ Compensation Claims for Amputations with Lost Work-Time*</td>
<td>No</td>
</tr>
<tr>
<td>OHI 6: Hospitalizations for Work-Related Burns*</td>
<td>No</td>
</tr>
<tr>
<td>OHI 8: Workers’ Compensation Claims for Carpal Tunnel Syndrome with Lost Work-Time*</td>
<td>No</td>
</tr>
<tr>
<td>OHI 9: Hospitalizations from or with Pneumoconiosis*</td>
<td>No</td>
</tr>
<tr>
<td>OHI 10: Mortality from or with Pneumoconiosis*</td>
<td>No</td>
</tr>
<tr>
<td>OHI 11: Acute Work-related Pesticide-Associated Illness and Injury Report to Poison Control Centers</td>
<td>No - can adapt from Ex. 2</td>
</tr>
<tr>
<td>OHI 12: Incidence of Malignant Mesothelioma, Ages 15 and Older*</td>
<td>No</td>
</tr>
<tr>
<td>OHI 13: Elevated Blood Lead Levels among Adults*</td>
<td>Yes – Ex. 3</td>
</tr>
<tr>
<td>OHI 14: Workers Employed in Industries at High Risk for Occupational Morbidity</td>
<td>Yes – Ex. 4</td>
</tr>
<tr>
<td>OHI 18: OSHA Enforcement Activities</td>
<td>No</td>
</tr>
<tr>
<td>OHI 20: Work-Related Low Back Disorder Hospitalizations*</td>
<td>No</td>
</tr>
<tr>
<td>OHI 21: Asthma Among Adults Caused or Made Worse by Work* †</td>
<td>No</td>
</tr>
<tr>
<td>OHI 22: Work-related Severe Traumatic Injury Hospitalizations</td>
<td>Yes – Ex. 5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Occupational Health Measures (Non-OHIs)</th>
<th>Example Guidance is provided?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Work-related Poisonings Reported to Poison Control Centers</td>
<td>Yes – Ex. 2</td>
</tr>
<tr>
<td>Work-related Injuries and Illnesses Reported to Worker’s Compensation</td>
<td>Yes – Ex. 6</td>
</tr>
</tbody>
</table>

*For some states, small case numbers prevent analysis and presentation with a single year of data and at certain geographic areas. Combining multiple years of data or presenting at larger geographic areas may be necessary to protect confidentiality.

† Check with your state’s Behavioral Risk Factor Surveillance System (BRFSS) coordinator to determine if data are weighted at sub-state levels.
Example 1. Work-related inpatient hospitalizations and/or outpatient hospitalizations (emergency department visits)¹

Numerator: Inpatient and/or outpatient hospitalizations with primary payer coded as workers’ compensation. Geography assigned based on patient residence.

Denominator: Employed population within county or region, age 16 years or older, for the same calendar year

Measures of Frequency:

- Annual county or regional-specific number of inpatient and/or outpatient hospitalizations for persons age 16 years or older (numerator)
- Annual county or regional-specific crude rate of inpatient and/or outpatient hospitalizations per 100,000 employed persons age 16 years or older

Data resources:

- Inpatient and/or outpatient hospitalizations by county of residence - numerator
- BLS’ Local Area Unemployment Statistics (LAUS) - denominator
- Note: Both data sets base geography on patient/respondent residence.

Significance: County or regional specific counts and rates can be used to better define the pattern of work-related inpatient and/or outpatient hospitalizations throughout the state. Proportions of work-related hospitalizations from regions and counties can be compared to all of the state.

1.1 Annual number of inpatient and/or outpatient hospitalizations for persons age 16 years or older

Obtain from the State Health Department the number of cases meeting the following criteria from the inpatient and/or outpatient hospital discharge file:

- Primary payer = Workers’ Compensation.
- Limit age to those 16 years or older.
- Select for state of residence = ‘your state’.
- Select for county of residence = ‘county’.
- Exclude:
  - age unknown
  - out-of-state residents and unknown residence
  - out-of-state inpatient and/or outpatient hospitalizations

¹ Emergency department visits are referred to in this guidance as outpatient hospitalizations. These data may not be available in every state. Both inpatient and outpatient hospitalizations in this guidance follow the same inclusion criteria as the Occupational Health Indicators guidance for state-level inpatient hospitalization data, except for the addition of the ‘county’ variable. Users of this guide may wish to request an additional variable to identify inpatient and outpatient hospital discharges, or request two separate data files.
- Use data that has not been de-duplicated (no exclusions for deaths, readmissions).
- Use discharge date for calendar year, not fiscal year.
- Use all cases reported on the discharge file, regardless of length of stay.
- This will yield ‘Annual number of inpatient and/or outpatient hospitalizations for persons age 16 years or older in ________ county.’

1.2 Annual crude rate of inpatient and/or outpatient hospitalizations per 100,000 employed persons age 16 years or older

a) To obtain the denominator for the rate:

**Method A:** Instructions for using Bureau of Labor (BLS) Local Area Unemployment Statistics (LAUS) to collect Annual Average Employment Data by County:

- Go to: [http://www.bls.gov/lau/data.htm](http://www.bls.gov/lau/data.htm)
- Choose: One Screen Data Search (green button).

**Query Steps:**

- Select your state.
- Select “F” (counties and equivalent).
- Select the counties of interest - You can choose one or many. (Use the CTRL key to select more than one county). Regardless, you will get data for each county separately.
- Select “Not seasonally adjusted.”
- Click on “Get Data.”
- A new window will pop up that contains monthly and annual average data for all counties in the query for all years and months available.
- At the top of this view, click on “More Formatting Options.”
- Specify the years of interest, and set the time period to “Annual Data.”
- Click on the “Retrieve Data” button.
- A new view will display. From here, you can download each county’s data into a MS Excel file. (Note: The column that contains the denominator data is labeled “employment”).

**Method B:** To select multiple counties:

- Go to [http://www.bls.gov/lau/home.htm#cntyaa](http://www.bls.gov/lau/home.htm#cntyaa)
- Find table for “Labor Force data by county” for year of interest.
- Click on the XLS link to download the excel table of data. This will provide you with the labor force data by county for ALL of the U.S.
- Scroll down to find the rows that correspond to the counties for your state. Select the counties of interest and cut /paste into new excel worksheet. (For sample SAS code see Appendix B.)
- Use “Employed” in column H (7th column), which is the ‘**Number of employed persons age 16 years or older, by county of residence.**’
b) To calculate the rate

- Divide the numerator (number of work-related hospitalizations) by the denominator (average employment by county).
- Multiply this result by 100,000 to get the ‘Annual crude rate of inpatient and/or outpatient hospitalizations by county per 100,000 employed persons age 16 or older.’
Displaying the Data: Example from New York State

Figure 1.1. Work-related injury hospitalizations per 10,000 employed persons age 16 years and older by region and counties, New York, 2007-2009

Source: 2007-2009 SPARCS Data as of July, 2011

<table>
<thead>
<tr>
<th>Region/County</th>
<th>Work-related Injury Hospitalizations</th>
<th>Employed</th>
<th>Total</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reg-1 Western New York</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allegany</td>
<td>59</td>
<td>64</td>
<td>193</td>
<td>22,517</td>
</tr>
<tr>
<td>Cattaraugus</td>
<td>170</td>
<td>137</td>
<td>480</td>
<td>39,118</td>
</tr>
<tr>
<td>Chautauqua</td>
<td>297</td>
<td>261</td>
<td>887</td>
<td>63,423</td>
</tr>
<tr>
<td>Erie</td>
<td>1,216</td>
<td>1,203</td>
<td>3,623</td>
<td>445,328</td>
</tr>
<tr>
<td>Genesee</td>
<td>99</td>
<td>86</td>
<td>262</td>
<td>31,574</td>
</tr>
<tr>
<td>Niagara</td>
<td>403</td>
<td>386</td>
<td>1,154</td>
<td>106,083</td>
</tr>
<tr>
<td>Orleans</td>
<td>52</td>
<td>49</td>
<td>166</td>
<td>18,165</td>
</tr>
<tr>
<td>Wyoming</td>
<td>64</td>
<td>57</td>
<td>189</td>
<td>20,609</td>
</tr>
<tr>
<td>Region Total</td>
<td>2,396</td>
<td>2,263</td>
<td>2,243</td>
<td>6,956</td>
</tr>
<tr>
<td>Reg-2 Finger Lakes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemung</td>
<td>84</td>
<td>87</td>
<td>226</td>
<td>39,044</td>
</tr>
<tr>
<td>Livingston</td>
<td>56</td>
<td>66</td>
<td>188</td>
<td>30,513</td>
</tr>
<tr>
<td>Moneo</td>
<td>448</td>
<td>517</td>
<td>1,402</td>
<td>356,456</td>
</tr>
<tr>
<td>Ontario</td>
<td>89</td>
<td>82</td>
<td>269</td>
<td>53,898</td>
</tr>
<tr>
<td>Schuylert</td>
<td>23</td>
<td>32</td>
<td>83</td>
<td>9,542</td>
</tr>
<tr>
<td>Seneca</td>
<td>38</td>
<td>37</td>
<td>116</td>
<td>16,194</td>
</tr>
<tr>
<td>Steuben</td>
<td>113</td>
<td>116</td>
<td>356</td>
<td>43,016</td>
</tr>
<tr>
<td>Wayne</td>
<td>93</td>
<td>116</td>
<td>304</td>
<td>45,249</td>
</tr>
<tr>
<td>Yates</td>
<td>36</td>
<td>22</td>
<td>83</td>
<td>12,559</td>
</tr>
<tr>
<td>Region Total</td>
<td>972</td>
<td>1,075</td>
<td>2,027</td>
<td>506,471</td>
</tr>
<tr>
<td>Reg-3 Central New York</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cayuga</td>
<td>92</td>
<td>132</td>
<td>317</td>
<td>39,377</td>
</tr>
<tr>
<td>Cortland</td>
<td>39</td>
<td>49</td>
<td>134</td>
<td>22,563</td>
</tr>
<tr>
<td>Herkimer</td>
<td>74</td>
<td>55</td>
<td>186</td>
<td>29,321</td>
</tr>
<tr>
<td>Jefferson</td>
<td>91</td>
<td>93</td>
<td>269</td>
<td>45,792</td>
</tr>
<tr>
<td>Lewis</td>
<td>27</td>
<td>34</td>
<td>84</td>
<td>11,764</td>
</tr>
<tr>
<td>Madison</td>
<td>73</td>
<td>85</td>
<td>203</td>
<td>34,319</td>
</tr>
<tr>
<td>Oneida</td>
<td>207</td>
<td>235</td>
<td>674</td>
<td>105,016</td>
</tr>
<tr>
<td>Onondaga</td>
<td>331</td>
<td>433</td>
<td>1,089</td>
<td>222,786</td>
</tr>
<tr>
<td>Oswego</td>
<td>113</td>
<td>110</td>
<td>333</td>
<td>56,164</td>
</tr>
<tr>
<td>St Lawrence</td>
<td>64</td>
<td>133</td>
<td>309</td>
<td>46,594</td>
</tr>
<tr>
<td>Tompkins</td>
<td>36</td>
<td>34</td>
<td>98</td>
<td>53,750</td>
</tr>
<tr>
<td>Region Total</td>
<td>1,144</td>
<td>1,393</td>
<td>3,698</td>
<td>657,446</td>
</tr>
</tbody>
</table>

*Three years of data can be averaged to calculate rate
Figure 1.2. Work-related injury hospitalizations per 10,000 employed persons age 16 years and older by region and counties, New York, 2007-2009
### Figure 1.3. Work-related injury hospitalizations per 10,000 employed persons age 16 years and older for Albany County, New York, 2006-2008

#### Occupational Health Indicators - Albany County

**Albany County**

**2006-2008**

<table>
<thead>
<tr>
<th>Indicator</th>
<th>3 Year Total</th>
<th>County Rate</th>
<th>NYS Rate</th>
<th>Sig. Diff.</th>
<th>NYS Rate exc NYC</th>
<th>Sig. Diff.</th>
<th>Ranking Quartile</th>
<th>HP 2010 Goal</th>
<th>HP 2010 Goal Met?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incidence of Malignant Mesothelioma per 1,000,000 Persons Age 15+ (2002-08)</td>
<td>19</td>
<td>15.2&quot;</td>
<td>10.2</td>
<td>No</td>
<td>11.8</td>
<td>No</td>
<td>4th</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Hospitalization Rates per 100,000 Persons Age 15+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pneumonia (ICD-9 580-599)</td>
<td>150</td>
<td>16.0</td>
<td>13.6</td>
<td>No</td>
<td>18.8</td>
<td>No</td>
<td>3rd</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Heart Disease (ICD-9 420-429)</td>
<td>111</td>
<td>14.8</td>
<td>12.9</td>
<td>No</td>
<td>16.8</td>
<td>No</td>
<td>3rd</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Work Related Hospitalizations per 10,000 Employed Persons Age 16+</td>
<td>501</td>
<td>12.6</td>
<td>15.5</td>
<td>Yes</td>
<td>12.6</td>
<td>No</td>
<td>1st</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Elevated Blood Lead Levels per 100,000 Employed Persons Age 16+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;20 μg/dL</td>
<td>48</td>
<td>10.4</td>
<td>27.9</td>
<td>Yes</td>
<td>18.9</td>
<td>Yes</td>
<td>1st</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>&gt;25 μg/dL</td>
<td>4</td>
<td>0.9&quot;</td>
<td>4.5</td>
<td>Yes</td>
<td>3.0</td>
<td>Yes</td>
<td>1st</td>
<td>0</td>
<td>No</td>
</tr>
<tr>
<td>Fatal Work-related Injuries per 100,000 Employed Persons Age 16+</td>
<td>10</td>
<td>2.2&quot;</td>
<td>2.5</td>
<td>No</td>
<td>1.6</td>
<td>No</td>
<td>2nd</td>
<td>3.2</td>
<td>Yes</td>
</tr>
</tbody>
</table>

* Total suppressed for confidentiality

* Fewer than 20 events in the numerator; therefore the rate is unstable

---

**Note:**

- The data reflects hospitalizations per 10,000 employed persons age 16 years and older for Albany County, New York, from 2006 to 2008.
- Data includes various health indicators such as hospitalization rates, pneumonia, heart disease, work-related hospitalizations, elevated blood lead levels, and fatal work-related injuries.
- The table provides a comparison between county and NYS rates, indicating whether the difference is significant and whether the county meets the HP 2010 goals.
Displaying the Data: Example from Colorado

Figure 1.4. Work-related injury hospitalizations per 100,000 employed persons age 16+ years in Region 1 compared to statewide, Colorado, 2009-2011

![Chart showing work-related hospitalization rates per 100,000 employed persons age 16+ years in Region 1 compared to statewide, Colorado, 2009-2011.]

**Finding:** Work-related hospitalization rates are significantly higher in Region 1 of Colorado compared to the state for each year; 2009-2011.

Figure 1.5. Work-related emergency department injury rates per 100,000 employed persons age 16+ years in Region 1 compared to statewide, Colorado, 2011

![Chart showing work-related emergency department injury rates per 100,000 employed persons age 16+ years in Region 1 compared to statewide, Colorado, 2011.]

