

Evaluation and Revision of Historical ESSENCE Syndromic Surveillance Definition Used to Identify Work-Related Emergency Department Visits

Public Health Reports

1–11

© 2024, Association of Schools and
Programs of Public Health

All rights reserved.



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/00333549241287929

journals.sagepub.com/home/phr



Laurel Harduar Morano, PhD, MPH^{1,2} 
Aaron Kite-Powell, MPH³; and Sara E. Luckhaupt, MD⁴ 

Abstract

Objectives: Occupational health surveillance programs have rarely incorporated syndromic surveillance. We documented a process for evaluating and revising syndromic surveillance definitions, using the example of work-related emergency department (ED) visits.

Materials and Methods: Of 140 331 481 ED visits in the National Syndromic Surveillance Program from May 15, 2016, through October 31, 2017, we extracted 391 827 records for phase 1 that met the previously unevaluated original work-related definition or included a work-related employment status ICD-10-CM code (Y99.0/Y99.1; *International Classification of Diseases, Tenth Revision, Clinical Modification*). Also for phase 1, we evaluated the performance of the original work-related definition using ICD-10-CM work-related employment status codes as a reference standard. In phase 2, we extracted 99 109 records containing candidate work-related keywords. Finally, in phase 3, we examined individual keywords and developed and evaluated a revised definition.

Results: The revised work-related definition identified 259 153 visits by keyword alone, 104 342 visits by ICD-10-CM work-related employment status codes alone, and 16 385 ED visits by both methods. Sensitivity (13.6% vs 11.3%) and positive predictive value (5.9% vs 4.8%) were higher for the revised definition than for the original work-related definition. Manual review indicated that 91.3% (95% CI, 89.6%-93.0%) of visits identified by the revised definition keywords met the study criteria for work-related visits as compared with 75.5% (95% CI, 73.0%-78.4%) of visits for the original work-related definition.

Practice Implications: The original work-related definition performed suboptimally because keywords were not sufficiently specific. The revised definition is an improvement but requires testing in real-world scenarios. The presented methods for developing a syndromic surveillance definition are useful, not only for occupational health, but for identification of other subpopulations treated in the ED.

Keywords

injuries, syndromic surveillance, occupational, work related, epidemiology

¹ Epidemic Intelligence Service, Division of Field Studies and Engineering, National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention, Cincinnati, OH, USA

² Division of Occupational and Environmental Medicine, Department of Medicine, College of Human Medicine, Michigan State University, East Lansing, MI, USA

³ Immediate Office of the Director, Office of Public Health Data, Surveillance, and Technology, Centers for Disease Control and Prevention, Atlanta, GA, USA

⁴ Division of Field Studies and Engineering, National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention, Cincinnati, OH, USA

Corresponding Author:

Laurel Harduar Morano, PhD, MPH, Michigan State University, College of Human Medicine, Department of Medicine, Division of Occupational and Environmental Medicine, 909 Wilson Rd, Rm 120, West Fee Hall, East Lansing, MI 48824, USA.

Email: harduarm@msu.edu

Occupational injuries and illnesses place a substantial burden on the US population. The Bureau of Labor Statistics estimated that 1 062 700 nonfatal occupational injuries and illnesses requiring days away from work occurred in 2021.¹ The National Safety Council estimated the total cost of work injuries in 2021 as \$167.0 billion.² Surveillance is a key component in reducing the number and cost of work-related injuries and illnesses. In the United States, surveillance of nonfatal occupational injuries and illnesses occurs through a combination of sources, including morbidity data at the jurisdiction level (eg, state, county, city, groups of counties), such as workers' compensation claims and administrative datasets (eg, emergency department [ED] data, hospital discharge data).³ The Bureau of Labor Statistics Survey of Occupational Injuries and Illnesses captures data on injuries and illnesses reported on Occupational Safety and Health Administration logs from a yearly sample of employers and provides national and jurisdictional estimates.³ Many other systems collect data on specific health conditions or exposures among workers.⁴ Each data source captures a slightly different population and information. However, the lag between injury occurrence and data release is typically 3 months to 3 years.^{1,4-6}

The Centers for Disease Control and Prevention (CDC) National Syndromic Surveillance Program (NSSP) captures near-real-time data from EDs and other sources for situational awareness and rapid response by public health professionals.^{7,8} The NSSP includes preclinical free-text information such as chief complaints (ie, patient's reason for visit) and clinical admit/discharge diagnosis codes. As of October 2023, all 50 states reported data to NSSP, and 30 of the 73 sites reporting to NSSP used it as their primary syndromic surveillance system (personal communication with NSSP staff, email communication, August 2, 2023).

Traditionally, occupational health surveillance has not included near-real-time ED visit data. A barrier to use is difficulty in identifying work-related visits. The published literature includes only 2 prior studies, both from New Jersey, identifying work-related visits to EDs via syndromic surveillance.^{5,9} Because of the variability of data across the country, the identified keywords and New Jersey Department of Health (NJDOH) evaluation results may not be representative of other jurisdictions.

The NSSP captures data from jurisdictions across the country.¹⁰ Data are stored on the CDC BioSense Platform.¹⁰ The Electronic Surveillance System for the Early Notification of Community-Based Epidemics (ESSENCE) is a system within the BioSense Platform that allows for data access, analysis, and monitoring. A work-related definition is available in ESSENCE, but a systematic evaluation has not occurred. This study builds on work from the NJDOH by detailing the development process for a revised ESSENCE work-related definition to identify ED visits for a new work-related injury or the exacerbation of a prior one. We also present a brief description of the types of injuries and illnesses that occurred among ED visits captured by the revised definition. The analysis was completed in 2019.

