CDPH
California Department of Public Health

Using Workers' Compensation Claims for Occupational Health Surveillance in California, 2015–2017

Confidential - Low

Contents

Using Workers' Compensation Claims for Occupational Health Surveillance in California, 2015-2017 1

Key Findings 3

Background 4

Why Workers' Compensation Claims for Occupational Health Surveillance? 4

Methods 6

Data Sources 6

Matching Process 7

Descriptive Analysis 7

Results 9

Industry 10

Tenure 10

Seasonal Variation 11

Establishment Size 11

Evaluating Risk 12

High Rate, High Count Industries 12

High Rate, Low Count 12

Low Rate, High Count 12

Discussion 13

Appendix 1 – Matching Results 15

References 16

Key Findings

This project calculated accurate industry-specific rates of occupational injuries that lead to a California's workers compensation claim

by combining four datasets: Workers' Compensation Claims, the Base Wage File, the Quarterly Census of Employment and Wages, and US Census data.

1. **There were 3.3 injuries per 100 workers** between 2015-2017.
2. **Rates were higher** among men, workers aged under 25 years old, and workers who had been in their jobs for two years or less.
3. **Employers with more than 1,000 employees had the highest rates of injury.**
4. **Workers in the transportation, warehousing, and utilities sectors had high rates of injury.** Interurban and rural bus transportation and couriers had especially high rates.
5. **Strains were the most common type of injury and nature of injury.** Strains occur when workers are lifting, pushing or pulling, twisting, reaching or engaging in repetitive tasks.
6. **Workers' compensation claims are beneficial for occupational safety and health surveillance research** because they include all private sector industries and occupations and have coded fields for cause of injury, nature of injury, and part of body injured.

Using workers' compensation data in public health research increases our understanding of the risk factors for work-related injuries and illness. This information can be used by organizations and individuals involved in work-related injury prevention and occupational safety and health regulation to better prevent costly work-related injuries.

Background

Why Workers' Compensation Claims for Occupational Health Surveillance?

Work-related injuries exert an emotional, physical, and mental toll on workers, families, communities, and employers. In addition to the medical costs associated with injuries, lost work time often means a loss of economic support; mental and psychological effects of injuries may occur as well. Understanding and tracking the trends, magnitude, and risk factors of occupational illnesses and injuries is essential for preventing them. Public health surveillance of these injuries can identify risky industries and occupations, detect new hazards or worsening conditions, and highlight interventions that are effective at preventing injuries.

While the Bureau of Labor Statistics (BLS) Census of Fatal Occupational Injuries (CFOI) is a near-comprehensive source for information on fatal work-related injuries in the U.S., no comparable data source exists for comprehensive information on non-fatal work-related injuries or for work-related illnesses.[1] The standard source for identifying work-related injuries, the BLS Survey of Occupational Injuries and Illnesses (SOII), has significant limitations. It is known to undercount injuries and cannot generate detailed information on the types of injuries within a specific industry or occupation group because of small sample sizes and confidentiality requirements.[2,3]

In California, any injury that requires time off or medical treatment beyond first aid is eligible for workers' compensation. Workers' compensation data is electronically submitted by claims administrators, and California has the largest workers' compensation system in the country. Information from these claims has long been used by actuaries to assess job risk and set rates of premiums they charge companies for insurance policies. Over the last 25 years, claims data has become a prominent data source for occupational safety and health researchers and public health practitioners. Some of these topics include tracking work-related asthma, pesticide illness, Valley Fever, carpal tunnel syndrome (CTS), sharps injuries in non-healthcare workers, workload in janitorial workers, first responder health, risks faced by young workers, the scope of workplace violence, and COVID-19. These projects range in complexity and build our understanding of the utility, completeness, and challenges of using workers' compensation data for public health surveillance and prevention efforts.

However, since there is no information on the number of workers *within* California's workers' compensation system, **calculating rates of injury requires combining several different databases**. In addition, industry and occupation, two of the main risk factors of interest in occupational safety and health, are not consistently or accurately coded within the workers' compensation claim database. To further utilize workers' compensation for prevention, this project builds on previous work by matching multiple datasets to create accurate rates of workers' compensation claims.