**Finding:** The work-related emergency department hospitalization rate for Region 1 is 1,719.7 per 100,000 population and significantly higher than for Colorado (625.5).
**Figure 1.6. Work-related emergency department injury rates per 100,000 employed persons age 16+ years in Region 1 cities, Colorado, 2011**

**Finding:** Work-related ED injury rates for Logan (2202.9 per 100,000 population), Washington (2801.0) and Yuma (3525.1) Counties are significantly higher compared to Region 1 (1719.7). Morgan (682.5) and Phillips (844.6) Counties are statistically significantly lower compared to Region 1.

Region 1 includes the Northeast Colorado counties of Logan, Morgan, Phillips, Sedgwick, Washington and Yuma.

*Data suppression due to small sample size.

Error bar represents the 95% Confidence Interval

**Source:** Colorado Health and Hospital Association (Emergency Department Dataset) available from the Colorado Department of Public Health and Environment.

**Numerator:** ED visits in the calendar year, ages 16 and over, workers compensation is the payer.

**Denominator:** Employed population, ages 16 and over, available from the Bureau of Labor Statistics (BLS) Local Area Unemployment Statistics (LAUS)

Websites: [www.bls.gov/lau/home.htm](http://www.bls.gov/lau/home.htm), [www.chd.dphe.state.co.us/HealthIndicators/](http://www.chd.dphe.state.co.us/HealthIndicators/)
Example 2. Work-related poisonings by county or zip code

**Numerator:** Reported cases of work-related poisonings reported to a poison center.

**Denominator:** Employed persons age 16 years and older for the same time period, as reported by the Bureau of Labor Statistics (BLS) Local Area Unemployment Statistics (LAUS) for county level rates, or zip code census population data for zip code level rates.

**Measures of Frequency:**

- Annual number of work-related poisonings for persons age 16 years or older (numerator).
- Annual crude rate of work-related poisonings per 100,000 employed persons age 16 years or older by county or by zip code.

**Significance:** The guidelines provided below are adapted from the CSTE Occupational Health Indicator (OHI #11) guidelines which provide instructions for occupational pesticide illness and injury surveillance using poison center data. These instructions differ from the original guidelines because they include occupational exposures to all substances reported to a poison center. This approach may be necessary to increase sample size for adequate county and regional analysis and to better understand variations in the burden of exposures by county or state region.

### 2.1 Annual number of work-related poisonings among persons age 16 years or older:

Obtain from the American Association of Poison Control Centers (AAPCC) National Poison Data System (NPDS)\(^2\) or your state’s Poison Center, a list of cases meeting the following criteria:

- Call Type = (exposure)
- Exposure Site = (workplace) OR Exposure Reason=(occupational), with duplicate cases removed
- For Exposure Site = (workplace), EXCLUDE: Exposure reason = 9 (suspected suicide); 11 (intentional abuse); 12 (intentional action but specific intention unknown); 14 (malicious); 18 (unknown reason)
- Medical outcome = 201 (minor effect); 202 (moderate effect); 203 (major effect); 204 (death); 206 (not followed, minimal clinical effects possible); 207 (unable to follow, judged as a potentially toxic exposure)
- Single substance exposure only = No (meaning, cases with exposure to multiple substances are also included)
- Age ≥16 years

---

\(^2\) NOTE: Poison center data access and points of contact vary by state. Visit [www.aapcc.org](http://www.aapcc.org) to find a local poison center. Online data requests can also be submitted through that website, although authors of this document have not attempted that process.
• EXCLUDE exposures to the following product categories: Bites and Envenomations, Food Products/Food Poisonings, Information Calls, Radiation

In addition to the data elements above, request at least the following data elements for each case:

• Caller state & zip code or county (name or FIPS)
• Product or substance generic code

A Note about Single- and Multiple-substance Exposures:

The CSTE OHI guidelines for occupational pesticide illness and injury surveillance (OHI # 11) suggest including only cases exposed to one substance, which is common practice when trying to identify or analyze health effects caused by a particular product. However, in order to capture all cases of occupational poisonings, both single- and multiple-substance exposures can be included for analysis. In the examples provided below, cases with multiple-substance exposures were classified according to the first-listed substance.

The following SAS code provides an example for reclassifying single- and multiple-substances cases into one of several product or substance categories for analysis. These steps have been excerpted from a larger analysis program and will require additional steps to import and clean case data prior to grouping.

Example data before & after recoding

<table>
<thead>
<tr>
<th>BEFORE recoding</th>
<th>AFTER recoding</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Multiple Substance Description</strong></td>
<td><strong>Generic Code</strong></td>
</tr>
<tr>
<td>Unknown Types of Paint, Varnish or Lacquer - 0254000 [PAINT - 3091579] ; Ketones - 0226000 [Methyl amyl ketone - 3700600] ; Other Chemicals - 0077260 [MESITYLENE - 2399130] ; Toluene and/or Xylene (Excluding Adhesives) - 0191500 [XYLENE - 3338947]</td>
<td>0254000</td>
</tr>
</tbody>
</table>

---

3 The Safe States Alliance (SSA) Injury Surveillance Workgroup 7 (ISW7) Consensus Recommendations for National and State Poisoning Surveillance ([www.safestates.org](http://www.safestates.org)) excludes exposures to bites and stings without envenomations, infections from food and waterborne sources, and exposure to external radiation in the definition of poisoning. Because environmental or internal radiation exposure could not be determined with the provided data, exposures to the “Bites and Envenomations” and “Radiation” product categories should be omitted from the final selected cases, as well those exposed to food-related categories.

4 Visit [www.aapcc.org/data-system/npds-elements](http://www.aapcc.org/data-system/npds-elements) for a more complete listing of additional data elements that may be available for analysis.

5 The AAPCC/NPDS regularly updates the product generic code list and groups all products/substances into major and minor categories. Contact your local poison center or the AAPCC for an updated pharmaceutical and non-pharmaceutical generic code index list.
*To parse generic code and substance type from substance description;

DATA clean;
   SET messy;
   type_genco=SCAN(substance, 1, '[]');
   generic_code=SCAN(type_genco, -1);
   type=SCAN(substance,1,',');
RUN;

*To group like generic codes into summary categories, identified by a current generic code list for the dataset utilized. Repeat string as needed;

DATA codecats;
   SET clean;
   LENGTH category $40;
   IF generic_code = ' ' THEN category='Missing';
   *example category1;
   ELSE IF generic_code IN ('0000001', '0000002', '0000003', '00000004', '00000005') THEN category='Category1';
   *example category2;
   ELSE IF generic_code IN ('0000006', '0000007', '0000008', '00000009', '00000010') THEN category='Category2';
   ELSE category='Other/Unknown';
RUN;

2.2 Annual crude rate of work-related poisonings per 100,000 employed persons age 16 years or older by county or zip code.

a) To obtain the denominator for calculating a COUNTY-level rate:

**Method A:** Instructions for using Bureau of Labor (BLS) Local Area Unemployment Statistics (LAUS) to collect Annual Average Employment Data by County:

- Go to: [http://www.bls.gov/lau/data.htm](http://www.bls.gov/lau/data.htm)
- Choose: One Screen Data Search (green button).

Query Steps:

- Select your state.
- Select “F” (counties and equivalent).
- Select the counties of interest - You can choose one or many. (Use the CTRL key to select more than one county). Regardless, you will get data for each county separately.
- Select “Not seasonally adjusted.”
- Click on “Get Data.”
• A new window will pop up that contains monthly and annual average data for all counties in the query for all years and months available.
• At the top of this view, click on “More Formatting Options.”
• Specify the years of interest, and set the time period to “Annual Data.”
• Click on the “Retrieve Data” button.
• A new view will display. From here, you can download each county’s data into a MS Excel file. (Note: The column that contains the denominator data is labeled “employment.”)

Method B: To select multiple counties:
• Go to http://www.bls.gov/lau/home.htm#cntyaa
• Find table for “Labor Force data by county” for year of interest.
• Click on the XLS link to download the excel table of data. This will provide you with the labor force data by county for ALL of the U.S.
• Scroll down to find the rows that correspond to the counties for your state. Select the counties of interest and cut/paste into new excel worksheet. (For sample SAS code see Appendix B.)
• Use “Employed” in column H (7th column), which is the ‘Number of employed persons age 16 years or older, by county of residence.’

b) To calculate the COUNTY rate

• Divide the numerator (number of work-related poisonings) by the denominator (average employment by county).
• Multiply this result by 100,000 to get the ‘Annual crude rate of work-related poisonings by county per 100,000 employed persons age 16 or older.’
Displaying the Data: Example from Florida (county-level)

Appendix E lists some resources for geographic information system (GIS) mapping.

Figure 2.1. Crude rates of occupational exposure calls reported to a poison center, by county of exposure, age 16 and older, Florida, 2003-2013 (N= 15,745)

Data source: Florida Poison Information Center Network (FPICN)
Numerator: Reported cases of work-related exposure from Florida FPICN case-level data, Age 16 and older
Denominator: County-level populations based on 2003-2013 Local Area Unemployment Statistics data
Rates are per 100,000 employed population in county
Prepared by: Florida Department of Health using ArcGIS
To obtain the denominator for calculating a ZIP CODE-level rate:

There are numerous private market products that provide population data by zip code, as well as various software systems to summarize geographic information. This example uses publicly available ZIP Code™ Tabulation Areas (ZCTAs™) from the United States Census Bureau. According to the Census Bureau website “ZIP Code Tabulation Areas (ZCTAs) are generalized areal representations of United States Postal Service (USPS) ZIP Code service areas. The USPS ZIP Codes identify the individual post office or metropolitan area delivery station associated with mailing addresses. USPS ZIP Codes are not areal features but a collection of mail delivery routes. The term ZCTA was created to differentiate between this entity and true USPS ZIP Codes. ZCTA is a trademark of the U.S. Census Bureau; ZIP Code is a trademark of the U.S. Postal Service.” For more information see https://www.census.gov/geo/reference/zctas.html

- Go to the United States Census Bureau, American FactFinder, http://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml
- Choose ‘Advanced Search’. A page similar to the one below should open.

( Accessed April 2016)

- On the left side of the screen choose ‘Geographies’. In the ‘Select Geographies’ window, under ‘Select a geographic type:’ choose “5-Digit ZIP Code Tabulation Area- 860.”

Note: 5-Digit ZIP Code Tabulation Area-860 is available from the 5 year-American Community Survey (ACS) estimates starting with the year 2011.

- Select your state.
- Select your ZCTAs. (You can select all ZCTAs in your state by selecting the first option or multiple individual ZCTAS by holding down the Ctrl key when clicking on the ZCTA).
- Select the ‘ADD TO YOUR SELECTIONS’ button.
• Close the ‘Select Geographies’ window.

(Accessed April 2016)

• In the ‘Refine your search results:’ box type in ‘B23025’ under ‘topic or table name’ to obtain the “EMPLOYMENT STATUS FOR THE POPULATION 16 YEARS AND OVER” tables. Submit the query by selecting ‘GO.’

(Accessed April 2016)

• Choose the table that corresponds to your year of interest.
- Under ‘Download’, download the data as a comma-separated values (.csv) file. The csv will download as a zipped file.
- In the CSV file the ZCTA is stored in column B ‘GEO.id2) and the employed persons estimate is in column J, ‘HDO1_VD04’ (Estimate; in labor force: - Civilian labor force: - Employed).

To calculate the rate by ZIP CODE

- Divide the numerator (annual number of work-related poisonings by zip code) by the denominator (2010 Census Population, Age 15-84, by zip code).
- Multiply this result by 100,000 to get the ‘Annual crude rate of work-related poisonings by county per 100,000 persons age 15 or older.’

Displaying the Data: Example from Colorado (zip code-level)

Figure 2.2. Crude rates of occupational exposures to chemicals reported to a poison center, by caller zip code, Age 16 and older, Colorado, 2000-2010 (N=1,459)

Source: Rocky Mountain Poison and Drug Center (RMPDC) data from the National Poison Data System (NPDS)
Numerator: Reported cases of work-related exposure from Colorado RMPDC case-level data in NPDS, Age 16 and older
Denominator: ZIP code populations based on 2010 ESRI, multiplied by 11 to reflect number of poisoning data years. Includes all ages.
Rates are per 100,000 population in ZIP code
Prepared by: Colorado Occupational Health and Safety Surveillance Program and Health Statistics Section, Colorado Department of Public Health and Environment (CDPHE)
“Chemicals” refers to those products included in the study grouping called “Chemicals”. See Appendix B of the CDPHE report on work-related exposures reported to a poison center, 2000-2010 available online: www.colorado.gov/pacific/cdphe/workplace-safety-data-and-reports.

Figure 2.3. Crude rates of occupational exposures to Fumes/Gases/Vapors¥ reported to a poison center, by caller zip code, Age 16 and older, Colorado, 2000-2010 (N= 956)

Data source: Rocky Mountain Poison and Drug Center (RMPDC) data from the National Poison Data System (NPDS)
Numerator: Reported cases of work-related exposure from Colorado RMPDC case-level data in NPDS, Age 16+
Denominator: ZIP code populations based on 2010 ESRI, multiplied by 11 to reflect number of poisoning data years. Includes all ages.
Rates are per 100,000 population in ZIP code
Prepared by: Colorado Occupational Health and Safety Surveillance Program and Health Statistics Section, Colorado Department of Public Health and Environment (CDPHE)
“Fumes/Gases/Vapors” refers to those products included in the study product grouping called Fumes/Gases/Vapors”. See Appendix B of the CDPHE report on work-related exposures reported to a poison center, 2000-2010 available online: www.colorado.gov/pacific/cdphe/workplace-safety-data-and-reports.
Figure 2.4. Crude rates of occupational exposures to Pesticides/Fertilizers\(^¥\) reported to a poison center, by caller zip code, Age 16 and older, Colorado, 2000-2010 (N=713)

Data source: Rocky Mountain Poison and Drug Center (RMPDC) data from the National Poison Data System (NPDS)
Numerator: Reported cases of work-related exposure from Colorado RMPDC case-level data in NPDS, Age 16 +
Denominator: ZIP code populations based on 2010 ESRI, multiplied by 11 to reflect number of poisoning data years. Includes all ages.
Rates are per 100,000 population in ZIP code
Prepared by: Colorado Occupational Health and Safety Surveillance Program and Health Statistics Section, Colorado Department of Public Health and Environment (CDPHE)
\(^¥\) "Pesticides/Fertilizers" refers to those products included in the study product grouping called “Pesticides/Fertilizers”. See Appendix B of the CDPHE report on work-related exposures reported to a poison center, 2000-2010 available online: www.colorado.gov/pacific/cdphe/workplace-safety-data-and-reports. Products identified by NPDS as fertilizers comprise less than 3% of this grouping.

