**Enhanced Surveillance of Work-Related Motor Vehicle Crash Injuries and Deaths in Massachusetts**

**Massachusetts Department of Public Health**

**Bureau of Community Health and Prevention**

**Injury Surveillance Program**

**Occupational Health Surveillance Program**

**Authors:**

Franchesca Molina, MPH,

Jeanne Hathaway, MD, MPH, &

James Laing

June 30, 2025

**Table of Contents**

[Acknowledgments iv](#_Toc199768655)

[Overview 1](#_Toc199768656)

[**Background 1**](#_Toc199768657)

[**Definitions 2**](#_Toc199768658)

[**Methods 3**](#_Toc199768659)

[Step 1. Selection of MVC Injury Cases 4](#_Toc199768660)

[Step 2. Identification of Potential Work Indicators 4](#_Toc199768661)

[Step 3. Assessment of Potential Work Indicator Data Quality 5](#_Toc199768662)

[Step 4. Identification of Potential Work-Related MVC Injury Cases 6](#_Toc199768663)

[Step 5. Assessment of Case Overlap Between Data Sources 6](#_Toc199768664)

[Step 6. Assessment of the Sensitivity of the Surveillance System 7](#_Toc199768665)

[Step 7. Data Source Evaluation 7](#_Toc199768666)

[**Results 8**](#_Toc199768667)

[Completeness of Potential Work Indicators 8](#_Toc199768668)

[Potential Work-Related Motor Vehicle Crash Injury Case Frequencies 10](#_Toc199768669)

[Potential Work-Related Motor Vehicle Crash Injury Case Overlap 14](#_Toc199768670)

[Sensitivity Assessment of the Surveillance System 16](#_Toc199768671)

[**Data Source Evaluation 17**](#_Toc199768672)

[**Conclusions 20**](#_Toc199768673)

[Appendix A 22](#_Toc199768674)

[**Potential ICD-10-CM Work Indicators 22**](#_Toc199768675)

[Appendix B 23](#_Toc199768676)

[**Potential Work Indicators in 2019 & 2020 MA Police Crash Report Data 23**](#_Toc199768677)

[Appendix C 29](#_Toc199768678)

[**Potential Work Indicators in 2020 MA Ambulance Trip Record (MATRIS) Data 29**](#_Toc199768679)

[Appendix D 33](#_Toc199768680)

[**Potential Work Indicators in 2019 & 2020 MA Acute Care Hospital Record Data 33**](#_Toc199768681)

[Appendix E 34](#_Toc199768682)

[**Potential Work Indicators in 2019 MA Trauma Registry Data 34**](#_Toc199768683)

[Appendix F 35](#_Toc199768684)

[**CDC Surveillance System Attributes of Each Data Source 35**](#_Toc199768685)

[MA Police Crash Report Data 36](#_Toc199768686)

[MA Ambulance Trip Record Information System (MATRIS) Data 40](#_Toc199768687)

[MA Acute Care Hospital Record Data 43](#_Toc199768688)

[MA Trauma Registry Data 47](#_Toc199768689)

# **Acknowledgments**

We consulted with the following programs in the development of this project and thank them for providing valuable direction and support:

* National Institute for Occupational Safety and Health (NIOSH), Center for Motor Vehicle Safety (CMVS)
* Minnesota Department of Health, Injury and Violence Prevention Program
* Minnesota Department of Health, Occupational Health and Safety Program
* Nebraska Department of Health and Human Services, Nebraska Occupational Health Surveillance Program
* University of Kentucky, Kentucky Injury Prevention and Research Center
* University of New Hampshire, New Hampshire Occupational Health Surveillance Program

We thank the following Massachusetts Department of Public Health (MA DPH) staff for their guidance on this project and review of this document:

* MA DPH, Occupational Health Surveillance Program:
  + Emily Sparer-Fine
  + Kathleen Grattan
  + Kathleen Fitzsimmons
  + Angela Laramie

Additionally, we would like to acknowledge the following organizations for providing a space for us to present work on the development of this surveillance system:

* Council of State and Territorial Epidemiologists (CSTE)
* Association of Transportation Safety Information Professionals (ATSIP)
* NIOSH, Center for Motor Vehicle Safety (CMVS)

This project was supported in part by an appointment to the Applied Epidemiology Fellowship Program administered by the Council of State and Territorial Epidemiologists (CSTE) and funded by the Centers for Disease Control and Prevention (CDC) Cooperative Agreement Number 1NU38OT000297-03-00.

**For further information, contact** [MDPH-ISP@mass.gov](mailto:MDPH-ISP@mass.gov) or [MDPH.OHSP@mass.gov](mailto:MDPH.OHSP@mass.gov).

# **Overview**

## **Background**

Motor vehicle crashes (MVCs) are the leading cause of fatal injuries[[1]](#footnote-2) and one of the leading causes of nonfatal injuries to workers in the United States.[[2]](#footnote-3) Surveillance of work-related MVC injuries comes with many barriers that may impact our ability to assess the true risk of such injuries within different workforces. For example, the Occupational Safety and Health Administration (OSHA) is the premier agency for enforcing standards affecting worker safety but does not have enforcement authority over the open public roadways where most crashes occur. Commonly used surveillance sources, such as death data, health care data, or news coverage, do not always identify crashes as work-related, even if work vehicles are involved. Other systems and indicators traditionally relied upon for conducting surveillance of worker health, such as Workers’ Compensation insurance, may not be available or utilized by driving workforces of particular interest. It is especially difficult to identify work-related injuries among emerging types of workers like gig workers[[3]](#footnote-4) because they are generally self-employed, and this type of work may not be captured in many state or local data sources.

The purpose of this project is, therefore, to improve statewide surveillance of fatal and nonfatal work-related MVC injuries among motorists and non-motorists[[4]](#footnote-5) using multiple data sources in the Massachusetts Crash-Related Injury Surveillance System (MA CRISS). One of the goals of this project is to identify gig workers injured in work-related MVC events. MA CRISS data sources used in this project included police crash report data, Massachusetts Ambulance Trip Record Information System (MATRIS) data, acute care hospital record data (which consists of inpatient hospital discharge, outpatient observation stay, and emergency department discharge data), and Trauma Registry data. In addition to identifying and testing indicators for work-related MVC injuries in each data source, we assessed the completeness of the indicators in each data source, compared the overlap in case identification between data sources, and evaluated the strengths and limitations of each data source’s ability to identify work-related MVC injuries.

## **Definitions**

***Work-Related MVC Injury Cases:***We defined work-related MVC injury cases as motorists or non-motorists who were involved in motor vehicle crashes in Massachusetts while working and sustained fatal or nonfatal injuries. We included traffic and non-traffic crashes,[[5]](#footnote-6) as many work-related crashes may occur in work zones, parking lots, or other locations that are not public roadways and thus are defined as “non-traffic” crashes in many data sources. We excluded incidents that did not involve a crash (“non-collision”), did not involve a motor vehicle, and, if distinguishable, crashes that occurred while commuting to work.

***Motorists:***Motorists included drivers or passengers of mopeds, motorcycles, cars, trucks, buses, and other heavy vehicles, including agricultural and specialized industrial vehicles. For most commercial vehicles, including emergency vehicles[[6]](#footnote-7) and vehicles that appeared to be involved in deliveries, we included injured drivers or passengers. For vehicles involved in transporting people, such as buses, taxis, and rideshare vehicles, we only included injured drivers as cases. This definition generally aligns with definitions for “motor vehicle occupants” in the ICD-10-CM Tabular List of Diseases and Injuries[[7]](#footnote-8) used in many health care data sources and the Model Minimum Uniform Crash Criteria (MMUCC)[[8]](#footnote-9) guidelines used in police crash report data. We defined people boarding or alighting a vehicle as MV occupants when this could be distinguished. Such individuals may have been documented as pedestrians in some data sources, however.

***Non-Motorists:***Non-motorists included pedestrians, pedal cyclists, and people on other motorized or non-motorized micro-mobility devices. Drivers or passengers of vehicles who were outside of their vehicle when struck by another motor vehicle (but not boarding or alighting), were also identified as “pedestrians” for purposes of categorizing injured persons. This definition generally aligns with the definitions for non-motorists in the ICD-10-CM Tabular List of Diseases and Injuries and MMUCC.

***Gig Workers:***We defined gig workers as persons who get paid for work by the job or task rather than by the hour/year. This includes rideshare drivers, electronic app-based food delivery persons, and certain private-company package delivery persons, if we had information that these companies used gig workers for some deliveries. This categorization may have resulted in some employed delivery persons being categorized as gig workers.

## **Methods**

**Data Sources:** We used the following statewide data sources in the Massachusetts Crash-Related Injury Surveillance System (MA CRISS) to identify work-related MVC injuries that occurred in Massachusetts: police crash report data, Massachusetts Ambulance Trip Record Information System (MATRIS) data, acute care hospital record data, and Trauma Registry data. Linkage and analysis of these data sources in MA CRISS are covered by an MA DPH IRB and 24A agreement, which additionally protects the data from subpoena. The following will describe the overall content of each of these data sources:

* **Police crash report data** include reports by state, municipal, and other law enforcement agencies on all motor vehicle crashes (MVCs) occurring in MA in which any person is killed or injured, or which result in property damage of over $1,000. Police crash report data are generally based on the MMUCC guidelines, which are published by the National Highway Transportation Safety Administration (NHTSA). In MA, police crash report data are owned and managed by the MA Department of Transportation (MassDOT) Registry of Motor Vehicles (RMV). The MA DPH Injury Surveillance Program has access to these data, including personal identifiers, through a data use agreement between MA DPH and the MassDOT RMV.
* **MATRIS data** include information on all ambulance trips completed by municipal or private ambulance companies in MA. The MA DPH Injury Surveillance Program only includes records from 911 emergency calls in the MA CRISS data to limit records to acute crash events only. MATRIS data are based on the National Emergency Services Information System (NEMSIS) criteria. In MA, MATRIS data are owned and managed by the MA DPH Bureau of Health Care Safety and Quality (BHCSQ), Office of Emergency Medical Services. The MA DPH Injury Surveillance Program has access to these data, including personal identifiers, through a data use agreement with the MA DPH BHCSQ.
* **Acute care hospital record data** are administrative billing records and include inpatient hospital discharge data, outpatient observation stay data, and emergency department discharge data from all MA acute care hospitals. We combined these three hospital record data sources in this surveillance project, as they shared many of the same variables across data sources. In MA, these data are collected by the Center for Health Information and Analysis (CHIA), which publishes hospital data submission guidelines. The MA DPH Injury Surveillance Program has access to these data, with minimal personal identifiers, through a data use agreement between MA DPH and CHIA.
* **Trauma Registry data** are submitted by all acute care hospitals in MA but only include data on patients who meet specific ICD-10-CM diagnosis criteria for a traumatic injury and are admitted for hospitalization or observation stay, transferred from a hospital emergency department (ED), or died in a hospital ED. Trauma Registry data include a different set of variables than the other acute care hospital record data sources. Hospitals designated as “trauma centers” submit more information than non-trauma centers. National Trauma Data Bank (NTDB) standards published by the American College of Surgeons describe requirements for Trauma Center data submissions. Like MATRIS data, Trauma Registry data in MA are owned and managed by the MA DPH Bureau of Health Care Safety and Quality (BHCSQ), Office of Emergency Medical Services. The MA DPH Injury Surveillance Program has access to Trauma Registry data, with minimal personal identifiers, under a data use agreement with the MA DPH BHCSQ.

**Data Years:** When possible, we used 2019 data for analysis as we felt it would be more representative of the “true” burden of total work-related MVC injuries than pandemic-era (2020) data. However, at the time of this project, only 2020 MATRIS data were available in MA CRISS. To assess the overlap in case identification between MATRIS and other main data sources, we also analyzed 2020 police crash report and acute care hospital record data. Two versions of NEMSIS data (V2 and V3) were included in 2020 MATRIS data, as Emergency Medical Service (EMS) agencies were transitioning from V2 to V3 NEMSIS criteria that year. For the sake of brevity, we will refer to these as V2 and V3 MATRIS data in this document.

