Journal of Safety Research 86 (2023) 21–29

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Using motor vehicle crash records for injury surveillance and research in agriculture and forestry

Eva M. Shipp [a,1,*], Amber B. Trueblood [b,2], Hye-Chung Kum [c,3], Marcie Perez [a,4], Shubhangi Vasudeo [a,5], Nishita Sinha [a,6], Ashesh Pant [a,7], Lingtao Wu [a,8], Myunghoon Ko [a,9]

[a] Texas A&M Transportation Institute, Center for Transportation Safety, United States
[b] CPWR-The Center for Construction Research and Training, United States
[c] Texas A&M School of Public Health, Population Informatics Lab, United States

ARTICLE INFO

Article history:
Received 15 November 2022
Received in revised form 29 March 2023
Accepted 5 June 2023
Available online 14 June 2023

Keywords:
"Agriculture
"Logging"
"Occupational Injuries" or "Work-related injuries"
"Motor vehicle crashes"
"Surveillance"

ABSTRACT

Problem: Fatal injuries in the agriculture, forestry, and fishing sector (AgFF) outweigh those across all sectors in the United States. Transportation-related injuries are among the top contributors to these fatal events. However, traditional occupational injury surveillance systems may not completely capture crashes involving farm vehicles and logging trucks, specifically nonfatal events. *Methods:* The study aimed to develop an integrated database of AgFF-related motor-vehicle crashes for the southwest (Arkansas, Louisiana, New Mexico, Oklahoma, and Texas) and to use these data to conduct surveillance and research. Lessons learned during the pursuit of these aims were cataloged. Activities centered around the conduct of traditional statistical and geospatial analyses of structured data fields and natural language processing of free-text crash narratives. *Results:* The structured crash data in each state include fields that allowed farm vehicles or equipment and logging trucks to be identified. The variable definitions and coding were not consistent across states but could be harmonized. All states recorded data fields pertaining to person, vehicle, and crash/environmental factors. Structured data supported the construction of crash severity models and geospatial analyses. Law enforcement provided additional details on crash causation in free-text narratives. Crash narratives contained sufficient text to support viable machine learning models for farm vehicle or equipment crashes, but not for logging truck narratives. *Discussion:* Crash records can help to fill research and surveillance gaps in AgFF in the southwest region. This supports traffic safety's evolution to the current Safe System paradigm. There is a conceptual linkage between the Safe System and Total Worker Health approaches, providing a bridge between traffic safety and occupational health. *Practical Applications:* Despite limitations, crash records can be an important component of injury surveillance for events involving AgFF vehicles. They also can be used to inform the selection and evaluation of traffic countermeasures and behavioral interventions.

* Corresponding author.
E-mail addresses: e-shipp@tti.tamu.edu (E.M. Shipp), atrueblood@cpwr.com (A.B. Trueblood), kum@tti.tamu.edu (H.-C. Kum), m-perez@tti.tamu.edu (M. Perez), shubhangi.s.maindale@gmail.com (S. Vasudeo), n-sinha@tti.tamu.edu (N. Sinha), a-pant@tti.tamu.edu (A. Pant), l-wu@tti.tamu.edu (L. Wu), m-ko@tti.tamu.edu (M. Ko).
[1] Senior Research Scientist, Center for Transportation Safety, Texas A&M Transportation Institute, 1111 RELLIS Parkway Bryan, TX 77807.
[2] Director, Data Center, CPWR- The Center for Construction Research and Training, 8484 Georgia Ave #1000, Silver Spring, MD 20910.
[3] Professor & Presidential Impact Fellow, Director Population Informatics Lab, Department of Health Policy and Management, Texas A&M School of Public Health.
[4] Center for Transportation Safety, Texas A&M Transportation Institute, 1111 RELLIS Parkway Bryan, TX 77807.
[5] Consultant, Center for Transportation Safety, Texas A&M Transportation Institute, 1111 RELLIS Parkway Bryan, TX 77807.
[6] Assistant Research Scientist, Center for Transportation Safety, Texas A&M Transportation Institute, 1111 RELLIS Parkway Bryan, TX 77807.
[7] Lead Data Analyst, Center for Transportation Safety, Texas A&M Transportation Institute, 1111 RELLIS Parkway Bryan, TX 77807.
[8] Associate Research Engineer, Center for Transportation Safety, Texas A&M Transportation Institute, 1111 RELLIS Parkway Bryan, TX 77807.
[9] Research Engineer, Center for Transportation Safety, Texas A&M Transportation Institute, 1111 RELLIS Parkway Bryan, TX 77807.

https://doi.org/10.1016/j.jsr.2023.06.004
0022-4375/

1. Introduction

Fatal occupational injury rates in the Agriculture, Forestry, Fishing and Hunting (AgFF) sector are among the highest in the United States. Based on Census for Fatal Occupational Injury (CFOI) data, the fatal injury rate in AgFF was 23.1 per 100,000 full time equivalents (FTE) versus a rate of 3.5 per 100,000 FTEs for all workers in 2019 (BLS, 2020a). Transportation events including motor vehicle crashes contribute heavily to this overrepresentation (CDC, 2011; NIOSH, 2016a; Perritt et al., 2017). In 2018, transportation events accounted for 47.7% of all fatalities in AgFF (BLS, 2019). In terms of the number of workers at risk, there are an estimated 2.4 million salary and wage, self-employed, and unpaid workers in agriculture and related industries (BLS, 2020b). Motor-vehicle and transportation events also burden youth in AgFF. Among youth under 18 years who were fatally injured in the agriculture production industry from 1994 to 2013, the most frequent type of fatal injury event was transportation (>60%) and the most frequent injury source was vehicle (>50%) (Perritt et al., 2017). With respect to nonfatal injuries, motor vehicles were a leading injury source for AgFF youth working on their family farm (NIOSH, 2016a). Five states comprise the southwest region, the focus of the National Institute for Occupational Safety and Health (NIOSH)-sponsored Southwest Center for Agricultural Health, Injury Prevention, and Education at The University of Texas at Tyler, including Arkansas, Louisiana, New Mexico, Oklahoma, and Texas. Historically, the rate of occupational highway transportation fatalities in the southwest region also exceeds national rates based on an in-depth analysis of occupational highway transportation fatalities (CDC, 2011).

Regarding common types of motor vehicles and equipment, there are nearly 3.5 million trucks including pickups, over 4 million tractors, and over 400,000 self-propelled pieces of equipment (e.g., grain and bean combines, cotton pickers and strippers, and forage harvesters) being operated on farms in the United States (USDA NASS, 2019). Approximately 14.7% of the nation's farm trucks, tractors, and self-propelled vehicles are in the five-state southwest region (USDA NASS, 2019). Farm vehicles have an increased crash risk since they are slow moving, difficult for other drivers to anticipate with respect to speed differentials and how they use the roadway compared to other roadway users, often lack seatbelts, and travel in rural areas, where high speeds prevail (Aarts & Schagen, 2006). The width of some of these vehicles and equipment also extend beyond the lane width, which may make it difficult for AgFF drivers to pull over onto a shoulder to allow another vehicle or roadway user to safely pass (TDI, 2019).

In terms of logging trucks, they are the primary transport for forestry products and account for more than half of logging business expenses (Roberts et al., 2005; Shaffer & Stuart, 2005). Illustrating this point, over 473 million tons of logs and other wood in the rough were moved by trucks in 2017 (FHWA, 2021). Crashes account for almost one-third of occupational fatalities in the logging industry in the United States (BLS, 2019a). Due to their increased risk for occupational injury, truck drivers are a priority population for the NIOSH Center for Motor Vehicle Safety (NIOSH, 2020a). Nationally, logging crashes have increased since 2012 and are expected to continue along this trajectory (Baker & Tyson, 2017; Cole, 2018; Conrad, 2018). The trend is a concern given the shortage of truck drivers in logging, along with an aging workforce (Costello & Karickhoff, 2019).

