Establishing Methodology for Occupational Injury Surveillance in Agriculture and Logging

By

Erika E. Scott

A Dissertation

Submitted to the University at Albany, State University of New York

In Partial Fulfillment of

The Requirements for the Degree of

Doctor of Philosophy

School of Public Health

Department of Environmental Health Sciences

2016

ABSTRACT

Developing a low-cost surveillance system for occupational injury in agriculture and logging has long been a challenge. These data are needed to prioritize public health interventions in the agriculture and logging industries, and are crucial to evaluating changes in injury rates, as well as the effectiveness of safety interventions. Recently, the transition to electronic data reporting created new opportunities for the development of sustainable, low-cost surveillance systems. Pre-hospital care reports (PCR) and hospital data (emergency department, inpatient, and outpatient) are datasets of particular interest.

The Northeast Center for Occupational Health and Safety in Agriculture, Forestry, and Fishing (NEC), was established by National Institute of Occupational Safety and Health (NIOSH) in 1990 to address hazards and health in the nation's most dangerous industries: Agriculture, Forestry, and Commercial Fishing. Maine and New Hampshire are the focus of this dissertation, as they are the first states to work with NEC in delivering data necessary for the establishment of such a surveillance system.

This dissertation explored the ability to harmonize disparate data sources and identify agriculture and logging-related injuries. In addition, much work was done to understand the ability to code these data using the Occupational Injuries and Illnesses Classification System (OIICS), which is widely used among other occupational health surveillance systems, as the granularity to which these data could be coded directly effects the usefulness of the system. Lastly, identified injuries were summarized, and several hypotheses were tested with these data.

Acknowledgements

Many exceptional people have helped me complete this dissertation. It is difficult to find adequate words to express all they have done for me; but in the end, I just hope to make them proud. First and foremost, I am indebted to my committee members Drs. Erin Bell, Paul Jenkins, Kitty Gelberg, Lloyd Wilson, and Edward Fitzgerald.

As my committee chair and academic advisor, Erin has provided ongoing guidance and support throughout the PhD process. I am forever grateful for her encouragement and advice, as she has helped me become a better writer and scientist. Her teaching and advisement style always put me at ease, but gave me the confidence I needed to forge ahead. Paul has been a role model to me in many ways, and he is one of the primary reasons I decided to pursue my doctorate. I feel honored to call him my mentor and friend, and look forward to many years of productive research, hilarious digressions, and pearls of wisdom.

Kitty, Ed and Lloyd have been invaluable in this process, with each of their respective expertise balancing out the committee. Kitty's background in occupational injury surveillance has been a huge asset, Ed has provoked me to think about surveillance in the context of epidemiology, and Lloyd has encouraged me to think about the practical application of surveillance within the framework of public health. Additionally, I would like to express my appreciation for the staff of the Environmental Health Sciences Department office, Danielle Grasso and Laura Hilton.

The New York Center for Agricultural Medicine and Health (NYCAMH)/Northeast Center for Occupational Health and Safety in Agriculture, Forestry, and Fishing (NEC) is a truly special place, filled with extraordinary people. I am inspired by the passion my colleagues bring to the mission of promoting health and safety among farmers, loggers, and commercial fishermen in the Northeast. This charge is led by two researchers whose compassion, dedication, tirelessness, and brilliance are unmatched - Drs. Julie Sorensen and John May. They define the essence of good leadership and scientific inquiry and I am honored to know them, and learn from them.

I thank Nicole Krupa, our invaluable data manager, for all her hard work; this research could have not been done without her. I owe a debt of gratitude to Liane Hirabayashi, NYCAMH research coordinator, as she is not only an integral part of the surveillance research team, but has become a dear friend. Her positive attitude has helped me get through many a stressful day. I especially want to thank Melissa Horsman, Alissa Kent, Deb Dalton, Emma Garrison, Marybeth Vargha, Pam Tinc, and Rebecca Russell for their assistance in record review and coding.

My family and friends have supported me throughout this journey in many ways. Tjibbe deserves special recognition for keeping me focused and motivated through the writing process; his love and support have been steadfast, and I am blessed to have him in my life. Debbie, my wonderful stepmom, proofread my dissertation chapters, and as any of you who read this document will see, that is certainly a labor of love.

My family has always supported my educational pursuits, and I am forever grateful – Mom, Dad, Grandma, Grandpa, Tante Fia, Nicholas, Michael, Aunt Doreen, Uncle Tom, Lauren, Andrew and the Van der Riet family. My friends have been incredibly encouraging throughout this process, and I thank them (especially Jennie, Jenny, Jen, Sam and Meg). Lastly, I recognize the love and unwavering support of my Opa, who passed away in April (2016), and I dedicate this dissertation in his memory - I love you more.

TABLE OF CONTENTS

List of Tables

Chapter

5. INJURY TRENDS

6. CONSIDERATIONS AND FUTURE DIRECTIONS

APPENDICIES

List of Figures

Chapter

CHAPTER 1
INTRODUCTION

SURVEILLANCE

Public health surveillance is defined by the Centers for Disease Control and Prevention (CDC) as "the ongoing systematic collection, analysis, and interpretation of health-related data essential to the planning, implementation, and evaluation of public health practice, closely integrated with the timely dissemination of these data to those who need to know. The final link in the surveillance chain is the application of these data to prevention and control".[1] Public health surveillance describes and quantifies the burden of disease, condition, or exposure among a certain population. The population under surveillance can vary substantially, depending on the goal of the system. For example, a tracking system might capture individuals living in a particular state or region (e.g., vital records, US census), be specific to a sub-population such as workers,[2] or relate to patients at a certain hospital.

By monitoring trends in these data over time, public health officials can take programmatic action, knowing they are targeting areas of high impact. An example would be implementing an advisory on fish consumption due to elevated blood mercury levels reported to the New York State Department of Health (NYSDOH) Heavy Metals Registry (HMR). Additionally, surveillance allows researchers to understand program efficacy over time. From the previous example, program efficacy could be demonstrated if fewer cases or lower test results (among a similar number of people) were reported to the HMR after the advisory was in place for a time, potentially indicating a decrease in overall blood mercury levels in the population under surveillance. Efficacy could be demonstrated provided there were no other changes to the reporting standard (e.g., definition), and analysis was conducted with care. This could include documenting possible changes in reporting that might falsely convey a true change in the rate of disease or exposure. While surveillance systems have been established nationally for public health issues such as infectious disease tracking[3] or cancer registries[4] for decades, a specialized approach to injury surveillance for farming or forestry, has not garnered attention on the same scale.

Classifying Surveillance Systems

Surveillance systems can be classified in multiple ways; however, a common classification focuses on passive and active systems.

Passive Systems

Passive systems rely on routinely collected data. A specific list of conditions or combination of variables are constructed to identify a case of interest.[5] When a condition or combination of variables are flagged, the case is reported as part of the surveillance system. A benefit of the passive system is that relatively little effort or financial resources are needed to continually capture cases once the system is established. While passive systems capture specific cases, they can be limited by incomplete primary data (such as missing data, or incorrect coding) or fluctuations in the primary data (e.g., reporting lags, changes in variables or coding). An examples of a passive surveillance system is the National Notifiable Diseases Surveillance System (NNDSS), maintained by CDC.[3] This system uses electronic methods of extracting data from healthcare records and does not require a person to actually report the disease.

Active systems

Active surveillance systems depend on data actively gathered in a repeated and timely fashion, such as contacting a person(s) whose job entails recording particular diseases or conditions within the scope of their daily work. Individuals who provide information for an active surveillance system may include health care providers, laboratory personnel, environmental technicians, public health officials and, in certain cases, the patient themselves. Results from active surveillance systems are sometimes used to validate those from passive, low-cost systems. Often, these active systems are not intended to provide a long-term means of surveillance, as they can cost substantially more than passive systems. There are occasions however, when ongoing active systems play an important, continuing role in public health. For example, the Surveillance, Epidemiology, and End Results (SEER) Program, maintained by the National Cancer Institute, which began in 1972, continues to provide active surveillance related to cancer outcomes by enlisting coders to extract information from hospitalization data. The accuracy and timeliness of these data have contributed to improvements in cancer treatment.[6] Other examples of active surveillance include the National Health and Nutrition Examination Survey (NHANES),[7] the Survey of Occupational Injuries and Illnesses (SOII),[8] the Behavioral Risk Factor Surveillance System (BRFSS), [9] past agricultural health surveys such as the Occupational Injury Surveillance of Production Agriculture (OISPA), and tracking for sudden disease outbreaks, such as the Ebola outbreak of 2014.[10] Active surveillance systems also include cancer registries[11] and registries for reportable diseases and conditions such as congenital malformations, pesticide poisonings, and occupational lung diseases.[12]

Follow-up

The terms "passive" and "active" may also be applied to following up on information once it has been recorded in the surveillance system.[13] For example, active follow-up may include notifying a patient of their health outcome and connecting them to appropriate care.

Passive follow-up indicates that the focus is on the input of information into the surveillance system, rather than on subsequent dissemination.

Analyses may be conducted on individual case reports (case-based surveillance) or on rates, if population data (denominator data) are available. Individual case reports are useful snapshots of what a single patient may experience, and can be linked to patient demographics.[13] Typically, case reports are the first indicator of a larger public health issue. For example, sentinel reports of heavy metal poisoning within certain ethnic groups alerted authorities to the importation of medication and spices tainted with dangerous heavy metals.[14] Population-based surveillance can detect changes over time when it accounts for changes in the overall population from which surveillance cases are identified.

Essential Elements of Surveillance Systems

The essential elements of passive and active surveillance systems are the same, and several defined steps must be in place to successfully establish either system (Table 1 and Appendix A). In 2001, the CDC updated the *Guidelines for Evaluating Surveillance Systems*,[15] which was originally created in 1988.[16] Beyond the logistics of establishing a surveillance system, several attributes are important to its overall success. Both the structure and operation of the system should be relatively simple.[15]

Table 1. Considerations for Surveillance Systems[17]

Attribute	**Description**
Justification	frequency/severity of problem direct/indirect costs
	public interest? interventions possible?
Objectives	measuring scope of problem underlying population monitoring changes feasible? isolated events or communicable disease
Case Definition	specific reproducible
Operational Considerations Data Acquisition	staff how often are data analyzed? frequency of reporting? data use agreements
Data Volume	seasonal or constant data? what variables are needed? maintaining timelines

Well-designed surveillance systems follow well-established guidelines, regardless of the type of surveillance. These attributes ensure that the system runs smoothly and efficiently (see Table 2 and Appendix A).

Table. 2 Attributes of Surveillance Systems (adapted from Lee et al.)[5,15]

Attribute	**Definition**
Simplicity	Structure and ease of operation
Flexibility	Ability to adapt to changing needs, both technical and information
Data quality	Completeness and validity of data
Acceptability	Willingness of stakeholders to participate in system
Sensitivity	Proportion of cases detected, or the ability to detect outbreaks over time
Predictive value positive	Proportion of reported cases that have the event under surveillance
Representativeness	Ability to accurately describe the event over time and the distribution in the population
Timeliness	Speed between steps in the system, including receiving data
Stability	Reliability and availability

Evaluation of the surveillance system is necessary to ensure the effective tracking of public health issues. This ensures quality data are being included in the system and that it remains effective. Important components of an evaluation include direct and indirect costs to maintain the system, the purpose of the system, and the identification of ways to improve the system.

WORKPLACE INJURY

Injuries are the leading cause of death for people ages one to forty-four in the United States.[18] Since many of our most productive years are spent working, occupational injury (injury occurring at work) can have a significant impact of the quality of life and health. The National Safety Council estimates that an occupational injury requiring medical attention in the United States costs on average $42,000, and the cost of an occupational fatality exceeds $1.4 million dollars.[19] While traumatic for families, coworkers and employers of the victim, these incidents also harbor significant societal and economic costs,[20] such as future lost income, insurance benefits, and lost home production, which includes home repairs and child rearing.[21]

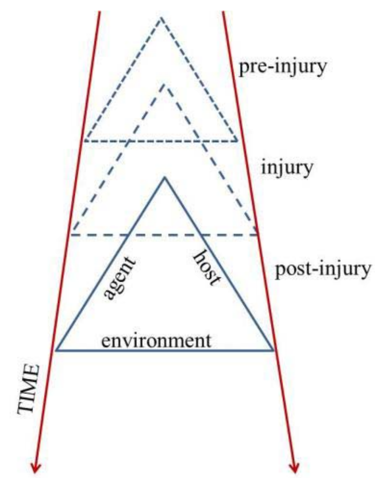


Figure 1. Haddon matrix as an extension of the epidemiologic triad (adapted)[17]

An injury is defined as "harm to the body resulting in impairment or destruction of health, which results from acute exposure to thermal, mechanical, electrical, or chemical energy that exceeds a threshold of tolerance in the body or from the absence of such essentials as heat or oxygen."[22] Injury can refer to unintentional or intentional harm (self-harm or violence). Dr. William Haddon's method to classify injuries based on the agent, host, and environment through time (Figure 1) is one of the basic pillars of the science.[23] The *agent* describes etiologic factors (e.g., a medium that causes injury or disease); the *host* encompasses essential elements of the person (e.g., behaviors, medical conditions, etc.), and the *environment* constitutes outside factors that affect the situation (e.g., workplace, community, rules, etc.). Haddon also described factors that impact the severity of injury, which include the type of energy transfer (mechanical, heat, radiation), amount of energy transferred, and whether or not safety equipment was used to reduce exposure.[23]

There are many instances where reporting the occupational injury event is required by law. Federally mandated reporting is processed through the US Department of Labor's Bureau of Labor Statistics (BLS) according to their record keeping regulations.[24] Beginning in 2015, establishments under federal OSHA jurisdiction must report all work-related fatalities within eight hours of their occurrence, and all work-related inpatient hospitalizations, amputations and eye losses within 24 hours. Self-employed, immediate family members of farm employees who do not hire outside employees, and businesses regulated by special federal agencies such as the Mine Safety and Health Administration [MSHA][24] are excluded from these requirements. Workers' compensation insurers also receive reports of work-related incidents for the employers and employees they cover.

Current Data Sources for Agricultural and Logging Related Injury

While past research[25-32] and surveillance systems[32-34] have informed public health professionals of injury patterns in agriculture and logging, the primary challenge is that these data become dated, and no ongoing systems are able to provide stable estimates of non-fatal injuries over time. Another problem in these two sectors is that specific information on causal agents is often lacking. The section below highlights existing occupational health surveillance programs, their uses, and their limitations.

OSHA Records

OSHA records provide data on injuries for major industry;[35] though they have limited use in agriculture, especially for small family owned farms, and for small logging operations. Businesses that fall within the reporting criteria are required to maintain logs of workplace injury and illness using standardized forms. With many small farms not required to fill out OSHA 300 forms (log of work-related injuries and illnesses), and a lack of inspectors (one inspector for every 59,000 US workers), enforcement of recordkeeping is difficult.[36] The OSHA local emphasis programs (LEP) have targeted agriculture and logging on several occasions. These programs focus inspections on high-risk industries. Currently, logging-related OSHA LEP programs are active in West Virginia, Idaho, Oregon, and Washington State.[37] Similarly, there is an OSHA LEP program active in New York State related to dairy farming.

Though reporting mechanisms exist through OSHA and workers' compensation, not all occupational injury and illness are reported. [38,39] Exemptions to OSHA reporting exist; if a company has ten or fewer employees, it is typically not required to report injuries to OSHA.

Another limitation of this data system is that limited information is available for certain vulnerable populations, such as undocumented immigrants and migrant workers.[40] Industries that employ such workers often have hazardous workplaces, and undocumented workers frequently feel pressured to not report injuries.[41] Though some injuries go unreported, researchers occasionally use certain techniques to fill in knowledge gaps. Due to the sparseness of data in some industries, correction factors may be used to more closely estimate the actual burden of disease or injury.[42]

Workers Compensation

Workers compensation claims have been explored for their use in surveillance systems.[43] While helpful, exemptions also exist for required WC coverage. Many farmers do not carry workers compensation insurance (sole proprietor), and even when they have coverage they might be hesitant to file a claim. According to a Canadian study, workers in agriculture, fishing, and forestry have high rates of work-related injuries, yet low rates of workers compensation claims.[44] Some hesitation might be due to fear of increased insurance costs, or negative repercussions on the jobsite. It is common for loggers to be covered by workers' compensation insurance, with the exception of sole proprietors; however, these data are not always accessible. Recent conversations with several regional workers' compensation insurers indicate they often have policies in place to prevent data sharing with researchers.