Methods

Data Sources

Workers' Compensation Information System (WCIS)

All private sector workers' compensation claims in the California Division of Workers' Compensation Workers' Compensation Information System (WCIS) with a date of injury between 1/1/2015 and 12/31/2017 were eligible for inclusion in the industry and demographic rate calculations (local, state, and federal workers were excluded).

Base Wage File (BWF)

The Base Wage File (BWF) is compiled by the Employment Development Department from quarterly employer reporting of employees along with an employer identifier or key. Using the BWF ensures that the same industry code is assigned to both the numerator and the denominator.

Quarterly Census of Employment and Wages (QCEW)

The Employment Development Department maintains microdata that are the primary source of information for the Quarterly Census of Employment and Wages (QCEW). QCEW microdata provides a comprehensive tabulation of employment and wage information for workers and serves as a near census of monthly employment and quarterly wage information. It was used for counts of workers by industry. The Employment Development Department database includes the BWF employer key that links workers' compensation claims to the correct employer.

American Community Survey (ACS)

The American Community Survey (ACS) is a household survey by the U.S. Census Bureau that asks about jobs and occupations, educational attainment, demographics, and other social and economic aspects. The primary job of all employed individuals is industry coded to a Census Industry Code (CIC code), which was cross-walked to a 3-digit North American Industry Classification System (NAICS) code. These 3-digit codes were used to calculate the full-time equivalent (FTE) for each 4-digit NAICS code it contained. Free, publicly available Public Use Microdata files (PUMS) were used to calculate the ratio of FTE to total workers by industry in California for 2015, 2016, and 2017.

Bureau of Labor Statistics (BLS) Current Population Survey

Publicly available data from the BLS Current Population Survey (CPS) January 2018 supplement national data provided counts of workers by tenure category.

Matching Process

Private sector claims were selected from WCIS and matched to worker records in the BWF in the quarter of injury. Ninety-two percent of private sector claims matched at least one BWF record. Of those, 80% had one job and 20% had multiple jobs. We identified an industry code for each worker based on the available information. If we could not identify the correct industry by matching using the employee key, we used the industry code from the BWF records when there was a single industry code across all BWF records, or we used the industry code in the workers' compensation records (See Appendix A). Among workers with multiple jobs, an employer was selected 85% of the time.

Identifying Industry Among Workers at Multi-Establishment Firms

A multi-establishment firm is an employer where work occurs at several different locations, or within different industries at the same location. According to QCEW data, multi-establishment firms make up less than 1% of California private sector employers but employ 37% of workers; these numbers did not meaningfully change during the study period. We assigned a single NAICS code for each multi-establishment firm if at least 80% of employees were in a single NAICS code. Among claims that linked to a multi-establishment employer, we were able to identify a NAICS code for 85% of claims.

Descriptive Analysis

Rates were calculated by age group, sex, and tenure groups. Rates for age and sex categories were calculated using ACS data alone, because QCEW does not have any information on age or sex. For tenure, the distribution of tenure with current employer was taken from the BLS Current Population Survey (CPS) January 2018 supplement national data, counts of workers in each tenure category were derived from the product of time in a job and total FTE.

Industry and quarter combinations were suppressed in the results for any 4-digit industry with less than 11 claims or any industry where a single employer made up 80% or more of the employment or there were fewer than 5 employers. Claims by industry

were summed, then divided by the FTE by industry, and multiplied by 100 to calculate the rate per 100 FTEs. Rates were calculated first by 4-digit NAICS.

Counts by quarter were examined to identify seasonal variation. Industries with greater than 100 injuries per year, a rate greater than 6 injuries per 100 FTE, and where the coefficient of variation was in the top 10% were included. Rates by size class were calculated by aggregating all establishments by size category. No data exist on hours worked in the state by size class, so these calculations are for number of employees directly from QCEW data, and were not adjusted by FTE. Because of aggregating higher and lower risk employers, these rates were presented by 1,000 employees.