Example Zip Code Rate Interpretation:

Occupational exposures to the study product groupings Chemicals, Fumes/Gases/Vapors, and Pesticides/Fertilizers were more commonly called in from Colorado’s urban areas, but rates were higher in rural areas. Zip codes located in western, southwestern, southern-central, and eastern Colorado had the highest rates of Chemical and Gas/Fume/Vapor exposures. Geographic distribution differed for reporting of Pesticide/Fertilizer exposure, which had higher rates along Colorado’s eastern state lines. Reasons for higher reporting rates in rural areas are unclear and need to be further investigated.
Example Zip Code Rate Limitation Description:

Though these data reflect the best possible geographic representation of reported occupational poison exposure in Colorado at the time, geo-coding with zip code information presents some challenges to interpretation. Zip codes are assigned according to the caller’s address and may not reflect site of the workplace exposure. Zip code boundaries are a construct of the United States Postal Service used to facilitate mail delivery and frequently change. Although occupational exposures occurred over a period of 11 years, the maps presented were generated using 2010 zip code boundary estimates. Furthermore, the only available population denominator data for the 2010 zip code boundaries included all ages, while cases in the numerator represented only those age 16 and older. This discordance in numerator-denominator populations would likely result in calculated crude rates that are lower than true rates, but may also skew rates if age distribution significantly differs across zip codes.
Example 3. Elevated blood lead levels among adults

**Numerator:** All reported county residents age 16 years or older, with a blood lead level \( > 10 \mu g/dL \), \( > 25 \mu g/dL \), and \( > 40 \mu g/dL \)

**Denominator:** Employed population, age 16 years or older, for the same calendar year

**Measures of Frequency:**
- Annual number of county residents with elevated blood lead levels (numerator)
- Annual prevalence rate per 100,000 employed persons age 16 years or older
- Annual number of incident cases of county residents with elevated blood lead levels
- Annual incidence rate per 100,000 employed persons age 16 years or older

**Data resources:**
- Reports of elevated blood lead levels from laboratories - numerator
- BLS’ Local Area Unemployment Statistics (LAUS) - denominator

**Recommendations/Limitations:** Reported elevated blood lead levels may not be considered an accurate reflection of population trends in workplace lead poisonings because a high percentage of blood lead testing and detection of overexposures relies on companies complying with the OSHA lead standard and testing their employees, which is often not done. However, groups of elevated blood lead levels can help identify where and how individuals were exposed. For example, a company with a good lead screening program will do regular testing so the county in which that company is located will have a higher rate of elevated BLL cases than those counties with companies that do NOT regularly test employees. Also county of residence may not necessarily indicate county of exposure because many people travel out of county for their work. Additionally, it is estimated that approximately 10-15% of all elevated BLLs among adults can be caused by non-occupational exposures. Not all states may be able to distinguish occupationally exposed individuals from non-occupationally exposed individuals so both could be potentially included in the calculations. This may lead to an overestimate of work-related cases.

3.1. Annual number of county residents with elevated blood lead levels \( (\geq 10 \mu g/dL, \geq 25 \mu g/dL \ or \geq 40 \mu g/dL) \), age 16 and older

- Contact your state’s Adult Blood Lead Surveillance program. A list of state contacts can be found at: [http://www.cdc.gov/niosh/topics/ABLES/ables.html](http://www.cdc.gov/niosh/topics/ABLES/ables.html)
- Request data according to the following criteria for the calendar year:
  - Blood lead level (BLL) \( \geq 10 \mu g/dL \) (or \( \geq 25 \mu g/dL \) or \( \geq 40 \mu g/dL \))
  - Age 16 years and older
  - Earliest date of either draw date, date laboratory received sample, or date laboratory analyzed sample
If a person is reported more than once during the time period, count that person one time only, at his/her highest BLL
- Include all cases, both occupationally and non-occupationally exposed
- Include all residents of state and unknown residence, but exclude out-of-state residents
- Request county of residence

- Once data are sorted accordingly by county, this will yield the ‘Annual number of _______ (name of county/region) residents with blood lead levels ≥ 10 μg/dL, age 16 and older.’

**Note:** if the cell size of the individual county data is too small, the data can be combined in a few ways; by year (i.e. - 2008-2010), or by region (i.e. - in New York State Region 2 consists of 4 counties, Livingston, Monroe, Steuben, and Yates). The goal for combining multiple years or counties is to yield data that is large enough to release without confidentiality issues.

### 3.2. Annual prevalence rate per 100,000 employed persons

a) **To obtain the denominator for the rate:**

**Method A:** Instructions for using Bureau of Labor (BLS) Local Area Unemployment Statistics (LAUS) to collect Annual Average Employment Data by County:

- Go to: [http://www.bls.gov/lau/data.htm](http://www.bls.gov/lau/data.htm)
- Choose: One Screen Data Search (green button).

**Query Steps:**

- Select your state.
- Select “F” (counties and equivalent).
- Select the counties of interest - You can choose one or many. (Use the CTRL key to select more than one county). Regardless, you will get data for each county separately.
- Select “Not seasonally adjusted.”
- Click on “Get Data.”
- A new window will pop up that contains monthly and annual average data for all counties in the query for all years and months available.
- At the top of this view, click on “More Formatting Options.”
- Specify the years of interest, and set the time period to “Annual Data.”
- Click on the “Retrieve Data” button.
- A new view will display. From here, you can download each county’s data into a MS Excel file. (Note: The column that contains the denominator data is labeled “employment.”)

**Method B:** To select multiple counties:

- Go to [http://www.bls.gov/lau/home.htm#cntyaa](http://www.bls.gov/lau/home.htm#cntyaa)
- Find table for “Labor Force data by county” for year of interest.
• Click on the XLS link to download the excel table of data. This will provide you with the labor force data by county for ALL of the U.S.
• Scroll down to find the rows that correspond to the counties for your state. Select the counties of interest and cut/paste into new excel worksheet. (For sample SAS code see Appendix B.)
• Use “Employed” in column H (7th column), which is the ‘Number of employed persons age 16 years or older, by county of residence.’

b) To calculate the rate
• Divide the numerator (number of cases of elevated blood lead levels by county) by the denominator (average employment by county).
• Multiply this result by 100,000 to get the ‘Annual prevalence rate by county per 100,000 employed persons age 16 years or older.’

3.3. Annual number of incident cases
Use the ABLES definition of an incident case: Case with a blood lead level of >=10 μg/dL reported in the calendar year, but was not reported in the immediately preceding year with a blood lead level of >=10 μg/dL (may appear in earlier years with a BLL of 10 μg/dL or greater).

3.4. Annual incidence rate per 100,000 employed persons
a) To obtain the denominator for the rate:
• Follow directions for 3.2.

b) To calculate the rate:
• Divide the numerator (3.3) by the denominator (3.2a).
• Multiply this result by 100,000 to get the ‘Annual incidence rate per 100,000 employed persons age 16 years or older’.

Note: the same procedure can be used to calculate cases for ≥25 μg/dL and ≥40 μg/dL.
*As explained earlier, data can be combined in different ways if individual cell sizes are too small.
Displaying the Data: Example from New York State

Figure 3.1. Elevated blood lead levels (≥ 10 μg/dL) per 100,000 employed persons age 16 years or older in Albany County, New York, 2003-2012

<table>
<thead>
<tr>
<th>Year</th>
<th>Single Year</th>
<th>3-Year Average</th>
<th>NYS exc. NYC</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>22.4</td>
<td></td>
<td>34.2</td>
</tr>
<tr>
<td>2004</td>
<td>16.9</td>
<td>20.9</td>
<td>28.6</td>
</tr>
<tr>
<td>2005</td>
<td>23.3</td>
<td>19.1</td>
<td>30.8</td>
</tr>
<tr>
<td>2006</td>
<td>16.9</td>
<td>17.8</td>
<td>28.1</td>
</tr>
<tr>
<td>2007</td>
<td>13.0</td>
<td>13.9</td>
<td>25.7</td>
</tr>
<tr>
<td>2008</td>
<td>11.7</td>
<td>13.0</td>
<td>24.0</td>
</tr>
<tr>
<td>2009</td>
<td>14.2</td>
<td>14.4</td>
<td>22.4</td>
</tr>
<tr>
<td>2010</td>
<td>17.3</td>
<td>14.0</td>
<td>25.7</td>
</tr>
<tr>
<td>2011</td>
<td>10.5</td>
<td>13.6</td>
<td>24.6</td>
</tr>
<tr>
<td>2012</td>
<td>13.1</td>
<td></td>
<td>20.7</td>
</tr>
</tbody>
</table>

*NYS exc. NYC: New York State excluding New York City

† The 3-year average is a running average that includes the surrounding 2 years of data (for example, the 2010 3-year average would include data from 2009, 2010 and 2011). It helps smooth out short-term fluctuations in the data.

Figure 3.2. Elevated blood lead levels (≥ 10 μg/dL) per 100,000 employed persons age 16 years or older in Albany County, New York, 2003-2012

*NYS exc. NYC: New York State excluding New York City
The 3-year average is a running average that includes the surrounding 2 years of data (for example, the 2010 3-year average would include data from 2009, 2010 and 2011). It helps smooth out short-term fluctuations in the data.

Example 4. Percentage of workers employed in industries at high risk for occupational morbidity by county

**Numerator:** Employed persons age 16 years or older, in private sector industries at high risk for occupational morbidity.

**Denominator:** Employed persons age 16 years or older, in all private industries for the same calendar year

**Measures of Frequency:**
- Number of employed persons in high morbidity-risk NAICS industries
- Percentage of employed persons in high morbidity-risk NAICS industries

**Data resources:**
- U.S. Census Bureau County Business Patterns (CBP) (numerator and denominator)

**Significance:** Work-related injuries and illnesses are preventable, and control of occupational hazards is the most effective means of prevention. Identifying where higher concentrations of workers in high-risk industries are employed can help prioritize prevention efforts and limited resources.

**Recommendations:** High-risk industries are provided in the OHI How-To Document for OHI 14, available at [http://www.cste.org/group/OHIndicators](http://www.cste.org/group/OHIndicators). These high morbidity-risk industries are industries with non-fatal injury and illness rates at least two times higher than the U.S. rate, based on the BLS Survey of Occupational Injury and Illness. Due to differences in regional industrial practices, high-risk industries within a specific state might differ from those identified from national data. **Therefore, states should consider identifying their own list of state-specific high-risk industries when calculating the county-level percentage of workers in high morbidity-risk industries.**

**Limitations:** The CBP is based on mid-March payrolls of all employers in the U.S. but does not cover farms, public administration, or the self-employed. Exact employment counts for a particular NAICS industry may not be provided within a State because of confidentiality issues. It is possible that some new employers are not counted in the CBP mid-March survey.

**4.1 Number of employed persons in high morbidity-risk industries by county.**

a) [Download the denominator data](http://www.cste.org/group/OHIndicators)

In the top bar under the page header, click DOWNLOAD CENTER.

Make sure the radio button is selected for **I know the dataset or table(s) that I want to download**. Click next.

In the “Select a program” box, select **Business Patterns**, then select the dataset year you want. Click **ADD TO YOUR SELECTIONS** then click next.

In the geography dropdown box, select **County - 050** and then select the desired state. Select **All Counties within [state]**. Click **ADD TO YOUR SELECTIONS** then click next.

You should see two table names. Click on the checkbox for the **Geography Area Series: County Business Patterns**. Once the table is checked, click on Next. In the download popup, uncheck the “Include descriptive data element names” box. Select OK.

The download could take a several minutes. Once the download is ready, save the .zip file and extract the zip file using WinZip or other unzipping utility tool.

You should have the following four files downloaded:

- aff_download_readme_ann.txt
- BP_year_00A1.txt
- BP_year_00A1_metadata.csv
- BP_year_00A1_with_ann.csv

**Summarize/prepare denominator data in SAS:**

**Note:** This method uses SAS 9.3. It is possible to generate this county-level data using other statistical software or Excel; however these directions are not shown.

Use Proc Import to import the data file called “BP_year_00A1_with_ann” into SAS and name dataset ‘c1’. Do not use the import wizard, as it might truncate variables and lead to import errors.

```
proc import datafile="PATH WHERE DATA ARE SAVED\FILE NAME.csv"
  dbms=csv
  replace
  out=c1;
  guessingrows=32767;
  getnames=yes;
run;
```

Create new dataset and new variables for use in later steps. Separate the GEO_id2 variable into 2 columns (state fips code and county fips code).

```
data c2;
  set c1;
  NAICS=NAICS_id;
  state_fips=substrn(GEO_id2,1,2);
  county_fips=substrn(GEO_id2,3,3);
```
county_name=substr(GEO_display_label,1,index(GEO_display_label,'County')-1);
if county_fips ne "000";

- For a subset of industries, the EMP (number of paid employees) variable is represented as a categorical variable representing a range. The ranges are defined in a text document included in the downloaded zip file called “BP_year_00A1.text”. In SAS, amend the SAS code below according to those reference values and ranges. Create a new variable (EMP2) as the midpoint of the range.

```sas
if EMP="a" then EMP2=(0+19)/2;
else if EMP="b" then EMP2=(20+99)/2;
else if EMP="c" then EMP2=(100+249)/2;
else if EMP="e" then EMP2=(250+499)/2;
else if EMP="f" then EMP2=(500+999)/2;
else if EMP="g" then EMP2=(1000+2499)/2;
else if EMP="h" then EMP2=(2500+4999)/2;
else if EMP="i" then EMP2=(5000+9999)/2;
else if EMP="j" then EMP2=(10000+24999)/2;
else if EMP="k" then EMP2=(25000+49999)/2;
else if EMP="D" then EMP2=0;
else if EMP="S" then EMP2=0;
else EMP2=EMP;
if EMP2 ne .;
run;
```

- Subset the dataset by selecting only high-risk industries.

```sas
data c3;
set c2;
if NAICS in ("enter each NAICS code here within quotations");
run;
```

- Sum the number of employees (EMP2) for each county by using PROC MEANS or another equivalent procedure. The resulting table should give you the “Number of employed persons in high morbidity-risk industries by county.”

```sas
proc means data=c3 sum;
output out=c4 (where=(_type_ = 3) drop=_freq_) sum=EMP2;
var EMP2;
class county_fips county_name;
run;
```

### 4.2 Percent of employed persons in high morbidity-risk industries by county

a) To obtain the denominator for the rate:
• The denominator can be calculated using the County Business Patterns dataset already downloaded in step 1.1a, as this dataset includes the number of employees in all sectors by county. The total number of employees in each county is identified in rows where the NAICS code is “00.”

• To summarize the total number of employees in each county, use the c2 dataset created in step 1.1b.

• In SAS, create a new dataset (d1) and a new variable (emp_total) that equals EMP2. Only keep the following variables: naics, county_fips emp_total, county_name. This new dataset (d1) includes the total number of employers in each county.

```sas
data d1;
  set c2;
  if NAICS='00';
  emp_total=EMP2;
  keep county_fips emp_total county_name naics;
run;
```

b) To calculate the percent of high-risk industries in each county at once:

• Sort the d1 (denominator) and c4 (numerator) datasets by county_fips.

```sas
proc sort data=c4;
  by county_fips;
run;
proc sort data=d1;
  by county_fips;
run;
```

• Merge numerator and denominator tables by county_fips. Create calculated variable “percent” that equals (emp2 / emp_total)*100.

```sas
data m1;
  merge c4 d1;
  by county_fips;
  percent=((EMP2/emp_total)*100);
  drop _type_;
run;
```

• The new merged table should show the ‘Percent of employed persons in high morbidity-risk industries by county.’

• Export the table to Excel for further formatting and preparation for presentation or mapping.

```sas
proc export data=m1 outfile= 'PATH WHERE YOU WANT THE DATA SAVED/m1.xlsx';
run;
```
Example 5. Work-related severe traumatic injury hospitalizations

**Numerator:** County or regional-specific inpatient hospital discharges with (1) primary payer coded as workers’ compensation and (2) first-listed diagnosis contained in the specified list of severe traumatic injuries. Geography assigned based on patient residence.

