Exposure and Illness Among Workers Filing for COVID-19–Related Workers' Compensation—California, 2020

Ximena P. Vergara, PhD, Matt Frederick, BS, Kathryn Gibb, MPH, Jessie Wong, MPH, and Robert Harrison, MD

Objectives: The aim of the study is to characterize the incidence of SARS-CoV-2 exposure and COVID-19 infection among California workers' compensation claims in 2020 and examine risk factors for exposure. **Methods:** Using a case-control approach, we combined machine learning techniques and job exposure matrices to assess associations between exposures and illness claims and clusters of claims within specific worksites over a 3-week period. **Results:** Of the 117,125 COVID-19 claims, most were primarily among younger groups, of shorter tenure, and from health care occupations. Illness claims were among older groups with longer tenure. Jobs with very close physical proximity and high physical activity, along with transportation and warehousing industries, were associated with being part of a cluster of claims. **Conclusions:** The findings merit further study but indicate respiratory viral transmission and support efforts to systematically incorporate work-related variables into other California data sources.

Keywords: COVID-19, occupational health, workers' compensation, industry, exposure risk factors

LEARNING OUTCOMES

- Characterize COVID-19–related workers' compensation claims filed in California during 2020 by claim characteristics and risk factors for exposure to SARS-CoV-2
- Describe the distribution of clusters by industry and occupation
- Understand which risk factors for exposure to SARS-CoV-2 are associated with claims clustered in space and time

COVID-19 was recognized as an occupational disease early in the pandemic.[1] However, owing to a lack of standardized data reporting, its workplace burden is slowly coming to the fore. In May 2020, California's workforce comprised over 16 million people,[2] 30% of whom were able to work from home.[3] Workers who were unable to telework ensured the continued operation of critical functions by performing duties in hospitals, driving trucks or vehicles, producing within manufacturing facilities, and working in warehouses. The highest proportion of coronavirus 2019 (COVID-19) outbreaks in California were reported among health care and social assistance, retail trade, and manufacturing industries.[4] Specific risks of SARS-CoV-2 exposure differ between these occupations and industries; for example, nurses may be at risk because of interacting with infected patients, whereas a manufacturing worker may work close to other infected co-workers.

Workers' compensation (WC) data offer unique insights into workplace COVID-19. Italian COVID-19 claims comprised nearly 20% of all WC claims and provided a first indicator of the burden of workplace-acquired SARS-CoV-2 infection.[5] Since then, several US-based COVID-19 WC studies[6–9] have demonstrated workers from health care industries have filed the most claims. The California Workers' Compensation Information System (WCIS), the largest compensation data system in the United States, provides a near-complete collection of eligible occupational injuries and illnesses. Therefore, the California WC claims related to SARS-CoV-2 infection and COVID-19 illness are a powerful data source.

Expansive workplace protections were needed during the COVID-19 pandemic, such as access to WC. In September 2020, Senate Bill 1159 codified the presumptions set in place by California lawmakers to ensure access if workers were infected with COVID-19 at work.[10] Under the California WC temporary presumption law, a COVID-19 illness was presumed to be work-related if a worker was at a worksite during the early phases of the pandemic, had contact with COVID-19 patients in a first responder or health care capacity, or tested positive during a workplace outbreak. During the 2020 pandemic year, 676,919 claims were submitted to California Department of Industrial Relations, 30.0% of which were self-insured.[11] California WC with state-level COVID-19 presumptions may allow identification of potentially SARS-CoV-2 exposed workers, an identified gap in the literature.[12]

We aimed to characterize the incidence of SARS-CoV-2 exposure and COVID-19 infection among California workers with injury claims during the 2020 year. To do so, novel machine learning techniques were created classify SARS-CoV-2 exposures and outcomes from administrative narrative text fields. In addition, we used SARS-CoV-2 risk factor tools to examine claims and clusters of claims and then assessed the extent to which certain industries or occupations were associated with large clusters of claims (21+ claims). These analyses capture workplace COVID-19 impacts and can inform target workplace prevention strategies for future pandemics.

METHODS

We extracted COVID-19–related records from all the California WC claims for injuries occurring in 2020 using any of the four criteria: 1) keywords within the injury description (eg, "COVID" or

From the California Department of Public Health-Occupational Health Branch, Richmond, California (X.P.V., M.F., K.G., J.W., R.H.); Heluna Health, City of Industry, California (X.P.V.); and Public Health Institute, Oakland, California (M.F., K.G., J.W.).

Funding sources: This work was partially funded by the Centers for Disease Control and Prevention National Institute for Occupational Safety and Health (contract no. 75D30121P10123). The Centers for Disease Control and Prevention, Epidemiology and Laboratory Capacity for Infectious Diseases (Cooperative Agreement Number 6 NU50CK000539) also supported XPV, KG, and JW.
Conflict of interest: None declared.
Disclaimer: The findings and conclusions in this article are those of the authors and do not necessarily represent the views or opinions of the California Department of Public Health or the California Health and Human Services Agency.
Authors' contributions: XPV was responsible for conception and design of the approach, interpretation of the data, and drafting and revising for content, and is accountable for all aspects of the work. KG and JW were collaborators and were responsible for data acquisition and analysis, devising clustering approaches, interpretation of data, and drafting and revising content. MF was responsible for data acquisition and preprocessing, interpretation of data, and drafting and revising content. RH is the project investigator and was responsible for interpretation of data and drafting and revising content.
Ethical considerations and disclosures: This workers' compensation project was part of routine surveillance activities carried out by the California Department of Public Health in response to the COVID-19 pandemic and considered exempt by the California Health and Human Services Agency's Committee for the Protection of Human Subjects.
Supplemental digital contents are available for this article. Direct URL citation appears in the printed text and is provided in the HTML and PDF versions of this article on the journal's Web site (www.joem.org).
Address correspondence to: Ximena P. Vergara, PhD, California Department of Public Health Occupational Health Branch, 850 Marina Bay Parkway, Richmond, California 94804 (ximena.vergara@cdph.ca.gov).

DOI: 10.1097/JOM.0000000000003163

"CORONA*VIRUS*" or "SARS" or "CV"); 2) COVID-19 International Classification Disease code version 10 (ICD-10) within the medical billing data (eg, U07.1 or U07.2); 3) nature of injury codes related to COVID-19 (83); and 4) cause of injury codes related to pandemic (83). All records from an extract taken August 2021 were included in initial searches for SARS-CoV-2 exposure and COVID-19 regardless of claim denial status or lost time. The dataset contained the following variables: date of injury, demographics of claims including age, sex, date of hire (used for tenure), zip code where the injury occurred (used for region), employer, employer location, injury description text field, and medical billing information. We identified claimants using name, Social Security Number, and date of birth.