In 2009, the USDA reported that "*Many details of public road crashes involving agricultural machinery and motor vehicles are unknown or lack sufficient detail to aid prevention efforts*" (CASHRE, 2009). Over a decade ago, this statement remains relevant despite growth in the area of research of AgFF crashes (Bunn et al., 2008; Costello et al., 2003; Costello et al., 2009; Gerberich et al., 1996; Gkritza et al., 2010; Glascock et al., 1995; Harland et al., 2014; Karimi & Faghri, 2021; Lacy et al., 2003; Luginbuhl et al., 2003; Mehlhorn et al., 2015; Peek-Asa et al., 2007; Ramirez et al., 2016; Ranapurwala et al., 2016, 2019; Scott et al., 2020; Shipp et al., 2019; Toussaint et al., 2017). Research gaps are difficult to address partially because current occupational surveillance systems are not fully capturing AgFF crashes and injuries (Leigh et al., 2014).

In the past, NIOSH and other agencies sponsored national agricultural injury surveys to help supplement more traditional surveillance systems such as the Survey of Occupational Injuries and Illness (SOII; BLS, 2019b). However, NIOSH stopped supporting these activities as of 2015 (NIOSH, 2020b). In 2012, an independent panel reviewed the NIOSH AgFF program and recommended actions to improve surveillance (NIOSH, 2012). Subsequently, the RAND Corporation assessed the recommendations and released their findings in 2017, which included listing the need to "Identify and evaluate the potential of existing data sources for illness and injury surveillance of agricultural workers" (Chari et al., 2017). The National Academies of Science, Engineering, and Medicine (2018) released a related report with a recommendation focused on the creation of a "…coordinated system for surveillance of both fatal and nonfatal occupational disease using multiple data sources" (NAS, 2018). The primary goal of the present study was to address research and surveillance gaps in AgFF by evaluating how crash records can be used to capture events occurring on public roadways that involve farm vehicles and equipment or logging trucks in the southwest region.

2. Methods

The primary data sources for this assessment were: (1) state-based crash records for Arkansas, Louisiana, New Mexico, Oklahoma, and Texas and (2) fatal crash records from the National Highway Traffic Safety Administration's (NHTSA) Fatality Analysis Reporting System (FARS) (NHTSA, 2022). Core analytical methods include traditional descriptive and inferential statistics followed by natural language processing, a machine learning approach, to assess the utility of the crash narrative in identifying and understanding AgFF crashes.

2.1. Primary data sources

For this study, we obtained a complete census of crash records for the southwest region. Each of these states maintain crash records as recorded by law enforcement officers and use electronic systems to collect and house these data. Although, some states may still allow limited paper-submission of reports, especially in rural areas or small jurisdictions with limited resources. In general, the states lag one to two years in releasing their annual crash records. Some of these states readily release public files, while others require data use agreements even for deidentified records. Across the United States, some states will release files with identifiers for research purposes given appropriate data use agreements and other stipulations (e.g., receipt of institutional review board approval).

There is no standardized system that each state uses to record their crash data, although it is common for states to store data elements or variables within files labeled as crash (e.g., crash configuration, weather and surface conditions, posted speed limits), the vehicle or unit involved (e.g., vehicle body style, vehicle age, crash contributing factors), and the person involved including the driver (e.g., sex, age, restraint use), vehicle occupants and others (e.g., pedestrians, bicyclists), at a minimum. Crash contributing factors are key for identifying behavior-based countermeasures. Examples of contributing factors include impairment by alcohol or drugs, speeding behaviors, distraction, and fatigue.

Once data are collected, states usually perform quality control checks and may recode data to improve its accuracy. In Texas, for example, each report is manually reviewed for completeness and additional interpreted data fields are added based on the reported information. Examples of interpreted fields include manner of collision and first harmful event. In addition, the Texas Department of Transportation assigns roadway characteristics based on the entered crash location (i.e., latitude and longitude) using their roadway inventory data. This also allows for various geospatial information system (GIS) tools to be used during analysis along with the inclusion of engineering-based variables into the crash report that are not readily knowable by law enforcement on the scene of the crash. These roadway characteristics can influence driver behavior and safety (e.g., number of lanes, if and how lanes going opposite directions are divided, roadway and/or shoulder width, intersection relationship, traffic control type). Roadway characteristics are common factors examined during engineering-based safety assessments. In addition to the structured data elements, the state crash reports include a free-text narrative field where law enforcement can record additional details regarding the crash and its causation. There also is a location for the entry of a crash diagram. Although both the narrative and the diagram offer qualitative details that can provide rich details about the crash and its outcomes, they are not readily coded or analyzable beyond manual assessment by a human.

States update their crash report or reporting system, but this does not occur at regular or necessarily predictable intervals. Major changes to the crash report, beyond minor modifications to the business rules or the addition of a category to a variable (e.g., finer distinction of the form of cell phone use such as talking hands-free or texting to a driver distraction variable) can be very expensive and may also require training law enforcement.

In addition to maintaining their own crash data systems, states submit their data on a regular basis to the Fatality Analysis Reporting System (FARS). FARS is a national surveillance system that originated in 1975 with the most recently available public data typically lagging one to two years behind the current year. It covers all 50 states, the District of Columbia and Puerto Rico. FARS includes all crashes involving a motor vehicle traveling on a public trafficway that results in a death of a vehicle occupant or non-occupant within 30 days of the crash (NHTSA, 2022). The National Center for Statistics and Analysis (NCSA), under the National Highway Traffic Safety Administration (NHTSA), maintains FARS.

Approximately 140 data elements are collected by FARS. The requested data elements are standardized with the coding rubric published annually along with any updates or changes in the coding or definitions. Overall, the data elements are similar or the same as the information collected by each state, although it may be formatted differently (e.g., larger number of more specific categories). Like state-based records, FARS organizes files and variables at the crash, vehicle, and person level as well as additional files with more detailed information on subtopics (e.g., obstruction of driver vision, in-depth data on distraction).

States populate FARS fields through abstracting data from various sources including police state crash reports, death certificates, state vehicle registration files, coroner/medical examiner reports, state driver licensing files, state highway department data, emergency medical service reports, vital statistics, and other state records. The use of multiple data sources improves data quality and completeness compared to using state crash records alone. Other strengths of FARS are that detailed information is available at the national level, and the records are deidentified with public use promoted. Consequently, these data are readily accessible via the NHTSA query system or through their file transfer protocol. All coding manuals are updated annually and posted to the website.

2.2. Statistical analysis

The structured data fields in the crash report formed the basis of traditional crash severity models (Lord et al., 2021). Crash severity models help to identify factors associated with higher severity versus lower severity crashes. The current project involved using multiple logistic regression and multinomial logistic regression to identify driver, vehicle, and crash variables associated with higher severity crashes in logging truck crashes. The intent is to identify factors to prioritize for countermeasure application or other forms of intervention. These analyses are published elsewhere and ongoing (Shipp et al., 2019; Vasudeo, 2022). States often provide the crash location in terms of the geospatial coordinates (i.e., latitude, longitude). These data are used to identify where crashes concentrate along the roadway system and to establish potential relationships between concentrated AgFF activity and crashes using GIS approaches.