Methods

We used data on ED visits from the NSSP ESSENCE Chief Complaint Query Validation data source, May 15, 2016, through October 31, 2017, which contains anonymized data on ED visits. As such, it was not possible to obtain informed patient consent. ESSENCE has built-in definitions that use weighted keywords (ie, single word or phrase) found in the chief complaint for automated subsyndrome classification.¹¹ The inclusion and exclusion of keywords have positive and negative numerical weights, respectively. For a record to meet the occupational subsyndrome definition (ie, ESSENCE's work-related definition), the sum of inclusion and exclusion keywords (Table 1) must be ≥ 6 points. For instance, the following phrase has an exclusion keyword and an inclusion keyword: "blood work (-10) needed for needle-stick while working (10)" resulting in a weight of 0; therefore, the record would not be captured by the syndrome definition. Subsyndromes are the basis of syndrome groups.¹¹ For example, a record with keywords that meet the subsyndrome definition for "fall" is included in the injury syndrome group. In ESSENCE, preprocessing of the original chief complaint occurs via an automated process, called the chief complaint processor, which removes punctuation stop words (eg, as, the, and, in) and expands common abbreviations (eg, SOB=shortness of breath).¹² The remaining information is stored in the parsed chief complaint field. For analysis, we used the following fields: original and parsed chief complaints, discharge diagnosis codes, syndromes, subsyndromes, and time variables.

This study received a nonresearch determination by CDC. It was conducted as part of public health surveillance, consistent with applicable federal law and CDC policy (eg, 45 CFR part 46; 21 CFR part 56; 42 USC §241[d], 5 USC §552a, 44 USC §3501 et seq).

The analysis had 3 phases (eFigure in the Supplement): phase 1 evaluated the original ESSENCE keyword-based work-related definition and identified candidate inclusion and exclusion keywords; phase 2 examined candidate inclusion and exclusion keywords to inform a revised work-related definition; and phase 3 evaluated the revised work-related definition.

For phase 1, we extracted all records of interest that met the original work-related definition (Table 1) or the reference standard, which we defined as all records with a work-related employment status ICD-10-CM code (*International Classification of Diseases, Tenth Revision, Clinical Modification*) in the discharge diagnosis field: Y99.0, civilian work done for income or pay, or Y99.1, military activity.¹³ When ESSENCE applies the definition, the chief complaint processor uses "fuzzy word" matching to account for letter insertions, deletions, substitutions or inversions, and negations.¹² The original work-related definition uses only the parsed chief complaint field. The evaluation compared the original work-related definition (index test) with the reference standard. Measures of performance included sensitivity, specificity, and positive predictive value (PPV).

Table 1. Frequency of keywords and work-related employment status ICD-10-CM codes among ED visits (N=284 744) in the NSSP ESSENCE Chief Complaint Query Validation data that met the original ESSENCE work-related definition, United States, May 15, 2016, through October 31, 2017^a

| Occupational subsyndrome definition ^b | Keyword present, ^c no. | Only 1 keyword present, ^d no. (%) | Work-related employment status ICD-10-CM code present, no. (%) |
|--|-----------------------------------|--|--|
| Inclusion keywords (weight) | | | |
| Job (12) | 10 162 | 9227 (90.8) | 220 (2.2) |
| Work (10) | 236 498 | 232 349 (98.2) | 12 177 (5.1) |
| Worker (10) | 21 124 | 19 430 (92.0) | 691 (3.3) |
| While working (10) | 9705 | 9227 (95.1) | 455 (4.7) |
| Coworker (10) | 1520 | 929 (61.1) | 37 (2.4) |
| Workplace (10) | 480 | 426 (88.8) | 23 (4.8) |
| Employee (8) | 9788 | 8670 (88.6) | 269 (2.7) |
| Exclusion keywords (weight) | | | |
| Blood work (-10) | 127 | 126 (99.2) | 1 (0.8) |
| Work up (-10) | 66 | 66 (100.0) | 0 |
| Lab work (-10) | 46 | 46 (100.0) | 1 (2.2) |
| Social worker (-10) | 10 | 10 (100.0) | 0 |
| Dental work (-10) | 6 | 6 (100.0) | 0 |
| Social work (-10) | 3 | 3 (100.0) | 0 |
| Work of breathing (-10) | 2 | 2 (100.0) | 0 |
| Work breathing (-10) | 0 | 0 | 0 |
| Fire work (-10) | 0 | 0 | 0 |
| Drug (-4) | 1090 | 1089 (99.9) | 38 (3.5) |

Abbreviations: ED, emergency department; ICD-10-CM, *International Classification of Diseases, Tenth Revision, Clinical Modification*; NSSP ESSENCE, National Syndromic Surveillance Program Electronic Surveillance System for the Early Notification of Community-Based Epidemics.

^a The original ESSENCE keyword-based work-related definition was evaluated during phase 1 of the analysis. Data sources: Gould et al¹⁰ and National Center for Health Statistics.¹³

^b For an ED visit to meet the occupational subsyndrome criteria (ie, ESSENCE original work-related definition), the keywords found in the free-text chief complaint field must sum to ≥ 6 points. Point values for each keyword are listed in parentheses. If an inclusion keyword is negated (eg, denies work), the weight for that word is negative. If an exclusion keyword is negated (eg, no blood work), the weight for that word is positive.