Industry and Occupation

The National Institute for Occupational Safety and Health Industry and Occupation Computerized Coding System, a web-based software tool, translated industry and occupation text to standardized industry and occupation codes.[13] For industry, we used 2017 North American Industry Classification System codes provided on the claim as the industry text for coding the claims. The coding bases presented in the data analyses are 2010 US Census Occupation Codes and 2012 Census Industry Codes.

Exposure and Outcome

Using both natural language processing coupled with machine learning (ML) methods, we categorized into binary fields: 1) exposure to a known SARS-CoV-2 positive person, 2) outbreak at a facility, or 3) report of a positive test. After selecting a random sample of the claims, we manually coded these three variables within 21.1% (n = 24,656) of the original dataset to use as the training data. We cleaned and preprocessed the injury text field, tokenized (broke the words into tokens) and tagged (used to identify parts of speech), and set up our models using distributed bag of words with trigrams to predict text classifiers for the remaining data. For other exposure and outcome variables described below, we also used exact or variations of exact string text search within the text fields to create binary fields to apply, in addition to ML classifiers, for the final classification of exposure and outcome variables. If multiple exposure and outcome assignments were possible for the claim, we chose to use the highest level of evidence as outlined below.

We categorized exposures using the coded occupation and the information within the medical billing and injury description fields. First, the Census Occupation Codes were linked to existing job exposure matrices to characterize risk factors for possible exposure to SARS-CoV-2. The job exposure matrices characterize worker contacts over the course of the 2020 pandemic by worker group assignments (health care workers, public facing [PF], or other as defined Council of State and Territorial Epidemiologist [CSTE]),[14] assess how close workers are to others, for example, physical proximity (very close, close, somewhat close, or not close),[15] evaluate whether occupations require working in cold environments (cold or not), assess whether work involved high physical activity (high, medium, or low), and investigate the degree to which an occupation requires loud talking on the job (very loud, loud, somewhat loud, or not loud), resulting in potential exposure to aerosols.[16] The physical proximity and physical activity variables from the job exposure matrices (JEMs) were recoded into binary categories: very close/not very close and high/not high. Second, we searched information from the medical bill (ICD-10) and the injury description using string text searches and ML. Exposure was classified from lowest level of evidence of a SARS-CoV-2 exposure (E1) to highest (E4) as follows: E1) unknown exposures or no evidence of exposure, E2) close contact with someone of unknown SARS-CoV-2 status, for example, bathing a person or taking pulse or blood pressure, E3) involved in an outbreak or where at least three persons tested positive, or E4) specific ICD-10 codes to capture a potential exposure to SARS-CoV-2 or by ML, exposure at work to someone who tested positive (Supplemental Digital Content, Table S1, http://links.lww.com/JOM/B630).

We defined two outcome measures using the following approaches: one outcome as groups of claims known as clusters, the other using level of evidence of health outcomes from the injury description and billing information. To create clusters, we divided the year into 21-day (3-week) periods. We used employer name, employer address, injury zip code, and the 3-week period to group claims that occurred at the same worksite. A cluster of claims was defined as two or more claims from a specific worksite within a 3-week period. California had numerous large COVID-19 workplace outbreaks in 2020 involving 50 or more workers covered by the media or made public on county webpages. For three of these well-publicized or public outbreaks, we evaluated whether our WCIS clustering methods would identify relevant claims and classify them into clusters of at least 21 claims. The three outbreaks used for clustering method validation were 1) a carpet manufacturing company involving 81 workers occurring within 10/1–11/30,[17] 2) a meat processing outbreak involving 392 workers within 6/1–7/30,[18] and 3) a tire recycling company involving 87 workers occurring between 5/1–6/30.[17] We determined the total number of claims and claimants that we had for each outbreak within the specified time periods, then the number of clusters we detected using our clustering approach. To assess trends in community incidence during the time of the workplace outbreaks, we obtained COVID-19 incident case data for the counties.[19]

Second, we grouped health outcomes into several levels from least amount of evidence (O1) to the best level of evidence (O8): O1) reported negative test, O2) asymptomatic or unknown symptoms, O3) reported positive test or described "viral" exposure, O4) reported unspecified illness or symptoms or death within the injury description field, O5) at least two signs or symptoms (eg, "nausea," "vomiting," "headache," "fatigue," "fever," or a corresponding ICD-10 diagnosis codes (Supplemental Digital Content, Table S1, http://links.lww.com/JOM/B630) per CSTE COVID-19 case definition,[20] O6) a single sign or symptom (eg, "cough," "shortness of breath," alterations in "smell" or "taste" corresponding ICD-10 diagnosis codes, per CSTE COVID-19 case definition[20] (Supplemental Digital Content, Table S1, http://links.lww.com/JOM/B630), O7) acute respiratory disease syndrome, pneumonias (viral or unspecified) and unspecified viral infection (J06.9, J80, B34.9), and O8) COVID-19 illness diagnosis (U07.1). Deaths could be determined using the fatality indicator or through text within the injury description. Briefly, if a claim provided evidence of being symptomatic, having illness, or dying, (O4 or higher) we considered them a case for the claims analysis.

Statistical Analyses

We included 117,125 unique COVID-19–related claims with dates of injury between January 1, 2020 December 31, 2020, for data analysis. We provide descriptive data analyses for key demographic variables and describe claims by industry and occupation within claim classification subgroup (exposure or outcome).

In addition, we used logistic regression models for the case-control approach of claims with a codable occupation (91% claims) to explore several questions: 1) whether worker group assignment, close physical proximity, talking loudly, cold environments or physical activity were independently associated with being a case, 2) whether these exposures were associated with being part of a large claim cluster (21+), and 3) whether specific industry groups were associated with being part of a large cluster of claims. Several models were created to probe claims and assess sensitivity to model selection: the crude, an adjusted model for age, sex, tenure, region, and industry group, and the final adjusted with an addition of occupation group, using "manager" as the reference group and the inclusion of other risk factors. For the cluster analyses, the final adjusted model contained the same covariates except for the other risk factors. For the industry group, we employed Firth's penalized logistic regression models due to issues of sparse data and multicollinearity.