2.3. Analysis of free text narratives

In this study we evaluated data mining methods, specifically natural language processing (NLP), for identifying AgFF crashes. The analysis included developing a well curated keyword list comparing different data representations (e.g., bag of words [BoW], bag of keywords [BoK]) and document classification algorithms (e.g., support vector machines [SVM], and multinomial naïve Bayes classifier [MNB]). We present these methods in detail elsewhere (Kim et al., 2021). Briefly, for training and testing the models, we developed a large gold standard dataset based on manual narrative review of the crashes involving farm equipment/vehicles or logging trucks. Algorithmic performance was assessed to determine how well the models trained using the gold standard from earlier periods predicted later periods of data.

3. Results

3.1. Structured data fields

Table 1 displays the data fields used to identify AgFF crashes. There was some variation with respect to the exact variable definition and coding. Although used in agriculture, all-terrain vehicles (ATVs) were not included in the study definitions since ATVs as a vehicle could not be isolated in the structured data for the southwest region.

States release a data dictionary with the variable names and coding schemes. In addition, states typically release manuals used to train law enforcement officers in the completion of crash report forms. These manuals provide additional information useful for understanding nuances in how each state identifies relevant vehicles and equipment. The reportable crash definition also varies by state as shown in Table 1. Table 2 provides a count of records identified as AgFF from the initial project period. Although the proportion of total crashes that involve farm vehicles/equipment or logging trucks is low, the raw number of crashes remains notable in the southwest region.

States use the KABCO scale to rate the injury severity of each crash. The KABCO scale includes five categories: fatal (K), suspected serious injury (A), suspected minor injury (B), possible injury (C), and property damage only (O). Each category is described further in Table 3 (NHTSA, 2017). The law enforcement officer records the injury severity for all individuals involved in the crash, including pedestrians and bicyclists. Then, the overall

Table 1
Variables used to identify AgFF crashes in Arkansas, Louisiana, New Mexico, Oklahoma, and Texas.

State	Reportable Crash Criteria	Farm Vehicles/Equipment	Logging Trucks	ATVs
Arkansas[a]	$1,000 in property damage, injury or death	Vehicle Type = 21 (farm equipment other than trucks)	Cargo Body Type = 7 (log-since 2015)	Vehicle Type = 33 (ATV)
Louisiana[b]	$500 in property damage, injury or death	Vehicle Configuration = T (farm equipment)	Cargo Body Type = H (log truck / trailer)	Not applicable
New Mexico[c]	$500 in property damage, injury or death	Vehicle Use 1 = FV (farm vehicle / equipment)	Vehicle Cargo Body = LT (log truck)	Not applicable
Oklahoma[d]	$500 in property damage, injury or death	Vehicle Configuration = 18 (farm machinery) Towed Vehicle Type = 3 (farm trailer)	Cargo Body Type = 13 (log trailer)	Vehicle Configuration = 19
Texas[e]	$1,000 in property damage, injury or death	Vehicle Body Style = 101 / FE (farm equipment)	Cargo Body Style = 14 (log trailer)	Not applicable
FARS[f]	Death occurs within 720 hours of a motor vehicle traffic crash and is a direct result of the crash	Vehicle Body Type = 92 (farm equipment other than trucks)	vehicle with cargo_bt = 10 (log-since 2007) vin_body_type = LG (truck logger)	Vehicle Body Type = 90 (ATV)

[a] Arkansas State Police (ASP) (2015).
[b] *(LADPSC/LHSC, 2019).*
[c] *Trailer Type 1 may also be relevant in subsequent years (NMDOT, 2021).*
[d] (OkDPS, 2019).
[e] *(TxDOT, 2018).*
[f] *FARS: Fatality Analysis Reporting System (NHTSA, 2022).*
[g] Updates to each state's data systems may cause these to change from 2020 forward.

Table 2
Frequency of farm vehicle/equipment and logging truck crashes in the southwest region during the initial project phase.

State	Years Structured Data Fields	Farm Equipment / Vehicle Crashes	Logging Truck Crashes	Total Records	Narratives
Arkansas[a]	2010–2015	285	94	366,490	8,834
Louisiana[b]	2010–2018	1,437	1,428	2,225,102	1,005,959
New Mexico	2008–2018	347	117	482,680	Not applicable
Oklahoma	2010–2018	656	144	634,831	Not applicable
Texas[c]	2010–2020	3,152	1,115	5,314,983	9,974

[a] *Logging truck crashes only available from 2015 forward for Arkansas.*
[b] *Total records cover 2005–2018.*
[c] *Texas narratives were sampled from 2010-2018.*

crash severity is coded as the highest injury severity recorded across all involved individuals. For example, if one person was fatality injured, but the other three occupants sustained suspected serious injuries, the crash severity would be coded as fatal. The injury and crash severity assessment is limited to the law enforcement officer's observation at the scene unless the person suffered a fatal injury. Fatalities occurring up to 30 days post-crash are captured.

The remaining structured data elements are not all categorized or coded the same for the five states in the southwest region. To create a regional surveillance system, key variables required harmonization. The Model Minimum Uniform Crash Criteria (MMUCC) formed the basis of data element harmonization. The National Highway Traffic Safety Administration (NHTSA) in collaboration with the Governors Highway Safety Association (GHSA), developed the MMUCC to encourage states to adopt a minimum set of crash data elements and attributes. The MMUCC originated in 1998 and is currently in its 5th edition (NHTSA, 2017).

Table 4 includes general variables used to describe crashes in the southwest region and in the analysis to identify variables associated with higher severity crashes.

3.2. Crash narratives

Our assessment of the utility of crash narratives to inform surveillance in AgFF yielded three key findings. First, the narratives from Texas and Louisiana, overall, provided a sufficient amount of text in individual narratives to support training viable machine learning models. Second, narratives describing crashes involving logging trucks did not contain enough text, even individual words, to support a similar analysis. In many cases, the narrative was blank or included only a sentence or two with little to no meaningful information (e.g., vehicle 1 and vehicle 2 crashed). This was unlike the average amount of text in narratives overall for the farm vehicle/equipment crashes. Third, crashes can involve farm equipment or vehicles that were not involved in AgFF activities at the time of the crash. The most prevalent example was using tractors with cutters to mow roadway medians or shoulders. Complex machine learning algorithms are not required to identify these events. Simple key word searches can be used to identify these crashes while reducing the manpower needed to review narratives (Trueblood et al., 2019).

In addition, we assessed the availability of crash narratives. Currently, not all states store their crash narratives as searchable electronic free text. Rather, the narratives may be stored as an image or PDF (e.g., Texas). To address these issues, we developed a process for scraping the text from the PDF and storing the data as free text in spreadsheets. Then the narratives went through a manual de-identification process. Since the review required considerable personnel time, we reviewed all the farm vehicle and equipment, and logging truck narratives from Texas and only a sample of other types of crashes rather than a complete census.

4. Discussion

Overall, we found that crash records can help to fill research and surveillance gaps in AgFF in the southwest region. A major strength of crash records is that they are not protected health information. Consequently, they can be easier to access even when the records

Table 3
Definition of injury severity[a] used in transportation safety.

Code	Definition	Examples
	Fatal injury	"A fatal injury is any injury that results in death within 30 days after the motorvehicle crash in which the injury occurred."
A	Suspected serious injury (formerly incapacitating)	"Severe laceration resulting in exposure of underlying tissues/muscle/organs or resulting in significant loss of blood; broken or distorted extremity (arm or leg); crush injuries; suspected skull, chest or abdominal injury other than bruises or minor lacerations; significant burns (second and third degree burns over 10% or more of the body); unconsciousness when taken from the crash scene; paralysis
B	Suspected minor injury (formerly non-incapacitating)	"...lump on the head, abrasions, bruises, minor lacerations (cuts on the skin surface with minimal bleeding and no exposure of deeper tissue/muscle)"
C	Possible injury	"...momentary loss of consciousness, claim of injury, limping, or complaint of pain or nausea. Possible injuries are those that are reported by the person or are indicated by his/her behavior, but no wounds or injuries are readily evident'
O	Property damage only	"A crash that results in damage to the motor vehicle or other property, but without injury to any occupants or non-motorists/"
Not Not	applicable applicable	Unknown severity

[a] MMUCC 5th Edition (NHTSA, 2017).

contain personally identifiable information. This is a strength in terms of timely monitoring of health-related events.