There have been many studies citing the underreporting of occupational injury and illness to both OSHA and workers' compensation.[45,46] For these reasons, using workers' compensation data as a sole source of surveillance data is not advisable for certain industries, including agricultural and logging.

Census of Fatal Occupational Injuries (CFOI)

Fatal traumatic occupational injuries are captured by the Census of Fatal Occupational Injuries (CFOI)[47] collected by the US Department of Labor, Bureau of Labor Statistics. This census is designed to capture all fatalities regardless of the size of the business. Information is gathered from a combination of sources including newspaper clippings, coroner's reports, death certificates, OSHA reports, WC, and occasionally, reports from the public. This system captures fatal incidents involving agriculture and logging even if the victim was a volunteer worker (unpaid help). One limitation of the CFOI system for logging and agriculture is that it may exclude injuries that occur while doing logging or agricultural work that isn't work-related. This may include a child agricultural fatality if the child was not working at the time of the incident. Additionally, it can be difficult to identify the distinction between work and home for incidents occurring on family property.

Survey of Occupational Injuries and Illnesses (SOII)

Non-fatal traumatic injuries and illnesses are captured by the Survey of Occupational Injuries and Illnesses (SOII), which is administered by BLS. Respondents use their OSHA recordkeeping logs to complete the mailed paper survey. Unlike CFOI, SOII only captures events (injury and illness) for a random sample of 200,000 businesses. Excluded from the SOII are the military, self-employed individuals, farms with less than 11 employees and Federal Agencies.[8] As such, the SOII estimates of non-fatal injury rate are not representative of the rates for smaller enterprises within agriculture, forestry and fishing.[48-50] Ruser cited out-of-scope workers (which includes small farms) as one of four dimensions that contribute to a potential undercount of these injuries and illnesses.[50] The SOII has been tested using capture-recapture methodology by Rosenman and colleagues, who found that 16.5% of cases, comparing SOII and workers compensation data in Michigan, were not reported.[51] Similarly, Boden and Ozonoff found that missed cases ranged from three to thirteen percent (across several states) when the capture-recapture method was employed.[38]

State Based Occupational Health Surveillance Programs

As of 2015, twenty-six US states were funded to operate occupational health surveillance programs in their respective states.[52] These programs range from the fundamental program, which provides data on occupational health indicators, to expanded programs, which add modules such as administration of the FACE program, respiratory disease surveillance program, and pesticide poisoning registries. NIOSH administers the Fatality Assessment and Control Evaluation Program (FACE) in several states (currently NY, MA, KY, MI, CA, WA and OR) around the country. FACE goes beyond a simple census, and conducts in-depth investigations of certain fatalities to determine root causes of injury.[53] Pesticide poisoning registries (CA, MI, MA, and WA) and respiratory disease surveillance exist in only a few states each (CA, MA, MI, NY, and WA). While these programs are integral to supporting occupational health initiatives in their respective states, the focus on agriculture and logging are not equally represented among all states. Individual states may choose their areas of focus, which at times could include agricultural or logging related issues. However, these areas of focus are known to change due to the nature of grant funding, emerging issues, as well as top causes of disease, injury or fatality.

National Electronic Injury Surveillance System Work Supplement

The NIOSH sponsored National Electronic Injury Surveillance System Work Supplement (NEISS-Work) captures traumatic occupational injuries from a nationally representative sample of 67 hospital emergency departments. Though NEISS has been used to explore agricultural injuries, it is limited to cases involving consumer products.[27] In addition, NEISS-Work does not provide injury prevalence estimates within occupational strata.

This lack of information on non-fatal traumatic injuries is particularly important, as it has been shown that for every fatal traumatic injury there are multiple non-fatal events.[54] Surveillance systems that use multiple sources of data have been shown to be more effective in capturing not only case counts, but also detail that is useful for intervention planning.[55]

Injury in Agriculture and Logging

Agriculture, forestry and fishing (AFF) have the highest traumatic injury fatality rates of any industry in the nation, over seven times higher than the average worker fatality rate (23.9/100,000 full time employees (FTE) versus 3.3/100,000 FTE).[56] In the decade from 2004 to 2013, 2,719 people died from traumatic injuries while performing agricultural, forestry and fishing related work across the United States.[57]

The U.S. Bureau of Labor Statistics indicates that in 2013, work-related fatality rates for logging workers (91.3 per 100,000 full-time workers [FTE]) were the highest for any civilian occupation[56] and nearly twenty-eight times the all-worker fatality rate (3.3 per 100,000 FTE) in 2013. In the ten-year period from 2004-2013, the Census of Fatal Occupational Injuries identified 676 logging fatalities.[47] In 1991, Dr. Tim Kelsey, an agricultural economist, estimated the cost of each of these injuries, adjusted to 2015 dollars using the consumer price index,[58] to be approximately $769,348.[59] Non-fatal traumatic injury, cumulative trauma, and chronic disease also have the potential to pose a great burden to this population, though the extent of this remains mostly unknown, as documentation of agricultural and forestry injury and illness is a relatively new practice.

AGRICULTURE

The first agricultural injury research focused on hospitalization records, as physicians documented traumatic injury cases that presented unique surgical challenges. These cases were cited as being more surgically involved than other industrial injuries.[60] After analyzing 310 farm injury cases between 1929 and 1938, Powers indicated that farmers' inadequate equipment contributed to the elevated number of injuries, along with socioeconomic status, and lack of skilled hired labor.[61] Between 1929 and 1948, health care professionals treated 1,028 farm-related injuries at the Mary Imogene Bassett Hospital in Cooperstown alone.[62] Calandruccio and Powers emphasized the contribution of tractor-related events to the total number of injuries, citing carelessness of operation and lack of needed safety equipment (roll-over protective structures), even though these safety innovations existed at the time for other industries (e.g., construction).[62] Even today, tractor incidents remain the leading cause of injury among farmers.[42]

Agricultural health and safety began to garner greater national attention in the 1970's and 1980's. By the close of the 1980's, specific agricultural safety research programs were established across the country. These included the Institute for Agricultural Medicine (IAM) at the University at Iowa, the Institute of Rural Environmental Health (IREH) at Colorado State University, the National Farm Medicine Center (NFMC) at the Marshfield Clinic in Wisconsin, along with the Bassett Farm Safety and Health Project (BFSHP) at the Mary Imogene Bassett Hospital in Cooperstown, New York.[63] With substantial scientific research emphasizing farm hazards, national safety expertise, and the support of Senator Thomas Harkin of Iowa, national agricultural health and safety centers, under the National Institute of Occupational Safety and

Health (CDC-NIOSH), were created in 1990.[63] Since their inception, these centers have been funded through the federal government to provide research for agriculture, fishing and forestry health and safety across the country.

Surveillance specific to agriculture has been implemented for several decades, though significant gaps exist in these systems. Estimates of fatal and non-fatal injury rates vary geographically and temporally, in part because different methods may be used to gather and analyze data. Injury rate calculations can be based on the size of the workforce, the number of individuals living on a farm, number of farms, estimated work hours, or any combination of these factors.[63] Since no universal standard exists, comparing rates across states or regions becomes very difficult. In 1957, Pennsylvania became the first state to document its farm injury rate per 1,000,000 work hours. This initial comparison showed that farming had a higher injury rate than any other industry in Pennsylvania at the time.[64] Other states followed with their own rate calculations, and by the late 1960's and 1970's states were working cooperatively on survey and surveillance design.[65-67]

When the Farm Accident Survey Report was published in 1988, data from over 31,000 farms provided detailed information on the epidemiology of farm injury.[29] The report cited a national accident rate of 20.5 per million hours worked (range 9.4-24.6). The leading causes of farm injury were animals, machinery, and other vehicles. Machinery resulted in the highest number of permanent injuries.[29] Earlier work by Doss and Pfister showed that tricycle tractors were over fifty percent as likely to be involved in an injury event, versus a wide front-end tractor.[68] More recent data still finds tractor roll-overs, machinery entanglements, and animal incidents as leading sources of farm injury[33,69,70] and children, older workers, and those with damaged hearing to be at greater risk.[71-73] This was also found in the Farm Family Health and Hazard Surveillance Study conducted jointly by NYCAMH and NYSDOH, which found that hearing loss, joint issues, and younger age (youth) were risk factors for sustaining a farm injury.[74] Issues with an aging workforce are likely to compound some of these risks.

In addition to traumatic injury, farmers and farmworkers are at risk for cumulative trauma and musculoskeletal disorders.[75-77] Though a significant burden, these issues will not be discussed at length in this dissertation. The nature of cumulative trauma versus traumatic injury differ, and the ability to capture their occurrence in existing data sources presents unique challenges.

Agricultural Injury Surveys

The beginning of NIOSH's agriculture safety program in the early 1990's set the stage for gathering injury data for farming. Large surveys were conducted to collect detailed exposure, injury, and health data on US farmers, farm children, and farm workers.

Occupational Injury Surveillance of Production Agriculture Survey (OISPAS)

In 2001 and 2004, NIOSH and USDA collaborated on the Occupational Injury Surveillance of Production Agriculture Survey (OISPAS) administered in conjunction with the Childhood Agricultural Injury Survey (CAIS). For both years, a total of 32,883 farms were surveyed by telephone interview. Results from these surveys indicated that though the absolute number of injuries went down, 75,756 in 2001 versus 71,081 in 2004, the rate of injury increased over the same period (12.3 injuries/1000 workers to 13.4 injuries/1000 workers).[33]

Childhood Agricultural Injury Survey (CAIS)

The Childhood Agricultural Injury Survey (CAIS) was conducted in 1999, 2001, 2004, 2006, 2009, and 2012.[78] Children and youth who live on farms have long been considered at risk for injury. Since the farmstead is not only a home but also a workplace, risks associated with production agriculture can also pose danger to children. Children, both young and adolescent, often help their parents by doing chores and tasks related to animal care or crop production. This is typically a normal and expected part of farm life. Historically, children often do the same tasks that their parents and grandparents were assigned as children. While performing these chores, children may be near farm machinery, or even operating it if they are deemed mature enough to do so. Tending to animals is also a common chore. For children who are not working, just playing may pose a risk, if they do not have a designated play area on the farmstead.

While it is recognized that youth injury has decreased over the last few decades, rates are still unacceptably elevated. Some injuries are due to the work environment; however, the majority of children who are hurt on the farm are not working. In addition, a 2006 study showed that nearly forty percent of child farm injuries occurred to children who did not live on the farm.[79]

National Agricultural Workers Survey (NAWS)

The National Agricultural Workers Survey (NAWS) was the only surveillance system focused on farmworkers, including migrant and seasonal farmworkers, which covered the entire US. The survey sampled adolescents and adults 14 and older without inquiring about immigration status. This work was important in estimating injury rates in these populations, as marginalized workers were not often included in other surveillance systems. It was not that other systems purposely excluded marginalized workers, but rather the workers decline to participate due to fear of retaliation, legal action, deportation, and many other factors.[80] Using NAWS injury data, researchers estimated the overall injury rate to be 4.3 injuries per 100 full time hired crop workers.[81] This same study found overexertion, falls, and injuries involving hand tools to be common. As expected, the majority of injuries were reported to have occurred during planting and harvest season, with June having the largest percent of injuries (16 percent).[81]

Unfortunately, these long-standing NIOSH-funded occupational injury surveys, including the Occupational Injury Surveillance of Production Agriculture (OISPA, since 2001). The Childhood Agricultural Injury Survey (CAIS, since 1998), and the National Agricultural Workers Survey[82] (NAWS injury module, since 1999) were terminated in fiscal year 2015. This further increases the need for a low-cost surveillance system that will not burden a funding agency in the long term. NIOSH has acknowledged the need for new, low cost methodologies for surveillance in agriculture, forestry, and commercial fishing.[82,83]

Logging

The U.S. Bureau of Labor Statistics indicated that in 2013, work-related fatality rates for logging workers (91.3 per 100,000 full-time workers [FTE]) was the highest for any civilian occupation[56] and nearly twenty-eight times the all-worker fatality rate (3.3 per 100,000 FTE) in 2013. In the ten-year period from 2004-2013, the Census of Fatal Occupational Injuries identified 676 logging fatalities.[47] Though it is known to be more dangerous than agriculture, less attention has been paid to quantifying forestry- and logging-related injury and illness. Logging workers (felling trees, machinery operation) have significantly higher rates of injury when compared to their forestry worker (conservation, replanting) counterparts. Logging fatalities are aggregated at the state level, via death certificates, and used in federal estimations.[47,53] Analyses of OSHA injury data have shown significant injury issues, with the most common injury events being struck by and against objects, along with falls from heights. However, these data are known to exclude self-employed and firms with less than 11 employees, as mentioned in the earlier section referencing OSHA.[84]

Logging processes have become increasingly mechanized, using large machinery to harvest timber, which has introduced additional sources of injury. While hand felling (a very dangerous activity) is less common today, mechanization has brought its own health and safety issues to the forefront. Increasingly, falls from heights while boarding and alighting equipment are a significant issue.[84]

Similar to agriculture, the workforce in logging has grown older, and it is increasingly difficult to recruit young workers into the field. Physical declines in workers can contribute to rates of injury and illness. It has been suggested that recovery rates are slower in older adults, and sensory deterioration contributes to injury events.[85] On the opposite end of the spectrum, young inexperienced workers also show higher rates of injury, due to lack of training and practice.[86] Some evidence showed that very new employees (0-12 months on the job) had safer practices than workers on the job for one to three years, suggesting that with some familiarity, complacency in safety measures may take place.[87] Risk taking behaviors are evident in these occupations, as it is commonly cited that dangerous methods of felling are faster than using proper techniques.[85] Specific logging injury surveillance has been conducted in the Pacific Northwest[88] and South,[84,89] however such methodology has not been explored in the Northeast. A recent report to NIOSH suggests that less than 15 percent of public data systems for occupational health and safety information are considered relevant to the logging worker population,[90] further emphasizing the need for creative methods of data acquisition. This is because these public data systems do not have a method of identifying logging or forestry as an industry or occupation, therefore relevant workforce data cannot be extracted. This further emphasizes the need for new methods of data acquisition.

Agricultural and Logging Surveillance Research Still Needed

With such unique challenges in conducting injury surveillance in the logging and agricultural industries, especially in light of the blurred work-home distinction in agricultural, administrative datasets are an attractive option for occupationally-related injury data. Additionally, these data are typically available for minimal cost. These data are primarily collected for other purposes, but they are likely to be in electronic form, or transitioning to this format. Electronic data reduce the time required to review records because computer technology can assist in the process. Improving surveillance for agriculture and logging allows public health professionals to make informed decisions about improvements in these industries. The first step in moving this field forward is having accurate foundational data, therefore justifying public expenditures on research and safety and health programs. Lastly, tracking and evaluating interventions provides valuable information about the effectiveness of these efforts. Quality surveillance measures rates over time, and can therefore capture when such rates remain stable, increase, or decrease. In addition, emerging trends are often identified by such surveillance efforts.

This proposed research will develop a methodology to capture traumatic agricultural and logging injuries using existing administrative datasets, thereby establishing a system that is low-cost and can be run with minimal staff time. There is a recognized gap in agricultural and logging surveillance and this project aims to fill that void.[48,49] Not only will this research fill a gap in the understanding of farm and logging injuries, but it will increase knowledge as to the potential of these data sources for other public health purposes. While hospital discharge data has been routinely collected in electronic format,[91] it was only recently that the transition from paper to computerized data was made for EMS (pre-hospital care report) records. These data are now nationally aggregated in the National Emergency Medical Services Information System (NEMSIS). [92]

Specific Aims of this Dissertation

Specific Aim 1a: To evaluate the utility of linking records to provide a comprehensive dataset from which to extract agricultural and logging related injury.

Specific Aim 1b: To extract records for which either the PCR or the hospital record (or both) indicate agricultural or logging injury from this case matched file.

Specific Aim 2: To evaluate the ability to classify cases using the Occupational Injury and Illness Classification System (OIICS) coding scheme and evaluate the level of granularity that can be obtained from each data source.

Specific Aim 3: To test the usefulness of the developed system by identifying trends in agricultural and logging injury in Maine and New Hampshire through rate calculations of injury using both PCR and hospital data.