We used Python for ML algorithm methods (Python Software Foundation. Python Language Reference, version 2.7. Available at http://www.python.org) and R statistical software for descriptive and statistical analyses (R Core Team, R Foundation for Statistical Computing, Vienna, Austria, 2013).

The California Department of Public Health Occupational Health Branch has a Memorandum of Understanding with the DIR Division of Workers' Compensation to analyze claims for purposes of public health surveillance. This study is public health practice carried out as part of routine surveillance for COVID response, not research, and is exempt from institutional review board per the California Health and Human Services Agency's Committee for the Protection of Human Subjects. This observational work adhered to the Strengthening the Reporting of Observational Studies in Epidemiology Guidelines (Supplemental Digital Content, http://links.lww.com/JOM/B631).

RESULTS

The WC COVID-19–related claims primarily comprised younger workers (age 16–39 years, 53.4%), slightly more male (50.2%), short job tenure (0–4 years, 54.3%), and injured in Southern California (60.9%) (Table 1). About half of the claims (54.2%) were from injuries occurring in the final quarter of 2020. Claims classified as illness cases were older (50–59 years, 27.5%), slightly more female (50.2%), and longer tenure (10.1%) compared to noncases. Occupation and industry could not be coded for 9.0% and 0.4.% of claims, respectively.

Industry and Occupation

Focusing on industry major groups, we found over 30% of claims represented the health care and social assistance industry including hospitals, followed by public administration (19.4%), retail trade (8.6%), manufacturing (7.1%), and transportation and warehousing (5.5%) (Fig. 1). Health care, hospitals, and public administration groups were exposed (E3 or higher), but a high proportion of claims within education services and agriculture, forestry, and fishing were also exposed. By occupation groups, 17% of claims represented health care practitioners and technical support occupations, followed by protective services (13.4%), office and administrative support (9.2%), health care support (7.8%), transportation and material moving (7.4%), and food preparation and serving related (4.8%) (Fig. 1).

When assessed by worker group assignments, 28.2% of claims with a valid worker group assignment were among health care worker occupations and 31.8% of the claims were among PF occupations. The remainder of the claims was among other occupations. Most of the WC claims were characterized as working in very close physical proximity (72.7%) (Table 2). Claims from occupational groups categorized as not in close physical proximity, the lowest exposure level, were the following groups: 1) computer and math, 2) management, 3) business and financial operations, and 4) education and training. Over 15% were considered working in high physical activity occupations (16.2%), very loud (28.9%), and cold (16.8%) environments.

TABLE 1. Demographics of California COVID-19–Related WC Claims, January 1, 2020–December 31, 2020

		Epidemiologic Claim Classification of Illness Case	
	All Claims (*N* = 117,125)	Yes‡ (*n* = 11,381)	No (*n* = 105,744)
Age, median (interquartile range)	38.6 (29.4, 49.9)	40.0 (16.3, 86.5)	38.5 (16.0, 86.9)
Age (yr)*			
16–29	31,507 (26.9)	2,690 (23.6)	28,817 (27.3)
30–39	31,029 (26.5)	2,961 (26.0)	28,068 (26.5)
40–49	25, 494 (21.8)	2,527 (22.2)	22,967 (21.7)
50–59	20,027 (17.1)	2,087 (18.3)	17,940 (17.0)
60–69	8,061 (6.9)	943 (8.3)	7,118 (6.7)
70+	843 (0.7)	104 (0.9)	739 (0.7)
Tenure (yr)*			
0–4	63,545 (54.3)	6,089 (53.5)	57,456 (54.3)
5–9	17,554 (15.0)	1,719 (15.1)	15,835 (15.0)
10–14	11,663 (10.0)	1,215 (10.7)	10,448 (9.9)
15–19	8,673 (7.4)	871 (7.7)	7,802 (7.4)
20+	10,481 (9.0)	1,150 (10.1)	9,331 (8.8)
Sex*			
Female	56,408 (48.2)	5,709 (50.2)	50,699 (47.9)
Male	58,768 (50.2)	5,599 (49.2)	53,169 (50.3)
Region*,†			
Bay Area	18,468 (15.8)	1,403 (12.3)	17,065 (16.1)
Greater Sacramento	6,868 (5.9)	494 (4.3)	6,374 (6.0)
Northern California	1,139 (1.0)	131 (1.2)	1,008 (1.0)
San Joaquin Valley	15,523 (13.3)	1,342 (11.8)	14,181 (13.4)
Southern California	71,271 (60.9)	7,551 (66.3)	63,720 (60.3)
Quarters			
January–March (Q1)	3,906 (3.3)	760 (6.7)	3,146 (3.0)
April–June (Q2)	22,572 (19.3)	3,253 (28.6)	19,319 (18.3)
July–September (Q3)	27,222 (23.2)	2,658 (23.4)	24,564 (23.2)
October–December (Q4)	63,425 (54.2)	4,710 (41.4)	58,715 (55.5)

*Excludes claims with unknowns: 164 for age, 5,209 for tenure, 1,949 for sex, and 3,856 for region.

†Region of the employer location, not necessarily where the injury occurred.

‡Yes classification was treated as a case in the logistic regression models. A case is considered to be a claim with illness: O8: COVID-19, O7: ARDS, pneumonia, virus; O6: CSTE COVID-19 Case Definition (Defn) signs/symptoms (sx); O5: Council of State and Territorial Epidemiologists 20-ID-02 Case Defn (two signs or sx); O4: Illness or symptomatic or death.