Traditional occupational surveillance systems do not fully capture nonfatal injuries or the impact of AgFF crashes on the transportation system or other roadway users. Studies suggest that national surveillance systems may be missing over 75% of nonfatal occupational injuries and illnesses in agriculture due to small farm reporting exemptions along with other issues including not taking time off work even when injured or ill (BLS, 2019b; Leigh et al., 2014; Scott et al., 2020). During a limited comparison of Texas crash records from 2016 to 2020 to the BLS Survey of Occupational Injury and Illnesses, no injuries were reported in standard BLS profiles, possibly due to so few cases being captured (BLS, 2023). In contrast, there were 48 AgFF individuals and 81 non-AgFF individuals who sustained a suspected serious injury due to a crash with a farm vehicle or logging truck. The number of individuals with a suspected minor injury or possible injury included 203 AgFF and 623 non-AgFF individuals, respectively (TxDOT, 2023). Crash records and other non-traditional AgFF data sources such as AgInjuryNews, can shed light on these broader impacts since they collect information on all of those injured in a crash (Weichelt et al., 2023).

Crashes also negatively impact AgFF individuals even when there is only property damage. This "minor" event can still be very costly. The average estimated cost of a crash even when no injury is observed in 2020 was $12,800. This estimate considered "...wage and productivity losses, medical expenses, administrative expenses, motor-vehicle damage, and employers' insurance costs." (NSC, 2022). In Texas in 2020, there were 320 property damage only AgFF vehicle-involved crashes (TxDOT, 2023). Even with no resulting injuries, the estimated costs of these events was $4,096,000.

Table 4
Common data elements used to describe crashes in transportation safety.

Data element or Variable	MMUCC Data Category
Crash severity	Crash
Date	Crash
Time	Crash
Day of week	Crash
Location (GIS coordinates)	Crash
First harmful event	Crash
Manner of crash / collision impact	Crash
Weather conditions	Crash
Lighting conditions	Crash
Roadway surface condition	Crash
Type of intersection	Crash
Number of motor vehicles involved (Single versus multi-vehicle involvement)	Crash
Number of non-motorists	Crash
Roadway curvature	Roadway
Roadway functional class	Roadway
Annual average daily traffic	Roadway
Width of lanes, shoulders, medians	Roadway
Roadway lighting	Roadway
Pavement markings	Roadway
Vehicle body type	Vehicle/Unit
Motor Vehicle Posted/Statutory Speed Limit	Vehicle/Unit
Trafficway description	Vehicle/Unit
Direction of travel	Vehicle/Unit
Total lanes in roadway	Vehicle/Unit
Roadway alignment and grade	Vehicle/Unit
Traffic control device type	Vehicle/Unit
Motor vehicle maneuver/action	Vehicle/Unit
Vehicle damage	Vehicle/Unit
Most harmful event for this vehicle	Vehicle/Unit
Contributing circumstances (i.e., factors)[b]	Vehicle/Unit
Sequence of events	Vehicle/Unit
Vehicle configuration	Large Vehicle/ Hazardous Material
Cargo body type	Large Vehicle/ Hazardous Material
Motor vehicle automation	Dynamic
Age (calculated from date of birth)	Person
Sex	Person
Person type	Person
Injury status	Person
Seating position	Person
Restraint system/motorcycle helmet use	Person
Ejection	Person
Speeding-related	Person
Distraction / distracted by	Person
Law enforcement suspects alcohol/drugs	Person
Alcohol or drug test result	Person
Driver actions at time of crash	Person
Driver license status/restrictions	Person

[a]MMUCC 5th Edition (NHTSA, 2017).
[b] States also record driver states (e.g., impairment due to alcohol/drugs or fatigue, distraction, "road rage") as contributing factors.

Collaborating with state Traffic Record Coordinating Committees (TRCC) is one mechanism for accelerating the use of crash records for AgFF surveillance. We base this observation on our experience engaging with the Texas TRCC (Texas TRCC, 2022). TRCCs are state-based partnerships comprised of representatives from state agencies with a goal of improving states' traffic records systems. The core traffic records databases include crash, citation, motor-vehicle registration, driver licensing, injury surveillance, and roadway. NHTSA encourages assessing data quality of these databases using a framework comprised of six areas: timeliness, accuracy, completeness, uniformity, integration, and accessibility (NHTSA, 2018). Although the majority of states are not integrating all of these databases yet, there are states linking one or more data sources to crash records. From 1992 to 2013, NHTSA assisted states in linking crash records to other data sources, namely medical records through the Crash Outcome Data Evaluation System (CODES) (Cook et al., 2015). Not all states participated in CODES and some states continued to support CODES after NHTSA support ended. Crash records linked with trauma registry data are beneficial for AgFF surveillance since these data can be used to identify specific injury diagnoses and severity resulting from specific types of crashes and circumstances (Curry et al., 2021; Conderino et al., 2017; NAS, 2021).

Staff from this research project became involved with the Texas TRCC and found that collaborating with TRCCs is one mechanism researchers can use to improve data quality. For example, findings from AgFF surveillance projects could engage crash records and law enforcement stakeholders to improve crash reporting through developing quick reference materials or "tip cards" for law enforcement officers. We also found that involvement with the TRCC supports educating data providers on the ways their data may be used to improve health and safety beyond their primary purpose. Finally, we found that building a relationship with the TRCC can also help with remaining current on large database or system updates that could impact the availability of data or its coding, and subsequently the identification of crashes involving AgFF road users, as well as other vulnerable road users.

With respect to narrative texts, their use to inform injury surveillance is not a new approach (Bunn et al., 2008; McKenzie et al., 2010). However, the use of machine learning is an innovation afforded by technological advances in data storage, management, and computing. Consequently, developing machine learning methods to analyze crash narratives is an area experiencing rapid growth (Das et al., 2020; Fitzpatrick et al., 2017; Kitali et al., 2021; Vallmuur, 2015; Zhang et al., 2020). Little of this research has specifically focused on AgFF crashes. However, its use in occupational health surveillance is growing. Hirabayashi and colleagues (2020) used machine learning to analyze pre-hospital care reports to support occupational surveillance. In addition, NIOSH's Industry and Occupation Computerized Coding System is based on machine learning models to code occupation and industry in mortality files (NIOSH, n.d.).

4.1. Limitations

Although crash records can help to fill surveillance gaps in AgFF in the southwest region, their use for this purpose has limitations. For example, crash records exclude events occurring off the public roadways such as motor-vehicle related injuries in fields, pastures, and private roads. This is a concern because transportation incidents off public roadways are also prevalent in AgFF (BLS, 2019a). Similarly, ATVs are commonly used in AgFF and contribute to fatalities in this sector (Helmkamp et al., 2011; Weichelt et al., 2020). However, not every state has a pre-coded category for ATVs. This limitation could be addressed by encouraging states to include a pre-coded category when updating their crash forms. With respect to occupational research, most states do not have fields on their crash reports indicating that the roadway user was "at work" when the crash occurred. To resolve this issue, the occupational nature of the crash is inferred from variables describing vehicle type, trailer type, or the use of these vehicles. An exception is FARS data that contain an "at work" field, but it is often missing for records in the southwest region.