References

1. CDC. *Comprehensive Plan for Epidemiologic Surveillance.* Atlanta1986.
2. Estes CR, Marsh SM, Castillo DN. Surveillance of Traumatic Firefighter Fatalities: An Assessment of Four Systems. *Public Health Rep.* 2011;126(4):540-551.
3. CDC. NNDSS | Centers for Disease Control and Prevention. 2015; http://wwwn.cdc.gov/nndss/. Accessed November 14, 2015.
4. CDC. National Program of Cancer Registries (NPCR). 2015; http://www.cdc.gov/cancer/npcr/.
5. Lee LM, Teutsch SM, Thacker SB, Louis MES. *Principles and Practice of Public Health Surveillance.* 3rd Edition ed. New York, New York: Oxford University Press; 2010.
6. Institute NC. Surveillance, Epidemiology, and End Results Program. 2015; http://seer.cancer.gov/. Accessed November 16, 2015.
7. CDC. NHANES - National Health and Nutrition Examination Survey Homepage. 2015; http://www.cdc.gov/nchs/nhanes.htm. Accessed November 12, 2015.
8. BLS. Survey of Occupational Injuries and Illnesses (SOII) Respondents Home Page. 2015; http://www.bls.gov/respondents/iif/. Accessed June 25, 2016.
9. CDC - BRFSS. 2015; http://www.cdc.gov/brfss/. Accessed Feb 4, 2016.
10. WHO. Ebola surveillance in countries with no reported cases of Ebola virus disease. *WHO* 2015; http://www.who.int/csr/resources/publications/ebola/ebola-surveillance/en/. Accessed Feb 4, 2016.
11. NYSDOH. NYS Cancer Registry. 2015; http://www.health.ny.gov/statistics/cancer/registry/.
12. NYSDOH. Reportable Diseases and Conditions. 2015; http://www.health.ny.gov/professionals/reportable_conditions/reportable_conditions.htm. Accessed July 30, 2015.
13. Neil M. *Workplace Health Surveillance.* Oxford University Press; 2000.
14. MMWR C. Lead poisoning associated with ayurvedic medications--five states, 2000-2003. *MMWR. Morbidity and mortality weekly report.* 2004;53(26):582-584.
15. CDC. Updated Guidelines for Evaluating Public Health Surveillance Systems. *MMWR* 2001; http://www.cdc.gov/mmwr/preview/mmwrhtml/rr5013a1.htm. Accessed August 3, 2015.
16. CDC. Guidelines for Evaluating Surveillance Systems. 1988; http://www.cdc.gov/mmwr/preview/mmwrhtml/00001769.htm. Accessed August 3, 2015.
17. Nash DEZ. Public Health Surveillance. In: EPIC, ed. *Epidemiology and Population Health Summer Institute*. New York, New York: Columbia University; 2013.
18. CDC. Leading Causes of Death. *Injury Center* 2015; http://www.cdc.gov/injury/overview/leading_cod.html. Accessed August 5, 2015.
19. Council NS. *Injury Facts.* Itasca, IL: National Safety Council; 2015.
20. Biddle EA. Is the societal burden of fatal occupational injury different among NORA industry sectors? *J Safety Res.* 2013;44:7-16.
21. Leigh JP. Economic burden of occupational injury and illness in the United States. *Milbank Q.* 2011;89(4):728-772.
22. Baker SP; O'Neill B GM, Li G. *The Injury Fact Book.* Second ed. New York: Oxford University Press; 1992.
23. Haddon W. On the escape of tigers: an ecologic note. *Am J Public Health Nations Health.* 1970;60(12):2229-2234.
24. Injury and Illness - Recordkeeping. 2015; https://www.osha.gov/recordkeeping/RKforms.html. Accessed June 25, 2015.

25. Day L, Voaklander D, Sim M, et al. Risk factors for work related injury among male farmers. *Occup Environ Med.* 2009;66(5):312-318.
26. DeRoo LA, Rautiainen RH. A systematic review of farm safety interventions. *American journal of preventive medicine.* 2000;18(4 Suppl):51-62.
27. Fox S, Ricketts M, Minton JE. Worker Injuries Involving the Interaction of Cattle, Cattle Handlers, and Farm Structures or Equipment. *J Agric Saf Health.* 2015;21(1):3-12.
28. Hawk C, Donham KJ, Gay J. Pediatric Exposure to Agricultrual Machinery:. *Journal of Agromedicine.* 2010;1(1):57-74.
29. Hoskin A. F, T. A. Miller, W. D. Hanford, S. R. Landes. Occupational injuries in agriculture: a 35-state summary. *NIOSHTIC-2 Publications Search - 20039768* 1988; http://www.cdc.gov/niosh/nioshtic-2/20039768.html. Accessed June 25, 2015.
30. Pratt DS, Marvel LH, Darrow D, Stallones L, May JJ, Jenkins P. The dangers of dairy farming: the injury experience of 600 workers followed for two years. *Am J Ind Med.* 1992;21(5):637-650.
31. Pedersen DH, Wilkins JR, 3rd, Bean TL, Mitchell GL, Crawford JM, Jones LA. Agricultural hazard data from a population-based survey of cash grain farms: Ohio observations. *Applied occupational and environmental hygiene.* 1999;14(5):299-305.
32. Forst L, Erskine T. Farm injuries in Ohio, 2003-2006: a report from the emergency medical services prehospital database. *J Agric Saf Health.* 2009;15(2):171-183.
33. Goldcamp EM. Work-related non-fatal injuries to adults on farms in the U.S., 2001 and 2004. *J Agric Saf Health.* 2010;16(1):41-51.
34. Layne LA, Goldcamp EM, Myers JR, Hendricks KJ. Youth living on Hispanic-operated farms: injuries and population estimates in the U.S., 2000. *J Agric Saf Health.* 2009;15(4):377-388.
35. Olorunnishola OA, Kidd-Taylor A, Byrd L. Occupational injuries and illnesses in the solid waste industry: a call for action. *New solutions : a journal of environmental and occupational health policy : NS.* 2010;20(2):211-223.
36. OSHA. Partial exemption for employers with 10 or fewer employees. - 1904.1. 2015; https://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=STANDARDS&p_id=9632. Accessed June 25, 2015.
37. USDOL. OSHA Enforcement | Local Emphasis Programs. 2015; https://www.osha.gov/dep/leps/leps.html. Accessed November 25, 2015.
38. Boden LI, Ozonoff A. Capture-recapture estimates of nonfatal workplace injuries and illnesses. *Ann Epidemiol.* 2008;18(6):500-506.
39. Azaroff LS, Levenstein C, Wegman DH. Occupational injury and illness surveillance: conceptual filters explain underreporting. *Am J Public Health.* 2002;92(9):1421-1429.
40. Costich J. Who pays for agricultural injury care? *J Agromedicine.* 2010;15(1):54-57.
41. Rebecca Smith JD. Immigrant workers and worker's compensation: the need for reform. *Am J Ind Med.* 2012;55(6):537-544.
42. Scott EE, Earle-Richardson G, Krupa N, Jenkins P. A correction factor for estimating statewide agricultural injuries from ambulance reports. *Ann Epidemiol.* 2011;21(10):767-772.
43. Karttunen JP, Rautiainen RH. Characteristics of and risk factors for compensated occupational injury and disease claims in dairy farmers: a case-control study. *J Agric Saf Health.* 2013;19(3):191-206.
44. Fan ZJ, Bonauto DK, Foley MP, Silverstein BA. Underreporting of work-related injury or illness to workers' compensation: individual and industry factors. *J Occup Environ Med.* 2006;48(9):914-922.
45. Rosenman KD, Gardiner JC, Wang J, et al. Why most workers with occupational repetitive trauma do not file for workers' compensation. *J Occup Environ Med.* 2000;42(1):25-34.

46. Biddle J, Roberts K, Rosenman KD, Welch EM. What percentage of workers with work-related illnesses receive workers' compensation benefits? *J Occup Environ Med.* 1998;40(4):325-331.
47. BLS. Census of Fatal Occupational Injuries (CFOI) - Current and Revised Data. 2015; http://www.bls.gov/iif/oshcfoi1.htm#2012. Accessed June 25, 2015.
48. Leigh JP, Marcin JP, Miller TR. An Estimate of the U.S. Government's Undercount of Nonfatal Occupational Injuries. 2004;46(1):10-18.
49. NORA. National Agriculture, Forestry and Fishing Research Agenda. *National Occupational Research Agenda.* 2008. www.cdc.gov/niosh/nora/comment/agendas/AgForFish/pdfs/AgForFishDec2008.pdf. Accessed December.
50. Ruser J. Examining evidence on whether BLS undercounts workplace injuries and illnesses. 2008; http://www.bls.gov/opub/mlr/2008/08/art2full.pdf. Accessed February 19, 2015.
51. Rosenman KD, Kalush A, Reilly MJ, Gardiner JC, Reeves M, Luo Z. How much work-related injury and illness is missed by the current national surveillance system? *J Occup Environ Med.* 2006;48(4):357-365.
52. CDC-NIOSH. Extramural Research and Training Programs - State Surveillance Portfolio. 2015; http://www.cdc.gov/niosh/oep/statereports.html. Accessed November 25, 2015.
53. NIOSH. CDC - Fatality Assessment and Control Evaluation(FACE)Program: Main Page - NIOSH Workplace Safety and Health Topic. 2015; http://www.cdc.gov/niosh/face/. Accessed June 25, 2015.
54. WHO. The Injury Pyramid. 2015; http://www.who.int/violence_injury_prevention/key_facts/VIP_key_fact_5.pdf?ua=1. Accessed August 3, 2015.
55. Earle-Richardson GB, Jenkins PL, Scott EE, May JJ. Improving agricultural injury surveillance: a comparison of incidence and type of injury event among three data sources. *Am J Ind Med.* 2011;54(8):586-596.
56. BLS. Final Work Related Fatalities, 2013. 2013; http://www.bls.gov/iif/oshwc/cfoi/cfch0012.pdf. Accessed June 25, 2015.
57. BLS. Fatal occupational injuries by selected characteristics, 2003-2013. 2015; http://www.bls.gov/iif/oshwc/cfoi/all_worker.pdf. Accessed July 2015, 2015.
58. BLS. Consumer Price Index (CPI). 2015; http://www.bls.gov/cpi/. Accessed October 12, 2015.
59. Kelsey T. Fatal Farm Accidents in New York: Estimates of Their Costs. *Northeastern Journal of Agriculture and Resource Economics.* 1991;20(2):202-206.
60. Lloyd E. Agricultural Accidents. *Industrial Medicine.* 1933;2:240-243.
61. Powers J. The Hazards of Farming. *Journal of the American Medical Association.* 1939;113:1375-1378.
62. Calandruccio RPJ. Farm Accidents: A clinical and statistical study covering twenty years. *American Journal of Surgery.* 1949(November):652-660.
63. Murphy D. *Looking Beneath the Surface of Agricultural Safety and Health.* ASABE Technical Library; 2003.
64. Pennsylvania Department of Public Instruction. A Study of Accidents to Pennsylvania Farm People. *Bulletin* Harrisburg, PA1957.
65. Hanford WD CD. Eight state Survey: zeroing in on farm accidents. *Farm Safety Review.* 1972;30(30):Special Issue.
66. Pfister RG HK. A Study of Accidents Occuring to Farm People in Michigan. *Journal of Safety Research.* 1969;1(1):28-41.
67. Stuckey WE PA. *Accidental Injuries to Ohio Farm People, 1957-1972.* Columbus, OH1973.

68. Doss HJ PR. *Nature and extent of use of farm machinery in relation to frequency of accidents.* East Lansing, MI: Michigan State University;1972.
69. Myers JR, Hard DL. Work-related fatalities in the agricultural production and services sectors, 1980-1989. *Am J Ind Med.* 1995;27(1):51-63.
70. Murphy DJ AF. Pennsylvania farm fatalities during 1990–1994. *National Age Safety Database.* 1998.
71. Hwang SA, Gomez MI, Sobotova L, Stark AD, May JJ, Hallman EM. Predictors of hearing loss in New York farmers. *Am J Ind Med.* 2001;40(1):23-31.
72. Amshoff SK, Reed DB. Health, work, and safety of farmers ages 50 and older. *Geriatr Nurs.* 2005;26(5):304-308.
73. Lim GW, Belton KL, Pickett W, Schopflocher DP, Voaklander DC. Fatal and non-fatal machine-related injuries suffered by children in Alberta, Canada, 1990-1997. *Am J Ind Med.* 2004;45(2):177-185.
74. Hwang SA, Gomez MI, Stark AD, St John TL, May JJ, Hallman EM. Severe farm injuries among New York farmers. *Am J Ind Med.* 2001;40(1):32-41.
75. Rosecrance J, Rodgers G, Merlino L. Low back pain and musculoskeletal symptoms among Kansas farmers. *Am J Ind Med.* 2006;49(7):547-556.
76. Gomez MI, Hwang S, Stark AD, May JJ, Hallman EM, Pantea CI. An analysis of self-reported joint pain among New York farmers. *J Agric Saf Health.* 2003;9(2):143-157.
77. Back pain among persons working on small or family farms--eight Colorado counties, 1993-1996. *MMWR. Morbidity and mortality weekly report.* 1999;48(15):301-304.
78. CDC. Childhood Agricultural Injury Prevention Initiative: Childhood Agricultural Injury Survey Results - NIOSH Workplace Safety and Health Topic. 2015; http://www.cdc.gov/niosh/topics/childag/cais/default.html. Accessed October 12, 2015.
79. Hendricks KJ, Goldcamp EM, Hendricks KJ, Goldcamp EM. Injury Surveillance for Youth on Farms in the U.S., 2006. *Journal of Agricultural Safety and Health.* 2015;16(4):279-291.
80. The National Agricultural Workers Survey, Employment & Training Administration (ETA) - U.S. Department of Labor. 2015; http://www.doleta.gov/agworker/naws.cfm. Accessed June 25, 2015.
81. Wang S, Myers JR, Layne LA. Injuries to hired crop workers in the United States: a descriptive analysis of a national probability survey. *Am J Ind Med.* 2011;54(10):734-747.
82. CDC - Looking to the Future for Agriculture Injury Surveillance at NIOSH - Agruiculture, Forestry and Fishing Resources. 2015; http://www.cdc.gov/niosh/agforfish/aginjurysurv.html. Accessed June 29, 2015.
83. NORA Council. NORA AFF. 2008; http://www.cdc.gov/niosh/nora/comment/agendas/AgForFish/pdfs/AgForFishDec2008.pdf. Accessed June 25, 2015.
84. Smidt M. A description of forest industries and occupations with focus on forestry workers' jobs and injury and illness surveillance. *34th Council on Forest Engineering, June.* 2011.
85. Slappendel C LI, Kawachi I, Marshall S, Cryer C. Factors affecting work-related injury among forestry workers: A review. 1993;24(1):19–32.
86. Wolf C DG. Logging Work Injuries in Appalachia. *U.S. Department of Agriculture, Forest Service, Northeastern Forest Experiment Station.* 1978.
87. Kawachi I, Marshall S, Cryer, C., Slappendel, C., Laird, I., and Wright, D. The epidemiological diagnosis of injury to forest workers: the final and amalgamated report. *Forest Industry Research Group to the Accident Compensation Corporation.* 1991(*Milestone 6 report*).
88. Stover B, Wickizer TM, Zimmerman F, Fulton-Kehoe D, Franklin G. Prognostic factors of long-term disability in a workers' compensation system. *J Occup Environ Med.* 2007;49(1):31-40.

89. Lynch SM, Smidt MF, Merrill PD, Sesek RF. Incidence of MSDs and neck and back pain among logging machine operators in the southern U.S. *J Agric Saf Health.* 2014;20(3):211-218.
90. Analysis HRa. *Assessing the adequacy of fatality and injury reporting in forestry.* 2014. 200-2013-M-55833.
91. Schoenman JA, Mueller CD. Rural implications of Medicare's post-acute-care transfer payment policy. *J Rural Health.* 2005;21(2):122-130.
92. NEMSIS. History of NEMSIS. *National Emergency Medical Services Information System (NEMSIS)* http://www.nemsis.org/theProject/historyofNemsis.html. Accessed January 11, 2016.