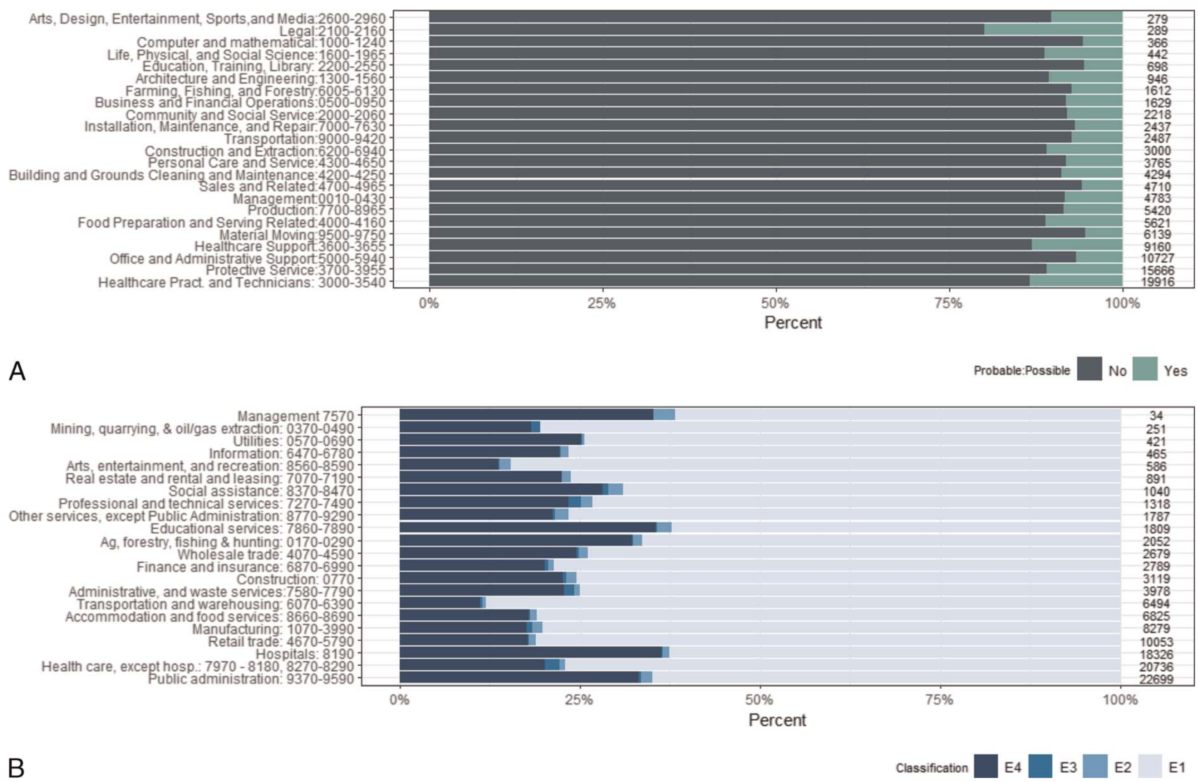

FIGURE 1. Distributions of occupation group by outcome tier (A) and industry group by exposure tier (B), California COVID-19–Related Workers' Compensation Claims January 1, 2020–December 31, 2020.

Nearly 70% of claims met the cluster definition of two or more claims (Table 3), with an average cluster size of five claims. When we examined by cluster size, over 70% of the clusters comprised 2–4 claims and 3% were clusters of 21+ claims. Most of all 18,133 claims part of large-sized clusters (21 + claims) were focused within health care practitioners and technical (13.3%), office and administrative support (11.6%), protective services (8.3%), and health care support occupation groups (8.6%). By industry, we observed similar patterns of distribution for the clusters as the claims.

Of the well-publicized outbreaks, we found each was captured within clusters of 21+ claims. We found 65 claimants among the carpet manufacturer (78% of county reported involved workers) as part of six clusters with a size distribution: three 2–4, two 5–10, and one 21+ claims. We identified 309 claimants within the meat processing plant (78.8% of county reported involved workers) as part of six clusters of size distribution: two 2–4, one 5–10, one 11–20, and three 21+ claims. We also found 88 claimants within the tire recycling (100% of reported) as part of two clusters, 2–4 and 21+ claims. Of note within the epidemic curves, the tire recycling and carpet manufacturing WC claim clusters predated the first incidence wave and the second incidence wave, respectively, in Los Angeles County, while the meat processing cluster coincided the community incidence within Merced County (Fig. 2).

In examining the distribution of the cluster sizes by occupation groups, we found these occupation groups had the highest proportion of their claims as large clusters (21+): protective service (27%), transportation and material moving (23%), and health care practitioners and technical workers (23%) (Fig. 3). Similar patterns were observed for industry groups; the industry groups with the highest proportion of their claims as large clusters (21+) were transportation and warehousing (29%), public administration (27%), and health care and social assistance (19%); notably, nearly half of the transportation and warehousing claims were part of clusters of 11+ claims (Fig. 3).

We examined the association between SARS-CoV-2 exposure risk factors and being a symptomatic, illness, or fatal claim. Most associations between very close proximity, very loud, high physical activity or cold and being a symptomatic, illness, or fatal claim were null or inverse associations in the final adjusted model. However, odds of being a health care worker with a symptomatic, illness, or fatal claim was 1.57 times that of other claims (Table 2), whereas the odds of being a PF worker for the same outcome was 1.16 higher compared to other claims.

When we examined the association between SARS-CoV-2 exposure risk factors and whether the claim was part of a large cluster (21 + claims) among a subgroup of clustered claims, we found inverse associations for health care workers, very loud, and cold work environments (Table 4). In contrast, we found worker claims in very close proximity (OR = 1.32, 95% CI = 1.22–1.42) and claims classified as high physical activity occupations (OR = 1.38, 95% CI = 1.30–1.46) to be associated with being part of a large cluster.

Finally, we evaluated industry group in relation to being was part of a large cluster (21+ claims) among a subgroup of clustered claims and found the highest associations among transportation and warehousing (OR = 11.6, 95% CI = 9.62–13.87), public administration (OR = 10.1, 95% CI = 8.49–12.1), health care and social assistance (OR = 6.03, 95% CI = 5.06–7.19), management of companies (OR = 7.37, 95% CID = 2.39–22.8) and manufacturing (OR = 4.79, 95% CID = 3.97–5.78).

DISCUSSION

In this examination of statewide demographics, industry, occupational, and SARS-CoV-2 risk factors for COVID-19 using 2020 California's WCIS claims, we found 17.3% of the 676,919 filed claims[21] to be COVID-19–related. The COVID-19–related claims were mostly among younger age groups versus all other, workers with

TABLE 2. Association of Workplace SARS-CoV-2 Exposure Risk Factors* and Epidemiologic Claim Classification, California COVID-19–Related WC Claims, January 1, 2020 – December 31, 2020