### ***Step 1. Selection of MVC Injury Cases***

Before identifying potential work-related MVC injury cases, we tried to limit each data source to cases involving acute MVC injuries only. Such cases were not always easy to identify. Where indicators were not clear, we erred on the side of broader criteria, so as not to exclude potential cases.

* **Police crash report data**: The MA DPH Injury Surveillance Program created individual person-level records from police crash report data for use in MA CRISS. This project included records of motorists and non-motorists who had an injury severity code indicating fatal, nonfatal, or “possible” injuries. We excluded records of crash witnesses and persons documented as having “no” injuries. We included persons with “possible” injuries as prior analysis of linked police crash report and acute care hospital record data revealed that many of these people received treatment for crash-related injuries in a hospital ED.
* **MATRIS data:** As noted above, we included only ambulance runs resulting from 911 calls in MA CRISS data. We used multiple variables to identify cases involving MVC injuries among motorists and non-motorists in V2 and V3 MATRIS data, as no single field contained reliable information to identify specific person types (i.e., drivers, passengers, motorcyclists, pedestrians, and bicyclists) involved in MVCs. These variables included dispatch complaint, cause of injury, vehicle and pedestrian risk factors, and narrative fields. We included ambulance run records even if the patient refused transport. The use of multiple fields to identify MVC injuries and the inclusion of ambulance runs where patients refused transport may have included some non-MVC injury cases.
* **Acute care hospital record and Trauma Registry data:** For all hospital data sources, we included injuries with specified ICD-10-CM external cause of morbidity codes for active treatment of traffic or non-traffic MVC injuries among motorists or non-motorists. That is, we searched all diagnosis fields for specified V-codes between V02 – V89 with a 7th character of “A” or blank, indicating active treatment. See [**Appendix A**](#_Appendix_A)for details.

### ***Step 2. Identification of Potential Work Indicators***

For each MA CRISS data source, we identified potential work indicators among the available variables and values in the data source. This was done by reviewing each data source’s data dictionary and the frequencies for each variable of interest. It was important to review the actual data rather than rely solely on the data dictionaries, as we found some variables were blank, had free text that required review to identify work-related terms, or had values that did not always correspond to those in the respective data dictionary.

We also used a supplemental list of ICD-10-CM codes from a 2021 article by Bush et al.[[9]](#footnote-10) to help us identify ICD-10-CM external cause codes that could be used as work indicators. ICD-10-CM codes were included in the MATRIS, acute care hospital record, and Trauma Registry data sources. These ICD-10-CM codes included external cause of morbidity V-codes, place of occurrence Y92 codes, activity at the time of the incident Y93 codes, and work status Y99 codes indicating civilian, military, or volunteer work at the time of the incident. Additionally, we included a Z-code for examination of a work-related incident. See [**Appendix A**](#_Appendix_A) for details.

One additional step we took to identify potential work indicators was to link previously identified MV fatality cases from the FACE[[10]](#footnote-11) program to each MA CRISS data source. These FACE cases were identified by the MA DPH Occupational Health Surveillance Program (OHSP) through media sources, death certificates, police crash reports (including state police fatal crash investigation summaries), and Workers’ Compensation data. Note that we could not link records for people who died at the scene and were removed by the medical examiner, as these cases are not included in MATRIS ambulance run or hospital data. For cases that we were able to link, we conducted a thorough review of all variables in these worker fatality cases to identify any additional work indicators, including additional text search terms. This process was successful in revealing that the vehicle owner fields in police crash report data could be searched for terms indicating commercial and government vehicle ownership. See **Appendices A** – **E** for the potential work indicators identified in each MA CRISS data source.

We considered indicators to be “potential” work indicators because some indicators may capture false positive cases. For example, the driver of a commercial vehicle injured in a crash may not have been using the vehicle for work-related purposes at the time. For another example, a driver and passenger of a commercial vehicle could both be injured in a crash while the driver was working at the time, but the passenger was a friend or family member, not a co-worker.

### ***Step 3. Assessment of Potential Work Indicator Data Quality***

We assessed the completeness and usability of potential work indicators by reviewing the frequencies of the indicators within MVC injury cases in each data source. “Complete” responses included valid responses providing specific information. Missing, invalid, or responses documented as “unknown”, “not reported”, or “not available” were considered “incomplete”. Narratives were considered “incomplete” only when they were blank. Some potential work indicators required minor revisions to assess their completeness and usability, such as reformatting character variables into numeric variables, aggregating similar variables into one variable, compiling typos of “unknown” or similar values into the “unknown” category, etc. See **Table 1** for a summary of the completeness of variable categories in each data source.

We had to deduplicate acute care hospital record and Trauma Registry data to account for the fact that some people may have had multiple hospital visits for the same incident. Police crash report and MATRIS data did not need to be deduplicated as these records already contained person-to-incident level data. In Trauma Registry data, which included the incident date, we were able to deduplicate cases to the person-to- incident level. Therefore, if a person was involved in two different MVCs that year, hospital visits for both incidents would be retained in the data. Since other acute care hospital record data did not include the incident date, we decided to deduplicate hospital records to the individual level using encrypted social security numbers and medical record numbers. This procedure may have led to the exclusion of records where a person was injured in more than one MVC that year. Given the nature of this project, we prioritized records that included a higher number of work indicators when deduplicating cases within each data source and year.

### ***Step 4. Identification of Potential Work-Related MVC Injury Cases***

Using the potential work indicators identified in **Methods *Step 2* and *Step 3***, we created algorithms to identify potential work-related MVC injury cases in each data source. Police crash report and MATRIS data were the only data sources that included narratives. Given the lack of specific variables identifying gig workers, we did extensive keyword searches for injured gig workers using these narratives. We also searched MATRIS narratives for other delivery persons and workers, given that there were relatively few additional work-related indicators in MATRIS data.

All algorithms made to identify potential work-related MVC injury cases were developed and implemented in SAS Studio Version 9.4 (SAS 3.81, Enterprise Edition). The process of creating the algorithms was iterative, in that we sometimes needed to refine the algorithms after reviewing the cases captured. We particularly double-checked cases that were identified by text searches or had only one work indicator to minimize capturing false positive cases. Cases solely identified through text searches of narratives were manually reviewed to refine search criteria and exclude duplicate narratives and false positives, such as non-workers injured in the same crash event as a worker. Cases were classified as true positives if their narrative clearly stated that the injured person was working at the time of the crash. Those who appeared to be commuting to or from work, unless related to a ridesharing or delivery gig, were considered false positives and excluded from total counts. Occupants of semi-trucks, box trucks, or trailer trucks (but not rental trucks) identified through narrative text searches were classified as true positives as we believed it to be less likely that these trucks would be driven for non-work-related purposes.

### ***Step 5. Assessment of Case Overlap Between Data Sources***

After creating and finalizing the algorithms to identify potential work-related MVC injury cases in each data source, we assessed overlap among the identified cases by linking and deduplicating cases from MA CRISS data of the same year. The 2019 potential work-related MVC injury case datasets linked were police crash report, acute care hospital record, and Trauma Registry data. The 2020 potential case datasets linked were police crash report, acute care hospital record, and MATRIS data. We deterministically linked the datasets using date of birth and crash/admission date, which we found to be sufficient to accurately match potential work-related MVC injury cases that had already been identified in their respective data sources. Duplicate cases were identified by assessing unique identifiers, such as state-unique crash ID and person number[[11]](#footnote-12) from crash report data, billing number from hospital record data, record ID from Trauma Registry data, and incident ID from MATRIS data. Cases were deduplicated using these unique identifiers per data source and with prioritization of higher injury treatment levels (more severe injuries) and a greater number of work indicators.

### ***Step 6. Assessment of the Sensitivity of the Surveillance System***

To assess the sensitivity of the surveillance system in identifying fatal work-related motor vehicle crash (MVC) injury cases, we utilized the FACE motor vehicle fatalities identified by OHSP as the gold standard for comparing cases identified in the four MA CRISS data sources. We linked fatal work-related MVC injury cases identified in each 2019 data source (police crash reports, acute care hospital records, and Trauma Registry data) with 2019 FACE cases to calculate the percentage of FACE cases identified by each data source in 2019. We also linked fatal work-related MVC injury cases identified in each 2020 data source (crash reports, hospital records, and MATRIS data) to 2020 FACE cases to calculate the percentage of FACE cases identified by each data source in 2020.

OHSP had also previously created a dataset of nonfatal work-related MVC injury cases that included 2019 and 2020 cases. OHSP identified these nonfatal injury cases through media reports and weekly police crash report summaries of fatal and near-fatal crashes. We attempted to use these data to conduct a second assessment of the sensitivity of the surveillance system in identifying nonfatalwork-related MVC injury cases. Unfortunately, this assessment was inconclusive, as there were insufficient personal identifiers common to the OHSP-identified cases and MA CRISS data sources to allow for data linkage.

### ***Step 7. Data Source Evaluation***

Throughout the development of algorithms to identify potential work-related MVC injury cases, we compiled a list of each data source’s strengths and limitations in its ability to identify such cases. We focused on the sensitivity of the data source to capture potential work-related injury cases, especially fatalities and serious injury cases, as well as cases involving gig workers. We also considered the number and quality of work-related indicators in each data source and the likelihood of capturing false-positive cases. Note that some of these attributes may vary in other states’ data.

We did not consider how well the data source captured other contextual factors and health outcomes, such as demographic data, crash circumstances, and injury diagnoses and dispositions. We also did not consider the ease of accessing these data sources or personal identifying information (PII) in this review. Under the MA CRISS IRB approval and our data use agreements with data owners, the MA DPH Injury Surveillance Program has full access to the data sources used in the work-related MVC injury surveillance system, including the PII necessary to link data sources, identify case overlap, and identify cases in narratives. Other states and organizations may have more limited access to this information.

## **Results**

### ***Completeness of Potential Work Indicators***

**Table 1** on the following page summarizes the completeness of each variable or category of variables used as work indicators by data source. Most of the variables or variable categories were at least 75% complete. Narratives in police crash report and MATRIS data respectively had a completion rate of 90% and 99.96%, however, some narratives were more detailed than others within each data source. The work-related injury/illness variable in MATRIS was less than 50% complete. Version 2 MATRIS data did not contain this variable, and it was only 27.5% complete in Version 3 MATRIS data. Several variable categories were less than 50% complete in acute care hospital records: ICD-10-CM place of occurrence codes, ICD-10-CM activity codes, and ICD-10-CM (work) status code. The ICD-10-CM activity codes, and the ICD-10-CM (work) status code and ICD-10-CM visit type codes were also less than 50% complete in Trauma Registry data. This may be because ICD-10-CM codes are not reimbursable and, therefore, requirements to complete these codes may vary across hospitals. For more detailed information on variable names, descriptions, and values used as potential work indicators, see [**Appendix A**](#_Appendix_A)for the ICD-10-CM code work indicators and **Appendices B** – **E** for other work indicators in each MA CRISS data source.