**Denominator:** Employed population within county or region, age 16 years or older, for the same calendar year

**Measures of Frequency:**
- Annual county or regional-specific number of inpatient hospitalizations for persons age 16 years or older (numerator)
- Annual county or regional-specific crude rate of inpatient hospitalizations per 100,000 employed persons age 16 years or older

**Data resources:**
- Numerator: Inpatient discharge data by county of residence
- Denominator: BLS’ Local Area Unemployment Statistics (LAUS)

**Significance:** County or regional specific counts and rates can be used to better define the pattern of work-related hospitalizations throughout the state. Proportions of work-related hospitalizations from regions and counties can be compared to all of the state.

5.1 Annual number of work-related inpatient hospitalizations for severe traumatic injury for persons age 16 years or older by county or region

For the numerator, obtain from the State Health Department the number of cases meeting the following criteria from the inpatient hospital discharge file:

- First listed diagnosis is a severe traumatic injury estimated to have an Abbreviated Injury Scale (AIS) severity of 3 or above or has a high probability of hospital admission. The list below includes diagnostic codes in the range 800–959.9.
- The list **excludes** the following diagnostic codes: 905.x–909.x (late effects of injury), 910.x–924.x (superficial injuries), 930.x–939.x (foreign bodies), 940.x-949.x (burns, which are not well-characterized by AIS-based severity measures), and 958.x (traumatic complications).

List of ICD-9-CM codes for severe traumatic injuries ("x" means that all subcodes are also included):

<table>
<thead>
<tr>
<th>800.x</th>
<th>807.18</th>
<th>824.5</th>
<th>874.5</th>
<th>926.x</th>
</tr>
</thead>
<tbody>
<tr>
<td>801.x</td>
<td>807.2</td>
<td>824.7</td>
<td>887.x</td>
<td>927.x</td>
</tr>
<tr>
<td>803.x</td>
<td>807.3</td>
<td>850.2</td>
<td>896.x</td>
<td>928.x</td>
</tr>
<tr>
<td>804.x</td>
<td>807.4</td>
<td>850.3</td>
<td>897.x</td>
<td>929.x</td>
</tr>
</tbody>
</table>
Sample Stata code for diagnosis codes of interest:

generate dxflag= ///
/*diagnoses specified to 3 digits*/ ///
inlist(substr(dx,1,3),"800","801","803","804","805","806","808","820","821") | ///
inlist(substr(dx,1,3),"851","852","853","854") | ///
inlist(substr(dx,1,3),"860","861","863","864","865","866","867","896","897") | ///
inlist(substr(dx,1,3),"900","901","902","925","926","927","928","929","952") | ///
/*diagnoses specified to 4 digits*/ ///
inlist(substr(dx,1,5),"807.2","807.3","807.4","807.5","807.6") | ///
inlist(substr(dx,1,5),"812.1","812.3","812.5","813.1","813.3") | ///
inlist(substr(dx,1,5),"823.1","823.3","824.5","824.7") | ///
inlist(substr(dx,1,5),"850.2","850.3","850.4","862.8","862.9","874.1","874.5") | ///
inlist(substr(dx,1,5),"904.0","904.1","904.2","904.3","904.4","904.5","904.3") | ///
/*diagnoses specified to 5 digits*/ ///
inlist(substr(dx,1,6),"807.03","807.04","807.05","807.06","807.07","807.08") | ///
inlist(substr(dx,1,6),"807.13","807.14","807.15","807.16","807.17","807.18")

Sample SAS code for diagnosis codes of interest:

dx3=substr(dx,1,3);
Inclusion and Exclusion Criteria for inpatient hospitalization count:

- Primary payer = Workers’ Compensation.
- Limit age to those 16 years or older (age at admission is preferred).
- Select for state of residence = ‘your state.’
- Select for county of residence = ‘county.’
- Exclude:
  - age unknown
  - out-of-state residents and unknown residence
  - out-of-state inpatient hospitalizations
- Use duplicated data (no exclusions for deaths, readmissions).
- Use discharge date for calendar year, not fiscal year.
- Use all cases reported on the discharge file, regardless of length of stay.
- This will yield ‘Annual number of work-related inpatient hospitalizations for severe traumatic injury for persons age 16 years or older in ________ county.’

5.2 Annual rate of work-related inpatient hospitalizations for severe traumatic injury for 100,000 employed persons age 16 years or older by county or region

a) To obtain the denominator for the rate, use either Method A or B.

**Method A**: Instructions for using Bureau of Labor (BLS) Local Area Unemployment Statistics (LAUS) to collect Annual Average Employment Data by County:

- Go to: [http://www.bls.gov/lau/data.htm](http://www.bls.gov/lau/data.htm)
- Choose: One Screen Data Search (green button).

Query Steps:
- Select your state.
- Select “F” (counties and equivalent).
Select the counties of interest - You can choose one or many. (Use the CTRL key to select more than one county). Regardless, you will get data for each county separately.

Select “Not seasonally adjusted.”

Click on “Get Data.”

A new window will pop up that contains monthly and annual average data for all counties in the query for all years and months available.

At the top of this view, click on “More Formatting Options.”

Specify the years of interest, and set the time period to “Annual Data.”

Click on the “Retrieve Data” button.

A new view will display. From here, you can download each county’s data into a MS Excel file. (Note: The column that contains the denominator data is labeled “employment.”)

**Method B:** To select multiple counties:

Go to [http://www.bls.gov/lau/home.htm#cntyaa](http://www.bls.gov/lau/home.htm#cntyaa)

Find table for “Labor Force data by county” for year of interest.

Click on the XLS link to download the excel table of data. This will provide you with the labor force data by county for ALL of the U.S.

Scroll down to find the rows that correspond to the counties for your state. Select the counties of interest and cut/paste into new excel worksheet. (For sample SAS code see Appendix B.)

Use “Employed” in column H (7th column), which is the ‘**Number of employed persons age 16 years or older, by county of residence.**’

b) To calculate the county or regional specific rates:

- Divide the numerator (number of work-related hospitalizations in the county) by the denominator (average employment in the county).

- Multiply these results by 100,000 to get the ‘**Annual crude rate of work-related inpatient hospitalizations for severe traumatic injury by county per 100,000 employed persons age 16 or older.**’
Example 6. Work-related injuries and illnesses reported to workers’ compensation

Numerator: Injury and illness cases with lost work-time filed with state workers’ compensation (WC) system within calendar year, stratified by place of work Public Use Microdata Areas (POWPUMAs).

Denominator: Estimated number of full-time equivalent (FTE) employees covered by state workers’ compensation system for the same calendar year, stratified by POWPUMAs.

Measures:

- Measure 1: Annual number of work-related injury and illness cases with lost work-time filed with workers’ compensation by POWPUMAs.
- Measure 2: Annual incidence rate of work-related injury and illness cases with lost work-time filed with workers’ compensation per 100,000 employed persons age 16 years or older by POWPUMAs.

Data Resources:

[疳 = Information or example specific to Massachusetts (MA)]

- Workers’ compensation system – numerator
- American Community Survey single-year PUMS file – denominator

Demographic: Workers covered by state WC system (i.e. does not include federal government, self-employed, working without pay, or unemployed workers).

Time period: Calendar year based on the date of injury (if date of injury is not available, use the date the claim was filed. If a trend analysis is to be performed for state data, claim-filing date is more appropriate to use than claim date of injury.6)

Geography: U.S. Census POWPUMAs, based on primary workplace location.

Definition and Background

- Public Use Microdata Areas (PUMAs) are non-overlapping areas that partition each state into areas containing about 100,000 residents. PUMAs were developed to be the most detailed geographic area available in the Public Use Microdata Samples (PUMS).7 8 In the American Community Survey (ACS) PUMS files, PUMA variables are available based on residential location (variable name: PUMA), primary place of work location (variable

---

name: POWPUMA), and place of migration from or where the respondent lived 1 year ago (variable name: MIGPUMA).

- Starting in 2012, the ACS PUMS files began to use the new U.S. Census Bureau 2010 PUMA boundaries for all PUMA-related variables, i.e. PUMA, POWPUMA, and MIGPUMA. When working with WC claims data prior to 2012, use the U.S. Census 2000 PUMA boundaries. For WC claims data 2012 and after, use the U.S. Census 2010 PUMA geographic boundaries. A crosswalk between the 2000 and 2010 PUMAs is available at http://mcdc2.missouri.edu/websas/geocorr2k.html. (More information about differences between the 2000 and 2010 PUMA geographic classification is available at: www2.census.gov/geo/pdfs/reference/puma/2010_puma_guidelines.pdf).

**Significance:** POWPUMA-specific work-related injury and illness counts and rates, based on WC claim data, can be used to identify sub-state areas with significantly higher or lower counts or rates of work-related injury and illness compared to statewide estimates. Local areas with high rates can be considered as priorities for prevention outreach.


**Geographic considerations:**

- Rates by POWPUMAs represent the average annual rate of WC claims filed for WC-eligible workers from that POWPUMA. They are not the actual risk of each individual WC-eligible worker or the average risk of workers at other sub-state geographies (e.g. city, town, county, or district).

- POWPUMAs were based on WC reported employer company’s city or town. A limitation of the WC data is that the reported employer city/town could be where the WC was filed (e.g. human resource or auditor’s office, company headquarters) and not the actual workplace location. This scenario would be an issue if there were a discrepancy in city/town of employment and where the WC claim was filed. If information about injury location is available in the WC data, then such misclassification can be checked and manually resolved by comparing the city/town reported for the employer company with that reported for the location of injury. Any discrepancies at the city or town level should not largely affect analyses at the POWPUMA level, as POWPUMA geographic boundaries are larger and encase multiple cities and towns.

---

9 U.S. Census Bureau. Place of Work (POW) and Migration (MIG) PUMAs: Their use in American Community Survey (ACS) Public Use Microdata Sample (PUMS) Files. Accessed at: http://www2.census.gov/geo/pdfs/reference/puma/Defining_POWPUMAs_MIGPUMAs.pdf
Use of the American Community Survey Public Use Microdata (ACS PUMS) data as a FTE denominator:

- The ACS is a household survey and is administered on an annual basis. Estimates from the one-year ACS PUMS file describe the average characteristics for that calendar year and include all the people interviewed that year, regardless of the reference period. ACS employment questions reference the last 12 months of the survey date, which can span across two calendar years. For example, employment from the 2011 ACS could represent a reference period that spans from January 2010 through November 2011. There is no record of which month the respondent took the survey and thus, there is no way to determine which records or which portion of any given record pertains to a given calendar year.
- As a result, FTEs calculated from the ACS are potentially subjected to higher recall bias and may be less reliable than those from the Current population Survey (CPS), which is conducted on a monthly basis. Additional differences between the ACS and CPS data can be found here: www.census.gov/hhes/www/poverty/about/datasources/factsheet.html.

- The approach described in this document to calculate FTEs from the ACS PUMS data serves as an example. There may be other, more conservative approaches\(^\text{10}\) that take into account the incongruence of the ACS reference period with the calendar year in which respondents are interviewed.

Some additional considerations/limitations of the WC data:

- WC claims may be denied.
- WC data can be poor: In MA, industry groups/NAICS codes are not available. Also, there is a lack of detailed information about the cause of work-related injuries and illnesses on many claims.
- WC data may not be complete, as under-reporting of work-related conditions and nationwide underutilization of WC benefits for work-related conditions has been well documented in the literature.\(^\text{11,12,13,14}\) Additionally, workers eligible for WC may file claims through other systems that do not get entered into the state’s WC system.
  - In MA, firefighters and police officers are covered by separate systems for compensating work-related health problems under state law
- In MA, only workers claiming five or more lost workday injuries or illnesses are eligible for lost-wage WC benefits, and therefore, injuries and illnesses that did not result in such

---

\(^{10}\) Another approach to calculate FTEs from the ACS PUMS data is to average ACS employment data over a two calendar year period. For example, if denominator information is desired for 2011, it would be best to use ACS data for 2011 and 2012 surveys. This approach was not taken for this guidance measure but is noted here for users to explore in their own analysis.


lost work-time could not be captured. WC claims for less serious injuries have generally been found to account for about 70% of all WC claims filed. 15

- In MA, it was not possible to distinguish claims filed from awarded claims in the Department of Industrial Accidents (DIA) WC database, which may result in an overestimate of the number of lost-wage WC claims awarded. However, it is more likely that the full extent of work-related injuries and illnesses workers would be underestimated.

- Times between the occurrence of an event and the submission of a claim can vary. The use of injury year could exclude claims for injuries that are filed after the calendar year of analysis (i.e. an injury occurred in 2011 but was not filed until March of 2012). As a result, total count and rate of WC claims would be an underestimate. One way around this is to use multiple years of WC claims and pull out claims from the same injury year. As a general rule, data users should check the average lag time between injury and filing of a claim, or the percent of claims in a given year that are for injuries in the preceding year.

METHODOLOGY:

( = Information or example specific to MA)

**Every state should follow their own protocols for processing and cleaning the WC data prior to analysis. The steps outlined below serve only as a guide and are based on MA’s WC data and protocols, which may not be similar to your state’s WC data and protocols. **

Step 1. Construct the WC database for analysis:

1a. Identify and obtain cases by date of injury for one year.
   • If date of injury is not available, use the date the claim was filed.
   • If a trend analysis is to be performed for state data, claim-filing date is more appropriate to use than claim date of injury.

1b. Include claimants of all ages and those with age unknown.

1c. Include out-of-state residents.

1d. Include only filed claims that resulted in lost workdays or ‘time loss.’
    In MA, only WC claims for injuries and illnesses resulting in 5+ lost workdays are available from the Department of Industrial Accidents (DIA) via employers’ First Reports of Injury (FROI), insurance company notifications, or employee claims. A “claim” refers to a report that has been filed, but not necessarily accepted. A claim also refers to a single incident.

1e. Exclude:
   • Records duplicated in the database because of administrative error.
   • Records with different board numbers that are determined to be duplicate records.
     o  In MA, this determination was made based on records with the same worker’s home zip code, date of birth, date of injury, nature of injury/illness code, body part code, regular occupation, company name, and narrative text description of the incident, regardless of whether there was identical information for weekly wage and injury source.

1f. De-duplicate and transpose file to create an incident-level dataset:
    In MA, a claim, represented by a unique board number, may have had more than one record in the database. This occurs when more than one nature of injury or illness was reported or when an injury affected more than one body part. For our analytical file, our unit of analysis was at the incident-level, i.e. one record per claim or unique board number.

Therefore, we took the following steps to prepare our analytical file:
Transpose the WC dataset from injury-level (i.e. each record represents a nature of injury-body part pair) to incident-level (i.e. each record represents a unique board number):

- For each year injury year, the injury-level dataset was transposed to an incident-level dataset that could have multiple injuries per claim (INJ_TYP1, BODY_PART1, INJ_TYP2, BODY_PART2, etc.).