	Epidemiologic Claim Classification of Illness Case? (N = 117,125)			Crude		Adjusted‡		Adjusted§	
	Yes† (n_1 = 11,381)	No (n_0 = 105,744)							
	n_1 (%)	n_0 (%)		Odds Ratio	95% CI	Odds Ratio	95% CI	Odds Ratio	95% CI
Worker group assignment									
Others	3,385 (29.7)	39,321 (37.2)	Other	—	—	—	—	—	—
Public facing	3,191 (28.0)	30,661 (29.0)	Public facing	1.21	(1.15, 1.27)	—	—	1.16	(1.04, 1.29)
Health care	4,006 (35.2)	26,040 (24.6)	Health care	1.79	(1.70, 1.88)	—	—	1.57	(1.21, 2.01)
Close physical proximity									
Not close	1,019 (9.6)	9,946 (9.4)							
Close	432 (4.1)	4,561 (4.3)							
Somewhat close	384 (3.6)	5,156 (4.9)	Not very close	—	—	—	—	—	—
Very close	8,747 (82.7)	76,359 (72.2)	Very close	1.23	(1.16, 1.29)	1.20	(1.13, 1.28)	0.95	(0.88, 1.04)
Talking loudly									
Not loud	274 (2.4)	2,428 (2.3)							
Somewhat loud	1,960 (17.2)	21,632 (20.5)							
Loud	4,938 (43.4)	41,555 (39.3)	Not very loud	—	—	—	—	—	—
Very loud	3,410 (30.0)	30,407 (28.8)	Very loud	1.03	(0.98, 1.07)	1.06	(1.01, 1.12)	1.09	(1.02, 1.16)
Physical activity									
Low	3,593 (31.6)	36,444 (34.5)							
Medium	5,434 (47.8)	42,185 (39.9)	Not high	—	—	—	—	—	—
High	1,555 (13.7)	17,393 (16.5)	High	0.78	(0.74, 0.82)	0.81	(0.76, 0.86)	0.81	(0.75, 0.88)
Cold environments									
Not cold	9,165 (80.5)	77,719 (73.5)	Not cold	—	—	—	—	—	—
Cold	1,417 (12.5)	18,303 (17.3)	Cold	0.66	(0.62, 0.70)	0.78	(0.73, 0.84)	0.77	(0.70, 0.83)

*Workplace exposure risk factors include worker group assignment, close physical proximity, talking loudly, physical activity, and cold environments. A total of 10,521 claims could not be assigned risk factor levels including 799 cases and 9,722 controls.

†Yes classification was treated as a case in the logistic regression models. A case is considered to be a claim with illness: O8: COVID-19, O7: ARDS, pneumonia, virus; O6: Council of State and Territorial Epidemiologists 20-ID-02 Case COVID-19 Case Definition (Defn) signs/symptoms (sx); O5: CSTE Case Defn (two signs or sx); and O4: Illness or symptomatic or death;

‡Adjusted for age, sex, tenure, region, and industry.

§Adjusted for age, sex, tenure, region, industry, occupation (with managers as a reference), and all other risk factors (including worker group assignment).

shorter tenure versus others, evenly distributed between males and females, and primarily associated with worksites in Southern California compared to other regions. However, claims classified as illness cases were older (50–59 years, 27.5%), slightly more female (50.2%), and longer tenure (10.1%) compared to noncases.

Workers classified as health care workers were more likely to file a possible illness claim compared to other, non-PF occupations. Most recently, Laskaris and Markowitz analyzed 29,814 New York state claims over a 2-year period (2020–2021) and found a similar distribution for health care and social assistance (39.8%).[9] In contrast, our proportion of the health care and social assistance industry claims was smaller compared to Modji et al who found 73.2% of indemnity claims from Wisconsin[8] and Bernacki et al who found 83.8% of the COVID-19–related claims from 11 Midwestern States were from the health care sector.[6] However, comparisons between each WC system findings may reflect the differences in the COVID-19 incidence, state-based presumptions, industries within each region, and subset of claims analyzed, rather than true differences.

We hypothesized some risk factors may place a worker at risk of being part of a large cluster of claims and assessed these questions using newly developed job exposure matrix tools. Health care and protective service workers were more likely to file claims, likely driven in part, by California presumptions for WC. Specific occupations within these groups were considered in close physical contact, another well-defined risk factor for exposure to SARS-CoV-2. Early in the pandemic, workers who interacted with infected patients seeking care or with infected prison populations were recognized to be at high risk of exposure to SARS-CoV-2 and COVID-19 illness.[22] Using WC data from Ontario, Smith et al found PF occupations, proximity at work, and work location were associated with risk of COVID-19 infections.[23] Lee et al examined 2010–2021 California WC claims among nursing care facilities finding 40.4% of the claims to be COVID-related.[24] The patterns for health care occupations mirror those observed in the California workplace outbreaks by counts and by incidence for industry, highest for the

TABLE 3. Distribution by Cluster Size and Quarter of California COVID-19–Related WC Claims Subgroup Classified as Clusters January 1, 2020–December 31, 2020

	Claims* (n_{clm} = 81,560) n_{clm} (%)	Clusters (n_{clu} = 16,057) n_{clu} (%)
Cluster size		
2–4	29,137 (35.7)	11,510 (71.7)
5–10	20,323 (24.9)	3,075 (19.2)
11–20	13,967 (17.1)	988 (6.2)
21–30	6,299 (7.7)	255 (1.6)
31–40	3,472 (4.3)	100 (0.6)
41+	8,362 (10.3)	129 (0.8)
Quarters		
January–March (Q1)	2,397 (2.9)	627 (3.9)
April–June (Q2)	15,218 (18.7)	3,277 (20.4)
July–September (Q3)	17,746 (21.8)	3,535 (22.0)
October–December (Q4)	46,199 (56.6)	8,618 (53.7)

*Table only includes claims part of clusters, defined as of two or more claims meeting the criteria from a specific worksite within a 3-week period.