**Table 1.** **Completeness of Work Indicator Variables or Category by MA CRISS Data Source**

(Green = 75%+ completion, Yellow = 50-74% completion,[[12]](#footnote-13) and Red = <50% completion. Dashes indicate no indicators of this type were available in the data source.)

| **Work Indicator Variable or Category** | **2019 MA Police Crash Reports** | **2020 MA Ambulance Trip Records (MATRIS)** | **2019 MA Acute Care Hospital Records** | **2019 MA Trauma Registry** |
| --- | --- | --- | --- | --- |
| Vehicle Type & Weight Variables | ✓ | - | - | - |
| Vehicle Owner Name Variables & Vehicle Registration Type | ✓ | - | - | - |
| Road Contributing Circumstance Code & Work Zone Relation Code | ✓ | - | - | - |
| Non-Motorist Activity Code | ✓ | - | - | - |
| Payer Source & Type Variables | - | - | ✓ | - |
| Work-Related Injury/Illness Variable | - | ✓ | - | ✓ |
| Narrative | ✓ | ✓ | - | - |
| ICD-10-CM External Cause of Morbidity Codes (V-codes) | - | ✓ | ✓ | ✓ |
| ICD-10-CM Place of Occurrence Codes (Y92) | - | - | ✓ | ✓ |
| ICD-10-CM Activity Codes (Y93) | - | - | ✓ | ✓ |
| ICD-10-CM External Cause (Work) Status Codes (Y99) | - | - | ✓ | ✓ |
| ICD-10-CM Visit Type Codes (Z-codes) | - | - | ✓ | ✓ |

### ***Potential Work-Related Motor Vehicle Crash Injury Case Frequencies***

**Table 2** shows the total number of motorists and non-motorists injured from any motor vehicle crash (MVC) in Massachusetts, the number and percentage of motorists and non-motorists injured from a potential work-related MVC (“potential work-related MVC injury cases”), and the number of injured gig workers among the potential work-related MVC injury cases by data source and year. In 2019, we used the indicators described above to identify potential work-related MVC injury cases in MA police crash reports, MA acute care hospital records, and MA Trauma Registry data. We identified a higher number and percentage of potential work-related MVC injury cases in police crash reports and acute care hospital records than in Trauma Registry data. Of the 43,106 injured motorists and non-motorists in 2019 MA police crash reports, 7.8% (n = 3,350) were identified as potential work-related MVC injury cases. Of those 3,350 potential cases, 1.0% (n = 32) were gig workers. In 2019 MA acute care hospital records, 3.7% (n = 2,712) of the 73,016 injured motorists and non-motorists were identified as potential work-related MVC injury cases. Of the 5,777 injured motorists and non-motorists in 2019 MA Trauma Registry data, 2.5% (n = 143) were identified as potential work-related MVC injury cases.

In 2020, we identified potential work-related MVC injury cases in MA police crash report, MA acute care hospital record, and MATRIS data[[13]](#footnote-14). Similar to 2019, we identified a higher number and percentage of potential work-related MVC injury cases in 2020 police crash reports and acute care hospital records than in MATRIS data. Of the 33,139 injured motorists and non-motorists in 2020 police crash reports, 5.8% (n = 1,926) were identified as potential work-related MVC injury cases. Of those 1,926 potential cases, 1.9% (n = 36) were gig workers. This percentage is nearly double the percentage identified in 2019 police crash report data (1.0%), despite there being a lower number and percentage of potential work-related MVC injury cases. The increase in injured gig workers is consistent with other studies that identified a rise in gig work throughout the pandemic era,[[14]](#footnote-15),[[15]](#footnote-16) and, therefore, a potential rise in injury cases among gig workers as well. In 2020 acute care hospital record data, 3.4% (n = 1,686) of the 49,531 injured motorists and non-motorists were identified as potential work-related MVC injury cases. The lower number of injured motorists and non-motorists and potential work-related MVC injury cases in 2020 than 2019 may be partly because there were fewer vehicles and workers on the road during the pandemic. In 2020 MATRIS data, 1.9% (n = 546) of the 28,403 injured motorists and non-motorists were identified as potential work-related MVC injury cases. Of these 546 potential cases, 2.0% (n = 11) were gig workers.

**Table 2.** **Fatal and Nonfatal Potential Work-Related Motor Vehicle Crash (MVC) Injury Cases Identified, by 2019-2020 MA CRISS Data Source**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Data Source** | **Data Year** | **Total Injured Motorists and Non-Motorists**  **#** | **Total Injured in Potential**  **Work-Related MVCs**  **# (%)** | **Injured Gig Workers[[16]](#footnote-17)**  **# (% of Potential Work-Related MVCs)** |
| MA Police Crash Reports | 2019 | 43,106 | 3,350 (7.8%) | 32 (1.0%) |
| 2020 | 33,139 | 1,926 (5.8%) | 36 (1.9%) |
| MA Acute Care Hospital Records | 2019 | 73,016 | 2,712 (3.7%) | N/A |
| 2020 | 49,531 | 1,686 (3.4%) | N/A |
| MA Trauma Registry | 2019 | 5,777 | 143 (2.5%) | N/A |
| MA Ambulance Trip Record Information System (MATRIS) | 2020 | 28,403 | 546 (1.9%) | 11 (2.0%) |

Please see the [**Data Source Evaluation**](#_Evaluation_of_the) section below for a discussion on the strengths and limitations of these data sources that may have contributed to differences in the percentages of potential work-related MVC injury cases identified.

***Narrative Review to Identify True Positive Cases***

As described in [**Methods**](#_Methods) ***Step 4*** above, cases identified as potential work-related MVC injuries solely through narrative text searches were manually reviewed to exclude duplicate cases and false positives. True positives included narratives that directly stated that the injured person was working at the time of the crash. Injured motorists and non-motorists who appeared to be commuting to or from work were considered false positives unless they were gig workers. Semi-truck, box truck, or trailer truck drivers were generally classified as true positives as we believed it to be less likely that these trucks would be driven for non-work-related purposes. The following results indicate the number of true positive cases found after manually reviewing cases. See **Table 3**.

In police crash report data, we only used narrative text searches to identify gig workers, as we found that text searches for non-gig workers yielded mostly false positive cases and there were many other potential work indicators in this particular data source. In 2019 police crash reports, text searches initially identified 82 cases involving injured gig workers, but only 39.0% (n = 32) of these were true positives. In 2020 police crash reports, text searches initially identified 62 cases involving gig workers, and 58.1% (n = 36) of these cases were true positives.

In 2020 MATRIS data, we used narrative text searches to identify MVC injuries among gig and non-gig workers due to the limited number of potential work indicators in this data source. Of the 135 cases initially identified as potential work-related MVC injuries in MATRIS narratives, a manual review found that 85.2% (n = 115) were true positives. Regarding gig workers, 78.6% (n = 11) of the 14 cases initially identified in narrative text searches were true positives. Regarding non-gig workers, 92.9% (n = 13) of the 14 cases initially identified in narratives as non-gig parcel delivery (referred to as mailing/shipping) workers were true positives and 85.1% (n = 91) of the 107 cases identified in narratives as other non-gig workers were true positives.

Based on cases identified as true positives, most of the gig workers identified in the 2019 crash report data were Uber drivers (n = 13), whereas most of the gig workers identified in the 2020 crash report data were Amazon delivery workers (n = 20). Most of the gig workers identified in the 2020 MATRIS data were food/grocery delivery people (n = 5). Continuing with 2020 MATRIS data, most of the non-gig mailing/shipping workers were United States Postal Service (USPS) delivery workers (n = 9), and most of the other non-gig workers were occupants of a work vehicle, work van, or work truck (n = 42). Further details on the numbers of true positive narratives by keyword and data source are shown in **Table 3** below.

We suggest that text searches of narratives only be used to identify cases when gig and other workers cannot be identified through other variables, given the large number of false positives that may be captured and the need to manually review narratives identified. Specific keywords and phrases used in narrative text searches can be found in [**Appendices B**](#_Appendix_B)and [**C**](#_Appendix_C).

**Table 3.** **Number of True Positive Work-Related MVC Cases Identified Using Narrative Text Searches after Manual Review, by Keywords and Data Source**

| **Keywords Used in Narrative Text Searching** | **Number of True Positive Narratives Identified** | | |
| --- | --- | --- | --- |
| **2019 Police Crash Report Data** | **2020 Police Crash Report Data** | **2020 MATRIS Data** |
| **Gig Work Keywords[[17]](#footnote-18)** | **32** | **36** | **11** |
| Rideshare driver | 2 | 2 | 0 |
| Uber driver | 13 | 4 | <5 |
| Lyft driver | 4 | 2 | <5 |
| Food / Grocery delivery | 5 | 2 | 5 |
| DoorDash delivery | 1 | 3 | 0 |
| Package delivery[[18]](#footnote-19) | 3 | 3 | <5 |
| Amazon delivery18 | 4 | 20 | <5 |
| **Non-Gig Mailing/Shipping Work Keywords** | **-** | **-** | **13** |
| FedEx delivery | - | - | <5 |
| UPS delivery | - | - | <5 |
| USPS / Mail / Postal delivery | - | - | 9 |
| **Additional Non-Gig Work Keywords[[19]](#footnote-20)** | **-** | **-** | **91** |
| Taxi driver | - | - | 0 |
| Tractor Trailer/ Box Truck / Semi driver[[20]](#footnote-21) | - | - | 41 |
| Work Vehicle / Work Van / Work Truck occupant | - | - | 42 |
| Working non-motorist / Working pedestrian / Working cyclist | - | - | 8 |
| **Total** | **32** | **36** | **115** |

### ***Potential Work-Related Motor Vehicle Crash Injury Case Overlap***

As described in [**Methods**](#_Methods) ***Step 5*** above, we assessed the overlap in work-related MVC injury cases by linking and deduplicating identified cases within each year. The 2019 data sources linked and deduplicated were police crash report, acute care hospital record, and Trauma Registry data. The 2020 data sources linked and deduplicated were police crash report, acute care hospital record, and MATRIS data. Venn diagrams showing case overlap between the 2019 and 2020 data sources are shown in **Figure 1** and **Figure 2** respectively.

**2019 Data Sources:** **Figure 1** shows that in 2019, police crash report, acute care hospital record, and Trauma Registry data identified 5,444 unique, potential work-related MVC injury cases. Of the 5,444 cases, 86.7% (n = 4,722) were identified in only one of the three data sources; 12.5% (n = 683) were identified in two of the three data sources; and 0.7% (n = 39) were identified in all three data sources. Police crash report and acute care hospital record data identified the most work-related MVC injury cases, while Trauma Registry data identified relatively few additional cases.

**Figure 1. Overlap in Potential Work-Related MVC Injury Cases in 2019 MA Police Crash Report, Acute Care Hospital Record, and Trauma Registry Data (N = 5,444)**

Diagram, venn diagram

AI-generated content may be incorrect.

In 2019 data, gig workers could only be identified as such in the police crash report narratives and comprise 0.6% (n = 32) of all unique 2019 cases (N = 5,444). Interestingly, 6 (19%) of the 32 gig worker cases identified in police crash report data were also identified as work-related MVC injury cases in acute care hospital record data. No gig worker cases identified in police crash report data were identified as work-related MVC injury cases in Trauma Registry data.

**2020 Data Sources:** **Figure 2** shows that in 2020, police crash report, acute care hospital record, and MATRIS data identified 3,539 unique, potential work-related MVC injury cases. Of these 3,539 cases, 85.6% (3,028) were identified in only one of the three data sources; 11.4% (403) were identified in two of the three data sources; and 3.1% (108) were identified in all three data sources. Similar to 2019 data, police crash report and acute care hospital record data identified the most unique cases. MATRIS data, however, identified a significant number of additional unique cases.

**Figure 2. Overlap in Potential Work-Related MVC Injury Cases in 2020 MA Police Crash Report, Acute Care Hospital Record, and Ambulance Trip (MATRIS) Data (N = 3,539)**

Diagram, venn diagram

AI-generated content may be incorrect.

In 2020 data, gig workers could be identified as such in both police crash report and MATRIS narratives. A total of 46 unique cases involving injured gig workers were identified between these two data sources, comprising 1.3% of all unique 2020 cases (N = 3,539). Of the 36 gig workers identified in police crash report narratives and the 11 gig workers identified in MATRIS narratives, there was only 1 case (2.8% and 9.1% respectively) who was identified as an injured gig worker in both data sources. Of the 36 gig worker cases identified in police crash reports, 6 (16.7%) were identified as work-related MVC injury cases in acute care hospital record data, and 2 (5.6%) were identified as non-gig work-related MVC injury cases in MATRIS data. Of the 11 gig worker cases identified in MATRIS narratives, 4 (36.4%) were identified in acute care hospital records as work-related MVC injury cases.