CHAPTER 2
METHODOLOGY

The Purpose of this Chapter

While information introduced in the first chapter outlines the need for improved surveillance in agriculture and logging, this chapter seeks to detail the population, materials, and methods to improve injury surveillance methodology for these industries. The focus on low-cost, ongoing methods has brought the use of existing electronic administrative datasets to the forefront. Pre-hospital care reports (PCR) and hospital data (emergency department, inpatient, and outpatient), will be at the center of building a system that can be successfully operated in years to come. The purpose of this research is to establish ongoing injury surveillance that will guide public health activities related to reducing morbidity and mortality in agriculture and logging.

Population at Risk

Farmers and loggers share some similarities; both work in the natural environment, use heavy machinery, are at the mercy of changing weather patterns, and provide the raw goods consumers need for life's necessities (food, shelter, heat). However, there are substantial differences between these industries, and the specific occupations within them.

Agriculture

The United States Department of Agriculture (USDA) conducts the Census of Agriculture every five years through the National Agricultural Statistics Service (NASS).[1] This extensive census covers all enterprises earning $1,000 or more per year from the sale of agricultural products.[2] The census is mandatory for all farmers who fall within these parameters. NASS staff conducts active follow-up for non-responders to encourage the highest response rate possible. The most recent data available from this census is from 2012. Conducting such census activities can be traced back to 1840, when agricultural questions were first posed in the national census.[3] In addition to the census of agriculture, NASS also conducts special surveys and censuses such as the Census of Aquaculture and the Organic Survey (both of which have been conducted three separate times).[4]

There are approximately 2.1 million farms in the United States.[5] Nationally, 914,527,657 acres of land is under use for farming, which is over 1.4 million square miles. Of these farms, 8.3 percent are located in the Northeast (ME, NH, VT, MA, RI, CT, NY, NJ, DE, PA, MD, WV). In the United States, those employed in agriculture account for less than one percent of the population, and those living on farms account for about two percent of the population. The most recent Census of Agriculture shows a national decrease in small and midsized farms, with an increase in the number of large farms (2,000 or more acres) to 82,207 from 80,393 in 2007.[5] For the Northeast from 2007 to 2012, half of the states saw an increase in the number of farms (ME, NH, VT, MA, RI and CT) and half saw a decrease (NY, NJ, DE, PA, MD and WV). The median size of Northeast farms is 134 acres, with the largest median size being in Delaware (208 acres) closely followed by New York with 202 acres, and the smallest in Rhode Island (56 acres).[5]

In the Northeast, Pennsylvania accounts for the most land in farm use, totaling 7,704,444 acres. While Pennsylvania has the most land in agricultural use, New York's median farm size is larger, at 202 acres, as opposed to Pennsylvania's 130 acres.[5] However, in number of farms, Pennsylvania ranks first for the Northeast, with over 59,309. New York and Pennsylvania account for the majority of Northeast agriculture, in both number of farms and agricultural acreage, though other states in the region contribute, in both number of farms and variety of operations. One unique aspect of Northeast agriculture versus other regions of the country (such as the mid-western states) is the wide variety of agricultural operations. Crops grown in the Northeast include corn, soybeans, alfalfa, wheat, vegetables (onions, lettuce, peppers, squash, herbs, tomatoes), and fruits (apples, pears, berries, peaches). Common livestock include dairy cows, beef cattle, hogs, chicken and other poultry, as well as horses, llamas, alpacas, sheep and goats.[5]

Overall, 61 percent of primary operators have a job off the farm.[5] Working off the farm is common in order to secure additional steady income, as well as fringe benefits, perhaps most importantly, health insurance. According to a 1982 survey conducted by the USDA, farmers felt as though full-time farming would not provide the financial security they desired (91% of respondents), and that farming was inherently risky (70% of respondents). Lack of pensions and insurance were also major factors.[6] Most farm households earn the majority of income from off-farm sources, an estimated $57,378 for 2013, according to the USDA Economic Research Service.[7] Median net farm income, on the other hand, was projected to have a loss of $2,799.[7] So, for many, working off the farm is an economic necessity.

Recent trends in agriculture show a shift to a more diversified workforce, including for principle operators. Nationally, there are more minority principle operators than ever before. The largest increase has been for principal operators of Hispanic origin, from 55,570 in 2007 to 67,000 in 2012.[5] Also on the rise are American Indian, Asian, and Black farm owners. Between 2007 and 2012, there was a six percent decrease in the number of women principal farm operators, to 288,264 in total.[5] According to the USDA, women-run farms tend to be more diversified than operations run by men. This also includes 'other livestock farms', many of which are horse farms and hay operations. Geographically, regions with the largest percentage of women farmers include the Northeast states, as well as the western portion of the country.[5] Generally, larger grain farms in the Midwest are still operated by men.

One of the most widely recognized changes in demographics is the aging of the farm population. The average age of a farmer in the 2012 census was 58.3, increasing by over one year from the 2007 census.[5] Older farmers are generally at greater risk for traumatic injury and also face problems from chronic disease. Research focused on the aging farm population has increased over the past few decades as the median age continues to climb. Age-related physical limitations often prove the most dangerous for farmers. Additionally, when injured, they are more likely to have a slower recovery or die from their injuries than their younger counterparts.[8] The most common fatal injuries in this older age group include tractor overturns, tractor run-overs, and falling from moving farm equipment.[9] The healthy worker effect has a strong influence in this population. Moreover, older farmers' cultural capital (also known as human capital) consisting of acquired knowledge, attitudes and values, greatly enhances their work, which allows many older farmers to successfully farm well into their seventies and eighties.[8] While older farmers are more likely to own their acreage, younger farmers (under 45 years old) increasingly work land that they do not own.[5]

Migrant Workers

While much data on farm owners and workers are gathered from the Census of Agriculture (conducted every five years), the same cannot be said of migrant farmworkers. Gathering information on migrant farmworkers for denominator data is challenging. Migrant farmworkers are often missed in official government counts due to legal status and immigration issues. While the census asks farm owners to report the number of migrant workers, it is widely acknowledged that significant underreporting occurs. Farm owners may be hesitant to count undocumented workers for fear of being caught harboring illegal aliens. The Department of Labor's National Agricultural Workers Survey (NAWS) was the only source of data that records the legal status of noncitizen farmworkers. Unfortunately, in 2015, the US government announced the elimination of the NAWS survey, along with several other longstanding agricultural worker surveys.[10] In 2009, about fifty percent of farmworkers were unauthorized (undocumented workers), another 20 percent were authorized non-citizens (e.g. H2A programs) and 30 percent were US citizens.[11]

Farm Size

Many differences exist between small family-owned operations and large corporate farms. One of the most striking of these is the age and safety of equipment. Financial constraints and slim profit margins make purchasing new equipment unrealistic for many farm families. Farms with eleven or more employees are under the jurisdiction of the Occupational Safety and Health Administration (OSHA), and must abide by strict safety rules, including mandated rollover protective structures for every tractor in use. Small farms (under eleven employees) are exempt from OSHA inspection and tend to have less safe equipment.[12] In general, planned obsolescence does not apply to farm equipment. It is common for farmers to use equipment over fifty years old. Often, older equipment was not manufactured to the same safety standards of today, and many have not retrofitted their equipment to include these features.

Logging

Generally, logging is classified in one of two ways, either manual or mechanical. Manual logging is characterized by the use of chainsaws to fell trees, and the use of tractors or skidders (and occasionally horses) to move logs out of the wood lot. It is generally accepted that manual logging is higher risk than mechanized logging. Mechanized logging can be defined as the use of heavy machinery to fell, limb, buck and load logs. The machinery operator is contained within the cab of the equipment, thereby benefitting from more protection than workers on the forest floor. Mechanized logging can further be broken down into cut-to-length logging and tree-length logging.[13] Cut-to-length logging involves the tree being cut down and processed (limbed, bucked, cut to saw lengths) by a highly specialized feller-processor. Tree-length logging involves additional pieces of specialized machinery, including a feller-buncher, which cuts down the tree and places them in piles, a grapple skidder, which moves the piles of trees at one time; and a slasher, which bucks the trees into lengths. Chipping operations function like other mechanical operations, with the exception of a large chipping machine instead of bucking logs into saw lengths.

Lastly, logging can be further classified by the type of tree that is being harvested. Hardwoods, also known as angiosperms, are trees that lose their leaves and are generally denser than softwoods, though exceptions to this rule exist. Softwoods, or gymnosperms, are needle bearing trees that do not lose their leaves, and are generally less dense than hardwoods.[14] Often such species grow separately; however, mixed forests exist and are logged as such.

Workforce size

While there is currently no census specifically related to logging, information regarding some basic logging demographics can be found in the American Community Survey.[15] There are an estimated 50,913 logging workers in the United States, with over 5,000 estimated to work in the Northeast.[16] Total annual wages for U.S. logging workers is $2,078,371,708.[16] These estimates are derived from the United States Census.[17] The average age of logging workers is 42.2 according to a 2013 BLS survey.[18] Only 2.8 percent of the total logging workforce is female.[19] This is substantially lower than the percentage of women in agricultural work where women make up one quarter of the working population for crop and animal production.[19] Details regarding the logging industry in the Northeast are sparse, as few studies have focused on this area.

Changes in Workforce

The transition to a mostly mechanized logging process has led to substantial decreases in workforce numbers; however, productivity has continued to increase.[20] This is primarily due to advancements in machinery design and capability. These advancements have generally made work safer, though it has been suggested that mechanization is correlated with higher rates of obesity and sedentary lifestyles.[21]

Industry Definition

The North American Industry Classification System (NAICS) is a commonly used standard by which the United States, Canada, and Mexico classify industries for economic analyses, and it is commonly used by occupational health professionals to standardize statistical analyses in workplaces. See Appendix B for more industry and occupational classification scheme information.

Agriculture

For the purpose of the injury surveillance system, agriculture is taken to mean the following: Activities that correspond to the North American Industry Classification System (NAICS) code 11.[22] This encompasses activities, locations and livestock/machinery typically associated with agricultural production for sale, and excludes gardening and landscaping, and similar non-commercial activities. Though more difficult to define, since farms can range from small hobby operations to large-scale industrial operations, a "farm" for the purpose of this surveillance system is defined as: a place where $1,000 or more of agricultural products are produced and sold (or normally would be sold)".[23]

Logging

"Logging" is defined as all activities necessary for the raising and felling of trees for commercial purposes, which corresponds to NAICS code 113310, a subset of agricultural NAICS code 11. There is no dollar value threshold set for forestry sales in order to receive this classification.[22]

Definition of Agricultural and Logging Injury

A "traumatic agricultural or logging injury" is defined as occurring when energy is transferred to an individual a) from an agricultural or logging source [e.g., tractor, bull, or chainsaw], b) while in an agricultural or logging location, c) while doing an agricultural or logging activity which is severe enough to require medical attention. Excluded are injury events that occur on the farm/logging site involving sources NOT associated with farm or logging activity at any time [e.g., skateboard, barbecue grill, etc.].

DATA SOURCES

Two primary administrative data have been chosen for inclusion in the surveillance study: pre-hospital care reports and hospital data (inpatient, emergency department, and outpatient). Both of these datasets are routinely aggregated by state health departments and may be requested for research use. In addition, these datasets are expected to be in use for many years to come.

PRE-HOSPITAL CARE REPORTS (PCR)

According to state law, whenever an ambulance travels to any location responding to an emergency call, the operators are required to complete a pre-hospital care report (PCR), also known as an ambulance report. These reports contain a standard set of variables, which include the date and county of event, type of injury, resulting injuries, and whether the call was made to a residence, a farm, an industrial site or another type of location. Previous research[24] has shown that nearly all "farm" records could be classified as to event type and source. In recent years, all states have formally agreed to transition to electronic reporting of these data through an arrangement with the National Emergency Medical Services Information System (NEMSIS).[25] The NEMSIS initiative has helped standardize the variables collected by EMS agencies throughout the country, which greatly benefits the injury surveillance system.

States are required to retain a standard set of variables, and have the option of adding additional fields as needed. One such additional field is the 'narrative'. The narrative is the responder's written (typed) account of the entire EMS response, from the chief complaint called into 911, to the impression of the patient, their care, and ultimately, their outcome. Regardless whether a patient is transferred to a hospital (most common), refuses transport (not common, but seen in minor injuries/illnesses), transferred to another EMS agency (air or ground), or transferred to the custody of a medical examiner or coroner (death), a PCR report is generated. Many, but not all, states retain the narrative text field in state aggregated PCR datasets. These narratives are often rich in detail, and provide greater insight into nature of the event. In addition, these data are captured at the scene and provide a firsthand account of the injury or illness event.

These data are typically available within 24-48 hours of the ambulance call being completed if transferred to the state's dataset electronically. Permission to obtain these data are granted by the Bureau of Emergency Medical Services, and an overarching Institutional Review Board (IRB) must also approve the request.

Hospital Data

Hospital data are maintained for a number of public health assessment purposes, and these data are available electronically in most states. The primary purpose of this dataset is to document everything about the patient and their condition that is necessary to facilitate payment to the hospital for services rendered. These data are also used extensively for research on many topics. Each record captures data on the patient from their initial entry into the facility through when they are discharged. Data elements include patient demographics, diagnosis and procedure codes, external cause of injury codes, and payment information. External cause of injury codes designate a location, event type, and source of injury in the patient's record. Information is retained for emergency department, outpatient, and inpatient visits. Hospital discharge data is typically available for research use within one to three years of the patient's visit.

Surveillance System Design

Electronic administrative data from 2008-2010 will be used in the study. Though PCR data is available up to 2014 (updated every twenty-four hours in state databases), hospital data typically lags behind due to data compilation (hospitals importing records to the state) and quality verification before the data can be released. The starting year of 2008 was chosen because it was one of the first years that PCR data was routinely collected electronically in the selected states. It is therefore in keeping with the principle that the surveillance system will use only existing electronic administrative data.

In order to identify the most effective methodology for linking pre-hospital care reports (PCR) and hospital records, researchers used four specific variables to match-merge the two datasets: gender, date of birth, patient zip code, and admission (ambulance run) date. This combination of variables lends a high degree of specificity to the case match. Prior to this matching, duplicate records in individual datasets were removed. Once matching is completed, the patient's course of treatment can be linked across datasets (e.g., ambulance run resulted in ED or inpatient visit).

Cases identified through keyword search were visually inspected to verify agricultural and logging relatedness. Cases identified thorough E-codes were further verified using ICD-9 diagnostic codes. The keyword algorithm was refined to reduce the volume of cases needing visual inspection. Cases were coded using the Occupational Injury and Illness Classification

System (OIICS) developed by the Bureau of Labor Statistics.[26] This coding scheme was revised in 2012 and includes hierarchical structure to code nature of injury, part of body, event or exposure, and source of injury.

Acquiring Data

Working with Outside Agencies

Data use agreements were established with the Maine Health Data Organization (MHDO), Maine Bureau of Emergency Medical Services, New Hampshire Bureau of Public Health Statistics & Informatics, and the New Hampshire Bureau of Emergency Medical Services. This process required IRB approval with each entity, since the data being applied for contained personal identifiers. Both EMS datasets were downloaded directly from the respective state's ImageTrend (EMS Software)[27] web portal using a password assigned specifically to our research group. Hospital data from New Hampshire was transferred using a secure file transfer website, again accessed with a private user name and password. Maine hospital data was received on a password protected CD via US mail. Annual renewals are required for continued use of these data.

IRB

This research has been reviewed by the Institutional Review Boards of both the Mary Imogene Bassett Hospital and the University at Albany. Additionally, as mentioned in the previous section, approval was also granted by each participating state's IRB or data use board. Approval for a waiver of informed consent has also been received, as we will be using secondary data and will have no contact with study subjects. All data is stored on secure servers at Bassett

Medical Center, with several layers of protection, including restricting access to appropriate study personnel.

Data Storage

NYCAMH/NEC uses appropriate and reasonable systems to ensure that only properly authenticated persons access its confidential records. Network components that provide pathways to information systems that contain sensitive information are of a commercial quality and configured to eliminate inappropriate network traffic to these systems. Operating systems are "hardened" before sensitive information is stored on the workstation or server. Such hardening includes installation of a secure disk file system, closing of unnecessary access ports, application of appropriate patches, and removal or disablement of unnecessary services so the information contained on the system is protected from unauthorized access, damage or deletion. Logging of system activity is enabled at the operating system level on all systems containing sensitive information such as confidential records. Data is kept both on secure SAS servers, as well as on restricted data folders (password protected) within NYCAMH/NEC's system.