Results

Overall

There were 1,386,948 workers' compensation claims included in this analysis, and counts and rates by age, sex, and job tenure are included in Table 1 below. The overall rate of injury among private sector workers was 3.3 per 100 FTE. The rate among men was 1.07 times higher than the rate among women. While most injuries occurred among workers aged 25-34, the highest rates of injury was among 15-18 year olds and 19-24 year olds (6.2 and 4.5 per 100 FTE, respectively).

Table 1. Counts and Incident Rates of Occupational Injury and Illness by selected demographic and job characteristics, private sector, California 2015-2017

	Claims	Percent	Annual FTE	Annual Rate per 100 FTE
Overall	**1,386,948**	**100%**	**13,825,229**	**3.3**
Sex*				
Female	584,328	42%	6,769,956	2.9
Male	789,763	57%	8,543,437	3.1
Unknown	12,857	1%	--	--
Age Category*				
15 - 18	16,704	1%	90,451	6.2
19 - 24	188,733	14%	1,392,625	4.5
25 - 34	345,481	25%	4,037,010	2.9
35 - 44	287,081	21%	3,567,361	2.7
45 - 54	300,015	22%	3,336,962	3.0
55 - 64	205,194	15%	2,325,629	2.9
65+	40,575	3%	558,438	2.4
Unknown	3,165	<1%	--	--
Job Tenure+				
Under one year	451,353	33%	3,414,887	4.4
1 to <2 years	184,565	13%	1,056,624	5.8
2 to <5 years	235,695	17%	3,552,707	2.2
5 to <10 years	175,411	13%	2,878,918	2.0
10 years or more	249,027	18%	4,410,257	1.9
Unknown	9,0897	7%	--	--

Industry

The 20 industries with the highest rates of injury are in Table 1. In three industries, more than 10 of every 100 FTE workers were injured each year: interurban and rural bus transportation (NAICS 4852, 13.1 per 100 FTE), building material and supplies dealers (NAICS 4441, 11.9 per 100 FTE), and couriers (NAICS 4921, 11.1 per 100 FTE), highlighting an increased risk in the transportation, warehousing & utilities sector.

Table 2. North American Industrial Classification System (NAICS) industries with the highest incidence rates of occupational injury and illness per 100 full-time equivalent (FTE), private sector, California 2015-2017

Industry	Claims	Rate per 100 FTE
Interurban and Rural Bus Transportation	262	13.1
Building Material and Supplies Dealers	34,565	11.9
Couriers	20,842	11.1
Animal Slaughtering and Processing	4,343	9.0
Waste Treatment and Disposal	2,767	8.7
Scheduled Air Transportation	11,357	8.3
Railroad Rolling Stock Manufacturing	92	8.0
Cattle Ranching and Farming	6,252	7.8
Other General Merchandise Stores	9,436	7.6
Sawmills and Wood Preservation	446	7.5
Poultry and Egg Production	586	7.5
Waste Collection	4,605	7.4
Amusement Parks and Arcades	8,457	7.3
Apparel Accessories and Other Apparel Manufacturing	442	7.3
Warehousing and Storage	11,773	7.3
Traveler Accommodation	42,603	7.2
Psychiatric and Substance Abuse Hospitals	1,367	7.1
Other Ambulatory Health Care Services	4,838	6.7
Used Merchandise Stores	2,640	6.64
Foundries	1,238	6.61

Tenure

One-third of claims occurred among workers with short job tenure, under one year. Workers with 1-2 years on the job made up 13% of claims and had the highest rate of injury (5.8 per 100 FTE).