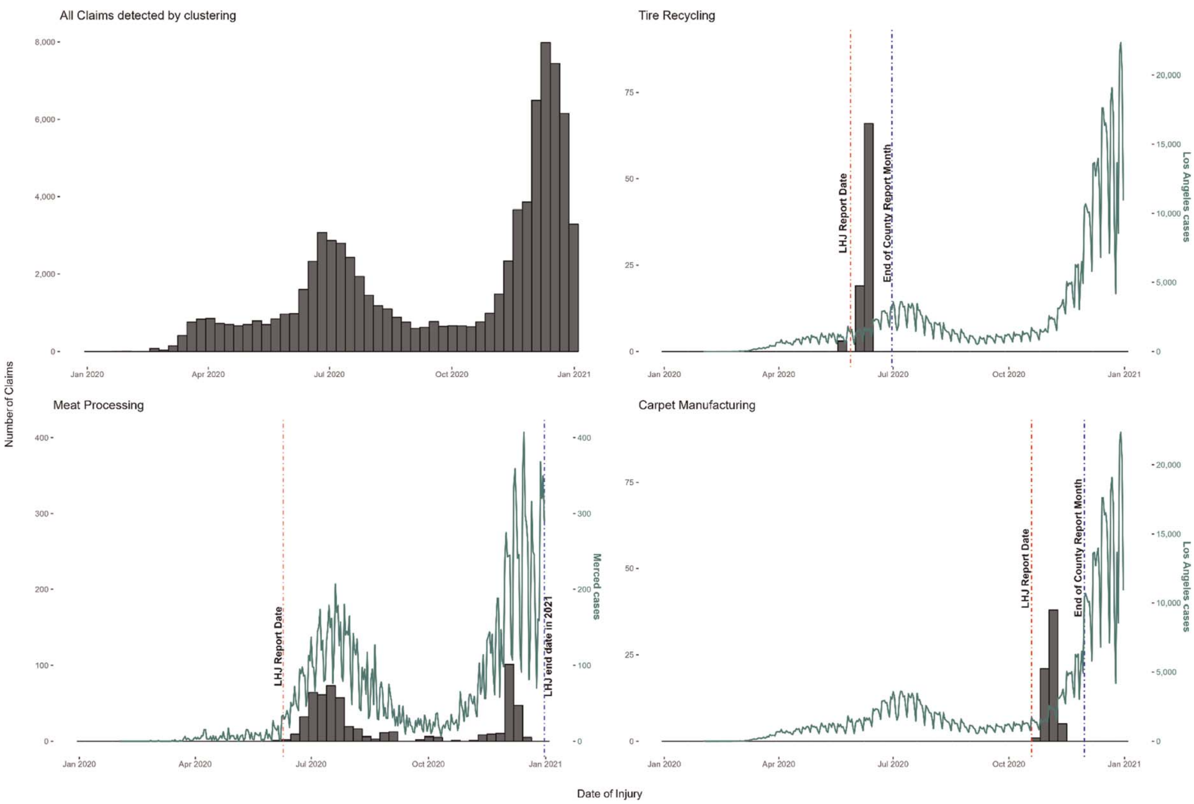

FIGURE 2. Epidemic curves for all clustered claims and outbreaks at a tire recycling, meat processing and carpet manufacturing facilities, California COVID-19–related Workers' Compensation Claims January 1, 2020–December 31, 2020.

public administration (78.1 outbreaks per 1000 establishments), which include protective service occupations.[4]

More attention is being focused on nonhealth care workers, specific workplace, and worker risk factors within the context of bioaerosols generated in the workplace.[25] Little is understood about COVID-19 risk and nonhealth care industries overall. Close contact services (eg, hairdressers, holistic or well-being locations, tailors) and restaurant/catering sectors had high attack rates among non-health care UK workplace COVID-19 outbreaks occurring in 2020.[26] To examine workplace risk factors in relation to clusters, Overton et al found physical proximity increased the risk of workplace clusters among 1.1 million UK workplaces.[27] Southern California, the region where most California

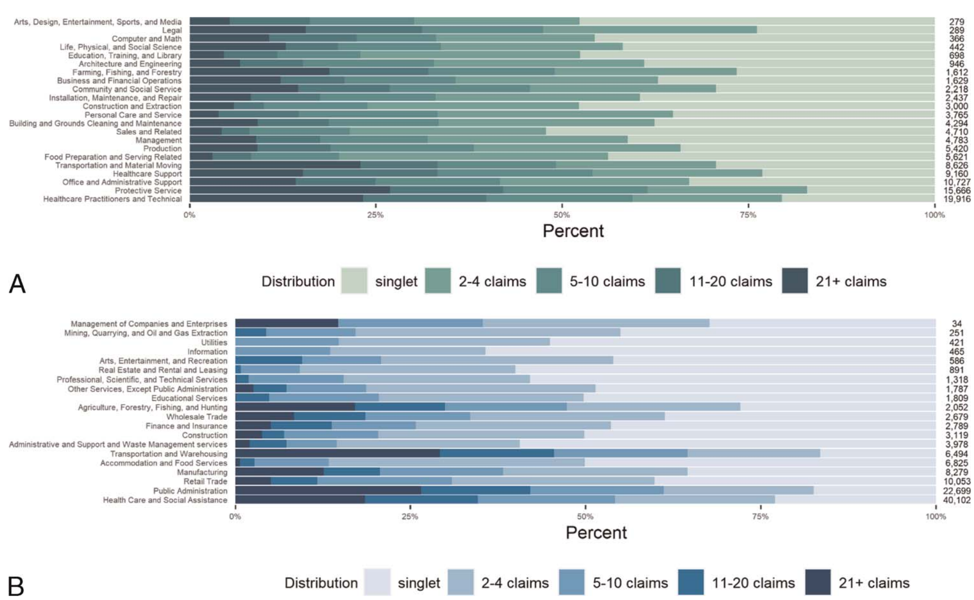

FIGURE 3. Distribution of occupation group (A) and industry group (B) by cluster size, California COVID-19–Related Workers' Compensation Claims January 1, 2020–December 31, 2020.

TABLE 4. Association of Workplace SARS-CoV-2 Exposure Risk Factors* and Large Cluster Status, California COVID-19–Related WC Claims Subgroup Classified as Clusters, January 1, 2020–December 31, 2020