**Example:** In MA, data on nature of injury and body parts affected were coded to ANSI Z16.2.

<p>| Injury-level dataset (i.e. each record represents a nature of injury-body part pair) |
|---------------------------------|------------------|--------------|</p>
<table>
<thead>
<tr>
<th>Record</th>
<th>Board_num</th>
<th>INJ_TYP</th>
<th>BODY_PART</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>01234</td>
<td>310</td>
<td>420</td>
</tr>
<tr>
<td>2</td>
<td>01234</td>
<td>170</td>
<td>318</td>
</tr>
<tr>
<td>3</td>
<td>01234</td>
<td>100</td>
<td>340</td>
</tr>
<tr>
<td>4</td>
<td>56789</td>
<td>160</td>
<td>315</td>
</tr>
<tr>
<td>5</td>
<td>56789</td>
<td>190</td>
<td>450</td>
</tr>
<tr>
<td>6</td>
<td>56789</td>
<td>170</td>
<td>200</td>
</tr>
</tbody>
</table>

<p>| Incident-level dataset (i.e. each record represents a unique board number) |
|---------------------------------|------------------|--------------|------------------|--------|</p>
<table>
<thead>
<tr>
<th>Record</th>
<th>Board_num</th>
<th>INJ_TYP1</th>
<th>INJ_TYP2</th>
<th>INJ_TYP3</th>
<th>BODY_PART1</th>
<th>BODY_PART2</th>
<th>BODY_PART3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>01234</td>
<td>310</td>
<td>170</td>
<td>100</td>
<td>420</td>
<td>318</td>
<td>340</td>
</tr>
<tr>
<td>2</td>
<td>56789</td>
<td>160</td>
<td>190</td>
<td>170</td>
<td>315</td>
<td>450</td>
<td>200</td>
</tr>
</tbody>
</table>

- When creating an incident-level dataset from an injury-level dataset, we also concatenated text values of the following variables into one if their values were different: narrative text description of the incident, regular occupation, company name, injury source, and injury source description. For example, regular occupation could now have the value “welder; driller” in the incident-level dataset that reflected the two different occupations that were initially listed per record.

**Example:** In MA, these variables were in text narrative format.

<p>| Injury-level dataset (i.e. each record represents a nature of injury-body part pair) |
|---------------------------------|------------------|--------------|</p>
<table>
<thead>
<tr>
<th>Record</th>
<th>Board_num</th>
<th>Reg_Occ</th>
<th>Company_name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>01234</td>
<td>Welder</td>
<td>Smith Comp</td>
</tr>
<tr>
<td>2</td>
<td>01234</td>
<td>Driller</td>
<td>Smith Comp</td>
</tr>
<tr>
<td>3</td>
<td>56789</td>
<td>Teacher</td>
<td>Springfield Public</td>
</tr>
<tr>
<td>4</td>
<td>56789</td>
<td>Teacher</td>
<td>Springfield Public</td>
</tr>
<tr>
<td>5</td>
<td>56789</td>
<td>Paraprofessional</td>
<td>Springfield Public Elementary</td>
</tr>
<tr>
<td>6</td>
<td>56789</td>
<td>Paraprofessional</td>
<td>Springfield Public Elementary</td>
</tr>
</tbody>
</table>
Incident-level dataset (i.e. each record represents a unique board number)

<table>
<thead>
<tr>
<th>Record</th>
<th>Board_num</th>
<th>Reg_Occ</th>
<th>Company_name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>01234</td>
<td>Welder; driller</td>
<td>Smith Comp</td>
</tr>
<tr>
<td>2</td>
<td>56789</td>
<td>Teacher; paraprofessional</td>
<td>Springfield Public; Springfield Public Elementary</td>
</tr>
</tbody>
</table>

Step 2. Obtain Measure 1: Annual number of work-related injury and illness cases with lost work-time filed with workers’ compensation by POWPUMAs.

2a. Cross-walk the numerator WC data into POWPUMAs, based on employer city/town location (or county):

i. Go to the U.S. Census American Community Survey (ACS) PUMS Technical Document Page:
https://www.census.gov/programs-surveys/acs/technical-documentation/pums/documentation.html

- Accessible on April 2016.

*If the link does NOT open, open link in another browser (e.g. Google Chrome).*

- Click on the year of interest: 2011
- Click on: **2011 1-year Code Lists**
ii. Download, Save, and Open the file (you’ll need Adobe Reader).

iii. Double-click on file called “Place of Work PUMA.xls” and Download. Then Open File.

iv. Your Federal Information Processing Standards (FIPS)\textsuperscript{16} state code is listed in the first column called ‘POWSP’. Scroll down to your state code (for example, MA’s FIPS = 025).

v. Use information provided as a crosswalk of the geographic level available in your WC numerator dataset (e.g. city/town or county) to POWPUMA code for your state.

POWS & POWSP = Place of work \textit{State}
POWC = Place of work \textit{County}
POWM FIPS Code: Place of work \textit{Town or City}
POWP FIPS Code: Place of work \textit{Place} (smaller than city/town)
POWPUMA = Place of work PUMA
Detailed Area Name = Place of work \textit{Detailed area} (smaller than Place).

\textsuperscript{16} The FIPS state codes are numeric and two-letter alphabetic codes to identify U.S. states and certain other associated areas.
• From the employer city/town name or FIPS code in your WC dataset for each claim, find the associated POWPUMA and add this information as a new variable/column in your WC dataset called ‘POWPUMA’. If employer city/town name or FIPS code are not available in your dataset, you can use county, although the crosswalk may not be a 1-to-1 match.

A crosswalk can be created manually or via a SAS program (see MA example - Appendix C).

vi. Once each WC claim has been assigned a POWPUMA, run a Proc Freq in SAS or a Query in MS Access to obtain the sum of WC claims by POWPUMA.

A sample SAS code is below:

```sas
libname num '.'; /* Enter folder location of your WC num dataset */

proc freq data=num.WClocal; /* Enter name of your WC num dataset */
tables powpuma/list missing nocum out=powpuma
(rename=(count=TotalClaims)) ;
title "Measure 1. Annual number of work-related injury and illness cases with lost work-time filed with state workers’ compensation by place of work Public Use Microdata Areas (POWPUMAs) (v2000) [ACS PUMS 2011]";
run;
```
Step 3. Obtain Measure 2: Annual number of work-related injury and illness cases with lost work-time filed with workers’ compensation per 100,000 employed persons, age 16 years or older, by POWPUMAs

**Steps below are adapted from methodology and SAS program developed by Dr. Martha Jones, Vanderbilt University, via Karen Louie, California Department of Finance, and Dale Garrett, US. Census Bureau.**

3a. Review background of methods (used in MA): The estimate of full-time equivalent (FTE) employees was calculated using the American Community Survey (ACS) Public Use Microdata (PUMS) files. FTEs were calculated based on the following ACS PUMS variables:

- WKHP: Usual hours worked/week in past 12 months [response is a discrete number]
- WKW: Weeks worked during the past 12 months [response is a range]

The average annual FTEs were calculated by multiplying WKHP by the median of the WKW variable, and then dividing by 2000 hours. The ACS was selected as the denominator data source, compared to other data sources of FTEs (e.g. Current Population Survey), due to its large sample size and sub-state geographic identifier of POWPUMAs.

The average annual claim rate, i.e. the number of WC claims per 100,000 FTE employees, and its 95% confidence interval (CI) were computed statewide and by POWPUMAs. Estimates for the number of FTEs were obtained from the ACS single-year 2011 PUMS using previously developed methods. Rates were restricted to claims and FTEs of currently employed (at work or not at work when surveyed) age 16 and older who worked in the past 12 months. Class of workers (ACS variable name: COW) that would typically not be WC eligible, such as the below categories, were excluded:

- Federal government employee
- Self-employed in own not incorporated business, professional practice, or farm
- Self-employed in own incorporated business, professional practice or farm
- Working without pay in family business or farm
- Unemployed and last worked 5 years ago or earlier or never worked

*Check that the WC exclusion criteria for your state are consistent with the above categories and revise SAS program in Step 3b as appropriate.*

Also, FTEs working in MA but residing in any of the five surrounding states (CT, RI, NY, NH, and VT) were included in the denominator. Additionally for MA, the Department of Industrial Accidents does not receive claims for firefighters and police officers, with the exception of claims for certain police forces such as college campus and transit, as they are covered by

---

17 A limitation of the ACS is that it is administered on an annual basis and respondents are asked to report usual hours worked in the past 12 months, whereas the CPS is a monthly survey and respondents are to report usual hours worked in the reference week. Other differences between the ACS and CPS can be found here: www.census.gov/hhes/www/poverty/about/datasources/factsheet.html
separate systems for compensating work-related health problems. Therefore in MA, WC claims and FTEs for all police officer and firefighter related occupations were excluded when calculating rates (not reflected in SAS program at end of Step 2).

The margin of error (MOE) of WC claim data was obtained using an exact Poisson 95% distribution, and the MOE of the FTEs was obtained using ACS replicate weights. The MOE for the average annual claim rate was then calculated using the U.S. Census Bureau’s methodology for derived ratios, which was then used to calculate the rate 95% confidence intervals.19

3b. Exclude WC claims with age unknown or less than 16 years old. This is important as ACS PUMS data for employment are based on workers age 16 and older.

3c. Obtain FTE estimates from the ACS PUMS file by POWPUMAs (example used is 2011 data):

i. Go to: http://www.census.gov/programs-surveys/acs/data/pums.html
   • Accessible on April 2016.

ii. Click on ACS 1-year PUMS file for year of interest. In this example, we selected 2011.

---

iii. Check the box for PUMS-SAS file “2011 ACS 1-year Public Use Microdata Samples (PUMS) - SAS format”.

iv. Click “View”.

v. Right-click on the Population Records for your state → Select “Save link as…” → Save the compressed file called “unix_pma” to a folder on your computer or network (you will need to reference this folder location later in your SAS program).
vi. Unzip/Extract the compressed file to the same folder location. Rename the SAS data set to “psam_p1”

The compressed file will contain a SAS Data Set called “psam_p25” and .pdf technical document called “ACS2011_PUMS_README” containing useful information about the data set, PUMAs, other documentation available, the PUMS file structure, and weights available. For this example, we will select Massachusetts Population Records.

vii. The ACS surveys individuals based on their residential location. Yet, workers can work in your state and thus be eligible for your state’s WC, while living in another state and commuting across the border for work.

In MA Therefore in MA, we also included MA workers who showed up as residents in ACS state samples of adjacent states: Connecticut, New Hampshire, Vermont, Rhode Island, and New York.

Therefore, repeat Steps v and vi above for states adjacent to your state.

In MA, Steps v and vi were repeated for all adjacent states: Connecticut, New Hampshire, Vermont, Rhode Island, and New York.

viii. Save and unzip these files to the same folder and give each SAS data set a numeric suffix in ascending order (i.e. your main state would end with “1”, then subsequent states with “2”, “3”, “4”, “5”, 6”, etc.)

In MA, these files were saved and unzipped to the same folder. These SAS datasets were then named as follows:

- Massachusetts: psam_p1
- Connecticut: psam_p2
- New Hampshire: psam_p3
- Vermont: psam_p4
- Rhode Island: psam_p5
- New York: psam_p6

ix. Run the SAS Program below to obtain Measure 2.

NOTE: Codes in this SAS program that are highlighted in yellow should be revised with your specific information, such as the folder location of data sets, specific variable names or text. Also, make sure your numerator dataset has the following variables in numeric type:

<table>
<thead>
<tr>
<th>Numerator data set field</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Numeric</td>
</tr>
<tr>
<td>Powpuma</td>
<td>Numeric</td>
</tr>
</tbody>
</table>
SAS Program to obtain FTEs by place of work PUMAS (POWPUMA) from the ACS PUMS data and to calculate rates.

Programmer: MyDzung Chu (mydzung.chu@state.ma.us), adapted from existing SAS codes (as noted)

I. DOWNLOAD ACS PUMS data:
2. click on "PUMS file (year(s) of interest)
3. check PUMS-SAS format
4. click on VIEW
5. Download Massachusetts Population Records
6. Save to folder location that will be referenced in LIBNAME.
7. Unzip and name SAS data file as psam_p# (or else Macro will NOT work!).
   For #, enter a number that is unique for that state.
8. Repeat items 1-7 above if pulling in data of adjacent states. For item 7 above, make sure to assign a unique ID # that is sequential to the previous file’s id (i.e. 1, 2, 3, 4…) in order for the array to work.

II. Run SAS Macro and Program below to pull in denominator and numerator data sets, obtain variance estimates, and calculate rates with 95% CIs. Refer to the ACS PUMS Data Dictionary for the respective data file year to select ACS variables.
   • Link to PUMS Technical Documentation: https://www.census.gov/programs-surveys/acs/technical-documentation/pums/documentation.2011.html
     (If this link does not work, Google “PUMS Technical Documentation”)

^^Current code is for ACS single-year estimates for 2011.
^^Areas highlighted in yellow should be modified with your state’s information

libname denom '.'; /*Enter location of denominator ACS PUMs data*/

/*STEP 1: Check, and if needed change, the following variables are in the correct format in all SAS data sets:
   numeric: wkw, esr, agep
   character: cow, puma, powsp, occp;*/

/*STEP 2: RUN MACRO*/

%macro importACS;
/* start do loop to run through each state*/
%do state = 1 %to 6; /*1=MA (p25), 2=CT(p09), 3=NH(p33), 4=NY(p-36), 5=RI(p44), 6=VT(p50)*/

/**************************************************
/*Step 1: Convert var to same format for macro*/
/**************************************************

data work.psam_p&state.;
set denom.psam_p&state.;
   wkwl=1*wkw;
esr1=1*esr;
drop wkw esr;
rename wkwl=wkw esr1=esr;run;
/*Step 2: Create FTE and POWPUMA variable*/

data p&state.;
  set work.psam_p&state.;
  length weeks $10;
  state=&state.; /*create state id for later concatenation*/
/*Entry criteria: Civilian employees (at work or with a job but not at work), age 16+, and worked 1+ week in last 12mo.
^Enter your state’s FIPS code for place of work state (powsp).
^^For class of worker (COW), included are those WC-eligible for MA:
  1. Employee of a private for-profit company or business, or of an individual, for wages,salary, or commissions
  2. Employee of a private not-for-profit,tax-exempt, or charitable organization
  3. Local government employee (city, county, etc.)
  4. State government employee
Revise as needed for your state.* /
  if (powsp='025' and agep ge 16 and wkw>0 and ESR in (1,2) and COW in ('1','2','3', '4')) then worker='Curr_Emp'; else worker='Other';
/*FTE calculations*/
  /* weeks=median of usual weeks worked in the past 12 mo. range (WKW)*/
  if wkw=1 then weeks='51';
  if wkw=2 then weeks='48.5';
  if wkw=3 then weeks='43.5';
  if wkw=4 then weeks='33';
  if wkw=5 then weeks='20';
  if wkw=6 then weeks='7';
  tothours=wkhp*weeks; /*total hours in past 12 mo*/
  FTE=(tothours/2000); /*added FTE calc*/
/*Class: Place of work PUMA v2000 (PUMA v2010 apply to ACS data 2012+)* /
  POWPUMA_NUM=(1*POWPUMA); /*converted to numeric*/
/*Take out negative #s in pop weights*/
array temp(*) pwgtp1-pwgtp8; /*Set negative weights to zero - necessary for later procedures to run*/
  DO i = 1 to dim(temp);
    IF temp(i) < 0 THEN temp(i) = 0;
  END;
  drop powwpuma; rename POWPUMA_NUM=powwpuma;
proc sort; by powwpuma;
/*Step 3. COMPUTE FTE and VARIANCE ESTIMATES using weights PGWT and PWGTP1-80
 * Reference: https://usa.ipums.org/usa/repwt.shtml
 * SAS PROC MEANS code to estimate SEs for FTEs were provided
* by Martha Jones (Vanderbilt University) via Karen Louie (CA Dept of Finance) and Dale Garrett, US Census Bureau */

```sas
proc means data = p&state. NOPRINT;
    where worker='Curr_Emp';
    var pwgtp pwgtp1-pwgtp80;
    weight FTE;
    class powpuma;
    output out = FTE_&state. (rename=(_freq_=n_&state.)) sum = est_FTE0-est_FTE80;

Data step1_&state. (KEEP=powpuma FTE_&state. CV_d_&state. MOE_fte_&state. n_&state.);
    set FTE_&state. ;
    ARRAY FTE_i (80) est_FTE1-est_FTE80;

*initialize the variance to 0;
variance_d=0;

*do the variance summations;
DO i=1 to 80;
    variance_d + (((FTE_i(i) - est_FTE0)**2)/20;
END;

std_error = SQRT(variance_d);* calc standard errors;
CV_d_&state. = std_error/est_FTE0; * calc coefficient of variation;
MOE_fte_&state. = 1.96*std_error; * calc MOE of estimated fte*;

rename est_FTE0=FTE_&state. ;
proc sort; by powpuma;
run;
%end;
%mend importACS;
%importacs /*POWPUMA = . is statewide estimate*/
```