^c Keyword categories are not exclusive. A chief complaint may have > 1 keyword.

^d Only 1 inclusion keyword was present in the chief complaint.

During the second part of phase 1, a single author (L.H.M.) performed a manual review of a simple random sample of records to determine consistency of the original/parsed chief complaint with a work-related visit (eFigure in the Supplement). We classified visits as work related if the injury or illness occurred at work or was caused or exacerbated by a workplace exposure. Work-exacerbated conditions included patients who arrived at the ED from work and patients who had a chronic condition (eg, diabetes) or whose illness symptoms started prior to the start of the patient's work day. We classified work-related injury/illness follow-up visits as not work related. One author (L.H.M.) manually reviewed a simple random sample of 1000 visits from each discordant pair group. Manual review occurred for all records with a work-related definition keyword that did not meet the weighted criteria. Additionally, 1 author (L.H.M.) used various methods to evaluate chief complaints: on the one hand, including computer-assisted review, such as the creation of bigram network graphs, which were used to quantify which 2 words most frequently appeared together; on the other

hand, leveraging a posteriori manual review to contextually group chief complaints (eg, all records with the keyword "recheck").¹⁴

The list of candidate keywords used in phase 2 originated from the phase 1 manual review and prior analyses of state-level syndromic surveillance data.^{5,14} During phase 2, we used the candidate keywords to extract a new dataset from the Chief Complaint Query Validation data source (eFigure in the Supplement-Phase 2A file). One author (L.H.M.) manually reviewed 1000 randomly sampled ED visits from each keyword group that had ≥ 3000 visits and did not meet the original work-related definition. To determine if keywords with < 3000 visits identified work-related visits, manual and computer-assisted review also occurred.

In phase 2, to determine how many additional records had an original work-related definition keyword but did not meet the weighted criteria and did not have a work-related employment status code, we reviewed records from the phase 1 file with the expanded keywords (eFigure in the Supplement-Phase 2B file). Keywords such as "work" and "job" from the

original work-related definition were expanded to “at work”, “on job”, and “job related”. Additionally, 1 author (L.H.M.) manually reviewed 1000 randomly sampled visits identified by the keyword “employee.”

For phase 3, we combined all records from phases 1 and 2 and then deduplicated them. Given the results from phases 1 and 2, we created a revised work-related definition using keywords found in the original and parsed chief complaint fields. We evaluated the revised work-related definition using the same methods described in phase 1, except that 1 author (L.H.M.) manually reviewed 1000 randomly sampled visits from those identified by only the revised work-related definition instead of discordant pairs.

To assess what types of injuries and illnesses the revised work-related definition may capture, we grouped all records that met the definition and had an ICD-10-CM code into ICD-10-CM chapters. Additionally, we examined the distribution of broad ESSENCE syndrome groups and 2 new ESSENCE syndrome definitions¹⁵ among records identified by the revised work-related definition or a work-related employment status ICD-10-CM code.

We analyzed all data using R version 3.4.4 (R Foundation for Statistical Computing) with the following packages: dplyr, ggraph, igraph, tidyverse, tidytext, and widyr. Extraction of keywords and phrases, including replicating the work-related definition, used regular expressions for pattern matching. During manual review, the evaluator was aware of the record source. The evaluator completed all manual reviews unblinded to method of ED visit identification.

Results

Phase 1

The NSSP contained 140 331 481 ED visits from 42 jurisdictions that occurred during the study period. Of these, the original work-related definition (index test) identified 271 100 visits; the ICD-10-CM work-related employment status codes (reference standard) identified 107 083 visits; and 13 644 visits met the index test and the reference standard criteria. The index test had a sensitivity of 11.3%, a specificity of 99.8%, and a PPV of 4.8%. However, the discharge diagnosis field was blank for 49.7% of records identified by only the index test.

Among the 284 744 visits that were identified by the index test, the most frequent inclusion keywords were “work” (83.1%; n=236 498), “worker” (6.9%; n=19 605), and “job” (3.3%; n=9321). Visits were counted in only 1 category, and most visits (98.4%; n=280 258) had only 1 inclusion keyword (Table 1). Records with the inclusion keyword “coworker” had the largest proportion of visits (38.9%; 591 of 1520) with another keyword present, usually “work” (91.2%; 539 of 591). Among visits identified by the index test, 0.5% had an exclusion keyword, the most frequent being “drug” (80.7%; 1090 of 1350).

Based on manual review of a random sample of 1000 visits, an estimated 75.5% (95% CI, 73.0%-78.4%) of the 271 100 index test positive/reference standard negative visits had descriptions consistent with work relatedness. According to computer-assisted review of chief complaints, non-work-related visits (66 420; 95% CI, 58 558-73 197) contained 1 of the following concept phrases: work or school excuse (n=23 306), recheck or follow-up (n=8317), request for a work physical (n=2467), working out (n=2154), or another concept (not explicitly reviewed).

Among the 107 083 visits identified by only the reference standard, 137 had an original work-related definition keyword but did not meet the weighted criteria. Manual review identified 49 visits in which the chief complaint was consistent with a work-related visit, 11 visits with a negation for workers' compensation (eg, denies workers' compensation), and 77 visits without enough information to make a determination. Within the latter group, 57 visits contained the exclusion keyword “drug.” Among the remaining visits identified by only the reference standard, manual review of a random sample revealed that not enough information was available to make a determination for an estimated 96.6% (95% CI, 95.5%-97.7%) of the 106 946 visits.