	Part of a Large Cluster (n = 21+)?			Crude		Adjusted‡		Adjusted§	
	Yes† (n_1 = 18,133)	No (n_0 = 63,427)							
	n_1 (%)	n_0 (%)		Odds Ratio	95% CI	Odds Ratio	95% CI	Odds Ratio	95% CI
Worker group assignment									
Others	5,343 (29.5)	21,796 (34.4)	Other	—	—	—	—	—	—
Public facing	5,538 (30.5)	18,356 (28.9)	Public facing	1.23	(1.18, 1.28)	1.12	(1.07, 1.17)	0.74	(0.66, 0.82)
Health care	6,155 (33.9)	17,444 (27.5)	Health care	1.44	(1.38, 1.50)	1.52	(1.45, 1.59)	0.86	(0.67, 1.09)
Close physical proximity									
Not close	1,300 (7.2)	6,651 (8.9)							
Close	416 (2.3)	2,422 (3.8)							
Somewhat close	488 (2.69)	3,019 (4.8)	Not very close	—	—	—	—	—	—
Very close	14,832 (81.8)	46,504 (73.3)	Very close	1.61	(1.53, 1.69)	1.57	(1.49, 1.65)	1.32	(1.22, 1.42)
Talking loudly									
Not loud	226 (1.3)	1,390 (2.2)							
Somewhat loud	2,776 (15.3)	12,136 (19.1)							
Loud	7,804 (43.0)	25,631 (40.4)	Not very loud	—	—	—	—	—	—
Very loud	6,230 (34.4)	18,439 (29.1)	Very loud	1.22	(1.18, 1.27)	1.23	(1.18, 1.28)	0.84	(0.80, 0.89)
Physical activity									
Low	4,455 (24.6)	20,787 (32.8)							
Medium	8,858 (48.9)	26,670 (42.1)	Not high	—	—	—	—	—	—
High	3,723 (20.5)	10,139 (16.0)	High	1.31	(1.25, 1.37)	1.23	(1.18, 1.29)	1.38	(1.30, 1.46)
Cold environments									
Not cold	13,937 (76.9)	47,296 (74.6)	Not cold	—	—	—	—	—	—
Cold	3,099 (17.1)	10,300 (16.2)	Cold	1.02	(0.98, 1.07)	0.99	(0.94, 1.04)	0.86	(0.81, 0.92)

*Workplace exposure risk factors include worker group assignment, close physical proximity, talking loudly, physical activity, and cold environments. A total of 6,928 claims could not be assigned risk factor levels including 1,097 cases and 5,831 controls.

†Yes classification was treated as the case in the logistic models. A case was defined as a claim part of a large cluster of 21 or more meeting the criteria from a specific worksite within a 3-week period.

‡Adjusted for age, sex, tenure, region, and industry.

§Adjusted for age, sex, tenure, region, industry, and occupation (with managers as a reference groups).

claims arose, had numerous workplace outbreaks occurring within manufacturing, transportation, warehousing, and retail trade industries.[28] Similarly, we found transportation and warehousing and manufacturing industries associated with large claim clusters. Meat processing facilities and warehousing employ workers within high density and close contact environments due to production line requirements.[28] German researchers found the lack of distance and low temperature environments increased odds of contracting COVID-19 within meat packing plants.[29] Our findings of very close proximity associated with large clusters of claims align with these results; moreover, our application of the new exposure risk factor tools revealed high physical activity jobs may also be associated with clustering of submitted WC claims.

Our WC analysis had several limitations. First, these claims are not generalizable to all workers. Several barriers to seeking WC such as recognition on the part of a worker that an exposure or illness is potentially work-related, lack of availability of sick time, and limited access to medical care for certain industries like agriculture, are well documented.[30] In part, we were limited by the terms of our data agreement with DIR Division of Workers Compensation and received only COVID-19–related claims. Moreover, we did not have access to laboratory data. We employed an epidemiologic approach to overcome these limitations by using asymptomatic claims as a comparison group instead of using laboratory confirmation for all submitted claims. Our epidemiologic claim definition of cluster also relied on a longer time window than CSTE definitions (21 days vs 14 days). In addition, we clustered by employer, a technique that does not capture multiple employer COVID-19 events, and we used administrative data subject to data entry issues. For example, it is possible that employees in roles such as nurses or janitors who may work as contractors within a worksite or claims that vary in employer spelling could not be identified as part of a workplace cluster. We found varying percentages of claims captured for three large workplace outbreaks in California documented in media or scientific reports. These outcomes may reflect the nature of employment within each of those settings, potential barriers to reporting a workplace exposure or illness, and the definitions of clusters interpreted by the various entities. Nonetheless, each of the three documented outbreaks was considered as part of large clusters. Despite limitations, our clustering approach gave us confidence we could detect large clusters (21+ claims). Finally, we did not account for time trends or race/ethnicity (variable not part of the WC claims) in our statistical analysis.

Despite these limitations, our analysis provides insights into California workers who filed claims during the first year of the COVID-19 pandemic. The 2020 California COVID-19–related claims were mainly among younger workers who have a lifetime of productivity in their future versus other age groups. Most literature has focused on industry or occupation, whereas we explored which exposure risk factors, both well-defined (eg, number of contacts, close physical contact) and less well-defined (eg, physical activity as a proxy for respiratory rates, cold environments), may be associated with large clusters of claims or symptomatic illness. We found both very close physical proximity and high physical activity associated with part of a large claim cluster. Interventions suggested for the general population such as social distancing, may not be feasible in a work environment; hence, other layered mitigation may be warranted for aerosol transmissible respiratory viruses. Furthermore, consideration of the type of work being performed, such as high physical activity jobs, could be crucial to understand potential catalysts for transmission of SARS-CoV-2. These findings merit further study using other sources of occupational COVID-19–related data; however, they indicate the importance of work for respiratory viral transmission and support efforts to systematically incorporate work-related variables into other data sources.

ACKNOWLEDGMENTS

The authors thank the California Department of Industrial Relations, Division of Workers' Compensation for providing the data for this analysis. The authors adhered to the Strengthening the Reporting of Observational Studies in Epidemiology Guidelines.

