Annual Report, July 2015—June 2016
Maximizing Use of Massachusetts Workers' Compensation Data for Surveillance and Prevention (U60OH010893-01)

SECTION I.
Project Summary: Workers' compensation (WC) claim records are an important source of information about work-related injuries and illnesses (WRII) sustained by workers across a range of industries and worker populations. However, their utility in providing information for action to protect MA workers has yet to be fully realized. The overarching aim of this project is to reduce the incidence of WRII among MA workers by maximizing the state's capacity to use available WC data to conduct surveillance and promote prevention. The intent is to create sustained improvements in the WC data system and develop methods to generate actionable information for preventing WRII on an ongoing basis. Specific Aims are to: 1) develop, implement, and document protocols for assigning standardized codes to key data elements, essential for surveillance; 2) compare alternative sources of statewide employment data for use in conjunction with WC claim data to generate statewide WRII rates and recommend the most appropriate denominator for ongoing use; and explore approaches to generating local area WRII rates to inform community health planning and practice; 3) prepare a technical evaluation report assessing completeness, usefulness and limitations of the current WC database and denominator data source, with recommendations for WC data system improvements; 4) analyze three years WC claim data (2014 – 16) and prepare and disseminate a descriptive report of WRII in MA that identifies health and safety priorities; and 5) increase opportunities for integrative WRII prevention and WC research.

Relevance: This project, still in its early phases, has the potential to substantially enhance our ability to document WRII and more effectively allocate limited prevention resources to protect workers in our state. Findings may be used by multiple state agencies as well as private sector partners. The collaborative nature of the project involving the multi-agency effort of the Massachusetts Departments of Public Health (lead agency), Industrial Accidents (DIA) and Labor Standard (DLS), that have a successful history of working together to protect the health of MA workers, will further strengthen the state infrastructure for workplace safety and health.

Key Personnel:
Letitia Davis, Director (PI); Kathleen Grattan, Epidemiologist (Project Coordinator)
Occupational Health Surveillance Program, MA Department of Public Health
250 Washington St. 4th flr., Boston, MA 02108, 617-624-5626, Letitia.Davis@state.ma.us

Yafa Rachmany-Arkin, Senior Applications Developer
MA Department of Industrial Accidents, 1 Congress St. #100, Boston, MA 02114
617-727-4900, Yafa.Rachmany-Arkin@state.ma.us

Imani K. Bishop, Program Coordinator II, Occupational Safety and Health Statistics Program
MA Department of Labor Standards, 19 Staniford St., Boston, MA 02114
617-626-6958, Imani.Bishop@state.ma.us

Prime Institution/Website Link: Occupational Health Surveillance Program, Massachusetts Department of Public Health; www.mass.gov/dph/ohsp

SECTION II.
Highlights of Major Accomplishments and Short Term Impact
Major accomplishments during this first year of the three year project are highlighted below.

Development of the research file. DPH has processed WC claim records provided by DIA for the first 18 months of the study period (2014-16) to create a research file of private sector only, indemnity claims for WRII resulting in five or more lost workdays (N= 49,556 incident claims). Beginning in 2013, insurers have had the option of submitting claim information (First Reports of Injury only) through the Electronic Data Interchange (EDI). Consequently, the research file that DPH received contained additional data elements for those claims originating through the EDI and accounting for approximately 54% of all claims in the 18-month research file.

Aim 1. Code key data elements essential for surveillance.
Coding Event (DPH lead with DLS) The MA WC data are not routinely coded for event. DPH has applied NIOSH's machine learning Autocoder to assign event codes to the claims in the research file. This Autocoder uses a regularized logistic regression model (trained on Ohio's WC data) to assign 2-digit event codes from the BLS Occupational Injury and Illness Classification System (OIICs) to a free text field describing the incident. The Autocoder did not perform as well as expected in predicting the 'true' Event codes when a small random sample (n = 1,338) of the WC claims auto-coded for event were compared with 'gold standard' OIICs event codes manually assigned by expert coders at DLS; however, the Autocoder is showing promise with some modifications (currently in progress).

DPH is also working to explore the validity of the Workers' Compensation Insurance Organization (WCIO) event codes which have been found to be available for almost 100% of the claims submitted through the EDI. A validation sample of approximately 50 - 100 claims per WCIO event category has been created (n= 3,629); DLS coders, blinded to the WCIO codes, have manually assigned 'gold standard' OIICs events codes to these records. A crosswalk of OIICs event to WCIO event codes has been developed for the validation analysis, currently underway.

Code Industry (NAICS) by linking MA WC data to MA Unemployment Insurance (UI) data (DIA lead with DPH). DPH has begun the process of refining the WC-UI linkage protocol developed by the Ohio Board of Workers' Compensation and performed metrics (e.g., percent missing values) on the key data variables to be used in the linkage. The WC-UI linkage will be carried out by DIA when the UI data become available (See below under Collaboration and Partnership Development).