Despite differences in the total number of unique work-related MVC injury cases in 2019 (5,444) and 2020 (3,539), it is noteworthy that about 85% of cases were identified in only one of the three data sources analyzed in both years. In addition, the majority of cases in both years were identified in police crash report and acute care hospital record data. MATRIS data captured many additional cases in 2020, while Trauma Registry data captured relatively few additional cases in 2019.

### ***Sensitivity Assessment of the Surveillance System***

As described in [**Methods**](#_Methods) ***Step 6*** above, we assessed the sensitivity of the surveillance system to identify work-related MVC injury cases by linking potential cases from each of the four MA CRISS data sources with FACE[[21]](#footnote-22) MVC fatalities identified by MA DPH OHSP and calculating the percentage of FACE fatality cases identified by each data source. Since only people transported to a hospital would be captured in acute care hospital record, Trauma Registry, or MATRIS data, sensitivity was based on the number of FACE cases documented to have been transported to a hospital prior to death.

Of the 16 fatalities in 2019 FACE data, 11 (68.8%) were identified as work-related MVC injury cases in one or more MA CRISS data sources. Specifically, 9 of the 16 (56.3%) FACE fatalities were identified as work-related cases in police crash reports. Of the 12 FACE cases documented as having been transported to a hospital prior to death, 2 (16.7%) were identified as work-related cases in acute care hospital records, and 2 (16.7%) in Trauma Registry data.

Of the 13 fatalities in 2020 FACE data, 7 (53.9%) were identified as work-related MVC injury cases in one or more MA CRISS data sources. Specifically, 7 of the 13 cases were identified as work-related cases in police crash reports. Of the 7 FACE fatalities documented to have been transported to a hospital by ambulance prior to death, only 1 (14.3%) was identified as a work-related case in acute care hospital records, and none were identified in MATRIS data.

Despite the moderate capture of 2019 and 2020 FACE fatality cases by MA CRISS data sources, MA CRISS data sources did capture a few additional potential work-related MVC fatalities that were not in OHSP’s FACE data. There were 24 total potential work-related MVC fatalities identified in 2019 MA CRISS data sources, of which 13 (54.2%) were not in the FACE data. There were 16 total potential work-related MVC fatalities identified in 2020 MA CRISS data sources, of which 9 (56.3%) were not in FACE data. OHSP staff investigated the cases that were not previously identified as FACE fatalities and found that 1 case in 2019 met the criteria for FACE fatalities and had not been identified in the usual FACE data sources. The other 12 cases in 2019 and the 9 cases in 2020 did not meet the FACE inclusion criteria but still met our criteria as a potential work-related MVC fatality case. The reasons these cases did not meet FACE inclusion criteria included that they were not using the vehicle for work at the time of the crash, their work status was undetermined, or the MVC and death from a medical event did not result in traumatic injury.

As noted in the **Methods**, we wanted to assess the sensitivity of MA CRISS data sources in identifying nonfatal cases by comparing a list of workers who suffered nonfatal MVC injuries identified by OHSP with the 2019 and 2020 nonfatal potentialwork-related MVC injury cases identified by MA CRISS data sources. However, the results were inconclusive due to insufficient personal identifiers common to the OHSP-identified cases and MA CRISS data sources to allow for data linkage. We, therefore, were not able to assess the sensitivity of MA CRISS data sources in identifying nonfatal work-related MVC injury cases.

## **Data Source Evaluation**

As stated in [**Methods**](#_Methods) ***Step 7***, we created a list of each data source’s strengths and limitations in its ability to identify work-related MVC injury cases. See **Table 4**. We focused on the sensitivity of each data source and considered the number and quality of work-related indicators in each data source and the likelihood of capturing false positive cases. Trauma Registry data was the only data source we utilized that included occupation and industry variables, but they were poorly completed.

**Table 4. Key Strengths and Limitations of each MA CRISS Data Source in Identifying Work-Related Motor Vehicle Crash (MVC) Injury Cases**

| **MA CRISS Data Source** | **Strengths** | **Limitations** |
| --- | --- | --- |
| **Police Crash Reports** | Captures many potential work-related MVC injury cases given the large number of variables related to vehicle type and owner.  Includes a variety of injury severity levels, including fatalities.  Work zone-related crashes can be identified through several variables.  Gig workers can be identified through narrative text searches. Appears to capture more gig workers than MATRIS narratives. | May include more false positive cases than other data sources due to characteristics of the crash such as work-zone relation being applied to all injured parties of said crash, the potential use of commercial vehicles outside of work, and the inclusion of “possible” injury cases.  Contains a limited number of variables to identify non-motorists working at the time of the crash.  Some variables are not well completed.  Open text fields for vehicle owner names require manual review and complex search criteria.  Narrative text searching is complicated and requires manual review to remove false positives.  Gig workers cannot always be differentiated from delivery persons working for wages in the narratives.  In MA data, Boston was not submitting crash reports to the statewide data system in 2019 or 2020. |
| **Massachusetts Ambulance Trip Record Information System (MATRIS)** | Likely to include more serious injury cases and some fatalities.[[22]](#footnote-23)  There is a yes/no variable indicating if the injury/illness was work-related.  Gig workers can be identified through narrative text searches. | It is difficult to identify MVC injury cases in general, which may contribute to false positive cases.  There are a limited number of work indicators in V2 and V3 MATRIS data.  Narrative text searching is complicated and requires manual review.  Gig workers cannot always be differentiated from delivery persons working for wages in the narratives. |
| **Acute Care Hospital Records**  (Hospital Discharge, Observation Stay, and Emergency Department Discharge data) | Captures a broad range of injury severity, including serious injuries and some fatalities.[[23]](#footnote-24)  Workers’ Compensation can be identified in payer source and type variables, and work status can be identified in the ICD-10-CM Y99 codes.  Likely to capture fewer false positives than police crash report data given the specificity of the work indicators. | The ICD-10-CM external cause of morbidity (V) codes only capture a few types of work vehicles.  The ICD-10-CM codes for place of occurrence (Y92), activity (Y93), work status (Y99), and visit type (Z) codes are often missing.  Gig workers cannot be identified due to the absence of narratives or other relevant variables. |
| **Trauma Registry** | Captures serious injuries and some fatalities.[[24]](#footnote-25)  There is a yes/no variable indicating if the injury/illness was work-related.  Likely to capture fewer false positives than police crash report data given the specificity of the work indicators. | Identifies few additional cases beyond those identified in acute care hospital record or police crash report data.  The ICD-10-CM external cause of morbidity (V) codes only capture a few types of work vehicles.  The ICD-10-CM codes for place of occurrence (Y92), activity (Y93), work status (Y99), and visit type (Z) codes are often missing.  The patient occupation and industry variables are mostly incomplete or unknown.  In MA data, the payment method variable does not include Worker’s Compensation.  Gig workers cannot be identified due to the absence of narratives or other relevant variables. |

We separately evaluated the strengths and limitations of each data source in terms of nine attributes recommended by the CDC to evaluate the performance of public health surveillance systems.[[25]](#footnote-26) These attributes include sensitivity, representativeness, predictive value positive, simplicity, data quality, acceptability, timeliness, stability, and flexibility. See [**Appendix F**](#_Appendix_F) for details.

## **Conclusions**

In summary, we found that the use of multiple data sources to identify potential work-related MVC injury cases was extremely valuable, as each system captured unique cases and there was relatively low case overlap. Our MA CRISS data sources identified approximately 5,400 potential work-related MVC injury cases in 2019 and approximately 3,500 cases in 2020 (pandemic era). Acute care hospital record and police crash report data identified the most potential work-related MVC injury cases, although crash report data may include more false positive cases than hospital record data.

For comparison, we looked at the number of work-related MVC injuries estimated by the Bureau of Labor Statistics (BLS) Survey of Occupational Injuries and Illnesses (SOII). This survey includes work-related injuries and illnesses that result in lost work time. SOII data estimate that there were 1,680 work-related MVC injuries in MA in 2019, and 1,170 such injuries in MA in 2020. These estimates are about one-third of the potential work-related MVC injury cases identified by MA CRISS data sources, which do not consider whether injuries resulted in lost work time. Although SOII is widely used and can offer a great level of detail, it has well-documented limitations.[[26]](#footnote-27) In addition to focusing on injuries and illnesses that result in lost work time, SOII is also based on logs maintained by employers for OSHA. For an MVC injury to appear in these records, the employer must acknowledge the MVC as work-related and add it to the injury log. Altogether, the rich data on potential work-related motor vehicle crash events available from MA CRISS data sources represent a great enhancement in surveillance and a complement to traditional systems.

We were able to identify some gig workers among potential work-related MVC injury cases in police crash report and MATRIS narratives, but the numbers of gig workers identified are likely to be greatly underestimated for several reasons. These reasons include the insufficiency of specific variables to identify gig workers and the need to use narratives to identify such workers. Narrative text searches for gig workers were complicated and required manual review to remove false positives. Additionally, gig workers could not always be differentiated from employed delivery people working for wages. Another barrier is that crash report or emergency medical service (EMS) narratives may not be available to researchers in all states.

For our next steps, we will analyze the 2019 and 2020 potential work-related MVC injury cases by person type, injury severity, race and ethnicity, age group, sex, and contextual factors, such as vehicle type and crash location, to identify workers at greatest risk of such injuries and potential areas of intervention. We will also analyze additional years of MA CRISS data sources to look at trends. Further research is needed to confirm the validity of the potential work-related MVC injury cases identified by the data sources and to refine the algorithms that identify these cases to reduce false positives. This is particularly true for police crash report cases, which included crash event-level indicators (such as crash occurrence in a work zone), “possible” injury cases, and cases only identified due to the crash involvement of a commercial vehicle, which may be used for non-work travel.

To improve these data sources for use in identifying MVC injuries among workers, we recommend advocating that police crash report, ambulance trip record, and hospital record data include the following specific work indicators: whether an injury/illness was work-related or had occurred at their workplace; if the worker was a civilian, military, or gig worker; and, if possible, the occupation and industry of the worker. We also recommend conducting outreach to data owners and record keepers about the value of documenting work-related injuries and illnesses. People involved in the documentation of the MA CRISS data sources we utilized include law enforcement officers, EMS providers, health care providers, trauma registrars, and health information management staff. Improving the identification of work-related MVC injuries and contextual information in statewide data sources would improve a state’s ability to determine the magnitude of these injuries and high-risk workers. This would ultimately inform overall prevention efforts, particularly for populations often excluded from existing efforts, such as gig workers, due to the limitations in identifying such populations.