/*********************************************/
* Step 4. Combine ACS data across the states
* Source: A-14 Appendix - Calculating MOEs for Aggregated Count Data.
* www.census.gov/content/dam/Census/library/publications/2009/acs/ACSPUMS.pdf*/

/*^Add or remove step1_# datasets depending on # of state data sets you’d pulled in and referenced at the start of the macro for %do state = # %to #;*/

data denom.ACSLocal (keep= powpuma fte_all MOEfte_all n_all cv_all var_all);
merge step1_1 step1_2 step1_3 step1_4 step1_5 step1_6;
by powpuma;
array change_numeric_;/*convert all missing to 0 to allow bellow sums*/
do over change;
    if change=. then change=0;end;
    fte_all=(fte_1+fte_2+fte_3+fte_4+fte_5+fte_6);
    MOEfte_all=SQRT((MOE_fte_1**2)+(MOE_fte_2**2)+(MOE_fte_3**2)+(MOE_fte_4**2)+(MOE_fte_5**2)+(MOE_fte_6**2));
    n_all=(n_1+n_2+n_3+n_4+n_5+n_6);
var_all=(MOEfte_all/1.96)**2;
cv_all=(MOEfte_all/1.96)/fte_all;
run;

/****************************************************************************
* Step 5: Upload count/num data. Calculate MOE of numerator using Poisson dist *
* SAS algorithm adapted from: Daly, Leslie,"Simple SAS macros for the
* calculation of exact binomial and Poisson Confidence limits"
libname num '.'; /*^Enter folder location of numerator dataset */
proc freq data=num.WCdataset2; /*^Enter name of your numerator dataset*/
where age gt 15; /*^To match ACS data for age*/
tables powpuma/list missing out=powpuma ; run;
data step1;
set powpuma end=eof;
tot + count;
  if eof then do;
  output;
  powpuma=0; /*statewide*/
  count=tot;
  output;
end; else output;
run;
data step2;
set step1;
if count lt 0 or not(0 lt (0.95) lt 1) then
  do;
    ll = .;
    lu = .;
  end;
else
  do;
    if count ne 0 then
      do;
        ll = gaminv((1 - (0.95))/2, (count));
        lu = gaminv((1 + (0.95))/2, (count) + 1);
      end;
    else if (count) eq 0 then
      do;
        lu = -log(1 - (0.95));
        ll = 0;
      end;
    end;
  MOE_n=(lu-l1)/2;
drop ll lu;
proc sort data=step2; by powpuma; run;

/****************************************************************************
* Step 6: Calculate rate, MOE and 95% CIs of rate using ACS formula for
Derived Ratios (A-14 Appendix).
Will give conservative estimate of the MOE*/

data step3;
merge step2 denom.ACSLocal; by powpuma;

Puma_rate= (count/fte_all); /*Injury rate*/
Puma_rate100000=100000*puma_rate; /*Injury rate per 100,000 FTEs*/

    MOErate=(SQRT((MOE_n**2)+((puma_rate**2)*(MOEfte_all**2))))/fte_all;
    LL_CIrate=Puma_rate-MOErate;
    UL_CIrate=Puma_rate+MOErate;
    LL_CIrate100000=100000*LL_CIrate;
    UL_CIrate100000=100000*UL_CIrate;

rename COUNT=Total_WCClaim    n_all=Total_ACSRespondents;
run;

/******************************************************************************
* Step 7: Print output*/

proc print data=step3 noobs round;  
    var powpuma Puma_rate100000 LL_CIrate100000 UL_CIrate100000 Total_WCClaim
    fte_all Total_ACSRespondents;
    title "Measure 2. Annual incidence rate of work-related injury and illness
    cases with lost work-time filed with workers’ compensation per 100,000 employed
    persons age 16 years or older by place of work Public Use Microdata Areas
    (POWPUMAs) (v2000) in Massachusetts [ACS PUMS 2011]";
    run;
Displaying the Data: Example from Massachusetts

Figure 6.1. Annual rate of WC claims among local government workers age 16+ by place of work Public Use Microdata Areas (POWPUMA), Massachusetts, 2009-2011 (N= 8,218)*

Numerator source: Massachusetts DIA Workers’ Compensation data, 2009-2011
Denominator source: American Community Survey single-year Public Use Microdata Sample data, 2009-2011

*Rates are restricted to MA local government workers age 16 and over and living in MA or any of its five surrounding states (i.e. CT, RI, NY, NH, and VT). 120 claims were excluded because age was missing (n=80) or the worker was a police officer (n=34) or under 16 years old (n=6).
Appendix A. Topics to Consider for Sub-State Analyses

When analyzing and reporting measures of workers’ health and safety, it is important to be aware of several potential issues that may be present in the data being used to generate a measure, the method(s) used to calculate an estimate of the measure, and the challenges that may exist in reporting and presenting any results to a variety of audiences. By highlighting a number of these issues within this appendix, we hope to encourage users of this guidance document to not only carefully consider these issues but then also work within their organization to ensure the issues highlighted in this appendix are handled in a way which is appropriate and consistent with their organization’s existing policies.

When creating and analyzing sub-state measures relevant to workers’ safety and health, it is useful to be cognizant of some of the aspects and challenges inherent to conducting analyses across locations within a given state. What follows is a brief list of some those issues and challenges.

<p>| Bias in case ascertainment, spatial | As in case-reporting and case-capture in general within public health surveillance, variation in the consistency of diagnoses and clinical reports may vary by location; potential bias in case-ascertainment across space may be a function of clinical sources, underreporting by industry type, or by employers, etc. |
| Change of support problem | Occurs when data which exists at one geographic level (e.g. point, census block, zip code) is changed (e.g. through aggregation, modeling, etc.) to a variable at a different geographic level (e.g. county, census tract); the data at these differing geographic levels, upon analyses, may not show the same spatial characteristics, associations, or statistics at different geographic levels; can be considered to be a more general form of the ecological fallacy. |
| Choosing values (cut points) to group and map continuous data in choropleth maps | When mapping a continuous variable in a choropleth map, specific values must be chosen to delimit the values for each grouping; the choice of those cut-points may affect how the map is interpreted by the viewer/reader; carefully consider the purpose of the map when choosing cut-points; calculated quantiles may be used to divide a continuous variable (e.g. 0-20%, 20-40%, etc.) into groupings with an equal number of observations; if identifying the areas with the lowest and highest 10% of values is important, the 0-10%, 10-50%, 50-90%, and 90-100% values could be used as cut-points (for example); note: some geographers may prefer using &quot;natural breaks&quot; (e.g. Jenks optimization) as cut-points because this approach minimizes the variation of values within groups--but this may not be appropriate for the purpose of your map(s); discuss the purpose of your map(s) with your geographer / GIS analyst. |
| <strong>Choropleth maps</strong> | A very common type of map which uses unique fill patterns (e.g. hatched) or colors to represent values for areal units (e.g. counties) within the map; this type of map may be easily interpreted by the viewer when reasonable color schemes (e.g. sequential or diverging color ramps) are used; however, this type of map may not work well when a large difference in the sizes of area units exists (e.g. very small census tracts for highly populated areas, very large ones for sparsely populated areas) across the area being mapped because the larger areal units will dominate the visual-frame and potentially limit the viewer's ability to appropriately interpret the data being mapped. |
| <strong>Colors in maps</strong> | Choice of colors within maps may influence how the information contained with the map is perceived by the viewer; color schemes chosen may make the map more easy or more difficult to interpret; as many as 1 in 12 men possesses a color-vision deficiency; consider using a color-scheme selection tool for mapping, like ColorBrewer 2.0, to choose appropriate color-schemes for a map (see: <a href="http://colorbrewer2.org/">http://colorbrewer2.org/</a>). |
| <strong>Confidentiality</strong> | Similar to non-spatial reporting of public health data, maintaining confidentiality is an important consideration when providing results of data by location or area; similar to dividing data and then reporting by categories (e.g. age, race, etc.), separating the data by geographic location (e.g. address, county) could unintentionally reveal information about individuals; a number of methods exist to preserve confidentiality when displaying information in the form of maps; data can be suppressed for those areal units (e.g. zip code, census tract) having a count of cases that fall below a set threshold (e.g. n=5); analogous to combining cells/groups for non-spatial data (e.g. age-groupings), data can be aggregated within adjacent areal units (e.g. counties next to each other) especially if the underlying spatial process (e.g. demographics, make-up of industries) is believed to be relatively homogenous across the combined areal units; for point data (e.g. street addresses) methods for geomasking have been developed and used to perturb the locations of the data points to help preserve confidentiality but also retain the overall spatial distribution of the observations. |
| <strong>Ecological fallacy</strong> | An error in inference wherein analyses and the ensuing results observed at the group-level (e.g. in ecological analysis) are used to interpret associations/processes at the individual-level; both risk factors and health outcomes are measured at the group-level in ecological analyses and no individual-level information is available/used; note difference between an ecological analysis versus a multilevel analysis (described below). |</p>
<table>
<thead>
<tr>
<th><strong>Edge effects</strong></th>
<th>A potential problem affecting the edges of a map when mapping data with a distinct boundary (e.g. an individual state)--even though an underlying process (e.g. location of clinical services used, worker’s residence) may cross the boundary.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Geocoding</strong></td>
<td>Process to transform a place name or street address into geographic coordinates (e.g. latitude and longitude) for use in a geographic information system (GIS).</td>
</tr>
<tr>
<td><strong>Geographic centroid</strong></td>
<td>The geographic center of an area (e.g. a county or state); sometimes used to represent or label an aerial feature.</td>
</tr>
<tr>
<td><strong>Modifiable Area Unit Problem (MAUP)</strong></td>
<td>Both a conceptual and practical problem where the results of an analysis of areas (e.g. for rate calculations) may be dramatically different if/when the boundaries (often arbitrary) within/across an area were to be changed; sub-units with apparent low or high values could vary if the boundaries of those sub-units were different; the implications of the MAUP are that the inferences derived from analysis based on set boundaries should not be extended to making inferences based on other possible boundaries for the study area.</td>
</tr>
<tr>
<td><strong>Multilevel analysis</strong></td>
<td>An analytic design which takes into account data being available at more than one &quot;level&quot; (e.g. individual-level, county-level, state-level, etc.); appropriate when data exists on individuals (e.g. an individual's health outcome, age, etc.) but an analyst also wants to consider group-level data (e.g. poverty by census tract, employer, or worksite) for those individuals in the analysis; the variance at level 1 (e.g. individual) can be calculated separately from the variance at level 2 (e.g. county) thereby overcoming the assumption of independent sampling (e.g. all individuals within a given census tract would have identical census-tract poverty rates); this design also makes possible an examination of level-2 effects versus level-1 effects.</td>
</tr>
<tr>
<td><strong>Projections, map</strong></td>
<td>Displaying a three-dimensional object (e.g. the Earth's surface) as a two-dimensional object (e.g. a map) requires distorting either the shape, area, direction, or distance within a resulting map; to address this challenge, geographers use projections to create maps of the Earth's surface; different projections may make mapped features look very different and separate data layers combined within a GIS need to use the same projection (and datum) to ensure the data layers are combined accurately within the map; consider deferring to your geographer / GIS analyst as to what projection is appropriate for your location / map(s).</td>
</tr>
<tr>
<td><strong>Proportional symbols maps</strong></td>
<td>Type of map which represents the values of features by displaying symbols that are proportional in size to the value of the variable at that location; larger values (e.g. for counts or rates) are represented with larger symbols, smaller values are displayed with proportionally smaller symbols; each symbol is typically centered on the areal unit's centroid; an alternative to choropleth maps.</td>
</tr>
<tr>
<td>Small number problems</td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td></td>
</tr>
<tr>
<td>Similar to small number issues that may occur within cells of R x C frequency tables, small number &quot;problems&quot; can be encountered when small numbers of cases are observed within specific areas; this often occurs in sparsely populated areas, especially for rare conditions; this type of problem can have implications for rate stability and/or confidentiality issues; one possible solution to this problem would be to aggregate the data over time for each area; it is worth noting that non-spatial analyses commonly aggregate data over space and spatial analyses can aggregate the data over time; heterogeneity in data may exist across both space and time but the limitations of cells or areas with small numbers may necessitate aggregating data by location, time, and/or combining cells or areas</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ZIP codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zone Improvement Plan (ZIP) codes are routes and points enumerated by the U.S. Postal Service for the purpose of delivering mail; ZIP codes are not polygons (areal units) and, unlike census units, are not designed to optimize homogeneity of the population within ZIP codes; importantly, the U.S. Census Bureau does not specifically calculate population estimates for ZIP codes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ZIP Code Tabulation Areas (ZCTAs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZCTAs are a creation of the US Census Bureau designed to serve as polygons in an attempt to represent ZIP codes from the US Postal Service; although ZCTAs represent the Census' best attempt to provide estimates of Census data similar to ZIP codes, ZCTAs and ZIP codes do not match each other perfectly; See <a href="https://www.census.gov/geo/reference/zctas.html">https://www.census.gov/geo/reference/zctas.html</a> for more information, including a description of the somewhat complicated approach of how the U.S. Census Bureau calculates estimates for ZCTAs</td>
</tr>
</tbody>
</table>