Phase 2

The candidate inclusion keywords identified 99 109 visits, 7.7% (n=7639) of which met the original work-related definition (Table 2). “WC”, “OJI”, “workman”, and “w/c” were the most frequently occurring new keywords; “WC” likely meant “workers' compensation” and “OJI” likely meant “on-the-job injury.” Almost all records with “workman” or “workmen” were consistent with work-related visits. Manual review of all 121 records with “not workman compensation” or “not workmen compensation” found 72 work-related and 49 non-work-related visits. The NSSP ESSENCE chief complaint processor expanded “w/c” to “wheelchair”. However, manual review of records with “w/c” revealed that the proportion of visits appropriately changed depending on placement of “w/c” in the chief complaint. If “w/c” was at the beginning or end of the chief complaint, for most visits (78.4%; 95% CI, 75.9%-80.9%), it meant “workers' compensation.” If “w/c” was in the middle of the chief complaint, for most visits, it meant “wheelchair” (50.9%; 95% CI, 48.1%-53.7%).

Examination of 3 keywords, “at work”, “on job”, and “job related”, created from an expansion of the original work-related definition inclusion keywords found that >97% of all records containing these keywords were captured by the original work-related definition (Table 2). Records that were not captured by the original work-related definition also contained an exclusion keyword, negation, or inclusion keywords within longer words (eg, “at working”, “on jobsite”, or “bite-job related” [original chief complaint=“bite.job related”]); note that records captured by the original work-related definition are based on the parsed chief complaint, which does not

Table 2. (continued)

| Inclusion keyword ^b | Keyword present, no. | Work-related definition present, no. (%) ^c | Work-related employment status ICD-10-CM code present, no. (%) | Manual review results, % (95% CI) ^d | |
|--------------------------------|----------------------|---|--|--|-----------------------|
| on job | 2741 | 2720 (99.2) | 116 (4.2) | not work-related= work-related= | 6.9 (5.4-8.4) 99.5 |
| job related | 1288 | 1270 (98.6) | 4 (0.3) | work-related= | 99.7 |

Abbreviations: ED, emergency department; ICD-10-CM, *International Classification of Diseases, Tenth Revision, Clinical Modification*; NSSP ESSENCE, National Syndromic Surveillance Program Electronic Surveillance System for the Early Notification of Community-Based Epidemics.

^a Individual inclusion and exclusion keywords were evaluated during phase 2 of the analysis.

^b Inclusion keywords based on ESSENCE chief complaint field. The parsed chief complaint field contains the chief complaint with punctuation removed. Data source: Gould et al.¹⁰

^c Met the weighted occupational subsyndrome criteria (ie, ESSENCE original work-related definition).

^d If 95% CI is not listed, then all records were either manually reviewed (ie, work related or not) or reviewed with a combination of manual and computer assistance.

^e Manual review of the ED visits that did not meet the weighted occupational subsyndrome^c definition and did not have a work-related employment status ICD-10-CM code found that 63 ED visits indicated that the individual was sent from or came from an occupational health clinic/service, 7 ED visits were work related, 4 ED visits indicated an occupational health nurse/physician, 2 visits were non-work related, and the chief complaint for 10 visits was equal to "occupational health."

^f From the phase 2A file (n = 99 109; see eFigure 1 in the Supplement), which contains records with the new inclusion keywords or prior state-based keywords.

^g Symbol (•) represents a blank space, which means the keyword is "w/c."

^h From the phase 2B file (n = 135 414; see eFigure 1 in the Supplement), which contains a subset of records identified by the original work-related definition restricted to "at work", "on job", "job related", or employee.

include punctuation). Review of records containing the keyword "employee" revealed that 75.1% (95% CI, 72.6%-77.6%) represented work-related problems, while 13.8% (95% CI, 11.8%-15.8%) described an employee physical or screening.

Phase 3

In the revised work-related definition (Figure), we determined inclusion keywords from the original definition to be too general and replaced them with more specific keywords. For example, we replaced "work" with "at work", "work injury", and similar phrases. We applied the same process to the candidate keywords examined in phase 2. Additionally, we used 60 phrases as exclusion keywords. For instance, to exclude visits in which "wc" most likely represented "wheelchair", the revised definition included exclusion keywords, such as "from wc" or "motorized wc."

The revised work-related definition alone identified 259 153 visits, the reference standard alone identified 104 342 visits, and 16 385 visits met both criteria. The sensitivity of the revised work-related definition was 13.6% with a specificity of 99.8% and a PPV of 5.9%. Manual review of a sample of visits identified by the revised work-related definition estimated that the chief complaint for 91.3% (95% CI, 89.6%-93.0%) was consistent with work-related visits.