Mechanics of Surveillance Database

The following section describes the process in which data were cleaned, match-merged, and subset to cases of interest. These processes were coordinated by the author in collaboration with the principal investigator, data manager, research assistant and coding team.

Eliminating Duplicates

These data were imported into SAS 9.3 to facilitate elimination of duplicates and case-matching. The first step in creating the surveillance database involved eliminating duplicates within both the EMS and hospital data sets. This was done by first sorting the file on the four available record identifiers (gender, DOB, zip code, and ambulance run date/admission date). SAS lag functions were then used to create a sub-file of potential duplicates. These were reviewed on a case-by-case basis by a research assistant to confirm whether they were true duplicates or not. Follow-up visits in hospital data (matching on DOB, gender, and zip code were reviewed in an analogous manner.

In cases of duplicates where both records contained identical information, one record was deleted at random. In cases where the duplicates contained different clinical information (e.g., differing ICD9 codes) the data were appended to create a single record that captured all available information.

Match-Merging

The EMS and hospital data sets were merged in SAS using the same four variables that were used to eliminate the duplicates. This matching process accounted for late night ambulance runs, where the patient may have been admitted early the following day, by performing a second match-merge with the EMS run date incremented by one.

Creating Keyword Search Algorithm

Keywords used to identify agricultural and logging cases were derived from the NIOSH NORA Dictionary of Agriculture, Forestry and Fishing (AFF),[28] in conjunction with input from a farmer, a safety specialist, and an occupational epidemiologist. The initial set of keywords contained 134 agricultural and logging terms, and was designed to encompass a broad spectrum of possible injury scenarios. Those keywords that were found in at least one record are listed in

Appendix C. To maximize the number of records available for examination of the keywords, data from both New Hampshire and Maine were processed through the keyword list.

For all records in which the keyword "field" was found, a random sample of 500 was visually inspected to determine if the injury event was truly related to agriculture or logging. This sampling was necessary due to the logistic impossibility of examining the large number of records (over 15,000) containing this keyword. For all other keywords, all records containing them were inspected. The proportion of cases containing each keyword that were actually related to agriculture or logging was then summarized. A keyword was deemed unnecessary for continued use in the search algorithm if it did not return any true agricultural or logging cases in at least 100 instances of the keyword being present in a narrative.

Sub-setting to Agricultural and Logging Injury Events

Using Existing Variables

For PCR records, if the farm location checkbox was marked, this was taken as a preliminary indication of agricultural relatedness. For hospital data, agricultural events were identified using ICD-9 external cause of injury codes. Specifically, ICD-9 codes E849.1 (farm location), and E919.0 (contact with agricultural machinery [provided the following location E-codes are not included: E849.3 (industrial location), E849.4 (recreation/sport), E849.6 (public building), E849.7 (residential institution)]) denoted agricultural injuries. There are currently no E-codes for logging in the ICD-9 system.

Using Free Text

Each PCR record that contained at least one of the 134 keywords was visually inspected by a research assistant. This was done in an excel spreadsheet where the narrative was read to determine if the case actually involved agriculture or logging. Each record was marked with a Y, N, or U with "Y" signifying that the record was a true agricultural or logging case, "N" signifying it was not an agricultural or logging case, and "U" denoting "undetermined". This process was completed at a rate of approximately 300 records per hour. The undetermined cases (U) were brought to the research team (principal investigator, research coordinator, research assistant) and reviewed until a consensus regarding agricultural or logging relatedness was reached.

Verifying Identified Cases

For injury events with match-merged PCR and hospital data, the linked record was inspected to check for coding mistakes and discrepancies in both databases. If the linked records were not congruent, they were reviewed by the research team in a manner analogous to free text review (see last section). In certain rare cases, it was noted that the hospital diagnostic codes in a matched PCR-hospital record contradicted the account of the incident in the PCR. In these cases the records were not considered to be a match. While keystroke errors are tolerated in the hospital record, if the majority of diagnostic, procedure and E-codes do not match the overall description in a linked PCR, the non-agricultural or logging linked record was removed.

Coding Cases

Occupational Injury and Illness Classification System (OIICS Coding)

Cases were coded separately by two experienced coders, and then compared for consistency. This process involved eight reviewers assigned to four sets of coding pairs. For a given text string, both reviewers attempted to assign each of the four OIICS levels. The two coders then discussed all discrepancies and attempted to resolve the differences. Any discrepancies that could not be resolved by the pair were brought before the entire group of eight coders for review, and consensus by majority. In order to maximize efficiency, this process was computerized wherever possible. Specifically, the text strings were electronically transferred to a Microsoft Access database along with a large number of template variables including age, gender, patient zip code, and general information on the location of the incident. However, the OIICS codes that are determined from visual inspection of the text string must be hand entered. This was accomplished by assignment of four key variables – OIICS number (up to 4 digits).

Farm and Agricultural Injury Classification Coding (FAIC Coding)

For PCR cases that are confirmed agricultural or logging injuries, two independent researchers examined the narrative in order to classify them into FAIC codes. These were "forestry or logging" (FAIC-2) or "agriculturally related" (FAIC-1, FAIC-4, FAIC-5, FAIC-6, FAIC-7, FAIC-8, FAIC-9, FAIC-10) based on definitions found in the ASABE Farm and Agricultural Classification (FAIC) Code.[29] Cases producing different classifications between the two observers were adjudicated by a third independent researcher.

Sample Size

The study attempted to collect information on the entire population of study subjects within Maine and New Hampshire. The probability of capturing every injury (the entire population) was considered unlikely; therefore, the analyses were performed as if the data had been obtained from a sample rather than the entire population. Means and proportions are therefore reported with their corresponding standard error and summarized using confidence intervals.

The composite file contained more injury records and more complete data than could be obtained through any single data file. When validated, this method could be applicable to the surveillance of a number of other occupational conditions in other settings. These analyses represent the first step in an evaluation of the proposed surveillance system.

The denominator for the purposes of calculating rates was taken from the 2007 Census of Agriculture. Although the 2012 Census of Agriculture has been released, the 2007 census numbers was used because 2007 is closer in time to the 2008-2010 injury data that was analyzed.

Following Chapters

The following chapters contain several manuscripts reporting detailed results of establishing the surveillance system. These include the relative number of cases identified in relation to the number of records reviewed, the overlap between the PCR and hospital datasets, the specificity to which data can be coded using the OIICS, advantages of this system, system limitations, summaries of logging and agriculture injury events in the Northeast, and suggested future directions.

Chapter 3: Data Processing and Case Identification in an Agricultural and Logging Morbidity Surveillance Study: Trends over Time

Chapter 4: Coding Data from an Agriculture and Logging Injury Surveillance System – Feasibility and Granularity using the Occupational Injury and Illness Classification System (OIICS)

Chapter 5: Trends in Non-Fatal Agricultural and Logging Injury in Maine and New Hampshire: Results from a Low-Cost Passive Surveillance System

References

1. USDA. NASS - National Agricultural Statistics Service. 2015; http://www.nass.usda.gov/. Accessed August 17, 2015.
2. Prager D. USDA Economic Research Service - Glossary. 2012; http://www.ers.usda.gov/topics/farm-economy/farm-household-well-being/glossary.aspx. Accessed August 20, 2015.
3. US Census Bureau CHS, ASCD. Census of Agriculture - History - U.S. Census Bureau. 2008; https://www.census.gov/history/www/programs/agriculture/census_of_agriculture.html. Accessed August 17, 2015.
4. USDA. USDA - NASS, Census of Agriculture - Publications. 2015; http://www.agcensus.usda.gov/Publications/. Accessed August 17, 2015.
5. USDA. *2012 Census of Agriculture Full Report.* 2015.
6. Bartlett P, Hallberg MC, Findeis JL, Lass DA. *Motivations of Part-Time Farmers.* Ames, Iowa: Iowa State University Press; 1991.
7. ERS. Farm Household Well-being. 2013; http://www.ers.usda.gov/topics/farm-economy/farm-household-well-being/farm-household-income.aspx#.Uf6RldLOvK0. Accessed August 8, 2013.
8. Cole HP, Donovan TA. Older farmers' prevalence, capital, health, age-related limitations, and adaptations. *J Agromedicine.* 2008;13(2):81-94.
9. Kelley J. Donham AT. *Agricultural Medicine: Occupational and Environmental Health for the Health Professions.* Wiley-Blackwell; 2006.
10. CDC - Looking to the Future for Agriculture Injury Surveillance at NIOSH - Agruiculture, Forestry and Fishing Resources. 2015; http://www.cdc.gov/niosh/agforfish/aginjurysurv.html. Accessed June 29, 2015.
11. Hertz T. USDA Economic Research Service - Background. 2011.
12. OSHA. 07/16/2007 - Small farming operations and exemption from OSHA enforcement activity under CPL 02-00-051. 2007; https://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=INTERPRETATIONS&p_id=25803. Accessed August 4, 2013.
13. Klase W, Steele TW. Logging Methods for Wisconsin Woodlands. 2011; http://dnr.wi.gov/topic/ForestManagement/documents/pub/FR-778.pdf. Accessed August 16, 2015.
14. Lee S, Ralfo AL, Gagnon JL. Trees and Water. *Virginia Cooperative Extension* 2015; http://pubs.ext.vt.edu/ANR/ANR-18/ANR-18NP_pdf.pdf. Accessed August 20, 2015.
15. Census U. American Community Survey. 2010; http://www.census.gov/acs/www/#. Accessed June 25, 2015.
16. BLS. Private, NAICS 1133 Logging, All States and U.S. 2014 Annual Averages, All establishment sizes Source: Quarterly Census of Employment and Wages - Bureau of Labor Statistics. 2015; http://www.bls.gov/cew/apps/table_maker/v4/table_maker.htm#type=0&year=2014&qtr=A&own=5&ind=1133&supp=0. Accessed June 18, 2015.
17. US Census County Business Patterns (NAICS). NAICS 1133 Paid Employees for Pay Period Including March 12. 2013; http://censtats.census.gov/cgi-bin/cbpnaic/cbpdetl.pl. Accessed October 1, 2015.
18. BLS. Employed persons by detailed occupation and age. 2015; http://www.bls.gov/cps/occupation_age.htm. Accessed August 20, 2015.

19. BLS. Employed persons by detailed industry, sex, race, and Hispanic or Latino ethnicity, 2014. 2015; http://www.bls.gov/cps/cpsaat18.htm. Accessed August 20, 2015.
20. Kelly MC GR. The Price of Freedom: Single-Operator Timber Harvests in the Northeast United States. 37th Council on Forest Engineering Annual Meeting; 2014; Moline, Il.
21. Befort CA, Nazir N, Perri MG. Prevalence of Obesity Among Adults From Rural and Urban Areas of the United States: Findings From NHANES (2005–2008). *J Rural Health.* 2012;28(4):392-397.
22. US Census Bureau. North American Industry Classification System. *US Census Bureau* 2002; http://www.census.gov/cgi-bin/sssd/naics/naicsrch?chart_code=11&search=2002. Accessed June 14, 2016.
23. NASS. *Census of Agriculture. United States Summary and State Data, 2007.* National Agricultural Statistics Service;2009.
24. Earle-Richardson GB, Jenkins PL, Scott EE, May JJ. Improving agricultural injury surveillance: a comparison of incidence and type of injury event among three data sources. *American Journal of Industrial Medicine.* 2011;54(8).
25. NEMSIS. History of NEMSIS. *National Emergency Medical Services Information System (NEMSIS)* http://www.nemsis.org/theProject/historyofNemsis.html. Accessed January 11, 2016.
26. BLS. Occupational Injury and Illness Classification System (OIICS). 2015; http://wwwn.cdc.gov/wisards/oiics/. Accessed June 25, 2015.
27. ImageTrend Software Development - Web-based Data Management Solutions. 2014; http://www.imagetrend.com/. Accessed June 25, 2015.
28. NORA Council. NORA AFF. 2008; http://www.cdc.gov/niosh/nora/comment/agendas/AgForFish/pdfs/AgForFishDec2008.pdf. Accessed June 25, 2015.
29. ASABE. Farm and Agricultural Injury Classification (FAIC) Code. *ASAE S575.1*: ASABE; 2002.

CHAPTER 3
MATCH MERGE

Data Processing and Case Identification in an Agricultural and Logging Morbidity Surveillance Study: Trends over Time

Erika Scott, Erin Bell, Nicole Krupa, Liane Hirabayashi, and Paul Jenkins

Abstract

Agriculture and logging are dangerous industries, and though data on fatal injury exists, less is known about non-fatal injury. Thus, establishing a non-fatal injury surveillance system is a top priority for occupational health researchers. Existing electronic data sources were explored as a sustainable and low-cost option for ongoing surveillance of occupational injury. Pre-Hospital Care Reports (PCR) and hospitalization data are two promising options. Using free-text and location codes from PCRs, along with ICD-9-CM external cause of injury codes (E-codes) from hospital data, we created a surveillance system that tracks farm and logging injuries. Using three years of data (2008-2010) from both Maine and New Hampshire, 1,585 agricultural and logging injury events were identified. This system is limited to traumatic injury for which medical treatment is administered, and is limited by the accuracy of coding and spelling. In conclusion, such a system has the potential to be both sustainable and low cost.

Background

Agricultural workers and loggers do some of the most dangerous work in the United States.[1-3] These workers provide the important natural resources for products we use every day, including our food, energy, building materials and paper. Farmers and loggers make up a small percentage of our overall working population (around two percent), but they account for a disproportionate percentage of workplace deaths, where approximately one in every 13 workplace fatalities are attributed to an agricultural or logging related injury.[4] According to the Bureau of Labor Statistics (BLS) Census of Fatal Occupational Injuries (CFOI), the fatality rate in 2014 was 26.0 deaths per 100,000 full-time equivalent workers (FTE) for agricultural workers [Standard Occupational Classification (SOC) 45-209X] and 109.5/100,000 FTE for logging workers (SOC 45-4020).[4] Both fatality rates contrast starkly with the 2014 all-worker fatal injury rate of 3.3/100,000 FTE.[4] While fatalities frequently make news headlines, non-fatal workplace injuries are usually more numerous, and account for physical suffering, lost work time, medical bills, and economic hardship for the employer and family. Developing an effective surveillance tool to address morbidity in these industries is the first step in understanding what potential safety interventions will have the greatest impact in reducing injury. In addition, such a surveillance tool would have the ability to measure the effect of an intervention over time. While programs exist to enumerate non-fatal workplace injury, they are not particularly effective in measuring the burden of injury for the agriculture, forestry, and fishing sector.[5-8]

The Survey of Occupational Injuries and Illnesses (SOII) is a survey for which employers submit information related to workplace injuries and illnesses. The survey is sent to over 200,000 businesses each year. SOII excludes self-employed workers, family workers, and agricultural employers with fewer than 11 employees; therefore, effectively excluding many Northeastern farms.[9] Leigh and colleagues estimate that the SOII missed an average of 77.6% of non-fatal farm injuries and illnesses across the United States in 2011, more than any other industry.[10] Therefore, the SOII is not a practical source of surveillance data for agriculture or small logging operations. Even more, several states do not have large enough sample sizes to generate stable estimates for SOII (NH, CO, ID, MS, ND, SD, OK, and RI).[11]

While many industries can rely on workers' compensation (WC) data for morbidity estimates, this is challenging for agriculture and for small logging businesses.[5] In agriculture, this is because many smaller farms have family workers that are not required to have coverage by workers' compensation.[5] In logging, many small businesses are purposefully run by solo operators so they can avoid the cost of insurance. This lack of data necessitates the creation of a low cost, sustainable surveillance system for agriculture and logging.

Hospital Data and Pre-Hospital Care Reports

Electronic health datasets are organized records that medical entities maintain for administrative purposes. For the purposes of injury surveillance, hospital records and pre-hospital care reports are particularly attractive because they contain the necessary variables to identify a logging or agricultural injury.[12,13] Hospital records (inpatient, outpatient, and emergency department) contain demographic data along with diagnostic, procedure, and external cause of injury (E-codes) codes. These records also contain data related to what entity is expected to pay for care (workers' compensation, private insurance, Medicare, Medicaid, self-pay, etc.).