**Additional Reading**


## Appendix B. Data Sources at Sub-state Level Relevant to Occupational Safety & Health

### Data Sources Containing Health Outcomes, Exposures, and/or Risk Factors

<table>
<thead>
<tr>
<th>Name</th>
<th>Brief Description</th>
<th>Collected By/ Sources</th>
<th>Highest Spatial Resolution Typically Available</th>
<th>Other geographies available</th>
<th>Constructs captured on workers</th>
<th>Relevant Websites</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult lead exposure surveillance, including Adult Blood Lead Epidemiology and Surveillance</td>
<td>Laboratory reported adult blood lead levels</td>
<td>States</td>
<td>Point (residence, worksite)</td>
<td>Point, County, or Zip Code</td>
<td>blood-lead level</td>
<td><a href="http://www.cdc.gov/niosh/topics/ables/state.html">http://www.cdc.gov/niosh/topics/ables/state.html</a></td>
<td>Data by individual inspection; unclear how to create analytic data sets by sub-state geography within OSHA system</td>
</tr>
<tr>
<td>Behavioral Risk Factor Surveillance System (BRFSS)</td>
<td>Collects information on health risk behaviors, preventive health practices, and health care access primarily related to chronic disease and injury.</td>
<td>State BRFSS programs</td>
<td>Varies by state, can be Metropolitan Statistical Area, county, or state</td>
<td>Sub-state information (e.g. county) is collected</td>
<td>Health risk factors, self-reported illnesses</td>
<td><a href="http://www.cdc.gov/brfss/">http://www.cdc.gov/brfss/</a></td>
<td></td>
</tr>
<tr>
<td>Census of Fatal Occupational Injuries (CFOI)</td>
<td>Looking at fatal workplace injuries in term of rates and counts, types of incident, occupational type and demographic</td>
<td>State</td>
<td>Metropolitan Statistical Areas</td>
<td>State</td>
<td>Industry, occupation, primary and secondary source of injury, worker characteristics</td>
<td><a href="http://www.bls.gov/iif/oshcfoi1.htm">http://www.bls.gov/iif/oshcfoi1.htm</a></td>
<td></td>
</tr>
<tr>
<td>Chemical Exposure Health Data</td>
<td>Industrial hygiene sample results for exposures to chemical hazards from OSHA inspections</td>
<td>OSHA</td>
<td>Zip Code</td>
<td>City</td>
<td>Personal, area, and bulk sampling results for airborne or physical agents</td>
<td><a href="https://www.osha.gov/opengov/healthsamples.html">https://www.osha.gov/opengov/healthsamples.html</a></td>
<td></td>
</tr>
<tr>
<td>Fire Fighter Fatality Investigation and Prevention Program (NIOSH Fire Fighter Face)</td>
<td>Fatality investigation</td>
<td>NIOSH</td>
<td>Point</td>
<td>None</td>
<td>Geographic location, department type (career and volunteer), department size, and population density, medical and trauma</td>
<td><a href="http://www.cdc.gov/niosh/fire/">http://www.cdc.gov/niosh/fire/</a></td>
<td></td>
</tr>
<tr>
<td>Hospitalizations and Emergency Department Visits</td>
<td>Inpatient and/or outpatient hospitalizations</td>
<td>States</td>
<td>Zip Code</td>
<td>County</td>
<td>Illness, injury, payer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Name</td>
<td>Brief Description</td>
<td>Collected By/ Sources</td>
<td>Highest Spatial Resolution Typically Available</td>
<td>Other geographies available</td>
<td>Constructs captured on workers</td>
<td>Relevant Websites</td>
<td>Notes</td>
</tr>
<tr>
<td>-----------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>-----------------------</td>
<td>-----------------------------------------------</td>
<td>---------------------------</td>
<td>------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------</td>
</tr>
<tr>
<td>MSHA Accident Injuries Data Set</td>
<td>Contains information on all accidents, injuries and illnesses reported by mine operators and contractors</td>
<td>MSHA</td>
<td>Mine ID</td>
<td>None</td>
<td>Mine's Controller/Operator, Contractor ID, <strong>Mine ID</strong>, Accident information including severity of injury, activity, equipment used, location within mine, shift, occupation</td>
<td><a href="http://arlweb.msha.gov/OpenGovernmentData/OGiMSHA.asp">http://arlweb.msha.gov/OpenGovernmentData/OGiMSHA.asp</a></td>
<td>Data covers January 1, 2000 to current; must link to Address of Record Data by Mine ID for sub-state analyses</td>
</tr>
<tr>
<td>MSHA Mine Address of Record Data Set</td>
<td>Contains information on individual mines beginning in 1998</td>
<td>MSHA</td>
<td>Address</td>
<td>City, State, Zip Code</td>
<td><strong>Mine ID</strong>, mine name/contact, type, status, SIC code, street address,</td>
<td><a href="http://arlweb.msha.gov/OpenGovernmentData/OGiMSHA.asp">http://arlweb.msha.gov/OpenGovernmentData/OGiMSHA.asp</a></td>
<td>Data covers 1998 to current</td>
</tr>
<tr>
<td>National Occupational Respiratory Mortality System (NORMS)</td>
<td>Industry-specific and occupation-specific mortality data on multiple respiratory conditions</td>
<td>CDC NCHS &amp; NIOSH</td>
<td>County</td>
<td>State</td>
<td>Hypersensitivity Pneumonitis, Malignant Mesothelioma, All Pneumoconiosis, Asbestosis, Byssinosis, Coal Workers' Pneumoconiosis, Silicosis, Number of deaths, crude death rates, age-adjusted death rates, and years of potential life lost (YPLL) by year, age group, race, gender, or Hispanic origin</td>
<td><a href="http://webappa.cdc.gov/ords/norms.html">http://webappa.cdc.gov/ords/norms.html</a></td>
<td>All results for states and counties with less than 10 deaths are suppressed</td>
</tr>
<tr>
<td>OSHA Data Initiative</td>
<td>OSHA’s Lost Workday Injury and Illness Database, Injury/Illness rate for establishments in high-hazard industries</td>
<td>OSHA</td>
<td>Address</td>
<td>City, Zip Code, State</td>
<td>Work-related injury and illness data from employers within specific industry and employment size specifications</td>
<td><a href="https://www.osha.gov/pls/odi/establishment_search.html">https://www.osha.gov/pls/odi/establishment_search.html</a></td>
<td></td>
</tr>
<tr>
<td>Poison Control Center Data</td>
<td>Data from calls to a poison center</td>
<td>American Association of Poison Control Centers &amp; state poison control centers</td>
<td>Zip Code</td>
<td>City, County, State</td>
<td>Exposure substance, medical outcome</td>
<td><a href="http://www.aapcc.org/">http://www.aapcc.org/</a></td>
<td>Geographic information is based on caller information, which can include individuals calling on behalf of the worker (e.g. health care provider).</td>
</tr>
<tr>
<td>Workers' Compensation</td>
<td>Data on injury, illness and fatalities that are filed as workers’ compensation claims</td>
<td>State Workers’ Compensation Claims</td>
<td>Address-level</td>
<td></td>
<td>Injury type, occupation, firm risk class and experience factor</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Not all data sources with sub-state data are included in this table. To be included in these tables, data sources: 1) contain data at sub-state level; 2) contain data from 2010 or more recent; and 3) contain data relevant to occupational safety & health including: health outcomes, estimates of workers, risk factors for workers, and be obtainable by state agencies.
* Data elements available may vary by state

## Data Sources Containing Estimates of Workers or Businesses

<table>
<thead>
<tr>
<th>Name</th>
<th>Brief Description</th>
<th>Collected By/ Sources</th>
<th>Highest Spatial Resolution Typically Available</th>
<th>Other geographies available</th>
<th>Constructs captured on workers</th>
<th>Relevant Websites</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Community Survey (ACS), Summary Tables</td>
<td>Collects and produces population and housing estimates for 1, 3, and 5 year periods; an ongoing statistical survey rather than a decennial census</td>
<td>US Census Bureau</td>
<td>Census block but varies based on sample size and variable</td>
<td>Block-groups, Census tracts, ZCTAs, Counties, etc.</td>
<td>Estimated worker counts, usual hours-worked, Census Industry Codes, NAICS, SOC, median income, class of worker, employment status, commuting patterns</td>
<td><a href="https://www.census.gov/acs/www/data/data-tables-and-tools/american-factfinder/">https://www.census.gov/acs/www/data/data-tables-and-tools/american-factfinder/</a></td>
<td>Three-year files for 2012-2014 and later periods are no longer available.</td>
</tr>
<tr>
<td>American Community Survey (ACS) Public Use Microdata Sample (PUMS) Files</td>
<td>Untabulated records about individual people or housing units allowing data users to create custom tables not available through pretabulated (or summary) ACS data products</td>
<td>US Census Bureau</td>
<td>Public Use Micro Area</td>
<td>Region, Division, State</td>
<td>Place of work, residence, residence previous year, estimated worker counts, usual hours-worked and weeks worked, median income, class of worker, employment status, commuting patterns</td>
<td><a href="https://www.census.gov/programs-surveys/acs/data/pums.html">https://www.census.gov/programs-surveys/acs/data/pums.html</a></td>
<td>Public Use Microdata Areas (PUMAs) are non-overlapping areas that partition each state into areas containing about 100,000 residents. 1, 3, and 5 year PUMS files are available. Three-year files for 2012-2014 and later periods are no longer available.</td>
</tr>
<tr>
<td>County Business Patterns (census)</td>
<td>Annual detailed geographic, industry, and other information for U.S. business establishments</td>
<td>Employer identification number, US Census Bureau, Business Register</td>
<td>Zip Code</td>
<td>Nation, State, County</td>
<td>6-digit NAICS, employment size class, number of establishment, payroll, most NAICS sectors</td>
<td><a href="http://www.census.gov/eco/cbp/overview.htm">http://www.census.gov/eco/cbp/overview.htm</a></td>
<td></td>
</tr>
<tr>
<td>Name</td>
<td>Brief Description</td>
<td>Collected By/ Sources</td>
<td>Highest Spatial Resolution Typically Available</td>
<td>Other geographies available</td>
<td>Constructs captured on workers</td>
<td>Relevant Websites</td>
<td>Notes</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Current Employment Statistics (CES)</td>
<td>Surveys ~145,000 business and government agencies, representing approx. 623,000 individual worksites, to provide detailed industry data on employment, hours, and earnings of non-farm workers</td>
<td>Bureau of Labor Statistics</td>
<td>Metropolitan Areas</td>
<td>None</td>
<td>All non-farm employees, production and supervisor employees, women employees, average hourly earnings, average weekly earnings, average weekly overtime hours (in manufacturing industries). Seasonally and not seasonally adjusted.</td>
<td><a href="http://www.bls.gov/ces/data.htm">http://www.bls.gov/ces/data.htm</a></td>
<td></td>
</tr>
<tr>
<td>Longitudinal Employer-Household Dynamics (LEHD)</td>
<td>LEHD offers two data products that characterize workforce dynamics for specific groups: Quarterly Workforce Indicator (QWI) and LEHD Origin-Destination Employment Statistics (LODES)</td>
<td>Census Bureau</td>
<td>Workforce investment areas</td>
<td>State, County, Metro Areas</td>
<td>Industry, sex, race and age of workers in the QWI, and census block detail of Origin-Destination (OD), Residence Area Characteristics (RAC) and Workplace Area Characteristics (WAC).</td>
<td><a href="http://lehd.ces.census.gov/">http://lehd.ces.census.gov/</a></td>
<td>OnTheMap for Emergency Management is an online application available to the public to use in local analysis.</td>
</tr>
<tr>
<td>Quarterly Census of Employment and Wages (QCEW)</td>
<td>Publishes quarterly counts of employment and wages reported by employers</td>
<td>Census Bureau</td>
<td>County</td>
<td>Metropolitan Statistical Areas</td>
<td>Data by industry, ownership, and time period. Establishment counts, monthly employment levels, total wages, average weekly wages.</td>
<td><a href="http://www.bls.gov/cew/home.htm">http://www.bls.gov/cew/home.htm</a></td>
<td></td>
</tr>
</tbody>
</table>

Note: Not all data sources with sub-state data are included in this table. To be included in these tables, data sources: 1) contain data at sub-state level; 2) contain data from 2010 or more recent; and 3) contain data relevant to occupational safety & health including: health outcomes, estimates of workers, risk factors for workers, and be obtainable by state agencies.

* Data elements available may vary by state
Appendix C. Sample SAS Code for LAUS denominator

Sample code provided by Emily Hall, Texas Department of State Health Services

Note: You must replace the Libname, file locations, dataset names and variables with your own state specific data (highlighted areas):

libname hdd "U:\Occ Health\Sub state pilot";

/*Import inpatient numerator data*/
PROC IMPORT OUT= HDDInPt2012 DATATABLE= "tbl AHospDisData" DBMS=ACCESS REPLACE;
DATABASE="S:\epidata\PUDF Access DBs\2012P Hosp Disch 14-12-03.accdb";
SCANMEMO=YES;
USEDATE=NO;
SCANTIME=YES;
RUN;

/*Import outpatient numerator data*/
%macro op (file,q);
data out&q;
  infile &file DLM='09'X DSD firstobs=2 missover truncover;
  input service_quarter $ record_ID $ THCIC_ID $ Provider_name $ spec_unit $ sex_code $ pat_cnty $ pat_state $ pat_zip $ pat_country $ PHR $ LOS $ Pat_age $ race $ ethnicity $ _st_Pmt_Src $;
  if _st_Pmt_Src='WC';
run;
%mend;

%op (file='U:\Occ Health\Sub state pilot\OP_PUDF_base1q2012_tab.txt', q=1);
%op (file='U:\Occ Health\Sub state pilot\OP_PUDF_base2q2012_tab.txt', q=2);
%op (file='U:\Occ Health\Sub state pilot\OP_PUDF_base3q2012_tab.txt', q=3);
%op (file='U:\Occ Health\Sub state pilot\OP_PUDF_base4q2012_tab.txt', q=4);
data HDDOutpt2012;
  set out1-out4;run;

/*Calculate inpatient and outpatient work-related hospital discharge case counts*/
%macro sum (type,year);
proc sql;
  create table hdd.sum&type&year as select
    distinct Pat_cnty, Pat_state, count(RECORD_ID) as Total&type
  from HDDInPt2012;
%mend;
DATA hdd.allhdd;
merge hdd.suminpt2012 hdd.sumoutpt2012;
by pat_cnty;
if total_inpt='.' then total_inpt=0;
if total_outpt='.' then total_outpt=0;
Total_cases=total_inpt+total_outpt;
rename pat_cnty=County;run;

DATA all;
merge hdd.allhdd hdd.denom (where=(state='48'));
by county;run;

/*Calculate OHIs*/
PROC SQL;
CREATE TABLE hdd.Measures AS
SELECT County, County_Name_State_Abbreviation AS County_name, Pat_state, total_ip, employed, ((total_IP/employed)*100000) AS Rate
QUIT;
Appendix D. Sample SAS Program to Crosswalk WC employer city/town into the 48 POWPUMAs (v2000 US Census boundaries) for Massachusetts – numerator data

/*********************************************************/
/* Creating PUMAS (2000 version) from employer city/town/county subdivisions, */
/* listed in WC database. (contact: mydzung.chu@state.ma.us) */
/* CAUTION: PUMA 2010 VERY different from PUMA 2000 version. */
/* ACS data prior to 2012 is in PUMA v2000. ACS 2012 and forward is PUMA */
v2010.
/* Crosswalk available at: http://mcdc2.missouri.edu/websas/geocorr2k.html */
/* Variable name: */
/* POWPUMA = place of work PUMA in this dataset */
/* COMPANY_CITY = name of employer */
/*********************************************************/
libname num '.'; /*^Enter location of numerator dataset */
data num.WCdataset2;
set num.WCdataset1;