An additional limitation is that law enforcement officers are essentially the "data collectors." However, it is not feasible for researchers and public health professionals to individually train them as they can train research staff for prospective studies. Furthermore, law enforcement officers may not be trained uniformly from state to state or even within the same state. This may be a greater issue in large states such as Texas, where the distance between jurisdictions can be hundreds of miles. Law enforcement officers also may not complete crash reports as accurately and completely as possible because they are not aware of the different uses of these data. Findings from a study of the collection and reporting of commercial motor-vehicle crash data in Texas indicated that law enforcement officers may not fully realize the extent to which crash data are used to allocate transportation safety resources including the identification and implementation of countermeasures (Trueblood & Shipp, 2019). Research also illustrates that the law enforcement officer's assessment of severity, using the KABCO scale, may not correlate strongly with the injury severity score as assessed by trauma clinicians with agreement weakening as the severity of the crash lessens (Shipp et al., 2015). Reclassifying injury severity into broader categories such as fatal, injury, and no injury categories or higher versus lower injury severity categories can help to minimize the negative impact of this issue (Farmer et al., 2003).

Finally in our study, although narratives were useful for identifying farm vehicle and equipment crashes, this was not true for logging truck crashes in the southwest region. Narrative data for logging truck crashes was often sparse or missing altogether. Reasons for the lack of narrative data describing logging truck crashes are not well understood. Researchers currently are assessing this issue by engaging key stakeholders including law enforcement officers. Given the relatively low frequency of farm vehicle/equipment and logging truck crashes versus other types of crashes in the southwest region, it is feasible to manually review the crash narratives to validate the classification of crashes as AgFF. However, not all states share crash narratives even with strict data use agreements in place. Therefore, validating the structured data with the crash narratives may not always be feasible. In the present study, it was not possible to verify the crashes coded as farm vehicle/equipment or logging truck related in Oklahoma or New Mexico. Consequently, cases may be over-counted if the structured data coded a crash as AgFF, but an AgFF road user was not involved (e.g., tractor driver/maintenance worker was just mowing grass on a median). Conversely, crashes may be undercounted if structured data fields were not coded correctly. In this case, machine learning models or even simple two-step key word searches, based on consideration of inclusion and exclusion keyword lists, can help to reduce this potential bias (Trueblood et al., 2019; Kim et al., 2021).

4.2. Practical Applications

Despite limitations, crash records are a viable source of key information, specifically for nonfatal crashes in the southwest region. Crashes involving farm vehicles and equipment can be isolated, temporal trends in their frequency established, and potential causal or contributing factors identified so that countermeasures can be identified and implemented. Analyses focused on logging truck crashes in Louisiana demonstrated this approach (Shipp et al., 2019; Vasudeo, 2022). Crash records support the key functions of worker health surveillance (NIOSH, 2021).

Areas in need of additional research include linking crash records with other data sources for the southwest region. Data sources of primary interest for the southwest include injury surveillance databases and law enforcement citation records. The former would help to characterize the types of injuries that result from AgFF crashes and their clinically assessed severity. The latter would enable characterization of high-risk driving behaviors among those using AgFF vehicles as well as other types of drivers. There is also a need to develop gold standard denominators for computing crash and injury rates in AgFF. In traffic safety, vehicle miles travelled (VMT) is the most widely used measure of exposure to crash and injury risk. However, VMT estimates specific to AgFF do not exist. Currently, rates would need to be based on the estimated size of the population of interest, the general population, vehicle miles travelled by the general population or truck fleet, or the amount of a commodity harvested such as logging truck crashes per units of wood harvested (Greene et al., 2007; Cole et al., 2018). Finally, gold standard keyword datasets are needed to support machine learning methods for the analysis of crash narratives that can be more openly shared across researchers.

From a historical traffic safety perspective, the ''Four E'' approach was a guiding force in crash and injury prevention (FHWA, 2011). The four E's stand for engineering, emergency response, education, and enforcement. Countermeasures were identified and implemented within these constructs. Despite a commitment to this framework, the United States continued, over the last decade, to experience increases in the number and rate of traffic fatalities. For example, from 2011 to 2021, the number increased from 32,479 to 42,915 fatalities per year. The rate per vehicle miles travelled increased overall from 1.10 fatalities per 100 million VMT in 2011 to 1.33 fatalities per 100 million VMT in 2021. For NHTSA Region 6 (i.e., Louisiana, Oklahoma, Mississippi, New Mexico, and Texas), nearly the same states as our study population, the fatality rate increased 16% from 2020 to 2021 alone (NCSA, 2022). Data from our project suggest an overrepresentation of fatal and serious crashes in AgFF. In Texas from 2011 to 2021, the proportion of farm equipment and logging truck crashes that were fatal or suspected serious injury increased from 6% to 10%. This proportion for 2022 is on track to surpass 2021. In comparison, for non-AgFF crashes in Texas, the proportion of crashes that were fatal or suspected serious injury was stable at 4% in 2012 and 2021 (TxDOT, 2022).

In response to continued high fatality rates in the United States, Secretary of Transportation Buttigieg, announced in January of 2022, the adoption of a long term, national goal of zero roadway fatalities and suspected serious injuries. This goal was outlined in the National Roadway Safety Strategy (USDOT, 2022). In pursuit of a goal of zero roadway fatalities, traffic safety researchers and practitioners in the United States adopted a Safe System Approach (SSA) to traffic safety. SSA promotes five key areas: safe road users, safe vehicles, safe speeds, safe roads, and post-crash care. In line with the new national goal, SSA emphasizes that deaths and serious injuries are unacceptable, humans make mistakes and are vulnerable, responsibility for safety is shared, safety is proactive, and redundancy is crucial (FHWA, n.d.). The Safe System approach acknowledges inequities in transportation safety and crash outcomes. Safe System roadway design acknowledges that humans and their behavior are imperfect. For example, lower speeds are posted in areas where motor vehicles and pedestrians are more likely to encounter one another. Speed is known to increase injury severity (Doecke et al., 2020). Reducing it can reduce injury severity. Even if a crash still occurs, a fatality or suspected serious injury may be prevented. From a rural perspective, improving post-crash care response time is high priority given the potential impact of prehospital time, specifically the time it takes for EMS to reach injured roadway users, on injury outcomes and survival (Adeyemi et al., 2022).

The Safe System approach is a holistic framework, which aligns well with the NIOSH Total Worker Health (TWH) approach (NIOSH, 2015). TWH emphasizes a hazard-free work environment, much like the Safe System approach seeks to minimize hazards in the traffic environment. TWH further recognizes that work factors impact not only workers, but their families and communities, and focuses on the design of work to improve acute health and safety issues, including mental health, much like the extended impact of AgFF crashes on the transportation system and broader community. These two approaches are not only similar in their respective realms, but they are also complimentary where those realms overlap in AgFF related transportation safety.

Vehicle related injury among AgFF workers is a complex issue with many potential stakeholders who could influence progress in the development of robust surveillance systems, and in reducing crashes and their severity. Cross-disciplinary collaboration that brings traffic safety and AgFF researchers and practitioners together is needed to accelerate progress in construction of injury surveillance systems. In the longer term, a cross-disciplinary approach can also help to overcome barriers to the rapid identification and deployment of countermeasures that benefit all roadway users, especially the high-risk AgFF workforce in rural areas with high agricultural and logging activity.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This paper was supported by CDC/NIOSH under Cooperative Agreement No. U50 OH07541 to the Southwest Center for Agricultural Health, Injury Prevention, and Education at the University of Texas Health Science Center at Tyler. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of CDC/NIOSH.

References

Aarts, L., & Van Schagen, I. (2006). Driving speed and the risk of road crashes: A review. *Accident Analysis & Prevention, 38*(2), 215–224.