Types of Injuries and Illnesses

Within each identification method (ie, revised work-related definition only, work-related employment status ICD-10-CM

code only, or both), the frequency of work-related visits by ICD-10-CM chapter varied; however, the rank order was similar (Table 3). The top 2 ICD-10-CM chapters associated with the largest number of work-related visits (n = 241 164) were injury, poisoning, and certain other consequences of external causes (S00-T88) and diseases of the musculoskeletal system and connective tissue (M00-M99); 59.5% of the 241 164 (n = 143 584) work-related visits had an injury/poisoning code, 10.8% (n = 25 979) had a musculoskeletal code, and 11.5% (n = 27 703) had both. According to the syndrome groups built into ESSENCE, 37.6% of the 379 880 visits (n = 142 859) identified by the revised work-related definition or an ICD-10-CM work-related employment status code were categorized into the injury syndrome group (Table 4). As with the revised work-related definition, the newest ESSENCE definitions use admit/diagnosis codes or a combination of admit/diagnosis codes and chief complaint text and do not use automated fuzzy word matching or positive/negative numerical weights. In addition to the built-in ESSENCE syndrome groups, we included 2 new syndromes based only on keywords and admit/diagnosis codes: a heat-related illness (n = 1762) and an opioid overdose syndrome (n = 197).

Practice Implications

Syndromic surveillance definitions can range from broad definitions that capture all cases but include noncases to narrow definitions that exclude all noncases but also miss cases.¹⁶ The original ESSENCE work-related definition is a broad definition, which we determined was not appropriate because it captured too many noncases. With the data

| | | | | | |
|---|--|---|-----------------------|--------------|--|
| Keywords derived from the occupational sub-syndrome definition (Phase 1) ^b | at work | injury employee | worker comp | at job | while working AND NOT (working out, working at home, working yard, working garden, working on, working with, working his yard, working her yard, working my yard, working his garden, working her garden, working my garden) |
| | work injury | employee injury | workers comp | job injury | |
| | work rela | employee rela | worker s comp | job rela | |
| | work acci | employee acci | worker mans comp | job acci | |
| | work comp | employee exposure | worker man s comp | job exposure | |
| | works comp | exposure employee | worker on duty | on job | |
| | work incident | is an employee | | | |
| | injury work | is employee | at worked | | |
| | from work | employee[end line] ^d | | | |
| | during work | employee complains of | coworker | | |
| | accident work | employee health | | | |
| sent by work | Employee AND (needle, bodily fluid, body fluid, on duty) | workplace | | | |
| Keywords derived from analysis of records identified by ICD-10-CM codes alone and state-based partners (Phase 2) ^b | workm[a e]n co ^c | wc AND NOT (MRI, CT, CTA) | occupational exposure | wrk injury | OJI |
| | workm[a e]ns co ^c | wc AND NOT (follow up, recheck) | occupational health | wrk acci | |
| | workm[a e]n s co ^c | wc AND NOT (wc bound, wc diff) | occupational injury | wrk rela | |
| | | wc AND NOT (from wc, by wc, in wc, to wc, via wc, of wc, his wc, her wc, own wc, bib wc, power wc, motorized wc, into wc, per wc, or wc, and wc) | (employee) | wrk comp | OTJ |
| | wcomp | | | @ wrk | |
| | | | | @wrk | |
| | @ work | w/c AND NOT (MRI, CT, CTA) | | | |
| | @work | w/c AND NOT (follow-up, follow up, recheck, f/u) | | | |
| | (work) | w/c AND NOT (w/c bound, w/c diff) | | | |
| | @ job | w/c AND NOT (from w/c, by w/c, in w/c, to w/c, via w/c, of w/c, his w/c, her w/c, own w/c, bib w/c, power w/c, motorized w/c, into w/c, per w/c, or w/c, and w/c) | | | |
| | @job | | | | |
| (job) | | | | | |

Figure 1. Inclusion and exclusion keywords^a that comprise a revised work-related syndrome definition displayed by related word groupings and created through evaluation of ED chief complaints within the NSSP ESSENCE Chief Complaint Query Validation data, United States, May 15, 2016, through October 31, 2017. Data source: Gould et al¹⁰ and National Center for Health Statistics.¹³ Abbreviations: CT, computed tomography scan; CTA, computed tomography angiography; ED, emergency department; ICD-10-CM, *International Classification of Diseases, Tenth Revision, Clinical Modification*; MRI, magnetic resonance imaging; OJI, on-the-job injury; OTJ, on the job.

^a Boolean operators are designated by capital letters (ie, AND; AND NOT), followed by keywords enclosed in parentheses. All other punctuation is intentional and part of the keyword unless otherwise noted.

^b The original ESSENCE¹⁰ keyword-based work-related definition was evaluated during phase 1 of the analysis. Phase 2 of the analysis evaluated candidate inclusion and exclusion keywords identified through review of chief complaints for visits identified solely by the ICD-10-CM work-related employment status codes and keywords from state-syndromic surveillance systems.^{5,14}

^c [a|e] = a or e; workman or workmen.

^d [end line] = the inclusion of an end of line operator.