While hospital data have generally been maintained electronically for a number of years, pre-hospital care reports (PCR) are newer to the electronic age. In the late 2000's, the National Emergency Medical Services Information System (NEMSIS) established a memorandum of understanding to have all states transition to electronic record keeping.[14] This change allows us to more easily review patient records. In addition, many states opt to retain narrative text fields that contain important bits of information about the injury event. While our past efforts have tested these methods to measure morbidity in a single state, over one or two years,[15-19] it is unknown whether this approach can be applied across more than one state and for multiple years. Therefore, we will assess the variability in these electronic datasets over a three-year period for two neighboring Northeast states, Maine and New Hampshire, and their ability to produce true agricultural and logging related injury cases.

METHODS

POPULATION AT RISK

The subjects are all persons at risk for agricultural or logging injury in the states of Maine and New Hampshire. By using this definition, certain individuals, such as visitors to a farm, may contribute events to the numerator despite the fact that they would not be included in the denominator used in calculating rates. The 2007 Census of Agriculture estimates the number of workers (all operators + hired labor) in agriculture in New Hampshire and Maine to be 12,042 and 28,697 respectively, for a total of 40,739 workers between the two states (Table 1).

DATA ACQUISITION

Permission to receive the hospital data was obtained from the New Hampshire Bureau of Public Health Statistics & Informatics and the Maine Health Data Organization. Access to the Emergency Medical Service (EMS) data was granted by the Bureau of Emergency Medical Services of both states. EMS data were downloaded directly from each state's EMS portal. The hospital data (inpatient, outpatient, and emergency department) were received via secure file transfer over the internet. All data were analyzed using SAS 9.3.1 (Cary, NC).

DEFINITION OF AGRICULTURAL AND LOGGING EVENTS

A "traumatic agricultural or logging injury" was defined as energy transferred to an individual a) from an agricultural or logging source [e.g., tractor, bull, or chainsaw], b) while in an agricultural or logging location, or c) while doing an agricultural or logging activity, which is severe enough to require medical attention. Excluded are injury events that occur on the farm/logging site involving sources NOT associated with farm or logging activity at any time [e.g., skateboard, barbecue grill, etc.]. This was considered appropriate, as injuries requiring medical care are typically more severe than those not requiring medical care, and are frequently work-related. This definition has also been used in other studies.[12,13] Injuries that occur on the farm but do not involve farm activities (e.g., fall from skateboard, burn from barbecue grill, etc.) are not considered to be injury events for the purposes of the surveillance system. This logic applies to non-work injury events on logging sites or woodlots, though these are expected to be uncommon.

In the hospital data, agricultural injuries were identified using ICD-9-CM external cause of injury codes E849.1 (farm location) or E919.0 (contact with agricultural machinery) in cases where the following location E-codes were not included: E849.3, E849.4, E849.6, E849.7 (non-farming locations). Diagnostic codes related to an injury (or injuries) must have been included in the patient record. Logging injuries cannot be identified using ICD-9-CM ecodes.

Two methods were used to identify agricultural and logging injuries in the EMS data. If the PCR's location code was marked as "farm" or if an agricultural or logging term (list available in Appendix C) was found in the PCR's narrative this was taken as an initial indication of an agricultural or logging injury. These preliminary cases were retained for further review.

For every case identified using a keyword, the entire text string was visually inspected for true agriculture or logging relatedness. The proportion of true agricultural or logging cases was then summarized for each keyword. Keywords that produced at least one true agricultural or logging case are shown in Table 2. For the purposes of this system, a true case is considered to be a record indicating agriculture or logging that has been visually inspected and verified to be a traumatic injury involving agricultural or logging events.

Treatment of Duplicates and Merging of Data Sets

Hospital records (emergency department, outpatient, and inpatient appended together) with duplicate data for the patient's gender, admission date, ZIP code, and date of birth were combined into a single record. PCR records were treated in an analogous manner. Each dataset was evaluated for follow-up visits for the same patient by linking date of birth, zip code and gender. Follow-up records for the same injury were not included in case counts.

These combined, de-duplicated, hospital data were then match-merged with the de-duplicated PCR dataset using gender, admission date, ZIP code, and date of birth. An exact match was required on all four variables (patient's gender, admission date, ZIP code, date of birth). To account for hospital admissions occurring early the following morning of the ambulance transport, the matching process was repeated allowing the hospital admission date to be one day later than the ambulance transport date. This protocol was approved by the Institutional Review Board of the Mary Imogene Bassett Hospital and The University at Albany.

Classifying Agriculture and Logging

For PCR cases that were confirmed to be of agricultural or logging origin, the primary author and a member of the research team examined the narratives independently in order to classify them into FAIC (Farm and Agricultural Classification) codes.[20,21] These codes were: "forestry or logging" (FAIC-2) or "agriculturally related" (FAIC-1, FAIC-4, FAIC-5, FAIC-6, FAIC-7, FAIC-8, FAIC-9, FAIC-10). Cases producing different classifications between the author and research team member were adjudicated by a third team member.

Data Analysis

The cumulative incidence of injury in the agricultural population was taken as the sum of the true agricultural cases divided by the total number of agricultural workers in Maine and New

Hampshire. This same methodology was used to estimate cumulative incidence of injury in the logging population. These morbidity estimates were also performed for each state and each year separately. Results are reported as the number of injuries identified per 1,000 workers. Though data from the 2012 Census of Agriculture are available, we chose to use data from the 2007 Census as it more closely aligns with the numerator data (2008-2010).

The ICD-9-CM system does not contain specific E-codes for logging related injury (woodlot location, or use of logging specific heavy machinery); therefore, logging related cases could not be identified in the hospital data. Because of this, estimating the number of logging injuries in these two states required using a correction factor. It is assumed that the proportion of ambulance transports that result in hospitalizations in agriculture and logging will be similar. This is supported by a few factors: both types of injuries occur from work in rural locations, distance to health care facilities is likely to be similar, and the types of injuries sustained both result from vigorous outside activities, often involving heavy equipment and tools. The correction factor was estimated using two methods: 1) by the overall ratio between hospital and PCR cases in agricultural cases, and 2) by the ratio between hospital and PCR cases for specific injury classifications (e.g. open wounds, intracranial injuries, traumatic injuries to bones, etc.).

RESULTS

HOSPITAL DATA

Between Maine and New Hampshire, a total of 18,259,580 hospital records (emergency department, inpatient, and outpatient) were received for the three years (5,297,453 for 2008, 5,600,282 for 2009, and 7,361,845 for 2010). A total of 562 or 0.0031% (562/18,259,580), true agricultural cases were identified over the three-year period (Table 3).

For these 562 true cases, 524 (93.2%) had the initial episode of care documented in only one of the three hospital files. The emergency department and outpatient file combined had the majority of the documented cases (529/562=94.1%), with the inpatient file having only 42 cases (42/562=7.5%). These percentages sum to more than 100% because of records found in both the inpatient and outpatient data.

Pre-Hospital Care Reports

For 2008-2010, 767,060 pre-hospital care reports were generated in Maine and New Hampshire, which resulted in the identification of 1,041 agricultural or logging cases (973 agricultural and 68 logging). Therefore, 0.136% of all PCRs contained a true case (Table 3). Of the 126 keywords, 96 were found in at least one narrative field with the other 30 not found (Table 1). A total of 25,945 of the 767,060 PCR records contained at least one of these 96 keywords. Fifty of the 96 keywords did not produce any true agricultural cases in the 3,645 records in which they were found.

Of the 1,041 true agricultural or logging cases, 744 (71.5%) were identified from the keyword/visual inspection process alone, 92 (8.8%) were identified by the farm location code alone. Thus, nearly twenty percent (205 cases) were identified by both the farm check box and keyword/visual inspection. Therefore, the keyword/visual inspection process identified a total of 949 true cases (true cases could contain more than one keyword). Using these 96 keywords without visual inspection would result in a false positive rate of (25,945-949)/25,945 = 96.3%. In addition to the 92 true agricultural cases identified by the farm location code alone and the 205 identified by the farm location code and keyword search combined, there were 303 other records with the farm box checked that were not true cases. Thus, the false positive rate for the farm check box was 303/600 = 50.5%. Conversely, of the 767,060 PCRs not having the farm box checked, 744 were true cases, resulting in a false negative rate of .00097 or .1%.

Data Overlap

Of the 562 hospital records indicating an agricultural injury, 46 (46/562=8.2%) could be linked to a PCR, and 18 of these 46 PCRs also indicated agricultural origin. Of the 1,041 PCR records indicating an agricultural or logging injury, 675 (675/1,041=64.8%) could be linked to a hospital record. Among those 675, only 18 hospital records (18/675=2.7%) indicated agricultural origin.

Of the 1,585 unique agricultural (1,517) and logging (68) cases, 734 (46.3%) would not have been identified without free text review of the PCR narrative. Minimal overlap was identified between PCRs and hospital records where both files indicated a farming incident. This was also true with regard to the overlap of the hospital datasets (Figures 1 and 2). Generally, more overlap was found between the Maine data files than those from New Hampshire.

Estimates of Injury Rates

With a total of 1,517 agricultural injuries identified (an average of 506 per year), the average annual cumulative incidence of injuries requiring medical attention in the New Hampshire and Maine agricultural industry is 506/40,739 = 0.0124 = 12.4/1,000 workers (Table).

Of the 1,517 agricultural cases, 973 could have been identified using only the PCRs. From this, it is estimated that combining the PCRs with the hospital data produces a count for agricultural cases that is 1.56 (1,517/973) times the count obtained from the PCR data alone. Applying this same multiplier to the count of logging cases obtained from the PCRs alone would yield an estimate of 35 logging injuries per year [(68 total injuries/3 years) × 1.56 multiplier]. Estimating the correction factor using the ratio of specific injury classifications yielded a multiplier range from 1.03 for nonclassifiable cases, to 5.20 for traumatic injuries to muscle, tendons, ligaments, etc. which would result in an expected 122 logging injuries, or 41 logging injuries per year [(122 total injuries/3 years) x (range 1.03-5.20)]. Thus, the overall correction factor aggregated over the specific injury categories is 122/68 = 1.79. (Table 4). Thus, the overall correction factor aggregated over the specific injury categories is 122/68 = 1.79. Data from the 2006-2010 American Community Survey estimates an average of 975 logging workers in New Hampshire, and 2,380 logging workers in Maine during the period 2008-2009. Therefore, the extrapolated average annual cumulative incidence of injuries requiring medical attention in the New Hampshire and Maine logging industry is 35 to 41/3,355 = 0.0104 to 0.0122 = 10.4 to 12.2/1,000 workers.

Discussion

We found that morbidity could be estimated for multiple years and states using hospital and PCR data. Our estimate of 12.4 injuries per 1,000 agricultural workers is close to a recent estimate made by Landsteiner and colleagues in Minnesota (14.0-18.5/1,000 individuals living and/or working on the farm).[13] Many of these injuries translate to lost work-time and significant medical bills. Leigh and colleagues have estimated that agricultural occupational injuries cost on average 30 percent more than other types of occupational injuries.[22] Considering the total cost of agricultural injury (in 1992) was estimated at $4.57 billion dollars, farm injuries pose a significant economic burden.[22] In addition, our estimates for morbidity are one order of magnitude larger than federal estimates for mortality, which we would expect according to previous injury research.[23] Further, our estimates remained relatively stable (as would be expected over three years without major safety interventions or changes in work practices) despite a large fluctuation in hospital record volume for New Hampshire in 2010. This fluctuation (2,456,398 in 2010 versus 748,611 in 2008 and 755,230 in 2009) was due to the fact that New Hampshire began to aggregate additional types of claims and discharges in 2010. This underscores the importance of our de-duplication procedures in maintaining a surveillance system over time. Further, verifying an actual change in injury counts must include examining the primary data that feeds the surveillance system. Changes in classifications and coding schemes could also artificially increase or decrease case counts.

There are certainly advantages to using more than one primary data source. Using both hospital data and pre-hospital care reports increases case catchment beyond what either single source could be used to identify, as had been shown in our previous studies.[15,18,19] Though some overlap did exist between PCR and hospital records, it was rare to find congruence in the indication of agricultural or logging origin.

The minimal overlap among agricultural and logging injuries between the PCR and hospital files has several implications. Initially, it was hoped that these combined (hospital + PCR) would provide a more thorough explanation of the injury than could be found in either file separately. However, if the number of these combined records is negligible, it may be advantageous to simply estimate the degree of overlap using a correction factor. This will forego the labor-intensive match-merging process in the interest of maximizing the cost-effectiveness of the system. Additionally, if the degree of overlap between hospital and PCR records in other states is also negligible, as we have seen for Maine and New Hampshire, a correction factor could be used to estimate morbidity in those states as well. Such a correction factor would also prove useful in states where public use datasets do not provide sufficient variables to perform a match-merge.

Strengths

A major strength of this methodology is that it takes advantage of administrative datasets that are likely to be available for many years to come. A system with longevity allows the assessment of trends over both short and long periods. Longer-term rate reductions might be linked to a particular safety intervention or policy change. An example of such a trend is the reduction in motor vehicle mortality and morbidity after the introduction of automobile safety features such as seatbelts and airbags.[24]

The system does not require any contribution on the part of the data releasing authority beyond providing the dataset. In addition, our past research, which combined active surveillance with passive data collection, showed that the passive system captured many more injury events than the active one.[15] Therefore, we feel using this type of passive system is a reasonable approach for low-cost surveillance in these industries.

Initial visual review of both PCR and hospital codes, along with PCR narratives, indicate that these data contain enough detail to classify the nature of the injury, part of the body, source of the injury, and the type of event. We plan to explore the extent to which these data can be coded using the Occupational Injury and Illness Classification System (OIICS) for our next analyses.

Limitations

This system has many advantages but it is not without limitations. For hospital data, the system relies on a coding scheme that is not optimized for occupational health. External cause of injury codes (E-codes) provide detail as to the nature or source of injury; however, the incentive for accurately recording these codes is not the same as for procedure codes. Hospital revenue is directly linked to procedure codes; therefore, these are typically under much more scrutiny.

Moreover, in New Hampshire, new data storage does not allow for the separation of an emergency department case from an outpatient visit, therefore limiting the ability to characterize an injury on that status.

Using the ICD-9-CM system (used in hospital coding pre-October 2015), it is not possible to identify a logging injury in hospital data because, while a farm location code is an option, an analogous logging or woods location does not exist. This is expected to change with the transition to ICD10-CM coding. The tenth revision of the International Classification of Diseases dramatically expands the choices for external cause of injury codes. Some codes will relate to logging injury, but in the meantime, logging injury counts must be supplemented by a correction factor developed from agricultural injuries.

We recognize that this system is likely to underestimate the true morbidity in agriculture and logging as not all serious injuries are captured by this system. Even when a patient receives care by emergency medical technicians or hospital staff for an agricultural or logging injury, there is a chance their records will not indicate farming or logging activities. Therefore, such a record could not be identified to be a case of interest. Secondly, farmers and loggers may seek treatment at local doctor's offices or self-treat, and these injuries would not be captured by the surveillance system. We intend to explore new administrative databases as they become available to fill these gaps, in addition to conducting other analyses to validate our findings.

When calculating injury rates, the denominator is likely to be smaller than the actual number of people at risk for agriculture and logging injuries. Several factors relate to this including changing population size (fluctuations between censuses), census underreporting, and the inability to accurately estimate farm bystanders, household members, and children living on the farm. Our data shows a decrease in the number of injuries in 2009 for both agriculture and logging, followed by slight increases in 2010, with rates for NH agriculture highest among the states and industries (Figure 3). However, until a sufficient number of data points are available to estimate annual random variability, it is not possible to say whether or not these fluctuations are due to real changes.

CONCLUSIONS

For 2008-2010, 1,585 agricultural and logging related injuries were identified in Maine and New Hampshire using our surveillance methods. Despite the restructuring of the New Hampshire hospital data file in 2010, which more than tripled the number of records received from this source, our methodology was able to account for this and identify a relatively consistent number of agricultural injuries over these three years. This underscores the ability of the system to account for large changes in its source data and still track changes over time. The cumulative incidence of injuries in Maine and New Hampshire was found to be 12.4/1,000 for agricultural workers, compared to 10.4 to 12.2/1,000 for logging workers. These estimates are consistent with other recent agricultural morbidity estimates.