# **Appendix A**

## **Potential ICD-10-CM Work Indicators**

We used the following ICD-10-CM codes for external cause, place of occurrence, activity, (work) status, and visit type as work indicators. These codes were largely based on ICD-10-CM work indicators developed by Bush et al. (2021).[[27]](#footnote-28) We searched all ICD-10-CM relevant fields for each code. Where ICD-10-CM external cause codes included a 7th character, we only included those with a 7th character of “A” or missing, to capture active treatment of injuries rather than routine follow-up care or treatment of injury sequelae.

|  |  |
| --- | --- |
| **Potential ICD-10-CM Work Indicators** | **Description** |
| ***External Causes of Morbidity (V60–V86)*** | |
| V60-V69 | Occupant of heavy transport vehicle injured in transport accident |
| V70-V78 (4th char = 0, 5)  V79 (4th char = 0, 4) | Driver of bus injured in transport accident |
| V83 | Occupant of special vehicle mainly used on industrial premises injured in transport accident |
| V84 | Occupant of special vehicle mainly used in agriculture injured in a transport accident |
| V85 | Occupant of special construction vehicle injured in transport accident |
| V86 (5th char = 1) | Occupant of ambulance or fire engine injured in transport accident |
| V86 (5th char = 4) | Occupant of military vehicle injured in transport accident |
| ***Place of Occurrence (Y92)*** | |
| Y92.6 | Industrial and construction area as the place of occurrence of the external cause |
| Y92.7 | Farm as the place of occurrence of the external cause |
| Y92.84 | Military training ground as the place of occurrence of the external cause |
| ***Activity (Y93)*** | |
| Y93.H3 | Activity, building and construction |
| Y93.H9 | Activity, other involving exterior property and land maintenance, building and construction |
| ***External Cause (Work) Status (Y99)*** | |
| Y99.0 | Civilian activity done for income or pay |
| Y99.1 | Military activity |
| ***Persons Encountering Health Services for Examinations (Z00–Z13)*** | |
| Z04.2 | Encounter for examination and observation following work accident |

# **Appendix B**

## **Potential Work Indicators in 2019 & 2020 MA Police Crash Report Data**

Note that many of the variables and values shown in the tablebelow were only used as work indicators when combined with specific person types, such as MV occupants or drivers alone. These parameters are included above the list of associated values/keywords and underlined, e.g., “Among Injured MV Occupants”.

| **Variable** (by category) | ***Name* – Description** | **Values/Keywords[[28]](#footnote-29)** |
| --- | --- | --- |
| **Vehicle Type & Weight Variables** | | |
| cargo\_body\_type\_code | *Cargo Body Type Code* – The type of body for buses and trucks more than 10,000 lbs in Gross Vehicle Weight Rating (GVWR). | Among Injured MV Occupants:  4 – Grain/gravel truck  5 – Pole-trailer  6 – Cargo tank  7 – Flatbed  8 – Dump  9 – Concrete mixer  10 – Auto transporter  11 – Garbage truck  12 – No cargo body (bobtail, light MV with hazardous materials (HM) placard, etc.)  13 – Log  14 – Intermodal container chassis  15 – Vehicle towing another vehicle  AmongInjured Drivers Only:  1 – Bus (seats for 16+ people, including driver)  2 – Bus (seats 9-15 people, including driver) |
| emer\_use\_code | *Emergency Use Code* – Whether the vehicle involved in the crash was an emergency response vehicle (such as police, ambulance, fire, military, etc.) involved in responding to an emergency with physical emergency signals in use. | Among Injured MV Occupants:  1 – Yes |
| gross\_vehc\_wght\_ratg\_code | *Gross Vehicle Weight Rating (GVWR) Code* – The amount recommended by the manufacturer as the upper limit to the operational weight for a motor vehicle and any cargo (human or other) to be carried. | Among Injured MV Occupants:  2 – 10,001-26,000 lbs  3 – More than 26,000 lbs |
| haz\_mat\_plcd\_code | *Hazmat Placard Code* – The presence of a placard indicating hazardous materials were being transported by the vehicle. | Among Injured MV Occupants:  1 – Yes [placards on all 4 sides of vehicle]  2 – No [transports hazardous materials without appropriate placard] |
| icc\_numb | *ICC (Interstate Commerce Commission Motor Carrier) Number* – Identifies a carrier or company that transports regulated items from another ownership or arranges their transport (for a fee or other compensation, in interstate commerce), or transports passengers in Interstate commerce (for a fee or other compensation, albeit direct or indirect compensation). | Among Injured MV Occupants:  All valid numbers  *Excluded* cases with missing, “none,” “unknown,” “not reported”, “not available” or “not applicable” values. |
| interstate | *Interstate Status* – The type of carrier involved in the crash. | Among Injured MV Occupants:  0 – Intrastate [within the state]  1 – Interstate [commercial vehicle that can cross state lines and needs a US DOT number]  3 – Not in Commerce (Government) [any government vehicle operated by local/state/federal government] |
| us\_dot\_numb | *US DOT (Department of Transportation) Number* – The identification number of an individual, partnership, or corporation responsible for the transportation of persons or property as indicated on the shipping manifest. | Among Injured MV Occupants:  All valid numbers  *Excluded* cases with missing, “none,” “unknown,” “not reported”, “not available” or “not applicable” values. |
| vehc\_confg\_code | *Vehicle Configuration Code* – Motor vehicle body type/configuration. | Among Injured MV Occupants:  6 – Single unit truck (2 axles, 6 tires)  7 – Single unit truck (3+ axles) [straight truck]  8 – Truck/trailer  9 – Truck tractor (Bobtail)  10 – Tractor/semi-trailer  11 – Tractor/doubles  12 – Tractor/triples [truck tractor pulling 1 semi-trailer and 2 full-trailers]  13 – Unknown heavy truck  Among Injured Drivers Only:  4 – Bus (seats for 16+, including driver)  5 – Bus (seats 9-15 people, including driver) |
| **Vehicle Owner Name Variables & Vehicle Registration Type Code** | | |
| vehc\_owner\_first\_name | *Vehicle Owner's First Name* – Vehicle owner's first name if a person.[[29]](#footnote-30) | Among Injured Drivers Only:  *Included* cases where vehc\_owner\_last\_name was missing, “Unknown”, or “Owner”.[[30]](#footnote-31)  *Excluded* cases where vehc\_reg\_type\_code values started with DL for Dealer, as these plates may be on cars at a dealership being test-driven by potential buyers.    *Excluded* cases with missing values, keywords such as “N/A”, “none”, and “unknown”, typos, and the following keywords in vehc\_owner\_first\_name which may indicate a car at a dealership or vehicle leased by a private individual:  Ean  Enterprise  U-Haul  Rent *or* Lease  Holding  Bank *or* Trust *or* Finance *or* Credit  Gelco  Hertz  Nissan  Ryder  Penske  Avis  Accord  Auto  Buick  Cadillac  Chevrolet  Flexdrive  Kia  Motor  Subaru  Toyota  Advantage  Balise  Herb Chambers  Capitol  Among Injured Passengers Only:  Same *inclusions* and *exclusions* as Among Injured Drivers Only.  Also *excluded* cases with these keywords in vehc\_owner\_first\_name:  Taxi *or* Cab  Limo  Car  Transport *or* Transit  Rideshare *or* Ride  Livery  Ambulance  Bus |
| vehc\_owner\_last\_name | *Vehicle Owner's Last Name* – Vehicle owner's last name if a person.[[31]](#footnote-32) | Among Injured Drivers Only:  *Included* cases where vehc\_owner\_first\_name was missing, “Unknown”, or “Owner”.[[32]](#footnote-33)  Same *exclusions* as vehc\_owner\_first\_name.  Among Injured Passengers Only:  Same *inclusions* and *exclusions* as vehc\_owner\_first\_name. |
| vehc\_owner\_nown | *Vehicle (No Individual) Owner Name* – Vehicle owner's company name if non-individual owner.[[33]](#footnote-34) | Among Injured Drivers Only:  *Included* cases where the first word in vehc\_owner\_nown did NOT= pers\_first\_name (the individual case’s first name), *or* if the first word in vehc\_owner\_nown = pers\_first\_name, then the values in vehc\_owner\_nown must have also *included* the keywords “co”, “inc”, or “llc”.  Same *exclusions* as vehc\_owner\_first\_name and vehc\_owner\_last\_name.  Among Injured Passengers Only:  Same *inclusions* and *exclusions* as vehc\_owner\_first\_name and vehc\_owner\_last\_name. |
| vehc\_reg\_type\_code | *Vehicle Registration Type Code* – The type of vehicle registration and associated plate, based on vehicle type and usage. | Among Injured MV Occupants:  *Excluded* cases with missing values, values such as “N/A”, “none”, “unknown”, and “other”, typos, and those that begin with:  PA – Passenger  AH – Camper  MC – Motorcycle  MP – Moped  OS – “Not found”  TP – Trailer  TR – Trailer (not used on Campers)  DL – Dealer  BU – Bus  LN – Livery  LV – Livery  SB – School Bus  SP – School Pupil  TA – Taxi  VP – Van Pool  Among Injured Drivers Only:  *Included* cases with values that begin with BU, LN, LV, SB, SP, TA, and VP. |
| **Road Contributing Circumstance & Work Zone Codes** | | |
| road\_cntrb\_code | *Road Contributing Circumstances Code* – Road conditions that may have contributed to the crash. | Among Injured MV Occupants and Non-Motorists:  7 – Work zone (construction, maintenance, utility)  8 – Non-highway work [work occurring in or near the trafficway but not related to the trafficway] |
| work\_zone\_reld\_code | *Work-Zone Related Code* – Indicating a crash that occurs in or related to a construction, maintenance, or utility work zone, identified by warning signs/signals/indicators, including those on transport devices, regardless of worker presence at the time of the crash. | Among Injured MV Occupants and Non-Motorists:  1 – Yes |
| **Non-Motorist Activity Code** | | |
| non\_mtrst\_act\_code | *Non-Motorist Activity* *Code* – Actions made by the non-motorist just prior to the crash based on verbal or physical evidence. | Among Injured Non-Motorists Only:  3 – Working |
| **Narrative** | | |
| narrative[[34]](#footnote-35) | *Crash Report Narrative* – Written narrative by reporting law enforcement officer detailing the crash circumstances, persons and vehicles involved, and disposition of injured persons and vehicles.[[35]](#footnote-36) | Among Injured MV Occupants and Non-Motorists:  *Included* keywords:  Rideshare  Uber Driver *or* Driving for Uber  Lyft Driver *or* Driving for Lyft  Food *and* Deliver  Uber Eats  Grubhub  Postmates  DoorDash  Grocery *and* Deliver  Instacart  Package *and* Deliver  Amazon  DHL |

# **Appendix C**

## **Potential Work Indicators in 2020 MA Ambulance Trip Record (MATRIS) Data**

Note that in 2020, Massachusetts Emergency Medical Service (EMS) agencies were transitioning from reporting data in the National Emergency Medical Services Information System (NEMSIS) Version 2 format to NEMSIS Version 3. Approximately 25% of MA motor vehicle crash (MVC) injury cases were reported using the V2 format, and 75% were reported using the V3 format in 2020.