Table 3. Distribution of ED visits in the NSSP ESSENCE Chief Complaint Query Validation data identified as work related per the revised work-related syndrome definition by ICD-10-CM chapter, May 15, 2016, through October 31, 2017^a

| ICD-10-CM | | No. (%) | | |
|----------------|---|--------------------------------------|--|---------------|
| Code | Chapter title | Revised work-related definition only | Work-related employment status ICD-10-CM code only | Both criteria |
| Overall | | 136 823 | 104 342 | 16 385 |
| S00-T88 | Injury, poisoning and certain other consequences of external causes | 63 317 (46.3) | 93 347 (89.5) | 14 623 (89.2) |
| M00-M99 | Diseases of the musculoskeletal system and connective tissue | 24 816 (18.1) | 25 586 (24.5) | 3280 (20.0) |
| R00-R99 | Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified | 22 379 (16.4) | 10 161 (9.7) | 1225 (7.5) |
| I00-I99 | Diseases of the circulatory system | 9567 (7.0) | 9315 (8.9) | 1239 (7.6) |
| F01-F99 | Mental, behavioral and neurodevelopmental disorders | 9109 (6.7) | 16 617 (15.9) | 1999 (12.2) |
| E00-E89 | Endocrine, nutritional and metabolic diseases | 7576 (5.5) | 6890 (6.6) | 994 (6.1) |
| J00-J99 | Diseases of the respiratory system | 4147 (3.0) | 2804 (2.7) | 346 (2.1) |
| G00-G99 | Diseases of the nervous system | 4140 (3.0) | 2975 (2.9) | 318 (1.9) |
| K00-K95 | Diseases of the digestive system | 3045 (2.2) | 1831 (1.8) | 209 (1.3) |
| N00-N99 | Diseases of the genitourinary system | 2884 (2.1) | 827 (0.8) | 84 (0.5) |
| L00-L99 | Diseases of the skin and subcutaneous tissue | 2704 (2.0) | 1579 (1.5) | 169 (1.0) |
| H00-H59 | Diseases of the eye and adnexa | 1751 (1.3) | 1791 (1.7) | 232 (1.4) |
| H60-H95 | Diseases of the ear and mastoid process | 1326 (1.0) | 1275 (1.2) | 144 (0.9) |
| D50-D89 | Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism | 1166 (0.9) | 492 (0.5) | 57 (0.3) |
| A00-B99 | Certain infectious and parasitic diseases | 995 (0.7) | 327 (0.3) | 32 (0.2) |
| O00-O9A | Pregnancy, childbirth and the puerperium | 846 (0.6) | 378 (0.4) | 77 (0.5) |
| C00-D49 | Neoplasms | 266 (0.2) | 119 (0.1) | 11 (0.1) |
| Q00-Q99 | Congenital malformations, deformations and chromosomal abnormalities | 135 (0.1) | 92 (0.1) | 10 (0.1) |
| P00-P96 | Certain conditions originating in the perinatal period | 9 (0) | 2 (0) | 0 |

Abbreviations: ED, emergency department; ICD-10-CM, *International Classification of Diseases, Tenth Revision, Clinical Modification*; NSSP ESSENCE, National Syndromic Surveillance Program Electronic Surveillance System for the Early Notification of Community-Based Epidemics.

^a A single record may have multiple ICD-10-CM codes and be counted in multiple chapters. Data sources: Gould et al¹⁰ and National Center for Health Statistics.¹³ Table includes only records in which the discharge diagnosis field was not blank.

available in the system (ie, chief complaint text or diagnosis codes), it is not possible to capture all work-related visits, negating the “all cases captured” benefit of the broad definition. For example, many records contain the nature of the injury or illness within the chief complaint (eg, laceration) and do not have information about the location where the incident occurred (eg, at work), associated exposures or activities (eg, forklift), or patient characteristics (eg, employee). Work-related visits, while constituting a small proportion of all ED visits,¹⁷ are not rare, and a large number of non-work-related visits captured with the broad definition would make responding to abnormal spatiotemporal patterns onerous (eg, the number of visits identified by the definition is greater than expected given a predefined historical period). Therefore, we aimed for a narrow definition to maximize the capture of work-related ED visits.

Manual review was important in the definition development process because discharge diagnosis codes, typically used as reference standards, are often not included in syndromic surveillance systems. In the NSSP data used for this study, 35% of visits did not have a discharge diagnosis code, while in the New Jersey study, 70% to 86% of facilities did not submit discharge diagnosis codes.⁵ Additionally, ICD-10-CM work-related employment status codes, the only available reference standard, are external cause codes, and the level of consistency in usage to document work-related visits is unknown and most likely varies by facility and jurisdiction.¹⁸ A limitation of the manual review process was that 1 author (L.H.M.) completed the review, resulting in potential misclassification. We believe that the bias is negligible due to the experience of the reviewer and the incorporation of computer-assisted review.

Table 4. Distribution of ED visits in the NSSP ESSENCE Chief Complaint Query Validation data identified as work related per the revised work-related syndrome definition by ESSENCE syndromes and selected subsyndromes, May 15, 2016, through October 31, 2017^a