Adeyemi, O. J., Paul, R., DiMaggio, C., Delmelle, E., & Arif, A. (2022). The association of crash response times and deaths at the crash scene: A cross-sectional analysis using the 2019 National Emergency Medical Service Information System. *The Journal of Rural Health*.

Arkansas State Police (ASP) (2015). *2015*. Crash Data: Data Dictionary. Arkansas Department of Public Safety.

Baker, S., & Tyson, J. (2017). The winding roads of log truck insurance. (Forisk Research Quarterly No. Q3). Forestry Resources Association. Retrieved from https://www.swpa.ag/wp-content/uploads/sites/4736/2017/08/Forisk_Feature-Article_The-Winding-Roads-of-Log-Truck-Insurance-Aug-2017.pdf.

Bunn, T. L., Slavova, S., & Hall, L. (2008). Narrative text analysis of Kentucky tractor fatality reports. *Accident Analysis and Prevention, 40*, 419–425.

Centers for Disease Control and Prevention (CDC) (2011). Occupational highway transportation deaths-United States, 2003–2008. *MMWR, 60*(16), 487–502.

Chari, R., Kress, A. M., & Madrigano, J. (2017). Injury and Illness Surveillance of U.S. Agricultural Workers: Assessment of Recommendations and Actions. Santa Monica, CA: RAND Corporation. https://www.rand.org/pubs/research_reports/RR1500.html.

Cole, N. (2018, April 25). *Regional Analysis of Log Truck Crashes in the United States between 2011 and 2015.* Virginia Polytechnic Institute and State University.

Committee on Agricultural Safety and Health Research and Extension (CASHRE) (2009). *Agricultural Equipment on Public Roads*. Washington, DC: USDA-CSREES. https://nasdonline.org/2065/section3/d001906/agricultural-equipment-on-public-roads.html.

Conrad IV, J. (2018). Costs and challenges of log truck transportation in Georgia, USA. (Forests No. 9). Basel: MDPI AG. doi:10.3390/f9100650.

Conderino, S., Fung, L., Sedlar, S., & Norton, J. M. (2017). Linkage of traffic crash and hospitalization records with limited identifiers for enhanced public health surveillance. *Accident Analysis & Prevention, 101*, 117–123.

Costello, T. M., Schulman, M. D., & Luginbuhl, R. C. (2003). Understanding the public health impacts of farm vehicle public road crashes in North Carolina. *Journal of Agricultural Safety and Health, 9*(1), 19.

Costello, T. M., Schulman, M. D., & Mitchell, R. E. (2009). Risk factors for a farm vehicle public road crash. *Accident Analysis & Prevention, 41*(1), 42–47.

Costello, B., & Karickhoff, A. (2019). Truck driver shortage analysis 2019. *American Trucking Associations.* https://www.trucking.org/sites/default/files/2020-01/ATAs%20Driver%20Shortage%20Report%202019%20with%20cover.pdf.

Cook, L. J., Thomas, A., Olson, C., Funai, T., & Simmons, T. (2015, July). Crash Outcome Data Evaluation System (CODES): An examination of methodologies and multi-state traffic safety applications. (Report No. DOT HS 812 179). Washington, DC: U.S. Department of Transportation, National Highway Traffic Safety Administration.

Curry, A. E., Pfeiffer, M. R., Metzger, K. B., Carey, M. E., & Cook, L. J. (2021). Development of the integrated New Jersey Safety and Health Outcomes (NJ-SHO) data warehouse: Catalysing advancements in injury prevention research. *Injury Prevention*, 1–7. https://doi.org/10.1136/injuryprev-2020-044101.