**Crosswalk employer city/town -> PUMA (v2000);**
if COMPANY_CITY IN ('STOCKBRIDGE', 'NEW ASHFORD', 'HANCOCK',
'CLARKSBURG', 'ADAMS', 'CHESHIRE', 'DALTON', 'GREAT
BARRINGTON', 'HINSDALE', 'LANESBOROUGH', 'LEE', 'LENOX', 'NORTH
ADAMS', 'PITTSFIELD', 'RICHMOND', 'WEST STOCKBRIDGE', 'WILLIAMSTOWN')
  then powpuma=100;
else if COMPANY_CITY IN ('ASHFIELD', 'ATHOL', 'BECKET', 'BERNARDSTON', 'CHARLEMONT', 'CHESTER', 'CHESTERFIE
LD', 'COLRAIN', 'CONWAY', 'CUMMINGTON', 'DEERFIELD', 'EGREMONT', 'FLORIDA', 'GRANVIL
LE', 'GREENFIELD', 'HARDWICK', 'HUBBARDSTON', 'LEVERETT', 'LEYDEN', 'MONTAGUE', 'NEW
BRAINTREE', 'NEW
MARLBOROUGH', 'NORTHFIELD', 'ORANGE', 'PERU', 'PETERSHAM', 'PLAINFIELD', 'SHEFFIELD
', 'SHELBURNE', 'WARWICK', 'WHATELY', 'WINDSOR', 'NEW SALEM', 'WESTHAMPTON',
'ERVING', 'OTIS', 'PHILLIPSTON', 'ROWE', 'PELHAM', 'ROYALSTON', 'SANDISFIELD',
'SAVOY', 'SHUTESBURY', 'TOLLAND', 'TYRINGHAM', 'WASHINGTON', 'WENDELL', 'WORTHINGTON
N', 'ALFORD', 'BLANDFORD', 'BUCKLAND', 'GILL', 'GOSHEN', 'HAWLEY', 'HEATH', 'LONGMEAD
OW', 'MIDDLEFIELD', 'MONROE', 'MONTEREY', 'MOUNT WASH', 'MOUNT WASHINGTON')
  then powpuma=200;
else if COMPANY_CITY IN ('ASHBURNHAM', 'ASHBY', 'FITCHBURG', 'GARDNER', 'LEOMINSTER',
'LUNENBURG', 'TEMPLETON', 'WESTMINSTER', 'WINCHENDON')
  then powpuma=300;
else if COMPANY_CITY IN ('AYER', 'BERLIN', 'BOLTON', 'HARVARD', 'HUDSON',
'LANCASTER', 'MARLBOROUGH', 'SHIRLEY', 'STOW', 'TOWNSEND')
  then powpuma=400;
else if COMPANY_CITY IN ('BILLERICA', 'DUNSTABLE', 'CHELMSFORD', 'DRACUT',
'GROTON', 'PEPPERELL', 'TEWKSBURY', 'TYNGSBOROUGH', 'WESTFORD')
  then powpuma=500;
else if COMPANY_CITY IN ('LOWELL')
  then powpuma=600;
else if COMPANY_CITY IN ('ANDOVER', 'LAWRENCE', 'METHUEN')
    then powpuma=700;
else if COMPANY_CITY IN ('BOXFORD', 'GEORGETOWN', 'GROVELAND', 'HAVERHILL', 'MERRIMAC', 'NORTH ANDOVER', 'WEST NEWBURY')
    then powpuma=800;
else if COMPANY_CITY IN ('AMESBURY', 'ESSEX', 'GLOUCESTER', 'IPSWICH', 'NEWBURY', 'NEWBURYPORT', 'ROCKPORT', 'ROWLEY', 'SALISBURY')
    then powpuma=900;
else if COMPANY_CITY IN ('WENHAM', 'DANVERS', 'HAMILTON', 'LYNNFIELD', 'MIDDLETOWN', 'PEABODY', 'TOPSFIELD')
    then powpuma=1000;
else if COMPANY_CITY IN ('BEVERLY', 'MANCHESTER', 'MANCHESTER-BY-THE-SEA', 'MARBLEHEAD', 'SALEM', 'SWAMPS'COTT')
    then powpuma=1100;
else if COMPANY_CITY IN ('LYNN', 'NAHANT', 'SAUGUS')
    then powpuma=1200;
else if COMPANY_CITY IN ('BURLINGTON', 'READING', 'WAKEFIELD', 'WILMINGTON')
    then powpuma=1300;
else if COMPANY_CITY IN ('ACTON', 'BEDFORD', 'BOXBOROUGH', 'CARLISLE', 'CONCORD', 'LITTLETON', 'MAYNARD', 'SUDSBURY', 'WAYLAND')
    then powpuma=1400;
else if COMPANY_CITY IN ('EAST BROOKFIELD', 'OAKHAM', 'BARRE', 'BOYLSTON', 'BRIMFIELD', 'BROOKFIELD', 'CLINTON', 'NORTH BROOKFIELD', 'NORTHBOROUGH', 'PRINCETON', 'RUTLAND', 'SPENCER', 'STERLING', 'WARREN', 'WEST BROOKFIELD', 'WESTBOROUGH')
    then powpuma=1500;
else if COMPANY_CITY IN ('AMHERST', 'GRANBY', 'HADLEY', 'HATFIELD', 'NORTHAMPTON', 'SOUTH HADLEY', 'SUNDERLAND', 'WILLIAMSBURG')
    then powpuma=1600;
else if COMPANY_CITY IN ('CHICOPEE', 'EASTHAMPTON', 'HOLYOKE')
    then powpuma=1700;
else if COMPANY_CITY IN ('AGAWAM', 'HUNTINGTON', 'MONTGOMERY', 'RUSSELL', 'SOUTHAMPTON', 'SOUTHWICK', 'WEST SPRINGFIELD', 'WESTFIELD')
    then powpuma=1800;
else if COMPANY_CITY IN ('SPRINGFIELD')
    then powpuma=1900;
else if COMPANY_CITY IN ('BELCHERTOWN', 'EAST LONGMEADOW', 'HAMPDEN', 'LUDLOW', 'MONSON', 'PALMER', 'WARE', 'WILBRAHAM')
    then powpuma=2000;
else if COMPANY_CITY IN ('CHARLTON', 'DOUGLAS', 'DUDLEY', 'HOLLAND', 'NORTHBRIDGE', 'OXFORD', 'SOUTHBRIDGE', 'STURBRIDGE', 'SUTTON', 'UXBRIDGE', 'WALES', 'WEBSTER')
    then powpuma=2100;
else if COMPANY_CITY IN ('auburn', 'GRAFTON', 'HOLDEN', 'LEICESTER', 'MILLBURY', 'PAXTON', 'SHREWSBURY', 'WEST BOYLSTON')
    then powpuma=2200;
else if COMPANY_CITY IN ('Worcester')
    then powpuma=2300;
else if COMPANY_CITY IN ('ASHLAND', 'HOLLISTON', 'HOPKINTON', 'MEDWAY', 'MILFORD', 'MILLIS', 'SOUTHBOROUGH', 'UPTON')
    then powpuma=2400;
else if COMPANY_CITY IN ('FRAMINGHAM', 'NATICK', 'SHERBORN')
    then powpuma=2500;
else if COMPANY_CITY IN ('DEDHAM', 'DOVER', 'NEEDHAM', 'WELLESLEY', 'WESTON', 'LINCOLN')
    then powpuma=2600;
else if COMPANY_CITY IN ('ARLINGTON', 'BELMONT', 'LEXINGTON', 'WALTHAM',
  'WATERTOWN')
  then powpuma=02700;
else if COMPANY_CITY IN ('MELROSE', 'STONEHAM', 'WINCHESTER', 'WOBURN')
  then powpuma=2800;
else if COMPANY_CITY IN ('CHELSEA', 'REVERE', 'WINTHROP')
  then powpuma=2900;
else if COMPANY_CITY IN ('MALDEN', 'MEDFORD')
  then powpuma=3000;
else if COMPANY_CITY IN ('EVERETT', 'SOMERVILLE')
  then powpuma=3100;
else if COMPANY_CITY IN ('CAMBRIDGE')
  then powpuma=3200;
else if COMPANY_CITY IN ('BOSTON')
  then powpuma=3300;
else if COMPANY_CITY IN ('BROOKLINE', 'NEWTON')
  then powpuma=3400;
else if COMPANY_CITY IN ('MEDFIELD', 'NORFOLK', 'NORWOOD', 'SHARON', 'WALPOLE',
  'WESTWOOD')
  then powpuma=3500;
else if COMPANY_CITY IN ('BELLINGHAM', 'BLACKSTONE', 'FOXBOROUGH', 'FRANKLIN',
  'HOPEDEALE', 'MENDON', 'MILLVILLE', 'PLAINVILLE', 'WRENTHAM')
  then powpuma=3600;
else if COMPANY_CITY IN ('BRAINTREE', 'CANTON', 'HOLBROOK', 'RANDOLPH',
  'STOUGHTON')
  then powpuma=3700;
else if COMPANY_CITY IN ('MILTON', 'QUINCY')
  then powpuma=3800;
else if COMPANY_CITY IN ('COHASSET', 'HANOVER', 'HINGHAM', 'HULL', 'NORWELL',
  'ROCKLAND', 'SCITUATE', 'WEYMOUTH')
  then powpuma=3900;
else if COMPANY_CITY IN ('ABINGTON', 'AVON', 'BROCKTON')
  then powpuma=4000;
else if COMPANY_CITY IN ('BRIDGEWATER', 'EAST BRIDGEWATER', 'EASTON', 'HALIFAX',
  'HANSON', 'LAKEVILLE', 'MIDDLEBOROUGH', 'PLYMPTON', 'RAYNHAM', 'WEST
  BRIDGEWATER', 'WHITMAN')
  then powpuma=4100;
else if COMPANY_CITY IN ('BERKLEY', 'DIGHTON',
  'MANSFIELD', 'NORTON', 'TAUNTON')
  then powpuma=4200;
else if COMPANY_CITY IN ('ATTLEBORO', 'NORTH ATTLEBORO', 'REHOBOTH', 'SEEKONK',
  'SWANSEA')
  then powpuma=4300;
else if COMPANY_CITY IN ('FALL RIVER', 'SOMERSET', 'WESTPORT')
  then powpuma=4400;
else if COMPANY_CITY IN ('ACUSHNET', 'DARTMOUTH', 'FAIRHAVEN',
  'FREETOWN', 'MARION', 'MATTAPOOSEITT', 'NEW BEDFORD', 'ROCHESTER')
  then powpuma=4500;
else if COMPANY_CITY IN ('CARVER', 'DUXBURY', 'KINGSTON',
  'MARSHFIELD', 'FEBROKE', 'PLYMOUTH', 'WAREHAM')
  then powpuma=4600;
else if COMPANY_CITY IN ('TRURO', 'WELLFLEET', 'AQUINNAH', 'CHILMARK',
  'GOSNOLD', 'BARNSTABLE', 'BOURNE', 'CHATHAM',
  'EASTHAM', 'EDGARTOWN', 'FALMOUTH', 'NANTUCKET', 'OAK
  BLUFFS', 'ORLEANS', 'PROVINCETOWN', 'TISBURY', 'WEST TISBURY')
  then powpuma=4700;
else if COMPANY_CITY IN ('BREWSTER', 'DENNIS', 'HARWICH', 'MASHPEE', 'SANDWICH', 'YARMOUTH')
  then powpuma = 4800;

if COMPANY_CITY IN ('OUT OF STATE', 'UNKNOWN', 'UNKNOWN MASSACHUSETTS')
  then delete;

run;
Appendix E. Resources for GIS Training and Mapping

Training:

- **ESRI ArcGIS Video Tutorials:** [http://video.arcgis.com/channel/11/tutorials/series](http://video.arcgis.com/channel/11/tutorials/series)
- **George Mason University Video Tutorials: Census Data and GIS Tutorial**
  [www.youtube.com/channel/UC4XEKGSQKeyKkaX0u8vOFCw](http://www.youtube.com/channel/UC4XEKGSQKeyKkaX0u8vOFCw)
- **CDC GIS Training:**
  - [https://www.cdc.gov/gis/](https://www.cdc.gov/gis/)
- **Michigan Center for Public Health: Geographical Information Systems (GIS) for Public Health: An Introduction to GIS and Hands-On Use of Epi-Map:**
- **US Census Geographic Areas Reference Manual:**
  [www.census.gov/geo/reference/garm.html](http://www.census.gov/geo/reference/garm.html)
- **Harvard School of Public Health: The Public Health Disparities Geocoding Project Monograph:**
  [www.hsph.harvard.edu/thegeocodingproject](http://www.hsph.harvard.edu/thegeocodingproject)
- **CSTE Webinar: Contextualizing Health Outcomes: A How-to for Spatial Data Analysis (11/21/2013):**

Tools:

- **QGIS: A Free and Open Source Geographic Information System:**
- **Northwestern University Library: Make a Map!: GIS Lite & Quick Map Tools: Free Map Tools:**
  [http://libguides.northwestern.edu/gislite](http://libguides.northwestern.edu/gislite)
  - **Tableau Public:** [https://public.tableau.com/s/download](https://public.tableau.com/s/download)
    - Create maps and other types of data charts; free download, link to training materials
  - **Open Heat Map:** [http://www.openheatmap.com/](http://www.openheatmap.com/)
    - Upload data and display on map - [documentation](https://www.openheatmap.com/) provided
    - This free downloadable application is produced by ESRI, creator of the popular ArcGIS software suite.
  - **DIYMaps.net:** [http://diymaps.net/](http://diymaps.net/)
    - Very basic. Download to your computer and create maps that color-code groups of counties within a state, or groups of states in the U.S., Canada, or Mexico.
  - **CartoDB:** [https://cartodb.com/](https://cartodb.com/)
○ Software as a service (SaaS) cloud computer platform that allows for GIS and web mapping tools within your web browser. Accounts are free up to a certain size with larger storage and features associated with fees. Tool provides the ability to quickly create maps from a spreadsheet that has geographic identifiers such as states, nations or X,Y coordinates.

- **Indie Mapper:** [http://indiemapper.com/app/](http://indiemapper.com/app/)

Appendix F. List of Workgroup Members

Karla Armenti, ScD, University of New Hampshire, karla.armenti@unh.edu
Juanita Chalmers, MPH, Florida Department of Health, juanita.chalmers@flhealth.gov
MyDzung Chu, MSPH, Massachusetts Department of Public Health, mydzung.chu@state.ma.us
Alicia Fletcher, MPH, New York State Department of Health, alicia.fletcher@health.ny.gov
Martha Jones, PhD, Vanderbilt University, martha.w.jones@Vanderbilt.edu
Amy Patel, MPH, Council of State and Territorial Epidemiologists, amypatel@cste.org
Jeff Shire, National Institute for Occupational Safety and Health, jshire@cdc.gov
Derry Stover, MPH, Nebraska Department of Health and Human Services, derry.stover@nebraska.gov
Amita Toprani, MD MPH, New York City Department of Health and Mental Hygiene, atoprani@health.nyc.gov
Meredith Towle, MPH, Wyoming Department of Workforce Services, meredith.towle@wyo.gov

For questions or if you would like to contribute to this document:
Contact Amy Patel at the CSTE National Office, amypatel@cste.org, 770-458-3811.