Aim 2. **Compare alternative sources of statewide employment (denominator) data for use in computing WRII rates; and explore approaches for generating local area rates** (DPH lead). DPH has generated Full Time Equivalent (FTE) estimates for major industry sectors in MA using data from four alternative data sources - the American Community Survey Public Use Microdata sample (ACS), Current Population Survey Basic (CPS), Current Employment Statistics survey (CES) and the U.S. Labor Productivity and Costs (LPC) - in conjunction with data from the Quarterly Census of Employment and Wages (QCEW). Employee counts from the QCEW were

adjusted using estimated ratios of hours worked per employee (FTE/EE) from each of the four data sources. DPH has also refined SAS-based protocols to derive standard errors (SEs) for FTEs from ACS[1], CPS and CPS Annual Social and Economic Supplement (ASEC) and computed confidence intervals for these FTE estimates overall and at the industry (NAICS) sector level. Findings have been shared with NIOSH staff, and project advisors and discussions are underway to select the most valid, reliable denominator for ongoing use.

A compendium of the potential denominator data sources with information on population covered, strengths and limitations in terms of comparability with the statewide WC data, etc., has been created. Also included in this compendium are methodological considerations and technical *How-to* instructions for generating WRII rates at various sub-state geographic areas (e.g. city/town, county, and statistical Public Use Microdata Area or PUMA). Some of this information was shared at the Annual Northeast Regional Occupational Disease and Injury Surveillance Conference in May 2016. Methods to summarize WC claims and claim rates at the sub-state level by both location of employment and residence are being piloted. As part of a related project, DPH has also conducted outreach to DPH programs and community partners to identify specific communities that may be interested in including occupational health data in their community health profiles.

Aim 3. Prepare a technical evaluation report (DPH lead with DIA and DLS).
An initial outline of the evaluation report has been developed, with input from project advisors. DPH has reviewed the research file to assess completeness of information on key data elements and become familiar with newly available data elements obtained through the EDI. In addition to WCIO event codes described above, 6-digit insurer-provided NAICS codes are available for one-half of the claims submitted through the EDI. These codes will be compared with the 'gold standard' UI NAICS codes and their validity will be assessed.

Aims 4 and 5. Conduct analyses of three years of WC claim data, prepare and disseminate a descriptive report of WRII and Increase opportunities for integrative WRII prevention and WC research.
A detailed three year project work plan has been developed. Activities to accomplish Aims 4 and 5 will be carried out later in the project period.

SHORT TERM IMPACTS:
Project activities have contributed to several enhancements in DIA's management of the WC claim data for use in surveillance and epidemiologic research. As a result of collaboration on this project, DIA has incorporated into the query they routinely use to pull records from the DIA database some of the steps that DPH uses to process/de-duplicate the WC claim records. Also, DPH's preliminary analysis of the WC data uncovered an issue related to the initial processing of claims – a number of claims should have been rejected by the EDI because they were missing information in the incident description field, a required field. This finding prompted DIA to

---

[1] ACS PUMS protocols adapted from methodology and SAS program developed by Dr. Martha Jones, Vanderbilt University, via Karen Louie, California Department of Finance, and Dale Garrett, U.S. Census Bureau.

look closely at the source of the claims missing incident description and is now working to resolve this issue.

Collaboration and Partnership Development

This project builds on a successful history of collaboration between DPH, DIA, and DLS and is serving to strengthen day to day working relationships. Interagency service agreements to channel project funds from DPH to DIA and DLS have been executed. A tri-agency study team has been established. Members are in regular communication and have met several times together with external experts who are serving as project advisors to review overall project plans, project findings, and methods for data linkage and event coding.

A significant advancement this year is the first executed Memorandum of Understanding (MOU) between DIA and a fourth agency - the MA Department of Unemployment Assistance (DUA) that maintains the UI data. This involved extensive negotiations and legal review on the part of all agencies involved. The MOU allows DIA to access data elements in the UI database essential for data linkage to abstract industry (NAICS) codes and size class (number of employees) codes to assign to the WC claim records. This MOU also allows DPH access to select UI data elements (e.g., two-digit industry code) needed for data analyses. A UI data file has been provided to DIA as of July 7, 2016. The DIA's recognition of its agency's role in prevention as well as the importance of enhancements to its data system which are necessary to inform prevention efforts played a key role in moving the MOU forward. DLS's contribution has also been substantial. To date, DLS expert coders have manually assigned OIICs event codes to almost 5,000 records to serve as 'gold standard' codes.

Knowledge and Capacity Improvement from the Research Activity

DPH has had access to DIA's WC claim data for many years for purposes of occupational health surveillance. This project has both expanded the range of variables now available to DPH for analysis (new data elements available through EDI) and substantially increased DPH's understanding of DIA's data management system. In turn, as described above, the project has led to several enhancements in DIA's management of their data for surveillance. Feedback from DLS coders as well as the Autocoder's less-than-optimal ability to code the event resulting in injury or illness point to needed enhancements to enrich the quality of narrative information included in the text field describing the incident. The new MOU between DIA and DUA will allow for efficient assignment of industry (NAICS) codes (as well as size class codes) to the full range of WC claim records, previously assigned manually by DPH for special studies only. Ongoing input from local project advisors with expertise in the policy-related and analytic issues surrounding WC as well as invaluable technical assistance from NIOSH and experts in other states are enhancing DPH's knowledge and capacity to conduct meaningful epidemiologic analyses of the WC data and develop recommendations for data system improvements.