Das, S., Le, M., & Dai, B. (2020). Application of machine learning tools in classifying pedestrian crash types: A case study. *IEEE Transactions on Software Engineering, 2*(2), 106–119.
Doecke, S. D., Baldock, M. R., Kloeden, C. N., & Dutschke, J. K. (2020). Impact speed and the risk of serious injury in vehicle crashes. *Accident Analysis & Prevention, 144* 105629.
Farmer, C. M. (2003). Reliability of police-reported information for determining crash and injury severity. *Traffic Injury Prevention, 4*(1), 38–44.
Federal Highway Administration (FHWA). (n.d.). Zero deaths – saving lives through a safety culture and a Safe System. Federal Highway Administration. Retrieved December 14, 2021, from https://safety.fhwa.dot.gov/zerodeaths/zero_deaths_vision.cfm.
Federal Highway Administration (FHWA). (2011). HSIP Noteworthy Practice Series:: Strategic Highway Safety Plan Get Stakeholders Involved. https://safety.fhwa.dot.gov/hsip/resources/fhwasa1102/flyr3_in.pdf.
Federal Highway Administration (FHWA) (2021). *Freight Analysis Framework Version 5*. Bureau of Transportation Statistics: National Transportation Research Center Oak Ridge National Laboratory. https://faf.ornl.gov/faf5/Default.aspx.
Fitzpatrick, C. D., Rakasi, S., & Knodler, M. A. (2017). An investigation of the speeding-related crash designation through crash narrative reviews sampled via logistic regression. *Accident Analysis & Prevention, 98*, 57-63.
Gerberich, S. G., Robertson, L. S., Gibson, R. W., & Renier, C. (1996). An epidemiological study of roadway fatalities related to farm vehicles: United States, 1988 to 1993. *Journal of Occupational and Environmental Medicine*, 1135–1140.
Gkritza, K., Kinzenbaw, C. R., Hallmark, S., & Hawkins, N. (2010). An empirical analysis of farm vehicle crash injury severities on Iowa's public road system. *Accident Analysis & Prevention, 42*(4), 1392–1397.
Glascock, L. A., Bean, T. L., Wood, R. K., Carpenter, T. G., & Holmes, R. G. (1995). A summary of roadway accidents involving agricultural machinery. *Journal of Agricultural Safety and Health, 1*(2), 93–104.
Greene, W. D., Baker, S. A., & Lowrimore, T. (2007). Analysis of log hauling vehicle accidents in the state of Georgia, USA, 1988–2004. *International Journal of Forest Engineering, 18*(2), 52–57.
Harland, K. K., Greenan, M., & Ramirez, M. (2014). Not just a rural occurrence: Differences in agricultural equipment crash characteristics by rural–urban crash site and proximity to town. *Accident Analysis & Prevention, 70*, 8–13.
Helmkamp, J. C., March, S. M., & Aitken, M. E. (2011). Occupational all-terrain vehicle deaths among workers 18 years and older in the United States, 1992–2007. *Journal of Agricultural Safety and Health, 17*(2), 147–155.
Hirabayashi, L., Scott, E., Jenkins, P., Krupa, N. (2020). Occupational injury surveillance methods using free text data and machine learning: Creating a gold standard data set. SAGE Research Methods Cases: Medicine and Health. doi:10.4135/9781529720488.
Kitali, A. E., Mokhtarimousavi, S., Kadeha, C., & Alluri, P. (2021). Severity analysis of crashes on express lane facilities using support vector machine model trained by firefly algorithm. *Traffic Injury Prevention, 22*(1), 79–84.
Karimi, K., & Faghri, A. (2021). Farm Vehicle Crashes on US Public Roads: A Review Paper. *Open Journal of Safety Science and Technology, 11*(2), 34–54.
Kim, J., Trueblood, A. B., Kum, H. C., & Shipp, E. M. (2021). Crash narrative classification: Identifying agricultural crashes using machine learning with curated keywords. *Traffic Injury Prevention, 22*(1), 74–78.
Lacy, J. K., Hunter, W., & Huang, H. (2003). Study of farm vehicle crashes in North Carolina. *Transportation Research Record, 1840*(1), 178–185.
Leigh, J. P., Du, J., & McCurdy, S. A. (2014). An estimate of the U.S. government's undercount of nonfatal occupational injuries and illnesses in agriculture. *Annals of Epidemiology, 24*(4), 254–259.
Lord, D., Qin, X., & Geedipally, S. R. (2021). *Highway Safety Analytics and Modeling*. Elsevier.
Louisiana Department of Public Safety and Corrections/Louisiana Highway Safety Commission (LADPSC/LHSC). (2019). Manual for use of the Uniform Traffic Crash Report.
Luginbuhl, R. C., Jones, V. C., & Langley, R. L. (2003). Farmers' perceptions and concerns: The risks of driving farm vehicles on rural roadways in North Carolina. *Journal of Agricultural Safety and Health, 9*(4), 327.
McKenzie, K., Scott, D. A., Campbell, M. A., & McClure, R. J. (2010). The use of narrative text for injury surveillance: A systemic review. *Accident, Analysis, & Prevention, 42*, 354–363.
Mehlhorn, S. A., Wilkin, H., Darroch, B., & D'Antoni, J. (2015). Physical characteristics of farm equipment crash locations on public roads in Tennessee. *Journal of Agricultural Safety and Health, 21*(2), 85–94.
National Center for Statistics and Analysis. (2022, April). Early estimate of motor vehicle traffic fatalities in 2021 (Crash•Stats Brief Statistical Summary. Report No. DOT HS 813 283). National Highway Traffic Safety Administration.
National Highway Traffic Safety Administration (NHTSA) (2022). *Fatality Analysis Reporting System Analytical User's Manual, 1975–2020.National Center for Statistics and Analysis*. National Highway Traffic Safety Administration DOT HS.
National Highway Traffic Safety Administration (NHTSA). (2018, August). Traffic records program assessment advisory, 2018 edition (Report No. DOT HS 812 601). Washington, DC: Author.
National Academies of Sciences, Engineering, and Medicine (NAS). (2018). A Smarter National Surveillance System for Occupational Safety and Health in the 21st Century. Washington, DC: The National Academies Press. https://doi.org/10.17226/248.
National Academies of Sciences, Engineering, and Medicine (NAS) (2021). Development of a Comprehensive Approach for Serious Traffic Crash Injury Measurement and Reporting Systems. Washington, DC: The National Academies Press. https://doi.org/10.17226/26305.
National Highway Traffic Safety Administration (NHTSA) (2017). *MMUCC Guideline: Model Minimum Uniform Crash Criteria* (5th Edition). D.C: U.S. Department of Transportation. Washington.
National Institute for Occupational Safety and Health (NIOSH). (n.d.). NIOSH Industry and Occupation Computerized Coding System (NIOCCS). Centers for Disease Control and Prevention. https://csams.cdc.gov/nioccs/.
National Institute for Occupational Safety and Health (NIOSH). (2012). Comments on the Agriculture, Forestry, and Fishing (AgFF) National Research Program Sponsored by the National Institute for Occupational Safety and Health (NIOSH), Centers for Disease Control and Prevention (CDC). Washington, D.C. https://www.cdc.gov/niosh/programs/agff/pdfs/120703oashreviewreport.pdf.
National Institute for Occupational Safety and Health (NIOSH). (2015). *Total Worker Health (TWH)*. Centers for Disease Control and Prevention. http://www.cdc.gov/niosh/twh/.
National Institute for Occupational Safety and Health (NIOSH). (2016). Analyses of the 2014 Childhood Agricultural Injury Survey (CAIS). Morgantown WV: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, Division of Safety Research. Unpublished. Posted on the Childhood Agriculture Injuries 2020 Fact Sheet. https://marshfieldresearch.org/Media/Default/NFMC/PDFs/ChildAgInjuryFactsheet2020.pdf.
National Institute for Occupational Safety and Health (NIOSH). (2020a). Center for Motor Vehicle Safety PPOP. https://www.cdc.gov/niosh/motorvehicle/default.html.
National Institute for Occupational Safety and Health (NIOSH) (2020b). Forestry and Fishing: Status of National Agriculture Injury Surveys at NIOSH. https://www.cdc.gov/niosh/agforfish/aginjurysurv.html.
National Institute for Occupational Safety and Health (NIOSH). (2021). Worker Health Surveillance. https://www.cdc.gov/niosh/topics/surveillance/default.html.
National Safety Council (NSC). (2022) Injury Facts: Costs. https://injuryfacts.nsc.org/all-injuries/costs/guide-to-calculating-costs/data-details/.
New Mexico Department of Transportation (NMDOT). (2021). Crash Database Data Dictionaries. The University of New Mexico Geospatial and Population Studies. https://gps.unm.edu/tru/dictionary.
Oklahoma Department of Public Safety (OkDPS). (2019). Oklahoma Traffic Collision Report Instruction Manual.
Peek-Asa, C., Sprince, N. L., Whitten, P. S., Falb, S. R., Madsen, M. D., & Zwerling, C. (2007). Characteristics of crashes with farm equipment that increase potential for injury. *The Journal of Rural Health, 23*(4), 339–347.
Perritt, K. R., Hendricks, K. J., & Goldcamp, E., M. (2017). Young worker injury deaths: a historical summary of surveillance and investigative findings. Morgantown, WV: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, DHHS (NIOSH) Publication No. 2017-168.
Ramirez, M., Bedford, R., Wu, H., Harland, K., Cavanaugh, J. E., & Peek-Asa, C. (2016). Lighting and marking policies are associated with reduced farm equipment-related crash rates: A policy analysis of nine Midwestern US states. *Occupational and Environmental Medicine, 73*(9), 621–626.
Ranapurwala, S. I., Mello, E. R., & Ramirez, M. R. (2016). A GIS-based matched case–control study of road characteristics in farm vehicle crashes. *Epidemiology, 27*(6), 827–834.
Ranapurwala, S. I., Cavanaugh, J. E., Young, T., Wu, H., Peek-Asa, C., & Ramirez, M. R. (2019). Public health application of predictive modeling: An example from farm vehicle crashes. *Injury Epidemiology, 6*(1), 1–11.
Roberts, T., Shaffer, R. M., & Bush, R. J. (2005). Injuries on mechanized logging operations in the Southeastern United States in 2001. *Forest Products Journal, 55*(3), 86 https://search.proquest.com/docview/214623508.
Shaffer, R. M., & Stuart, W. B. (2005). A checklist for efficient log trucking. *Va. Coop. Ext.*, 420–1094.
Scott, E., Hirabayashi, L., Jones, N., Krupa, N., & Jenkins, P. (2020). Characteristics of agriculture related motor vehicle crashes in rural New York State. *Journal of Agromedicine, 25*(2), 173–178.
Shipp, E. M., Wunderlich, R., del Junco, D. J., Holcomb, J., Fox, E. E., Tomasek, J., ... Maindale, S. (2015). *Assessing the feasibility of creating an integrated crash and injury data system for Texas*. College Station (TX): Texas Transportation Institute. Final Report TTI/SRP/15/161413-1..
Shipp, E. M., Vasudeo, S., Trueblood, A. B., & Garcia, T. P. (2019). Single vehicle logging-related traffic crashes in Louisiana from 2010–2015. *Journal of Agromedicine, 24*(2), 177–185.
Texas Department of Insurance (TDI). (2019). Farm equipment and rural road safety fact sheet. TDI Division of Workers' Compensation. HS19-002A (09-19). https://www.tdi.texas.gov/pubs/videoresource/fsfarmroad.pdf.
Texas Department of Transportation (TxDOT). (2018). 2018 Texas Peace Officer's crash report forms. https://www.txdot.gov/content/txdotreimagine/us/en/home/data-maps/crash-reports-records/forms-law-enforcement.html.
Texas Department of Transportation (TxDOT). (2022). Data extracted from the Crash Records Information System (CRIS) on November 8, 2022.
Texas Department of Transportation (TxDOT). (2023, March 13). Crash Records Information System (CRIS).
Texas Traffic Records Coordinating Committee (Texas TRCC). (2022). Texas TRCC website. Texastrcc.org.