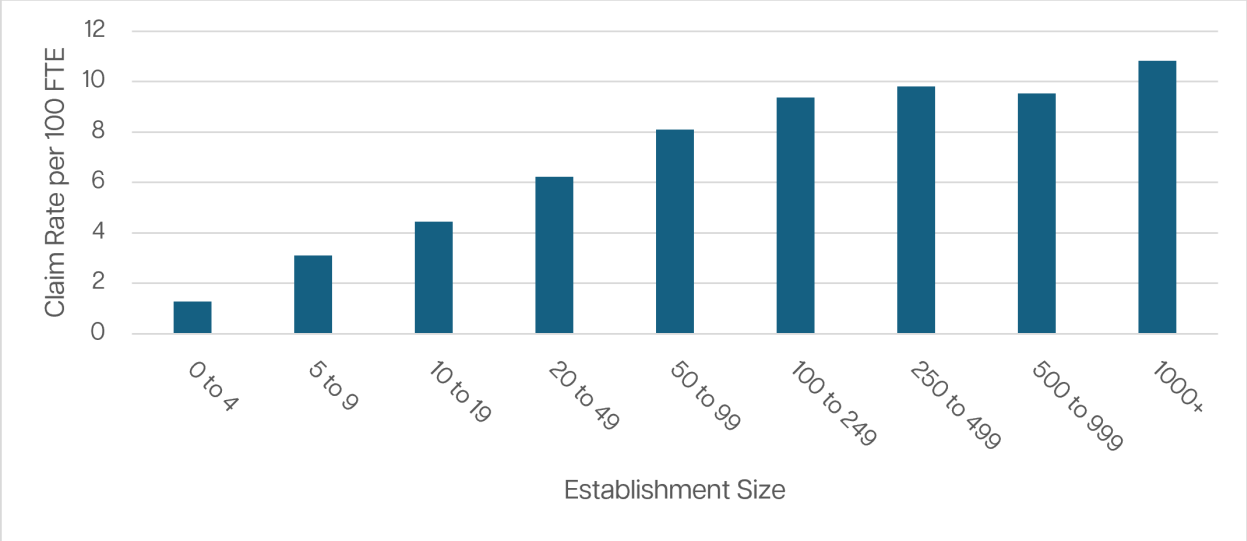
Seasonal Variation

The industries with the most seasonal variation in claim counts were fruit and vegetable preserving and manufacturing, and sugar and confectionary product manufacturing, where the most injuries occurred in the second and third quarters. For both industries, employment varied throughout the year resulting in a stable rate each quarter. For both spectator sports and warehousing and storage, injuries increased over the four quarters in the calendar year.

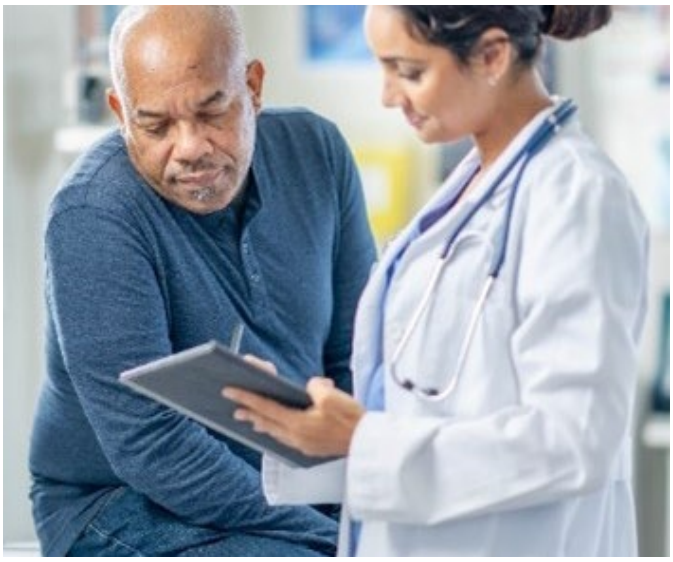
Establishment Size

The rate of injury increased with the firm size. Most employers (73%) had less than five employees and these firms had the lowest injury rate, 1.28 per 1000 workers. The highest rate was in firms with over 1,000 employees, 10.9 per 1,000 employees per year. These large firms make up less than one percent of all companies and employ 31% of workers.

Chart 1. Incidence rates of occupational injury and illness per 100 workers by employer size, private sector, California 2015-2017



Evaluating Risk



A high rate of incidents indicates a higher risk for workers, and a high claim count indicates a greater burden of injury. By looking at rate and count together, we identify industries where a public health intervention could have the greatest impact.