| ESSENCE syndrome | No. (%) (n = 379 880) | ESSENCE subsyndrome ^b | No. (% within syndrome group) |
|--------------------------------------|--------------------------|----------------------------------|----------------------------------|
| Injury | 142 859 (37.6) | Cut or pierce | 47 833 (33.5) |
| | | Fall | 44 367 (31.1) |
| | | Struck by | 13 609 (9.5) |
| | | Dizziness | 12 690 (8.9) |
| | | Overexertion | 11 606 (8.1) |
| Neurological | 18 637 (4.9) | Seizure | 2918 (15.7) |
| | | Altered mental status | 1694 (9.1) |
| | | Drowsiness | 548 (2.9) |
| Gastrointestinal | 15 225 (4.0) | Nausea vomiting diarrhea | 10 290 (67.6) |
| | | Abdominal pain | 5013 (32.9) |
| Respiratory | 11 548 (3.0) | Difficulty breathing | 6454 (55.9) |
| | | Shortness of breath | 4531 (39.2) |
| | | Upper respiratory infection | 936 (8.1) |
| Shock/coma | 10 696 (2.8) | Loss of consciousness | 10 619 (99.3) |
| | | Shock | 80 (0.7) |
| Exposure | 6762 (1.8) | | |
| Botulism like | 4664 (1.2) | Muscle weakness | 2919 (62.6) |
| | | Blurred vision | 1379 (29.6) |
| Fever | 2138 (0.6) | Fever or chills | 2115 (98.9) |
| | | Fever plus | 1724 (80.6) |
| Rash | 1906 (0.5) | Rash | 1420 (74.5) |
| | | Chickenpox | 28 (1.5) |
| | | Measles | 2 (0.1) |
| Heat-related illness VI ^c | 1762 (0.5) | | |
| Records of interest | 746 (0.2) | Infectious hepatitis | 205 (27.5) |
| | | CO poisoning | 176 (23.6) |
| | | Rabies | 176 (23.6) |
| | | Meningitis | 84 (11.3) |
| | | Pertussis | 29 (3.9) |
| Influenza-like illness | 589 (0.2) | Fever plus | 360 (61.1) |
| | | Influenza | 281 (47.7) |
| Opioid overdose VI ^c | 197 (0.1) | | |
| Hemorrhagic illness | 73 (0.0) | | |
| Other | 112 297 (29.6) | Subsyndrome = blank | 110 135 (98.1) |

Abbreviations: ED, emergency department; NSSP ESSENCE, National Syndromic Surveillance Program Electronic Surveillance System for the Early Notification of Community-Based Epidemics.

^a Records may be classified into multiple syndrome and subsyndrome categories; if the subsyndrome is blank, then the record is categorized only into the other syndrome category. Data source: Gould et al.¹⁰

^b Syndrome groups comprise potentially multiple subsyndrome groups.

^c These syndromes are created exclusively from keywords and are not based on subsyndrome groups, and fuzzy word matching is not used (ie, the process used to account for letter insertions, deletions, substitutions/inversions, and negations).¹⁵

In addition to manual review, measures of performance calculations provided a comparison between definitions. In phase 3, the revised work-related definition improved the sensitivity and PPV. The sensitivity (original definition, 11.3%; revised definition, 13.6%) and PPV (original definition, 4.8%; revised definition, 5.9%) for both work-related definitions evaluated in this study were higher than the New Jersey work-related definition (sensitivity, 5.4%; PPV, 2.8%).⁵

During manual review of the syndrome components, we identified 4 common problems with the original definition, which the revised work-related definition addressed. First, we identified work-related visits with an original definition exclusion keyword associated with procedures related to a work injury, illness, or exposure; for example, a needle-stick at work requiring blood work (exclusion phrase) or a work injury requiring a drug (exclusion word) screen. Two, the negation written in the chief complaint is not always

appropriate (eg, not workers' compensation). There are many reasons why a worker may decline or not want to file workers' compensation, ranging from too much paperwork to fear of retaliation.¹⁹ Third, punctuation may provide important context (eg, @ "work"). Fourth, expansion of abbreviations may be incorrect, such as "w/c" to wheelchair instead of workers' compensation.

Possible next steps include testing against visit records with triage notes, which generally contain more information, or initial complaint text from other data sources, such as urgent care or emergency medical services. In addition to keywords, a work-related definition should include work-related employment status discharge diagnosis codes, despite their limitations, because information on work relatedness or the situation surrounding the exposure is often not available in the chief complaint.

As part of definition development, it is important to determine how best to incorporate syndromic surveillance methodology into multifaceted surveillance programs, such as occupational health surveillance.²⁰⁻²² Not only can syndromic surveillance be used to capture a broad spectrum of work-related injuries and illnesses, but it may also be useful for identifying and responding to conditions associated with a specific occupational hazard.²² For example, the NJDOH refers cases of occupational chemical exposure identified through syndromic surveillance to the Occupational Safety and Health Administration. The NJDOH has also used syndromic surveillance to examine tree-related injuries.^{5,9}

The described methods for developing and revising a definition are useful, not only for occupational health, but for identification of other types of visits. Our revised work-related definition builds on the New Jersey study by Borjan et al.⁵ The revised definition may benefit from additional evaluation and refinement based on the perspectives of occupational and syndromic surveillance professionals from multiple jurisdictions.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Disclaimer

The findings and conclusions in this article are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

ORCID iDs

Laurel Harduar Morano, PhD, MPH  <https://orcid.org/0000-0002-9079-8612>

Sara E. Luckhaupt, MD  <https://orcid.org/0000-0002-3000-8818>

Supplemental Material

Supplemental material for this article is available online. The authors have provided these supplemental materials to give readers additional information about their work. These materials have not been edited or formatted by *Public Health Reports's* scientific editors and, thus, may not conform to the guidelines of the *AMA Manual of Style*, 11th Edition.