Toussaint, M., Faust, K., Peek-Asa, C., & Ramirez, M. (2017). Characteristics of farm equipment-related crashes associated with injury in children and adolescents on farm equipment. *The Journal of Rural Health, 33*(2), 127–134.
Trueblood, A. B., Pant, A., Kim, J., Kum, H. C., Perez, M., Das, S., & Shipp, E. M. (2019). A semi-automated tool for identifying agricultural roadway crashes in crash narratives. *Traffic Injury Prevention, 20*(4), 413–418.
Trueblood, A. B., Shipp, E. M. (2019). Best Practices for Commercial Motor Vehicle (CMV) Crash Reporting. Improving the Collection and Reporting of CMV Crash Data in CRIS. Federal Motor Carrier Association Grant No. FM-MHP-0300-17.
U.S. Department of Agriculture. National Agricultural Statistics Service (USDA NASS). (2019). 2017 Census of Agriculture: Summary and State Data. Geographic Area Series AC-17-A-51. U.S. Department of Agriculture, Washington, DC. https://www.usda.gov/media/press-releases/2019/04/11/2017-census-agriculture-data-now-available.
U.S. Department of Labor Bureau of Labor Statistics (BLS). (2019a). Census of Fatal Occupational Injuries: Table A-2: Fatal occupational injuries resulting from transportation incidents and homicides, all United States, 2018.
U.S. Department of Labor Bureau of Labor Statistics (BLS). (2019b). Survey of Occupational Injuries and Illnesses. https://www.bls.gov/respondents/iif/.
U.S. Department of Labor Bureau of Labor Statistics (BLS). (2020a). Census of Fatal Occupational Injuries, Fatal occupational injuries, total hours worked, and rates1 of fatal occupational injuries by selected worker characteristics, occupations, and industries, civilian workers, 2019.
U.S. Department of Labor Bureau of Labor Statistics (BLS). (2020b). Labor Force Statistics from the Current Population Survey. https://www.bls.gov/cps/aa2019/cpsaat15.htm.
U.S. Department of Transportation (USDOT). (2022). National Roadway Safety Strategy. https://www.transportation.gov/NRSS.
Vallmuur, K. (2015). Machine learning approaches to analyzing textual injury surveillance data: A systematic review. *Accident, Analysis & Prevention, 79*, 41–49.
Vasudeo, S. (2022). Driver characteristics and behaviors associated with higher injury severity large logging truck crashes on public roadways in Louisiana (2010-2018). [Doctoral dissertation, Texas A&M University School of Public Health].
Weichelt, B., Gorucu, S., Jennissen, C., Denning, G., & Oesch, S. (2020). Assessing the Emergent Public Health Concern of All-Terrain Vehicle Injuries in Rural and Agricultural Environments: Initial Review of Available National Datasets in the United States. *JMIR Public Health Surveillance, 6*(2), e15477.
Weichelt, B., Gorucu, S., Benny, C., Burke, R., Koshalek, K., Pilz, M., Wanat, C. (2023, March 3). AgInjuryNews: About Us. https://www.aginjurynews.org/AboutUs/Index.
Zhang, X., Green, E., Chen, M., & Souleyretter, R. R. (2020). Identifying secondary crashes using text mining techniques. *Journal of Transportation Safety & Security, 12*(10), 1338–1358.

Eva M. Shipp, Ph.D. Dr. Eva Shipp is a Senior Research Scientist and Manager of the Crash Analytics Team at the Texas A&M Transportation Institute and adjunct faculty at the Texas A&M School of Public Health. She holds an M.S. and Ph.D. in epidemiology from The University of Texas Health Science Center at Houston School of Public Health. Her research interests include injury surveillance methods, integrated traffic record systems, traffic safety, and the health and safety of occupational and agricultural populations.

Amber B. Trueblood, Dr.P.H. Dr. Amber Trueblood is the Data Center Director at CPWR – The Center for Construction Research and Training. Prior to this position, she was an Associate Research Scientist at the Texas A&M Transportation Institute in Crash Analytics. Her research interests including using surveillance datasets to monitor injury trends to improve worker safety, as well as improving data collection/quality. She holds an M.P.H from The University of Texas Health Science Center at Houston School of Public Health and a Dr.P.H. in epidemiology and environmental health from the Texas A&M School of Public Health.

Hye-Chung Kum, Ph.D., F.A.M.I.A. Dr. Hye-Chung Kum is a professor in the Department of Health Policy and Management at the School of Public Health with a joint appointment in Computer Science & Engineering and Industrial Systems & Engineering at Texas A&M University. She is a Fellow of the American Medical Informatics Association and a Presidential Impact Fellow. Dr. Kum directs the Population Informatics Lab and focuses on applying data science to integrated population data sets for social good. She is an expert in data science, record linkage, sequential analysis, health services research, information privacy and confidentiality, and use of government administrative data for timely decision making.

Marcie Perez, M.Ed. Marcie Perez is an Associate Research Scientist at the Texas A&M Transportation Institute. She holds a Master of Education in Educational Psychology from Texas A&M University. She has over 10 years of experience working with crash records. She is an expert in data visualization for diverse audiences and awarded the Association of Traffic Safety Information Professionals Best Data Visualization award in 2019. Her research interest include innovative applications of crash records to improve safety.

Shubhangi Vasudeo, Dr.P.H. Dr. Shubhangi Vasudeo is a consultant at the Texas A&M Transportation Institute. She holds a Dr.P.H. in epidemiology and environmental health from the Texas A&M School of Public Health. Her research interests include public health surveillance methods, transportation safety, and data science and computing.

Nishita Sinha, Ph.D. Nishita Sinha is currently an Assistant Research Scientist in the Crash Analytics Team at Texas A&M Transportation Institute (TTI). She holds a Ph.D. in Agricultural Economics from Texas A&M University. Her research interests lie at the intersection of behavioral economics and occupational safety. More specifically, recent projects focused on safety in agricultural aviation, the evaluation of racial over-representation in traffic crash data, analyzing relative risk to pedestrians using hospital records and using linkages between motor vehicle crash data and clinical records to gain a deeper understanding of injury profiles and crash characteristics.

Ashesh Pant, M.S., M.A. Ashesh Pant is a Lead Data Analyst at the Texas A&M Transportation Institute in the Crash Analytics Team. He holds a M.S. in data analytics from Harrisburg University of Science and Technology, and a M.A. in international affairs from The Bush School of Government at Texas A&M University. His research interests include data science and computing, passively collected data systems, integrating crash records.

Lingtao Wu, Ph.D., P.E. Dr. Wu is an Associated Research Engineer at the Texas A&M Transportation Institute in the Crash Analytics Team. He holds a Ph.D. in Civil Engineering from the Texas A&M University, a M.S. in Transportation Engineering from Research Institute of Highway. Dr. Wu's research interests include transportation safety, statistical modeling, safety effectiveness evaluation, and transportation data analyses.

Myunghoon Ko, Ph.D., P.E. Dr. Ko is a Research Engineer at the Texas A&M Transportation Institute in the Crash Analytics Team. He holds a Ph.D. in Civil Engineering from the Texas A&M University.His research interests include transportation safety statistical modeling and analysis, evaluation, and education with emphasis on vulnerable roadway users (young pedestrians, bicyclists, and older drivers and pedestrians) and specialty in the application of systemic/proactive safety approaches and driving environment simulation programming.