High Rate, High Count Industries

Building Material and Supplies Dealers

Couriers

Traveler Accommodation

High Rate, Low Count

Interurban and Rural Bus Transportation

Animal Slaughtering and Processing

Motor Vehicle Body and Trailer Manufacturing

Low Rate, High Count

Restaurants & Eating Places

General Medical & Surgical Hospitals

Services to Buildings and Dwellings

Discussion

This project combined four datasets to calculate the most accurate rates of workers' injury possible. Between 2015-2017, there were 3.3 workers' compensation claims per 100 FTE. Rates were higher among men, workers aged under 25 years old, and workers who had been in their jobs for two years or less. The rate of injury increased with the number of employees at an establishment; employers with more than 1,000 employees had the highest rates of injury, and employers with less than 5 employees had the fewest injuries. Industries in the transportation, warehousing, and utilities sectors had high rates of injury. In California, 4-digit NAICS rates can be calculated using workers' compensation claims for almost all industries each year, showing a greater level of detail than is otherwise available while maintaining employer and worker privacy.

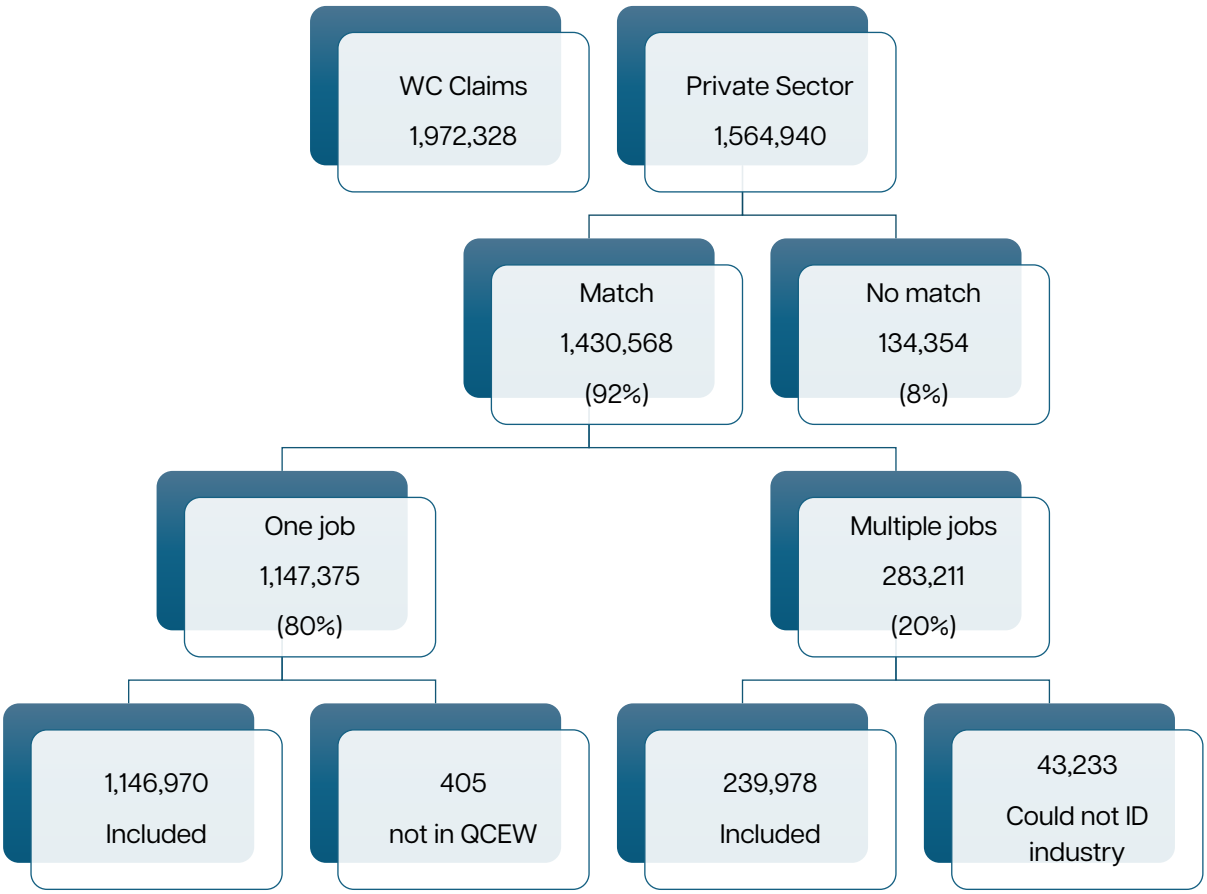
Workers' compensation claims are beneficial for occupational safety and health research because workers' compensation covers all private sector industries and occupations and have coded fields for the cause of injury, nature of injury, and injured part of body. This makes it a flexible database to use in many different applications. The same data can identify commercial bakers and bus drivers and gather the same information about the injuries both groups experience. While not used in these analyses, workers' compensation also includes both medical-only and time-loss claims, so it is possible to measure injury severity. In addition, in California all medical bills are collected in a medical billing database which gives researchers the ability to track medical services for claims, medical costs of claims, and prescription drugs prescribed to workers.