| **NEMSIS Version 2 Variable[[36]](#footnote-37)** | **NEMSIS Version 3 Variable37** | ***Name* – Description** | **Included Codes/Keywords[[37]](#footnote-38)** |
| --- | --- | --- | --- |
| N/A | Situation\_Work\_Related\_Illness | *Work-Related Illness/Injury* or *Work-Related* – Indicates whether or not the illness or injury is work-related. | Among Injured MV Occupants and Non-Motorists:  “Yes” |
| Cause\_Of\_Injury\_\_E10\_1\_  Cause\_Of\_Injury\_\_E10\_1\_ (cont.) | Injury\_Cause\_Of\_Injury\_Descript  Injury\_Cause\_Of\_Injury\_Descript (cont.) | *Cause of Injury* – ICD-10-CM External Causes of Morbidity description and code. | Note that the parentheses “()” included below were part of the variable value. The brackets “[]” indicate an array of numbers/letters used in the prxmatch function. The curly brackets “{}” indicate the required combinations of keywords and are not part of the variable value. We needed to include both keywords and codes due to truncation of the values in this field.  *Included* cases with keywords:  Occupant of Heavy *or* (Passenger) of Heavy *or* Driver of Heavy *or* Passenger of Heavy *or* V6  Driver of Bus *or* Bus Driver *or* V7[0-9].0 *or* V7[0-8].5 *or* V79.4  {Occupant of Special *or* (Passenger) of Special *or* Driver of Special *or* Passenger of Special} *and*  {Industrial *or* V83}  {Occupant of Special *or* (Passenger) of Special *or* Driver of Special *or* Passenger of Special} *and*  {Agricultural *or* V84}  Occupant of Special Construction *or* (Passenger) of Special Construction *or* Driver of Special Construction *or* Passenger of Special Construction *or* V85  Occupant of Ambulance *or* (Passenger) of Ambulance *or* Driver of Ambulance *or* Passenger of Ambulance *or* V86.[0-7]1 *or* V86.91  Occupant of Military *or* (Passenger) of Military *or* Driver of Military *or* Passenger of Military *or* V86.[0-7]4 *or* V86.94 |
| Incident\_Narrative\_Report\_\_E13\_1  Incident\_Narrative\_Report\_\_E13\_1 (cont.)  Incident\_Narrative\_Report\_\_E13\_1 (cont.) | Patient\_Care\_Report\_Narrative\_\_e  Patient\_Care\_Report\_Narrative\_\_e (cont.)  Patient\_Care\_Report\_Narrative\_\_e (cont.) | *Patient Care Report Narrative* or *Run Report Narrative* – The narrative of the patient care report (PCR).[[38]](#footnote-39) | Note that the parentheses “()” included below were part of the variable value. The brackets “[]” indicate an array of numbers/letters used in the prxmatch function. The curly brackets “{}” indicate the required combinations of keywords and are not part of the variable value.  *Included* cases with keywords:  {Pt Was Work *or* Non-Motorist *or* Pedestrian *or* Cyclist} *and*  {Working *or* For Work} *and*  {Motor Vehicle *or* MV *or* Vehicle *or* Auto} *and NOT*  {Working on Extrication *or* Motorcyclist *or* Working a Detail}  {Work Van *or* Work Vehicle *or* Work Truck *or* Driving for Work} *and*  {Involving Police Vehicle *or* Involving Police Cruiser} *and*  {Driver *or* Passenger *or* Occupant} *and NOT*  {By a Work}  Driver of Tractor *or* His Tractor *or* Trailer Driver *or* Driver of Box *or* Box Truck Driver *or* Driver of Semi *or* Semi Driver  {Ride-Share} *and*  {Drive} *and NOT*  {Passenger}  {Uber Driver *or* Driving for Uber *or* Driving for Work with Uber} *and NOT*  {Passenger *or* Back Seat *or* Rear Seat *or* Sitting in the Back *or* Ordered an Uber}  {Lyft Driver *or* Driving for Lyft *or* Driving for Work with Lyft} *and NOT*  {Passenger *or* Back Seat *or* Rear Seat *or* Sitting in the Back *or* Ordered a Lyft}  {Food *or* Grocery} *and*  {Deliver *or* Carrier *or* Courier *or* Messenger}  {Uber Eats} *and*  {Deliver *or* Carrier *or* Courier *or* Messenger}  {Grubhub} *and*  {Deliver *or* Carrier *or* Courier *or* Messenger}  {Postmates} *and*  {Deliver *or* Carrier *or* Courier *or* Messenger}  {DoorDash} *and*  {Deliver *or* Carrier *or* Courier *or* Messenger}  {Instacart} *and*  {Deliver *or* Carrier *or* Courier *or* Messenger}  {Package} *and*  {Deliver *or* Carrier *or* Courier *or* Messenger} *and NOT*  {Packaged *or* On Package}  {Amazon} *and*  {Deliver *or* Carrier *or* Courier *or* Messenger} *and NOT*  {Delivery Bag}  {DHL} *and*  {Deliver *or* Carrier *or* Courier *or* Messenger}  {Mail *or* Postal *or* USPS} *and*  {Deliver *or* Carrier *or* Courier *or* Messenger} *and NOT*  {Via U.S. Mail *or* Via Mail *or* (Delivered by USPS Mail) *or* Via USPS}  {FedEx} *and*  {Deliver *or* Carrier *or* Courier *or* Messenger} *and NOT*  {With FedEx *or* By FedEx}  {UPS} *and*  {Deliver *or* Carrier *or* Courier *or* Messenger} *and NOT*  {Follow Ups} |

# **Appendix D**

## **Potential Work Indicators in 2019 & 2020 MA Acute Care Hospital Record Data**

|  |  |  |
| --- | --- | --- |
| **Variable(s)[[39]](#footnote-40)** | ***Name* – Description** | **Included Codes/Keywords[[40]](#footnote-41)** |
| InjEcode1 – InjEcode15 | *ICD-10-CM Principal and Additional External Cause Codes* | Refer to **Appendix A.** Note that these fields were created by MA DPH from “ICD-10-CM associated diagnosis fields” to make it easier to identify injury-related external cause codes. |
| InjPlace1 – InjPlace3 | *ICD-10-CM Place of Occurrence Codes* |
| InjAct1 – InjAct3 | *ICD-10-CM Activity Codes* |
| InjStatus | *ICD-10-CM External Cause (Work) Status Codes* |
| PayrType, Pay2Type | *Type of Primary/Secondary Payer* | 2 – Worker’s Compensation |
| PayrSrce, Pay2Srce | *Standardized Payer Source Code* – Primary and Secondary payer for ED visit and Observation Stay | 146 – Worker’s Compensation (WOR) |

# **Appendix E**

## **Potential Work Indicators in 2019 MA Trauma Registry Data**

|  |  |  |
| --- | --- | --- |
| **Variable(s)** | ***Name* – Description** | **Included Codes/Keywords[[41]](#footnote-42)** |
| WorkRelated | *Work-Related* – If the injury was caused by a work-related incident. | 1 – Yes |
| ICD10PrimaryExternalCauseCode | *ICD-10-CM External Causes of Morbidity Codes* | Please see **Appendix A** for specific codes. We searched for external cause codes, place of occurrence codes, activity codes, external cause (work) status codes, and examination reason codes in all applicable ICD-10-CM code fields. |
| ICD10PlaceofOccurrenceExternalCa | *ICD-10-CM Place of Occurrence Codes* |
| ICD10AddECode1 – ICD10AddECode12  E\_V\_Codes1 – E\_V\_Codes28  Other\_ICD\_10Codes1 – Other\_ICD\_10Codes58  Morb\_ICD\_10Codes1 – Morb\_ICD\_10Codes37 | *Additional ICD-10-CM Code Variables[[42]](#footnote-43)* |

# **Appendix F**

## **CDC Surveillance System Attributes of Each Data Source[[43]](#footnote-44)**

This section evaluates the work-related motor vehicle crash (MVC) surveillance system developed by MA DPH in terms of the nine attributes specified in CDC guidelines for surveillance system evaluation: *sensitivity*, *representativeness*, *predictive* *value* *positive*, *simplicity*, *data quality*, *acceptability*, *timeliness*, *stability*, and *flexibility*. Our assessment of sensitivity and representativeness was limited by the absence of a “gold standard” for nonfatal work-related MVC injury cases.

**Surveillance System Attribute Definitions:**

* *Sensitivity* refers to the proportion of cases of a health-related outcome detected by the surveillance system or its “capture rate”.
* *Representativeness* refers to how accurately the data collected by the system reflects the true occurrence of a health event within a specific population, considering factors like demographics, location, and time, essentially indicating how well the surveillance system captures the overall picture of the health issue it's monitoring.
* *Predictive Value Positive* (*PVP*) is the proportion of reported cases that actually have the health-related outcome/event under surveillance, i.e., true positive cases.
* *Simplicity* describes the simplicity of the surveillance system’s structure and ease of operation while meeting its objectives, i.e., identifying work-related motor vehicle crash injuries.
* *Data Quality* reflects data completeness and validity within the system.
* *Acceptability* refers to the willingness of individuals and organizations, within and outside the sponsoring agency that manages the system, to document and submit information to the surveillance system.
* *Timeliness* describes the speed between steps in the surveillance system, such as the interval between the occurrence of the crash, the reporting of the crash, and the use of the data for surveillance purposes.
* *Stability* is the system’s ability to collect, manage, and provide data properly without failure (*reliability*) and its ability to be operational when needed (*availability*).
* *Flexibility* describes a surveillance system’s ability to adapt to necessary information changes or operating conditions with little personnel, additional time, or allocated funds, as well as using standard data formats that can be easily integrated with other systems.

### ***MA Police Crash Report Data***

| **Surveillance System Attributes** | **Strengths** | **Limitations** |
| --- | --- | --- |
| Sensitivity | MVCs involving commercial vehicles are likely to be identified as there are several variables on vehicle characteristics and ownership. This included vehicle type, weight, size, presence of hazardous materials, ICC number, US DOT number, registration number, name of the owner, etc.  2019 police crash reports identified as potential work-related cases 56.3% (n = 9) of the 16 work-related fatalities externally identified by the MA DPH Fatality Assessment and Control Evaluation (FACE) program.  2020 police crash reports identified as potential work-related cases 53.9% (n = 7) of the 13 work-related fatalities in 2020 that were externally identified by the MA DPH FACE program.  We identified fatalities in police crash report data that were not previously identified by the FACE program. Several of these cases were subsequently verified as cases that met FACE criteria but were not previously identified. | The incompleteness of many of the work-related indicators may lead to a lower sensitivity.  There were few variables available that indicated whether injured non-motorists were working, i.e., non-motorist activity, work-zone occurrence, road-contributing circumstances, and narratives, which likely lowered the overall sensitivity of the algorithm. |
| Representativeness | Police crash reports are more representative of the scope of fatal and non-fatal work-related MVC injury cases as police are required to report all unintentional injury crashes that occur on MA public roadways. EMS and hospital data sources may not capture fatalities taken directly to the morgue.  Gig workers could be identified through narrative text-searching.  About one-third (36%) of the injured gig workers identified through police crash report narratives were rideshare drivers. | There are few variables available that indicated whether injured non-motorists were working, i.e., non-motorist activity, work-zone occurrence, road-contributing circumstances, and narratives, making it likely that working non-motorists may be under-represented.  Gig workers were likely to be under-represented, however, as there were no specific indicators for gig workers, they were difficult to identify through text searches, and gig workers for private companies, such as Amazon, could not be differentiated from regular employees in the narratives. |
| Predictive Value Positive (PVP) | Of the 22 potential work-related MVC fatalities identified in 2019 police crash reports, 45.5% (n = 10) were verified as true positives by OHSP (as FACE cases or in subsequent review by OHSP). The other 12 cases still met our criteria for potential work-related MVC injury cases but did not meet the strict criteria for FACE cases.  Of the 14 work-related MVC fatalities identified in 2020 police crash reports, 50.0% (n = 7) were verified as true positives by OHSP (as FACE cases). The other 7 cases still met our criteria for potential work-related MVC injury cases, but did not meet the strict criteria for FACE cases. | Police crash reports may have more false positive cases due to 1) some work indicators were crash rather than person-level indicators (such as work zone related), 2) the commercial vehicles identified may have been being driven for non-work-related purposes, and 3) we included cases identified by police as “possible” injury cases to increase the sensitivity of the surveillance system.  There were initially 82 gig work cases identified in the narratives, but only 32 (39.0%) cases ended up being true positives (therefore yielding a low PVP).  Some non-gig-worker keywords, especially those related to USPS, FedEx, or UPS workers, also identified many narratives that turned out to be false positives, e.g., the commercial vehicle was not involved in the crash but obstructed the view of another driver; the commercial vehicle was almost involved in the crash; or the working driver or pedestrian was simply a witness to the crash. |
| Simplicity | This data source focuses on motor vehicle crash injuries and contains a variable that can differentiate the severity of injuries per case. | There is no stand-alone variable indicating if the illness/injury was work-related.  Text searching of narrative and other free text fields was complicated as it often required multiple inclusion and exclusion terms and needed to be adjusted based on the terms included in each year’s data. Manual review was necessary to double-check that narrative searches were true positives and to exclude duplicate cases. |
| Data Quality | Many work indicators had a high rate of completeness: work zone relation (99.7%), vehicle configuration code among injured occupants (99.5%), emergency use code among occupants (97%), road contribution code (96%), non-motorist activity code among injured non-motorists (91%), vehicle registration type code (90%) among injured occupants, individual vehicle owner first name (86%) among occupants, and individual vehicle owner last name (86%) among occupants. | Several work indicators were mostly unknown or incomplete, i.e., ICC number (0.04% complete), hazardous material placard presence (0.12%), interstate status (0.25%), cargo body type (0.28%), US DOT number (0.37%), gross vehicle weight rating (0.49%), and company vehicle owner name (11%).  We found some problems with internal consistency of the data in that some indicators had values that did not align with each other or with the narrative, e.g., several mopeds had a vehicle registration type for a vehicle for transporting school pupils, or work-zone related was documented, but the narrative did not mention a work-zone.  Some narratives did not have detailed information, e.g. stated “see supplemental report” or that the crash was still under investigation.  We found some duplicate crash reports that should have been removed by the state Crash Data System. |
| Acceptability | Most local and state law enforcement agencies in MA (N = 395) submit crash reports to the state Crash Data System. Most law enforcement agencies transitioned to submitting their crash reports electronically over the past few years. | Boston did not submit crash reports to the state Crash Data System for many years but began submitting cases in 2024. |
| Timeliness | Most law enforcement agencies in MA submit crash reports electronically. These can be immediately submitted to the RMV or saved and submitted as part of a larger package at a scheduled time depending on the local agency. Preliminary crash data may be available on the MA Crash Impact Portal[[44]](#footnote-45) within days or weeks of the crash. | While preliminary crash data is fairly timely, crash data may not be finalized by the RMV for several years. In 2024, the most recent final crash data available was for 2020. Fatality cases often take the longest to finalize. |
| Stability | The Crash Data System in MA is stable in that the RMV has been operating continuously and reliably since 2012. There is a routine data validation process in place and crash data are available to other agencies upon request when the necessary data use agreements are in place. |  |
| Flexibility |  | On the national level, changes to the data collected occurs through updates to the Model Minimum Uniform Crash Criteria (MMUCC) guidelines. States may adopt the new guidelines or adjust data collected on their own. Development of new state crash report forms and updating law enforcement agencies’ electronic reporting systems is difficult and time consuming. |