ACKNOWLEDGEMENTS

This research is funded by the Centers for Disease Control and Prevention National Institute for Occupational Safety and Health (CDC-NIOSH) Grant Number 2U54OH007542-11. We thank the Maine Emergency Medical Services Bureau, the New Hampshire Bureau of Emergency Medical Services, the Maine Health Data Organization, and the New Hampshire Department of Health and Human Services (DHHS) for the use of their data. We are grateful to Deb Dalton, Emma Garrison, Melissa Horsman, Alissa Kent, Pam Tinc, and Marybeth Vargha for their help with record verification and coding. Additionally, we appreciate the thoughtful review of Edward Fitzgerald, PhD, Kitty Gelberg, PhD, and Lloyd Wilson, PhD.

References

1. Goldcamp EM. Work-related non-fatal injuries to adults on farms in the U.S., 2001 and 2004. *J Agric Saf Health.* 2010;16(1):41-51.
2. Hayden GJ, Gerberich SG, Maldonado G. Fatal farm injuries: a five-year study utilizing a unique surveillance approach to investigate the concordance of reporting between two data sources. *J Occup Environ Med.* 1995;37(5):571-577.
3. Wolf C DG. Logging Work Injuries in Appalachia. *U.S. Department of Agriculture, Forest Service, Northeastern Forest Experiment Station.* 1978.
4. BLS. Fatal Work Injuries, 2014. 2015; http://www.bls.gov/iif/oshwc/cfoi/cfch0013.pdf. Accessed December 2, 2015.
5. Costich J. Who pays for agricultural injury care? *J Agromedicine.* 2010;15(1):54-57.
6. Leigh JP. Economic burden of occupational injury and illness in the United States. *Milbank Q.* 2011;89(4):728-772.
7. Boden LI, Ozonoff A. Capture-recapture estimates of nonfatal workplace injuries and illnesses. *Ann Epidemiol.* 2008;18(6):500-506.
8. Azaroff LS, Levenstein C, Wegman DH. Occupational injury and illness surveillance: conceptual filters explain underreporting. *Am J Public Health.* 2002;92(9):1421-1429.
9. BLS. Survey of Occupational Injuries and Illnesses (SOII) Respondents Home Page. 2015; http://www.bls.gov/respondents/iif/. Accessed June 25, 2016.
10. Leigh JP, Marcin JP, Miller TR. An Estimate of the U.S. Government's Undercount of Nonfatal Occupational Injuries. 2004;46(1):10-18.
11. BLS. Injuries, Illnesses, and Fatalities. 2015; http://www.bls.gov/iif/. Accessed June 25, 2015.
12. Forst L, Erskine T. Farm injuries in Ohio, 2003-2006: a report from the emergency medical services prehospital database. *J Agric Saf Health.* 2009;15(2):171-183.
13. Landsteiner AM, McGovern PM, Alexander BH, Lindgren PG, Williams AN. Incidence Rates and Trend of Serious Farm-Related Injury in Minnesota, 2000-2011. *J Agromedicine.* 2015;20(4):419-426.
14. NEMSIS. Goals & Objectives. 2015; http://www.nemsis.org/theProject/whatIsNEMSIS/goalsAndObjectives.html. Accessed June 25, 2015.
15. Earle-Richardson GB, Jenkins PL, Scott EE, May JJ. Improving agricultural injury surveillance: a comparison of incidence and type of injury event among three data sources. *Am J Ind Med.* 2011;54(8):586-596.
16. Scott EE, Earle-Richardson G, Krupa N, Jenkins P. A correction factor for estimating statewide agricultural injuries from ambulance reports. *Ann Epidemiol.* 2011;21(10):767-772.
17. Scott EE, Krupa NL, Sorensen J, Jenkins PL. Electronic merger of large health care data sets: cautionary notes from a study of agricultural morbidity in New York State. *J Agromedicine.* 2013;18(4):334-339.
18. Scott EE, Krupa NL, Horsman M, Jenkins PL. Estimation of agricultural and logging injury incidence in Maine using electronic administrative data sets. *J Agromedicine.* 2015;20(2):195-204.
19. Scott EE, Hirabayashi L, Krupa NL, Sorensen JA, Jenkins PL. Developing Surveillance Methodology for Agricultural and Logging Injury in New Hampshire Using Electronic Administrative Data Sets. *J Occup Environ Med.* 2015;57(8):866-872.
20. ASABE. Farm and Agricultural Injury Classification (FAIC) Code. *ASAE S575.1*: ASABE; 2002.
21. Yoder AM, Murphy DJ. Evaluation of the Farm and Agricultural Injury Classification Code and follow-up questionnaire. *J Agric Saf Health.* 2000;6(1):71-80.
22. Leigh JP, McCurdy SA, Schenker MB. Costs of occupational injuries in agriculture. *Public Health Rep.* 2001;116(3):235-248.

23. Baker SP; O'Neill B GM, Li G. *The Injury Fact Book.* Second ed. New York: Oxford University Press; 1992.
24. CDC. Achievements in Public Health, 1900-1999 Motor-Vehicle Safety: A 20th Century Public Health Achievement. Vol 48: MMWR; 1999:**369-374**.

TABLES AND FIGURES

Table 1. Denominator Data, Agriculture (2007) and Logging (2010)

	Maine	New Hampshire	Total
Agricultural Workers	28,697	12,042	40,739
Logging Workers	2,380	975	3,355

Table 2. Keywords/Abbreviations

Keyword	# True Cases	# PCR Records w/ Keyword	Hit Rate	Keyword	# True Cases	# PCR Records w/ Keyword	Hit Rate
BEATERS	0	1	0.000	CHUTE	1	61	0.016
DRIVE LINE	0	1	0.000	SPREADER	0	62	0.000
LIVESTOCK	0	1	0.000	PASTURE	10	67	0.149
SILAGE	1	1	1.000	CROP	0	69	0.000
SPRING POLE	1	1	1.000	LOADER	3	73	0.041
THREE POINT HITCH	1	1	1.000	HAY	10	78	0.128
TEDDER	0	1	0.000	SHEEP	1	82	0.012
CULTIVAT	0	1	0.000	RAM	0	92	0.000
BALING	0	2	0.000	HITCH	0	99	0.000
CHOKER	0	2	0.000	CHAINSAW	44	106	0.415
UNHITCH	0	2	0.000	TROUGH	0	112	0.000
DEBARK	0	3	0.000	IMPLEMENT	0	113	0.000
KICKBACK	0	3	0.000	AUGER	0	115	0.000
SHEAVE	0	3	0.000	BULL	2	117	0.017
SKIDSTEER	0	3	0.000	BIND	0	140	0.000
POULTRY	0	3	0.000	CABLE	0	154	0.000
MANURE	1	4	0.250	ARCH	0	165	0.000
HARROW	1	5	0.200	BUNKER	0	197	0.000
LIMBING	2	5	0.400	STALL	6	198	0.030
SKID STEER	1	5	0.200	TIMBER	0	198	0.000
KICKER	1	6	0.167	CHOPP	2	216	0.009
AMISH	0	7	0.000	COMBINE	0	223	0.000
BUGGY	0	7	0.000	IRRIGATION	1	223	0.004
FERTILIZER	0	7	0.000	HOG	3	225	0.013
TIE DOWN	1	8	0.125	BARN	26	242	0.107
PIPELINE	0	8	0.000	BUCK	0	257	0.000
METHANE	0	9	0.000	CALV	0	261	0.000

Table 2 Continued. Keywords/Abbreviations

SPRAYER	0	10	0.000	PLANT	0	298	0.000
CLEANSER	0	10	0.000	BALE	1	304	0.003
DIGGER	0	10	0.000	PLOW	0	324	0.000
BREEDING	0	12	0.000	SKIDD	11	351	0.031
SCRAPER	0	13	0.000	STRAW	1	351	0.003
PESTICIDE	0	19	0.000	WAGON	2	362	0.006
VACUUM PUMP	0	19	0.000	MOWER	0	367	0.000
PRUNING	3	20	0.150	FENCE	21	436	0.048
BOBCAT	1	20	0.050	FEED	3	579	0.005
SANITIZER	0	20	0.000	CHICKEN	1	626	0.002
PTO	1	21	0.048	CHAIN	63	669	0.094
ENTANGLEMENT	0	21	0.000	ANIMAL	1	684	0.001
GOAT	2	27	0.074	CART	5	904	0.006
RAKE	0	29	0.000	BLADE	5	945	0.005
WINCH	2	32	0.063	TRACTOR	75	967	0.078
PIG	0	33	0.000	HORSE	343	969	0.354
COW	5	40	0.125	FARM	49	1516	0.032
UDDER	0	40	0.000	PEN	6	1890	0.003
PENS	0	47	0.000	WOODS	42	2071	0.020
SILO	0	52	0.000	LOG	79	3163	0.025
WEDGE	3	57	0.053	TREE	105	3572	0.029

Table 3. Results from Case Identification, Maine and New Hampshire, 2008-2010

		Maine	New Hampshire	Both
Total Hospital Records				
	2008	4,548,842	748,611	5,297,453
	2009	4,845,052	755,230	5,600,282
	2010	4,905,447	2,456,398	7,361,845
	Total	14,299,341	3,960,239	18,259,580
Agricultural Injuries from Hospital Records				
	2008	154	30	184
	2009	108	39	147
	2010	190	41	231
	Total	452	110	562
Total Pre-Hospital Care Reports (PCR)				
	2008	101,745	101,309	203,054
	2009	155,023	119,138	274,161
	2010	157,570	132,275	289,845
	Total	414,338	352,722	767,060
PCR Injuries Identified Using Farm Location Only				
	2008	36	31	67
	2009	9	5	14
	2010	6	5	11
	Total	51	41	92
PCR Injuries Identified From Keyword Search Only				
	2008	93	151	244
	2009	114	110	224
	2010	142	134	276
	Total	349	395	744
PCR Injuries Identified From BOTH Farm Location and Keyword Search				
	2008	24	42	66
	2009	28	48	76
	2010	31	32	63
	Total	83	122	205
*TOTAL Farm and Logging Injuries (Hospital and PCR)**				
	2008	302	252	554
	2009	253	201	454
	2010	366	211	577
	Total	921	664	1,585

*Some injuries are identified in both hospital data and PCR or multiple records exist for one patient's injury event, therefore these totals do not reflect the total records identified, but individual injury events.

Table 4. Calculation of Logging Multiplier using Specific Injury Classifications

		Agriculture				Logging	
OIICS Injury Code	Code Meaning	Frequency PCR	Frequency Hospital	Frequency Total	Correction factor (total/#PCR)	Frequency PCR	Expected Frequency Total (using Correction Factor)
10	Traumatic injury	40	12	52	1.30	4	5
11	Traumatic injury to bones, spinal cord, nerves	93	66	159	1.71	10	17
12	Traumatic injury to muscle, tendons, ligaments, joints	20	84	104	5.20	2	10
13	Open wounds	168	140	307	1.83	15	27
14	Surface wounds and bruises	37	97	135	3.65	3	11
16	Intracranial injuries	36	7	43	1.19	6	7
18	Multiple traumatic injuries and disorders	64	89	149	2.33	11	26
19	Other traumatic injuries and disorders	401	49	447	1.11	13	14
99	Nonclassifiable	35	1	36	1.03	4	4

Table 5. Estimated Injury Rates for Agriculture and Logging, Maine and New Hampshire, 2008-2010

	Maine				New Hampshire			
	Agriculture	Injury Rate/1,000 Workers	Logging	Injury Rate Range/1,000 Workers	Agriculture	Injury Rate/1,000 Workers	Logging	Injury Rate Range/1,000 Workers
2008	290	10.1	12	7.9-9.0	238	19.8	14	22.4-25.7
2009	242	8.4	11	7.2-8.3	194	16.1	7	11.5-12.8
2010	351	12.2	15	9.8-11.3	202	16.8	9	14.4-16.5
	Average	10.2	Average	8.3-9.5	Average	17.5	Average	16.0-18.3

Figure 1. Maine True Case Overlap, 2008-2010

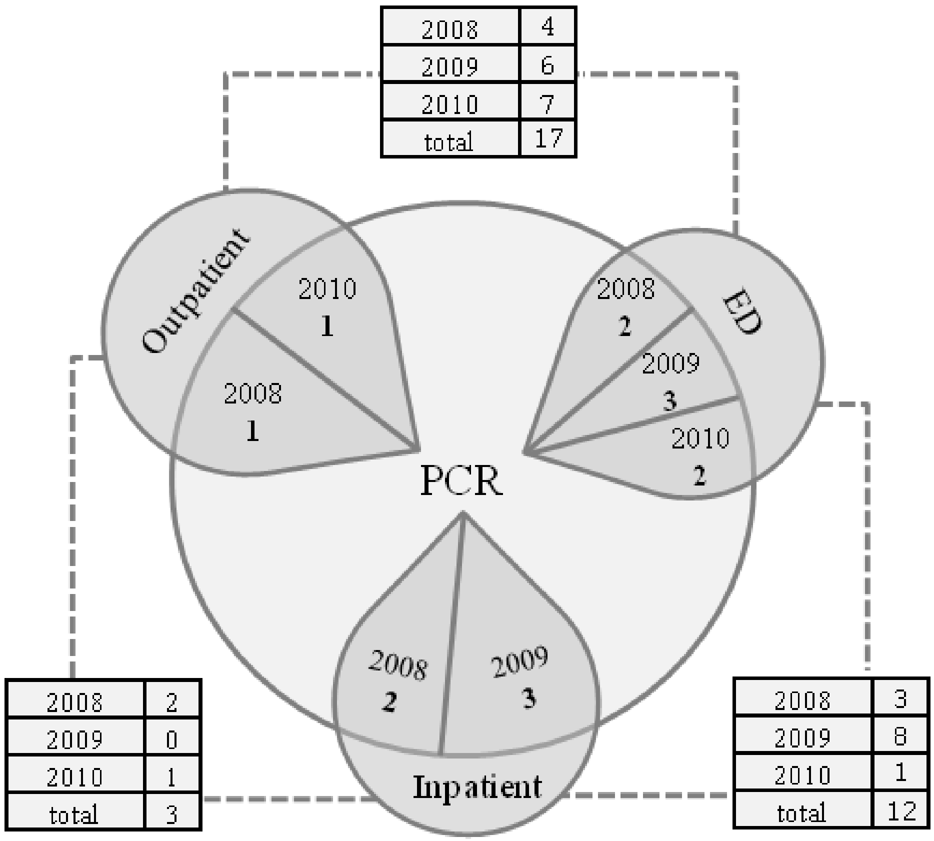
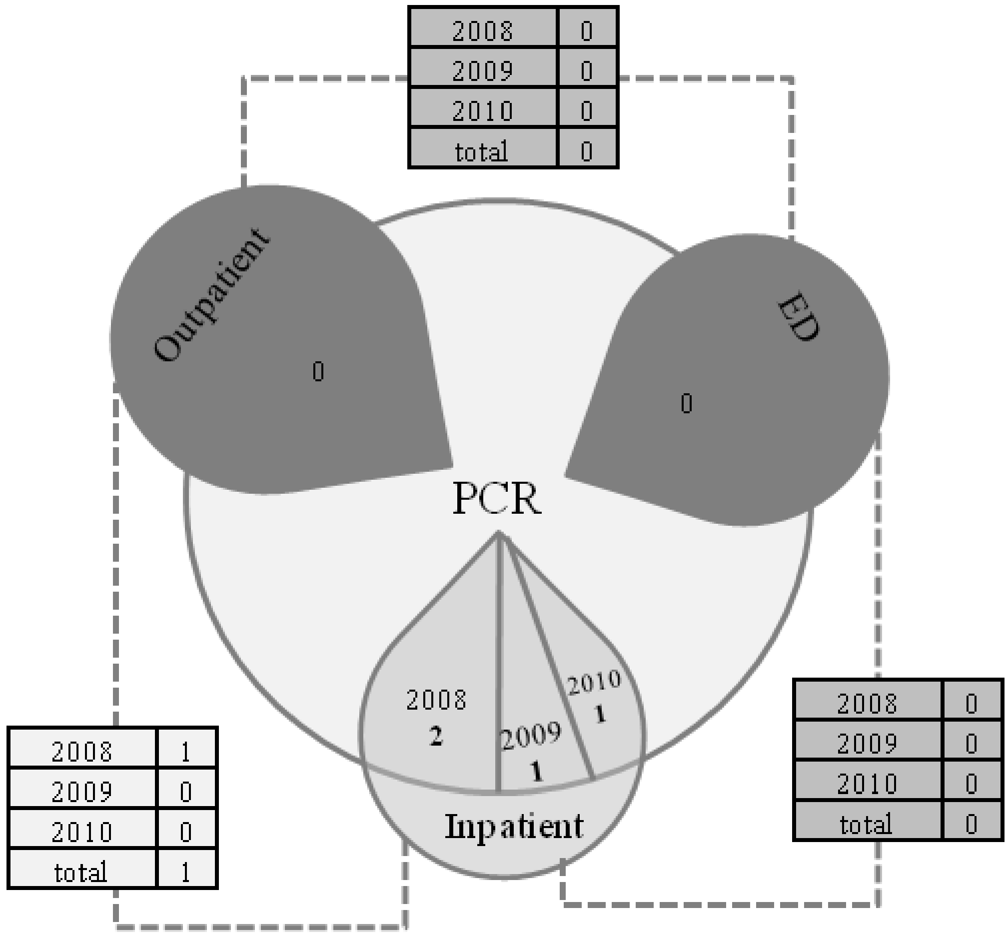