There are some limitations to this analysis. We are unable to classify low-wage workers to determine if they are at-risk workers. Work-arrangement is not well characterized, so it is difficult to identify workers in multi-employer relationships such as temporary workers or contract workers. There is no data on race or ethnicity in workers' compensation, so we were unable to calculate rates or magnitude to be able to track these important demographic variables. It is also difficult to compare results from one state to another because of substantial differences in eligibility and reporting requirements among states.

Perhaps the greatest limitation to worker's compensation rates is under reporting, which happens for many different reasons. Barriers to filing claims may affect the severity of injuries that get reported and are likely shaped by worker demographics. Lack of familiarity with workers' compensation, uncertainty about work-relatedness, and reliance on patients to identify their own work-related injuries all lead to under reporting of occupational injuries.[4] Studies examining the costs of group health medical claims comparing individuals with and without workers' compensation claims show that employer liability costs are shifted to the group health system.[5] However, this under reporting may still be less than in SOII, where research found that 20% of lost time claims were likely ineligible for SOII case reporting during the survey year.[6]

Using workers' compensation data for compliance and public health surveillance increases our understanding of the risk factors for work-related injuries and illness. This information could be provided to organizations and individuals involved in work-related injury prevention and occupational safety and health regulation to better prevent costly work-related injuries.

Appendix A – Matching Results



References

1. National Academies of Sciences, Engineering, and Medicine. 2018. A Smarter National Surveillance System for Occupational Safety and Health in the 21st Century. Washington, DC: The National Academies Press. doi.org/10.17226/24835.

2. Ruser JW. "Examining evidence on whether BLS undercounts workplace injuries and illnesses." Monthly Labor Review, August 2008. Accessed May 10, 2019.

3. Rogers E. "The Survey of Occupational Injuries and Illnesses Respondent Follow-Up Survey," Monthly Labor Review, U.S. Bureau of Labor Statistics, May 2020, https://doi.org/10.21916/mlr.2020.9. Accessed July 17th, 2026.

4. Azaroff LS, Davis LK, Naparstek R, Hashimot D, Laing JR, Wegman DH. Barriers to Use of Workers' Compensation for Patient Care at Massachusetts Community Health Centers. Health Services Research February 28. 2013. Volume 48, Issue 4. DOI: https://doi.org/10.1111/1475-6773.12045. Accessed May 7, 2019.

5. Bhattacharya A, Park RM. Excess healthcare costs associated with prior workers' compensation activity. Am J Ind Med. 2012 Nov;55(11):1018-27. doi: 10.1002/ajim.22112. Epub 2012 Sep 11. PubMed PMID: 22968954.

6. Wuellner SE, Adams DA, Bonauto DK. Workers' compensation claims not reported in the Survey of Occupational Injuries and Illnesses: Injury and claim characteristics. Am J Ind Med. 2017 Mar;60(3):264-275. doi: 10.1002/ajim.22685. Epub 2017 Feb 1. PubMed PMID: 28144976.