***MA Ambulance Trip Record Information System (MATRIS) Data***

| **Surveillance System Attributes** | **Strengths** | **Limitations** |
| --- | --- | --- |
| Sensitivity |  | MATRIS data are less likely to capture work-related MVC fatalities, as deaths on the scene may be transported to a morgue by the medical examiner rather than taken by ambulance to a hospital.  We were not able to identify any of the 13 work-related fatalities in the 2020 FACE data within the 2020 MATRIS data.  Work-related crashes involving passenger cars or light trucks are also less likely to be captured by MATRIS data, as the ICD-10-CM cause of injury codes focus on crashes involving specific types of commercial vehicles. |
| Representativeness | MATRIS data should be representative of most nonfatal injury cases that involved a serious enough injury that a 911 call was made.  It may be possible to identify gig workers through narrative text searches. | Workers with nonfatal injuries who were not transported to a medical facility by ambulance would not be included in MATRIS data.  Workers who died at the scene and were transported to a morgue by the medical examiner would not be included in MATRIS data. We were not able to identify any FACE fatalities in MATRIS data.  Work-related crashes involving passenger cars or light trucks are also less likely to be captured by MATRIS data, as the ICD-10-CM cause of injury codes focus on crashes involving specific types of commercial vehicles.  Gig workers may be under-represented in MATRIS data as there were no specific indicators for gig workers and in text searches of MATRIS narratives, gig workers in private companies could not be differentiated from employees. In addition, external organizations may not be able to access MATRIS narratives, which would limit their ability to identify work-related MV injury cases, especially among gig workers. |
| Predictive Value Positive (PVP) | Work-related cases identified in MATRIS are more likely to be true positives due to the specific nature of most indicators.  The majority of cases identified through text searches of MATRIS narratives were found to be true positives in manual review.  There were originally 14 postal/FedEx/UPS delivery cases identified in the narratives, of which 13 (92.9%) were true positives.  There were originally 107 work vehicle/truck/van occupant, trailer/box truck driver, and worker/non-motorist cases identified in the narratives, of which 91 (85.1%) were true positives.  There were originally 14 gig work cases identified in the narratives, of which 11 (78.6%) were true positives. | It was difficult to identify MVC injuries in general in MATRIS data, which may have led to the inclusion of some non-MVC injury cases.  Cases identified through text searches of MATRIS narratives may yield some false positives.  The 1 potential work-related MVC fatality identified in 2020 MATRIS could not be verified as a true positive by OHSP (as a FACE case or in subsequent review by OHSP). It still met our criteria for potential work-related MVC injury cases but did not meet the strict criteria for FACE cases. |
| Simplicity | MATRIS data includes a binary variable indicating whether the illness/injury was work-related.  Work indicators based on ICD-10-CM external cause codes were easier to identify than indicators that required text searches. | It was difficult to identify MVC injuries in MATRIS data. It required the use of multiple variables and inclusion and exclusion criteria.  Text searching of narrative and other free text fields was complicated as it often required multiple inclusion and exclusion terms. Manual review was necessary to double-check that narrative searches were true positives.  In MA, our 2020 MATRIS data included submissions from EMS agencies using Version 2 and 3 National Emergency Medical Services Information System (NEMSIS) criteria. This required the use of different software code to identify work indicators in each version. |
| Data Quality | Nearly all MVC injury cases had complete ICD-10-CM cause of injury variables (99.8%) and patient care report narratives (99.9%). | There was no work-related incident variable in V2 MATRIS data. The work-related incident variable in V3 MATRIS data was complete for only 27.6% of MVC injury cases.  There were variables in MATRIS data for suspected EMS work-related exposure/injury/death and the patient’s primary method of payment, but the data owners did not provide this information in the data we had access to given low reporting rates by agencies. |
| Acceptability | All public and private ambulance services in MA submit data on ambulance runs to the MA DPH Office of Emergency Medical Services. |  |
| Timeliness | Preliminary MATRIS data is fairly timely, as EMS agencies typically send reports to the state system electronically within 24 hours of the incident. | It may take two or more years for the MA DPH Office of Emergency Medical Services to validate, clean, deduplicate, and finalize MATRIS data for use by other programs and organizations. |
| Stability | MATRIS ambulance run data are stable in that the MA DPH Office of Emergency Medical Services has been collecting this data continuously and reliably since 2013. Routine data validation and cleaning processes are in place and MATRIS data are available to other programs and organizations upon request when the necessary data use agreements are in place. | External organizations may not have access to personally identifying information (PII) in MATRIS data, including narratives. |
| Flexibility |  | Changes in variables or values collected would occur through updates to the national NEMSIS criteria. |

### ***MA Acute Care Hospital Record Data***

| **Surveillance System Attributes** | **Strengths** | **Limitations** |
| --- | --- | --- |
| Sensitivity | Acute care hospital record data may capture a range of nonfatal injury cases and some fatalities, if the worker receives any hospital treatment. | Acute care hospital record data are less likely to capture fatal work-related MVC injuries, as deaths on the scene may have been transported to a morgue rather than a hospital. 2019 acute care hospital data only captured 12.5% (n = 2) of the 16 FACE fatalities that year and 2020 acute care hospital data only captured 7.7% (n = 1) of the 13 FACE fatalities in that year.  Work-related crashes involving passenger cars or light trucks are also less likely to be captured by acute care hospital record data, as the ICD-10-CM cause of injury codes focus on crashes involving specific types of commercial vehicles.  Gig workers cannot be identified due to the absence of narratives or other identifying variables. |
| Representativeness | Acute care hospital record data are most likely to be representative of work-related crash injuries involving specific types of commercial vehicles that are covered in ICD-10-CM external cause codes. | Workers who died at the scene and were transported to a morgue would not be included in acute care hospital data.  Work-related crashes involving passenger cars or light trucks are less likely to be captured by acute care hospital record data, as the ICD-10-CM cause of injury codes focus on crashes involving specific types of commercial vehicles.  Gig workers cannot be identified due to the absence of narratives or other identifying variables. |
| Predictive Value Positive (PVP) | Work-related cases identified in acute care hospital record data are more likely to be true positives due to the specific nature of most indicators.  The 2 fatalities identified in 2019 acute care hospital record data were both verified as fatal work-related cases by OHSP (as FACE cases).  Of the 2 fatalities identified in 2020 acute care hospital record data, 1 (50.0%) was verified as a fatal work-related case by OHSP (as a FACE case). The other case still met our criteria for potential work-related MVC injury cases but did not meet the strict criteria for FACE cases. |  |
| Simplicity | Work-related indicators based on the patient’s payer source and type, and ICD-10-CM external cause, place of occurrence, activity, external cause (work) status, and examination codes were relatively simple to identify. The 7th character of ICD-10-CM codes can also be used to distinguish treatment of acute injuries from routine follow-up treatment or treatment of injury sequelae. | MA acute care hospital record data consists of three separate datasets: inpatient hospital discharge, outpatient observation stay, and emergency department discharge datasets. Given that many variables in these data sources are the same, we decided to combine them prior to identifying work-related crashes. But then it is necessary to deduplicate cases to the individual level. This was complicated by missing unique identifiers (encrypted social security numbers). The lack of an incident date variable made it impossible to deduplicate records to the person-incident level. |
| Data Quality | The following relevant variables in MA acute care hospital record data were mostly complete: ICD-10-CM external cause codes (100%), primary payer source codes (81%), and secondary payer source codes (72%). | The following relevant variables in MA acute care hospital record data were mostly incomplete: ICD-10-CM place of occurrence external cause codes (42% complete), ICD-10-CM activity external cause codes (15% complete), ICD-10-CM work status codes (10% complete). This is partly because acute care hospital record data is primarily used for billing purposes and external cause codes are not “reimbursable”. |
| Acceptability | All MA acute care hospitals submit inpatient hospital discharge, outpatient observation stay, and emergency department discharge data to the Center for Health Information and Analysis (CHIA) in MA. All three types of data have been collected since 2012. These data are primarily for administrative billing purposes. |  |
| Timeliness | Hospitals submit records for all patients to CHIA on a quarterly basis. They are required to submit records within 75 days after the end of each quarter. | After receiving the acute care hospital records each quarter, CHIA conducts validation checks, data cleaning, and deduplication. Final data files are made available to the MA DPH approximately 12 to 18 months after the end of the federal fiscal year. |
| Stability | Acute care hospital record data are stable in that all three types of data have been collected from MA hospitals since 2012. CHIA has a standardized process for validating and cleaning the data. CHIA also has an application process by which government agencies and organizations can request the data. Due to HIPAA restrictions, only MA DPH can get acute care hospital record data with personally identifying information, with proper justification for public health purposes. |  |
| Flexibility |  | Changes in submission criteria for the acute care hospital data are made by the Center for Health Information and Analysis in MA. There is a process for end users to suggest changes in submission requirements, but this is lengthy, and suggested changes are not always approved. |