References

- Employer-reported workplace injuries and illnesses—2021. Report USDL-22-2139. News release. Bureau of Labor Statistics, US Department of Labor. November 9, 2022. Accessed July 7, 2024. https://www.bls.gov/news.release/archives/osh_11092022.pdf
- National Safety Council. Work injury costs. 2022. Accessed October 23, 2023. <https://injuryfacts.nsc.org/work/costs/work-injury-costs>
- Council of State and Territorial Epidemiologists. *Occupational Health Indicators: A Guide for Tracking Occupational Health Conditions and Their Determinants*. Revised July 2021. Accessed January 3, 2024. https://cdn.ymaws.com/www.cste.org/resource/resmgr/occupationalhealth/OHI_GuidanceManual_2018_FINA.pdf
- National Academies of Sciences, Engineering, and Medicine. *A Smarter National Surveillance System for Occupational Safety and Health in the 21st Century*. National Academies Press; 2018. doi:10.17226/24835
- Borjan M, Lumia M. Evaluation of a state based syndromic surveillance system for the classification and capture of non-fatal occupational injuries and illnesses in New Jersey. *Am J Ind Med*. 2017;60(7):621-626. doi:10.1002/ajim.22734
- Harduar Morano L, Jagger MA, Barrett EC, et al. *Syndromic Surveillance Climate and Health Guidance Document: How Jurisdictions Can Use Surveillance to Quantify and Track Climate-Related Health Impacts*. Council of State and Territorial Epidemiologists; September 2017. Accessed January 3, 2024. https://cdn.ymaws.com/www.cste.org/resource/resmgr/pdfs/pdfs2/Syndromic_surveillance_clima.pdf
- Johnson G, Ishikawa C, Zwickl R, Minami M, Kass-Hout T, Streichert L. Recommendations for syndromic surveillance using inpatient and ambulatory EHR data. *Online J Public Health Inform*. 2013;5(1):e53.
- Hopkins RS, Tong CC, Burkom HS, et al. A practitioner-driven research agenda for syndromic surveillance. *Public Health Rep*. 2017;132(1):116S-126S. doi:10.1177/0033354917709784
- Marshall EG, Lu SE, Williams AO, Lefkowitz D, Borjan M. Tree-related injuries associated with response and recovery from Hurricane Sandy, New Jersey, 2011-2014. *Public Health Rep*. 2018;133(3):266-273. doi:10.1177/0033354918766871
- Gould DW, Walker D, Yoon PW. The evolution of BioSense: lessons learned and future directions. *Public Health Rep*. 2017;132(1):7S-11S. doi:10.1177/0033354917706954
- Conway M, Dowling JN, Chapman WW. Using chief complaints for syndromic surveillance: a review of chief complaint based classifiers in North America. *J Biomed Inform*. 2013;46(4):734-743. doi:10.1016/j.jbi.2013.04.003
- Coletta MA. CDC NSSP ESSENCE in-person training workshop. Presented at: Public Health Informatics Conference;

- August 21, 2016; Atlanta, GA. Accessed January 3, 2024. <https://www.cdc.gov/nssp/documents/essence-training-presentation-phi-conference.pdf>
13. National Center for Health Statistics. *International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM)*. Updated June 29, 2023. Accessed June 25, 2024. <https://www.cdc.gov/nchs/icd/icd-10-cm>
 14. Silge J, Robinson D. *Text Mining With R*. O'Reilly Media; 2017.
 15. Heat Syndrome Workgroup, Climate Change Subcommittee, Council of State and Territorial Epidemiologists. *Heat-Related Illness Syndrome Query: A Guidance Document for Implementing Heat-Related Illness Syndromic Surveillance in Public Health Practice*. September 2016. Accessed January 3, 2024. https://knowledgerepository.syndromicsurveillance.org/sites/default/files/2022-10/SyndromicSurveillance_CSTEGuidanceDocument_HeatRelatedIllness.pdf
 16. Lall R, Abdelnabi J, Ngai S, et al. Advancing the use of emergency department syndromic surveillance data, New York City, 2012-2016. *Public Health Rep*. 2017;132(1):23S-30S. doi:10.1177/0033354917711183
 17. Groenewold MR, Baron SL. The proportion of work-related emergency department visits not expected to be paid by workers' compensation: implications for occupational health surveillance, research, policy, and health equity. *Health Serv Res*. 2013;48(6 pt 1):1939-1959. doi:10.1111/1475-6773.12066
 18. Annett JL, Fingerhut LA, Gallagher SS, et al. Strategies to improve external cause-of-injury coding in state-based hospital discharge and emergency department data systems: recommendations of the CDC Workgroup for Improvement of External Cause-of-Injury Coding. *MMWR Recomm Rep*. 2008;57(RR-1):1-15.
 19. Utterback DF, Schnorr TM, eds. *Use of Workers' Compensation Data for Occupational Injury and Illness Prevention: Proceedings From September 2009 Workshop*. National Institute for Occupational Safety and Health; 2010. Publication 2010-152. Accessed March 3, 2024. <https://www.cdc.gov/niosh/docs/2010-152/pdfs/2010-152.pdf?id=10.26616/NIOSH PUB2010152>
 20. Sosin DM, Hopkins RS. Monitoring disease and risk factors: surveillance. In: Pencheon D, Guest C, Melzer D, Gray M, Guest J, eds. *Oxford Handbook of Public Health Practice*. 2nd ed. Oxford University Press; 2006:112-118.
 21. Groseclose SL, Buckeridge DL. Public health surveillance systems: recent advances in their use and evaluation. *Annu Rev Public Health*. 2017;38:57-79. doi:10.1146/annurev-publhealth-031816-044348
 22. Council of State and Territorial Epidemiologists. *Syndromic Surveillance for Occupational Health Surveillance: Guidance Document and Examples*. March 2021. Accessed August 2, 2023. https://cdn.ymaws.com/www.cste.org/resource/resmgr/occupationalhealth/publications/OccSyS_WhitePaper_03_08_21.pdf