REFERENCES

1. Carlsten C, Gulati M, Hines S, et al. COVID-19 as an occupational disease. *Am J Ind Med* 2021;64:227–237.
2. US Bureau of Labor Statistics. California Occupational Employment and Wage Statistics [Online]. Washington, DC; US Bureau of Labor Statistics, 2020. Available at: https://www.bls.gov/oes/2020/may/oes_ca.htm. Accessed July 6, 2024.
3. California Center for Jobs and the Economy. California Workers: Modernized Telecommuting Policies to Build Equity and Reduce Costs. [California Center for Jobs & the Economy web site] October 2020. Available at: https://centerforjobs.org/ca/special-reports/california-workers-modernized-telecommuting-policies-to-build-equity-and-reduce-costs. Accessed July 26, 2022.
4. Heinzerling A, Nguyen A, Frederick M, et al. Workplaces most affected by COVID-19 outbreaks in California, January 2020-August 2021. *Am J Public Health* 2022;112:1180–1190.
5. Marinaccio A, Boccuni F, Rondinone BM, Brusco A, D'Amario S, Iavicoli S. Occupational factors in the COVID-19 pandemic in Italy: compensation claims applications support establishing an occupational surveillance system. *Occup Environ Med* 2020;77:818–821.
6. Bernacki EJ, Hunt DL, Yuspeh L, et al. What industrial categories are workers at excess risk of filing a COVID-19 workers' compensation claim? A study conducted in 11 Midwestern US states. *J Occup Environ Med* 2021;63:374–380.
7. Bernacki EJ, Hunt DL, Tsourmas NF, et al. Attributes of long duration COVID-19 workers' compensation claims. *J Occup Environ Med* 2022;64:e327–e332.
8. Modji KKS, Morris CR, Creswell PD, et al. Lost time: COVID-19 indemnity claim reporting and results in the Wisconsin workers' compensation system from March 12 to December 31, 2020. *Am J Ind Med* 2022;65:1006–1021.
9. Laskaris Z, Markowitz SB. Industry-specific profile of COVID-19–related workers' compensation claims in New York state. *J Occup Environ Med* 2023; 65:e319–e329.
10. (Hill) C. California SB 1159 Workers' compensation: COVID-19: critical workers. S.B. 1159, 2020, Regular Session (California 2020). Available at: https://legiscan.com/CA/text/SB1159/id/2208989. Accessed July 6, 2024.
11. DWC. Workers' compensation information system (WCIS) Table 4: Claims by Insurer Type, Market Share, and Year of Injury, 2010–2020. Department of Industrial Relations (DIR), Division of Workers' Compensation (DWC); 2021. Available at: https://www.dir.ca.gov/dwc/wcis/WCIS_tables/Table-4/WCIS_Reports-Table-4.html. Accessed March 23, 2023.
12. Quigley DD, Qureshi N, Gahlon G, Gidengil C. Worker and employer experiences with COVID-19 and the California workers' compensation system: a review of the literature. *Am J Ind Med* 2022;65:203–213.
13. NIOSH. NIOSH Industry and Occupation Computerized Coding System (NIOCCS). 2020. U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, Division of Field Studies and Engineering, Health Informatics Branch. Available at: https://wwwn.cdc.gov/nioccs/IOCode. Accessed August 21, 2021.
14. CSTE. Documentation of the development of the category of public facing workers (other than healthcare) for applying the CSTE surveillance case definition of work-related COVID-19 (2-9-21), Occupational Health Subcommittee - Working Group Definition. Council of State and Territorial Epidemiologists; 2021.
15. Vergara XP, Gibb K. Close physical proximity on the job—an exposure matrix. *Am J Ind Med* 2022;65:537–547.
16. Vergara XP, Gibb K Ullman E, Peerless P, Gebreegziabher E, Bui D. (2023, June 9 - 13) Sars-Cov-2/COVID-19 Job Exposure Matrix (SCOVJEM): talking loudly, physical activity and cold work environments. Conference presentation Abstract presented at: Council of State and Territorial Epidemiology Meeting; June 9–13, 2023. Salt Lake City, UT.
17. Los Angeles County. COVID-19 Data, Los Angeles County Case Summaries - Non-Residential Settings [dashboard online]. Los Angeles, CA; County of Los Angeles Public Health; 2021. Available at: http://publichealth.lacounty.gov/media/Coronavirus/locations.htm#nonres-settings. Accessed May 23, 2023.
18. Merced County Health Officer. Statement Regarding COVID-19 Outbreak at Foster Farms Facility in Livingston. Merced, CA: Merced Local Health Jurisdiction (LHJ); 2020. Available at: https://www.countyofmerced.com/DocumentCenter/View/25500/Updated-Statement-Regarding-Foster-Farms-with-Health-Order-8292020?bidId. Accessed May 23, 2023.
19. CDPH. COVID-19 Cases Deaths Tests [dashboard online]. California Department of Public Health (CDPH); 2023. Available at: https://data.chhs.ca.gov/dataset/covid-19-time-series-metrics-by-county-and-state/resource/046cdd2b-31e5-4d34-9ed3-b48cdbc4be7a?view_id=be65bada-5066-47cf-8781-04c58dbbc5b4. Accessed April 3, 2023.
20. CSTE. Update to the standardized surveillance case definition and national notification for 2019 novel coronavirus disease (COVID-19), Interim-20-ID-02, Infectious Disease Committee. Council of State and Territorial Epidemiologists; 2020.
21. DWC. Workers' compensation information system (WCIS) Table 9: FROI and SROI Data Summary, by Year of Injury, 2010–2020. Department of Industrial Relations (DIR), Division of Workers' Compensation (DWC); 2021. Available at: https://www.dir.ca.gov/dwc/wcis/WCIS_Reports.html. Accessed March 23, 2023.
22. OSHA. *Worker exposure risk to COVID-19 (OSHA 3993)*. US Department of Labor; 2020.
23. Smith PM, Liao Q, Shahidi F, et al. Variation in occupational exposure risk for COVID-19 workers' compensation claims across pandemic waves in Ontario. *Occup Environ Med* 2024;81:171–177.
24. Lee SJ, Yun Y, Hwang J, Jong S. Impact of COVID-19 on occupational injuries and illnesses among nursing care facility workers: analysis of California workers' compensation data, 2019–2021. *Am J Ind Med* 2023;66:965–976.
25. Acke S, Couvreur S, Bramer WM, Schmickler M-N, De Schryver A, Haagsma JA. Global infectious disease risks associated with occupational exposure among non-healthcare workers: a systematic review of the literature. *Occup Environ Med* 2022;79:63–71.
26. Chen Y, Aldridge T, Ferraro C, Khaw FM. COVID-19 outbreak rates and infection attack rates associated with the workplace: a descriptive epidemiological study. *BMJ Open* 2022;12:e055643.
27. Overton CE, Abbey R, Baird T, et al. Quantifying the risk of workplace COVID-19 clusters in terms of commuter, workplace, and population characteristics. *arXiv preprint arXiv* 2023;230508745.
28. Contreras Z, Ngo V, Pulido M, et al. Industry sectors highly affected by worksite outbreaks of coronavirus disease, Los Angeles County, California, USA, March 19-September 30, 2020. *Emerg Infect Dis* 2021;27:1769–1775.
29. Pokora R, Kutschbach S, Weigl M, et al. Investigation of superspreading COVID-19 outbreak events in meat and poultry processing plants in Germany: a cross-sectional study. *PLoS One* 2021;16:e0242456.
30. Schofield K, Ramos A, Weichelt B, Bent B. Realities of workers' compensation coverage for agricultural workers: before, during, and into the future of the COVID-19 pandemic. *J Agromedicine* 2023;28:69–72.