### ***MA Trauma Registry Data***

| **Surveillance System Attributes** | **Strengths** | **Limitations** |
| --- | --- | --- |
| Sensitivity | Trauma Registry cases are most likely to capture moderate to severe MVC injury cases. | Trauma Registry Acute data are less likely to capture fatal work-related MVC injuries, as deaths on the scene may have been transported to a morgue rather than a hospital. Trauma Registry data only identified 12.5% (n = 2) of the 16 FACE fatalities in 2019.  Trauma Registry data also would not include work-related cases involving minor injuries that did not require a hospital stay or transfer of the patient to another hospital.  Work-related crashes involving passenger cars or light trucks are also less likely to be captured by Trauma Registry data, as ICD-10-CM cause of injury codes focus on crashes involving specific types of commercial vehicles.  In 2019 MA Trauma Registry data, the primary method of payment variable did not include workers’ compensation as a value. |
| Representativeness | Trauma Registry data are most likely to be representative of moderate to severe MVC injury cases involving specific types of commercial vehicles that are covered in ICD-10-CM external cause codes. | Workers who died at the scene would not show up in the Trauma Registry as they would have been transported to a morgue.  Trauma Registry data also would not include work-related cases involving minor injuries that did not require a hospital stay or transfer of the patient to another hospital.  Work-related crashes involving passenger cars or light trucks are also less likely to be captured by Trauma Registry data, as the ICD-10-CM cause of injury codes focus on crashes involving specific types of commercial vehicles.  Gig workers could not be identified due to the absence of narratives or other relevant variables. |
| Predictive Value Positive (PVP) | Work-related cases identified in Trauma Registry data are more likely to be true positives due to the specific nature of most indicators.  The 2 work-related MVC fatalities identified in 2019 Trauma Registry data were both verified as fatal work-related cases by OHSP (as FACE cases). |  |
| Simplicity | Trauma Registry data includes a binary variable indicating whether the illness/injury was work-related.  Work-related indicators based on ICD-10-CM external cause, place of occurrence, activity, and external cause (work) status codes were relatively simple to identify. The 7th character of ICD-10-CM codes can also be used to distinguish treatment of acute injuries from routine follow-up treatment or treatment of injury sequelae. | Trauma Registry data may include more than one record for each person-incident if the person received treatment for the injury in more than one hospital or was treated at the same hospital multiple times within two weeks after the incident. Therefore, the data needs to be deduplicated to the person-incident level. This may be difficult if personal identifiers and the incident date are not available. |
| Data Quality | The following relevant variables in MA Trauma Registry data were mostly complete: ICD-10-CM primary external cause code (100%), ICD-10-CM place of occurrence external cause code (85%), and work-related injury code (82%). | The following ICD-10-CM variables in Trauma Registry data were mostly incomplete or unknown: place of occurrence, activity code, (work) status code, and type of examination code.  Although not used as identifiers for work-related MVC injury cases, variables indicating the patient’s occupational industry and occupation were mostly unknown or incomplete (about 1% each complete). |
| Acceptability | Trauma Registry data has been collected by the MA DPH Bureau of Health Care Safety and Quality since 2009. In earlier years, only designated Trauma Centers submitted data, but in recent years all MA acute care hospitals submitted data on patients that meet the trauma criteria. Non-trauma centers submit fewer variables than Trauma Centers, however. |  |
| Timeliness |  | Although hospitals must submit Trauma Registry data to MA DPH quarterly, it may take two or more years for the MA DPH Bureau of Health Care Safety and Quality to validate, clean, deduplicate, and finalize Trauma Registry data for use by other programs and organizations. |
| Stability | Trauma Registry data are stable in that these data have been collected from MA hospitals since 2009. The MA DPH Bureau of Health Care Safety and Quality implemented new data management systems in the past five years to improve the data collection and validation processes. Trauma Registry data are available to other programs and organizations upon request when the necessary data use agreements are in place. |  |
| Flexibility |  | Changes in variables or values are based on revisions in National Trauma Data Bank criteria. |

1. The National Institute for Occupational Safety and Health (NIOSH) (2022, Aug 30). *Motor Vehicle Safety At Work: Burden, Need, and Impact*. Centers for Disease Control and Prevention (CDC). <https://www.cdc.gov/niosh/motorvehicle/ncmvs/BNI.html> [↑](#footnote-ref-2)
2. U.S. Department of Labor: Bureau of Labor Statistics (BLS) (2023, Nov 8). *Employer-Reported Workplace Injuries and Illnesses – 2021-2022*. <https://www.bls.gov/news.release/pdf/osh.pdf> [↑](#footnote-ref-3)
3. Typically defined as people who work and get paid by the task/job rather than by the hour/year. [↑](#footnote-ref-4)
4. Motorists include car and truck occupants (drivers and passengers) and motorcyclists. Non-motorists include cyclists and pedestrians [↑](#footnote-ref-5)
5. Traffic crashes occur on public roadways. Non-traffic crashes occur off public roadways, which include driveways, private properties, parking lots, etc. [↑](#footnote-ref-6)
6. The ICD-10-CM code for ambulances also applied to fire engines, prompting us to include all occupants in emergency vehicles. [↑](#footnote-ref-7)
7. Centers for Medicare & Medicaid Services (CMS) (2024, Feb 1). *ICD-10-CM TABULAR LIST of DISEASES and INJURIES*. <https://www.cms.gov/medicare/coding-billing/icd-10-codes/2024-icd-10-cm> [↑](#footnote-ref-8)
8. National Highway Traffic Safety Administration (NHTSA) (2024, Jan). *MMUCC guideline: Model minimum uniform crash criteria, 6th edition* (Report No. DOT HS 813 525a). <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813525> [↑](#footnote-ref-9)
9. Bush, A. M., Bunn, T. L., & Liford, M. (2021). Identification of work-related injury emergency department visits using International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) codes. *Injury Prevention, 27*(Suppl 1), i3–i8. <https://doi.org/10.1136/injuryprev-2019-043507> [↑](#footnote-ref-10)
10. FACE (Fatality Assessment and Control Evaluation) is the NIOSH cooperative agreement under which MA DPH monitors and documents workplace deaths. [↑](#footnote-ref-11)
11. The “number for the motor vehicle occupant in the motor vehicle they occupied, or for each non-motorist, in consecutive order” – National Highway Traffic Safety Administration (NHTSA) (2024, Jan). *MMUCC guideline: Model minimum uniform crash criteria, 6th edition* (Report No. DOT HS 813 525a). <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813525> [↑](#footnote-ref-12)
12. No results fell within this range of completion. [↑](#footnote-ref-13)
13. MA CRISS only had 2020 MATRIS data available for use at the time of this analysis. [↑](#footnote-ref-14)
14. Friedland, J., & Balkin, D. B. (2023). When gig workers become essential: Leveraging customer moral self-awareness beyond COVID-19. *Business horizons, 66*(2), 181–190. <https://doi.org/10.1016/j.bushor.2022.05.003> [↑](#footnote-ref-15)
15. Garin, A., Jackson, E., Koustas, D. K., & Miller, A. (2023). The Evolution of Platform Gig Work, 2012-2021. *National Bureau of Economic Research (NBER)* Working Paper No. 31273. Available at <https://www.nber.org/papers/w31273> [↑](#footnote-ref-16)
16. Gig workers could only be identified in narratives in police crash reports and MATRIS data. [↑](#footnote-ref-17)
17. We included keywords relating to Grubhub, Instacart, Postmates, Uber Eats, and DHL delivery in our text search, but they did not capture any cases in any data source. Therefore, they are omitted from this table. [↑](#footnote-ref-18)
18. At this time, Amazon and DHL were using some gig workers. A manual review of package delivery cases revealed they did not appear to be companies that used employees only (like FedEx, UPS, or USPS), so we classified them as gig workers. [↑](#footnote-ref-19)
19. These additional text searches were used with MATRIS narratives given the limited number of other work indicators in MATRIS data. [↑](#footnote-ref-20)
20. We chose to specify “driver” as the inclusion of “occupant” or “passenger” led to too many false positives. [↑](#footnote-ref-21)
21. FACE (Fatality Assessment and Control Evaluation) is the NIOSH cooperative agreement under which MA DPH monitors and documents workplace deaths. [↑](#footnote-ref-22)
22. Persons transported by the medical examiner to the morgue or transported to a medical facility without an ambulance would not be included, but persons transported to the hospital in an ambulance who subsequently died would be included in MATRIS data. [↑](#footnote-ref-23)
23. Includes persons who are dead on arrival or who die after being admitted to the Emergency Department/hospital. [↑](#footnote-ref-24)
24. Includes persons who are dead on arrival or who die after being admitted to the Emergency Department/hospital. [↑](#footnote-ref-25)
25. German, R. R., Lee, L. M., Horan, J. M., Milstein, R. L., Pertowski, C. A., Waller, M. N., Birkhead, G. S., & the Guidelines Working Group (2001, Jul 27). Updated Guidelines for Evaluating Public Health Surveillance Systems. *CDC MMWR, 50*(RR13), 1-35. <https://www.cdc.gov/mmwr/preview/mmwrhtml/rr5013a1.htm> [↑](#footnote-ref-26)
26. Gunter, M. M. (2016, Sep) An update on SOII undercount research activities. *Monthly Labor Review*, U.S. Bureau of Labor Statistics. <https://www.bls.gov/opub/mlr/2016/article/an-update-on-soii-undercount-research-activities.htm> [↑](#footnote-ref-27)
27. Bush, A. M., Bunn, T. L., & Liford, M. (2021). Identification of work-related injury emergency department visits using International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) codes. *Injury Prevention, 27*(Suppl 1), i3–i8. <https://doi.org/10.1136/injuryprev-2019-043507> [↑](#footnote-ref-28)
28. Keywords and ICD-10-CM codes were included and excluded using the prxmatch function in SAS Studio. You can request a copy of each algorithm’s SAS program. [↑](#footnote-ref-29)
29. You may need to modify keywords based on the values in your data. Although not the intended use of this field, police often documented a company name rather than an individual’s name here. [↑](#footnote-ref-30)
30. Company names in the first name fields often had the last name blank or “Unknown”. The last name “Owner” was also used to indicate the company in the first name field as the vehicle owner. [↑](#footnote-ref-31)
31. You may need to modify keywords based on the values in your data. Although not the intended use of this field, police often documented a company name rather than an individual’s name here. [↑](#footnote-ref-32)
32. Company names in the last name fields often had the first name blank or “Unknown”. The first name “Owner” was also used to indicate the company in the last name field as the vehicle owner. [↑](#footnote-ref-33)
33. You may need to modify keywords based on the values in your data. [↑](#footnote-ref-34)
34. Narratives provide crash event-level data rather than person-level data. Therefore, it is important to verify which person(s) injured in the crash was/were working. [↑](#footnote-ref-35)
35. You may need to modify keywords based on the values in your data. [↑](#footnote-ref-36)
36. Variable names were truncated in the MATRIS data files that were made available to the Injury Surveillance Program. [↑](#footnote-ref-37)
37. Keywords and ICD-10-CM codes were included and excluded using the prxmatch function in SAS Studio. You can contact the authors to request a copy of the SAS program used with each data source to identify work-related MVC injury cases. [↑](#footnote-ref-38)
38. You may need to modify keywords based on the values in your data. [↑](#footnote-ref-39)
39. The variables injecode1-injecode15, injplace1-injplace3, injact1-injact3, and injstatus were created for the Injury Surveillance Program by the MA DPH Office of Data Management and Outcomes Assessment (ODMOA) from the principal external cause code and associated ICD-10-CM diagnosis code fields provided in acute care hospital records. [↑](#footnote-ref-40)
40. Keywords and ICD-10-CM codes were included and excluded using the prxmatch function in SAS Studio. You can contact the authors to request a copy of the SAS program used with each data source to identify work-related MVC injury cases. [↑](#footnote-ref-41)
41. Keywords and ICD-10-CM codes were included and excluded using the prxmatch function in SAS Studio. You can request a copy of each algorithm’s SAS program. [↑](#footnote-ref-42)
42. It was not clear why there were four variable groups that included ICD-10-CM codes in the Trauma Registry data the Injury Surveillance Program received from the Bureau of Health Care Safety and Quality (BHCSQ). Some of the fields were created by a BHCSQ staff person who no longer works with MA DPH. [↑](#footnote-ref-43)
43. German, R. R., Lee, L. M., Horan, J. M., Milstein, R. L., Pertowski, C. A., Waller, M. N., Birkhead, G. S., & the Guidelines Working Group (2001, Jul 27). Updated Guidelines for Evaluating Public Health Surveillance Systems. *CDC MMWR, 50*(RR13), 1-35. <https://www.cdc.gov/mmwr/preview/mmwrhtml/rr5013a1.htm> [↑](#footnote-ref-44)
44. Massachusetts Department of Transportation (MassDOT) (2025). *IMPACT*. [Crash Data Dashboard]. Mass.gov. <https://apps.impact.dot.state.ma.us/cdp/home> [↑](#footnote-ref-45)