Figure 2. New Hampshire True Case Overlap, 2008-2010



CHAPTER 4
GRANULARITY

Coding Data from an Agriculture and Logging Injury Surveillance System – Feasibility and Granularity using the Occupational Injury and Illness Classification System (OIICS)

Erika Scott, Erin Bell, Liane Hirabayashi, Nicole Krupa, and Paul Jenkins

ABSTRACT

Little is known about traumatic non-fatal agricultural and logging injuries in the Northeast. Recently, the trend towards electronic storage of administrative datasets, especially pre-hospital care reports (PCR), has created new opportunities for occupational injury surveillance. Electronic hospital data files can provide additional detail on the injury outcome and treatment. We have been developing methods to identify agricultural and logging injuries in these datasets. To combine the information for the PCR and hospital datasets, it is necessary for the agricultural and logging injuries found in each file to be coded using a standard classification scheme, in this case, the Occupational Illness and Injury Classification System (OIICS). The two coders were instructed to assign the most specific OIICS code possible. A total of 570 hospital records and 1,041 PCRs were double coded using the OIICS scheme. Coder agreement was best among the part of body (division, level 1) in the hospital data (κ=0.8933) and for the type of event (division, level 1) in the PCR data (κ=0.8405). For both hospital data and PCRs, the percentage of unclassifiable data increased as coding became more specific. Maintaining a useful surveillance system depends on weighing tradeoffs between cost, efficiency, and the utility of the information produced. Hospital and PCR data can be coded using the OIICS coding scheme to provide useful information for injury surveillance in agriculture and logging. Based on these data, it is our recommendation that the type of event and source of injury be coded to level 4 wherever possible. For the nature of injury and body part, it may be most cost efficient to code only as far as level 2.

Background

Farms and woodlots are dangerous worksites. Recent estimates for fatal injury among agricultural workers (26.0 deaths per 100,000 full-time equivalent workers (FTE)) and loggers (109.5/100,000 FTE [Standard Occupational Classification (SOC) 45-209X]) are markedly higher than the all-worker fatal injury rate of 3.3/100,000 FTE.[1] Workplace deaths often garner news headlines in the Northeast; however, little is known about traumatic non-fatal agricultural and logging injuries in this region. These injuries often account for extensive medical treatment, lost work time, and economic hardship for the patient, the patient's family, and the employer. Understanding these injuries through data is the first step in the public health model to reduce such events.[2]

A primary challenge of past research in agricultural and logging injury is that data becomes antiquated, and systems are not able to provide estimates of non-fatal injury events in an ongoing fashion.[3-12] While OSHA records provide data on injuries for major industry,[13] they have limited use in agriculture, especially for small family owned farms, and for small logging operations.[14-16] This is partly due to limited information available for certain vulnerable populations, such as undocumented immigrants and migrant workers.[17] Workers compensation claims are not a particularly effective means of tracking injuries in agricultural due to its limited use on small family farms.[18] In addition, research has shown that workers in agriculture, fishing, and forestry have high rates of work-related injuries, yet low rates of workers compensation claims.[19]

Even the federal Survey of Occupational Injuries and Illnesses (SOII) has limitations in that the military, self-employed individuals, farms with less than 11 employees and Federal Agencies are excluded from the survey.[20] While twenty-six state based Occupational Health

Surveillance Programs are currently funded, the focus on agriculture and logging are not equally represented among all states.[21] Individual states may choose their areas of focus, which are not guaranteed to include agricultural or logging related issues. The NIOSH sponsored National Electronic Injury Surveillance System Work Supplement (NEISS-Work) captures traumatic occupational injuries but it is limited to cases involving consumer products.[5]

Recently, the trend towards electronic storage of administrative datasets, especially pre-hospital care reports (PCR), has created new opportunities for occupational injury surveillance.[22,23] Since many traumatic injuries warrant the use of ambulance transport, PCR records are a rich source of data for characterizing these injuries, as well as the course of pre-hospital treatment. Electronic hospital data files can provide additional detail on the injury outcome and treatment.[24,25] We have been developing methods to identify agricultural and logging injuries in these datasets. This has the potential to be a long-term, low cost method for injury surveillance in these industries.[26-29]

To best understand the information for the PCR and hospital datasets, it would be necessary for the agricultural and logging injuries found in each file to be coded using a standard classification scheme such as the Occupational Illness and Injury Classification System (OIICS). Coding may present a particularly challenging goal, given the very different formats in which the two electronic files are stored. Specifically, these codes must be assigned from the PCR files by reading text descriptions of the injuries and assigned in the hospital files by reviewing ICD-9-CM codes including External Cause of Injury Codes (E-codes). The ability to code these data using a standard scheme for both data sets allows us to calculate morbidity using the four individual OIICS coding scheme definitions. This information could prove useful to injury researchers in designing appropriate safety interventions.

We have recently refined methodology[26-29] for identifying non-fatal agricultural and logging injuries in Maine and New Hampshire using administrative datasets. This paper explores the extent to which these data sets (hospital records and pre-hospital care reports) can be coded using OIICS. In particular, the report seeks to explore the tradeoff between expending additional labor to increase coding specificity versus the practical utility of this increased specificity in informing interventions.

METHODS

The data sets used to identify agricultural and logging injuries were obtained from the New Hampshire Bureau of Public Health Statistics & Informatics, the Maine Health Data Organization and the Bureau of Emergency Medical Services of both states. EMS data were downloaded directly from each state's EMS portal. The hospital data (inpatient, outpatient, and emergency department) were received via secure file transfer over the internet. The process for identifying agriculture and logging injuries from these datasets has been described in detail in previous publications.[26,27] This protocol was approved by the Institutional Review Board of the Mary Imogene Bassett Hospital and The University at Albany.

DATA ELEMENTS

Hospital records contained patient demographics including age, gender, race, ethnicity, and zip code. Additional variables of interest for the surveillance system include diagnostic codes (up to 16 in Maine and up to 25 in New Hampshire) along with external cause of injury codes (E-Codes). Pre-hospital care reports contain similar patient demographics (age, gender, race/ethnicity) and a free text field containing a written description of the injury incident. As the subject of this report is to quantify agreement between coders and the percentage of classifiable cases, follow-up records (subsequent visits for an initial injury episode, n=8) were included in the analyses.

OIICS Coding of True Agricultural and Logging Injury Cases

In 1992, the Bureau of Labor Statistics developed the Occupational Injury and Illness Classification System (OIICS).[30] The OIICS provides a hierarchical coding structure, also referred to as coding trees, for each of the following: the nature of the injury, the part of the body, the type of event, and the source of injury. Secondary source of injury may be assigned if appropriate. For each definition, the hierarchical structure is composed of: division level (level 1), major groupings (level 2), and detailed (levels 3 and 4) levels. In some instances, the third level is the terminal level; however, some fourth levels are available to provide the highest degree of specificity. Detailed coding selection rules aid in the selection of the proper code. An example of this coding scheme can be found in Table 1.

The system was updated in collaboration with the National Institute of Occupational Safety and Health (NIOSH) in 2010, and again in 2012. OIICS coding trees are now readily available online or in downloadable software. This coding scheme is used by many occupational injury programs, including the Census of Fatal Occupational Injury (CFOI),[31] the Survey of Occupational Injuries and Illnesses (SOII),[20] many Workers' Compensation (WC) insurers, and by numerous occupational injury researchers.

While OIICS was designed to be compatible with the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM), no crosswalk exists between OIICS and ICD-9-CM.[32] However, OIICS coding has been used to characterize work-related injuries in emergency department data through the NIOSH Work-Related Injury Statistics Query

System (Work-RISQS).[33] While we are unaware of efforts to apply OIICS coding to PCRs, the coding scheme is frequently used for other free-text based datasets (such as the Census of Fatal Occupational Injury and Workers Compensation), as previously mentioned.[34,35]

A Microsoft Access database was created to facilitate the coding of these data. We provided extensive OIICS training using practice datasets before we began coding the actual cases identified from our surveillance system. Each record was assigned the four OIICS classification categories (nature of injury, part of body, source of injury, and type of event or exposure) by two independent coders. The secondary source of injury was assignable by both coders for 17.0% of all records. It was therefore not considered further.

The two coders were instructed to assign the most specific OIICS code possible. Results from the double coding process were then compared, and any discrepancies were set aside to be resolved. This resolution was completed in discussion between the coding pair. For any cases that could not be resolved by the coding pair, the discrepancy was discussed in a meeting of all eight study team members who assigned the final codes.

DATA ANALYSIS

All analyses described below were performed separately for PCR and hospital data. Wherever possible, the results of these separate analyses were contrasted. All data were analyzed using SAS 9.3.1 (Cary, NC).

The kappa coefficients (κ) quantifying the degree of agreement between the two coders were calculated for each of the four Coding Scheme Definitions (nature of injury, part of body, source of injury, and type of event or exposure). These coefficients were calculated separately for the level 1 (division level), level 2 (major grouping), and levels 3 and 4 (detailed). Following this, the two coders resolved all discrepancies between the codes that were assigned and a single database was created representing their common observations.

Kappa values were evaluated using the guidelines presented by McHugh.[36] Specifically, values greater than 0.90 were considered to indicate near-perfect agreement, 0.80 to 0.90 indicative of strong agreement, 0.60 to 0.79 indicative of moderate, and values below 0.60 indicate poor agreement.

Frequency distributions for each of the four coding scheme definitions were created for both the hospital and PCR dataset. Within each coding scheme a prevalence for each level of detail (levels 1-4) was also created. Unclassifiable cases refer to a record that cannot be assigned a code due to a lack of sufficient detail at a particular level. The percentages in these distributions were calculated under two different assumptions. First, it was assumed that the distribution of un-classifiable cases was the same as those that were classifiable, which resulted in the denominator being only those cases that were classified. Second, the percentages were calculated using all cases whether classified or not. The results from those two methods were contrasted in order to measure the implications of using these differing denominators given the large number of unassignable OIICS codes.

RESULTS

Using our surveillance methodology,[26,27] we identified 1,593 true agricultural (1,525) and logging (68) cases. A total of 570 hospital records and 1,041 PCRs were double-coded using the OIICS scheme. Slight overlap (18 records) exists between the datasets; however, results are presented here for each dataset separately. Coders processed hospital records at an average rate of thirty records per hour (30/hr), whereas PCRs were coded at an average of fifteen per hour

(15/hr). For both hospital records and PCRs, approximately 75% needed secondary review to resolve discrepancies (typically done by the coding duo). All hospital records could be resolved without the need for team review. Team resolution was required for 18 PCR records that could not be resolved by the coding pair. Coding hospital records and PCRs took approximately two working weeks (at 1.0 FTE).

Agreement between coders

Kappa coefficients can be found in Table 2.

Hospital data

Strong agreement was observed among coders for level 1 of body part (κ=0.89) and event type (κ=0.85). Agreement at level 1 for the source of injury (0.74) was moderate. In contrast the kappa value of 0.32 for level 1 of nature of injury indicated poor agreement. This low value may have been indicative of the high penalty for chance agreement for this outcome due to the relatively few categories that were chosen.

With the exception of the nature of injury category, all kappa values declined monotonically in the progression to the more specific detail levels. There was a sharp increase in kappa when going from the level 1 to level 2 of nature of injury, which again may be reflective of the greater number of categories selected, and consequent reduction in the penalty for chance agreement at these more detailed levels. At the fourth level of detail, only the source of injury category was above the threshold for moderate agreement.

PCR

PCR kappa values were generally higher than hospital values at all levels for source of injury, and indicative of strong agreement down to level 3. Agreement was poor when coding for nature of injury among PCR records (κ ranging 0.48 to 0.52). Agreement was moderate for the two highest levels of part of body (levels 1&2), but poor for the two more detailed levels (levels 3&4). For type of event, PCR kappa values were very similar to their hospital counterparts.

UNCLASSIFIABLE DATA

Details of the extent of unclassifiable data in the two separate data sets (hospital: n=570 and PCR: n=1,041) are shown in Table 3. Because these data are intended to be used as a single integrated database, the percentages presented below are from the two datasets combined, with reference to important differences between the two sub datasets where observed.

NATURE OF INJURY

For this category, over eighty percent of the data are available for level 1 (96.7%), level 2 (92.4%), and level 3 (80.1%). In contrast, only 44.1% are available for level 4.

PART OF BODY

As shown in Table 3, high percentages of the data are available for the level 1 (99.6%) and level 2 (88.5%). These percentages drop off precipitously for levels 3 (55.4%) and 4 (4.7%).

TYPE OF EVENT

For level 1, 93.0% of the records were classifiable. However, less than three quarters of the data are available for level 2 (71.2%), level 3 (62.6%) and level 4 (51.1%). As shown, for this category there is considerably more data available at levels 3 and 4 in the PCR data set than in the hospital dataset.

SOURCE OF INJURY

In source of injury, a monotonic decline from level 1 (85.5%), to level 2 (84.2%), to level 3 (63.2%) to level 4 (55.6%) was observed. As was the case with type of event, there was considerable variability in the proportion of data available between the hospital and PCR data sets, with far more data being available at levels 3 and 4 for the PCR data.

Granularity of the Data Set

Percent estimates for the source of injury, nature of injury, part of body, and type of event are shown in Tables 4 through 7. For reasons explained previously, the percentages discussed are from the single integrated data set, with notable differences between hospital and PCR estimates referred to as necessary.

Nature of Injury

In general, there is greater specificity with regard to the nature of the injury in the hospital data set as opposed to PCR. One strong theme that emerges is that of open wounds, which constitutes about 1 in 5 injuries (Table 4).

Part of Body

These data indicate a relative balance in the body part affected between the head (13.4%), the trunk (20.9%), the upper extremities (24.2%), and the lower extremities (22.1%). Involvement of multiple body parts (16.7%) was also common. Almost one-third (31.5%) of the injuries were distributed across the back (8.1%), the hands (9.7%) and the legs (13.7%) (Table 5).

Source of Injury

Approximately half of the source of injury codes fall into the machinery and animal categories, with animal-related injuries accounting for double the percentage of machinery injuries. Depending upon the choice of denominator (all records versus all classifiable records), these two categories account for between 48.9% and 58.1% of all injuries (Table 6).

Type of Event

Transportation incidents and contact with objects and equipment were the two main types of event and accounted for at least two-thirds of all injuries regardless of the choice of denominator. Further, the majority of transportation incidents were due to being thrown from an animal, mainly horses (Table 7).

Discussion

Maintaining a useful surveillance system depends on weighing tradeoffs between cost, efficiency, and the utility of the information produced. The greater the time investment, the greater the detail in the cases identified. However, the extent to which this detail adds to the utility of the system is the key factor to be considered.

One of the main contributors to the cost of this system is the person-time required for visual inspection in the coding process described above. This process was labor intensive, but was necessary in order to obtain the detail whose value was to be judged. Clearly, if this additional effort leads to the case being 'unclassifiable', then the time and money involved cannot be justified. In contrast, if detailed information can be obtained from visual inspection that may inform interventions or identify emerging problems then it may indeed be considered cost effective.

In discussing the cost-effectiveness of employing visual inspection to acquire detail, one must also consider what use will be made of each of the four categories. In planning interventions, the source of injury and type of event are the main drivers. In quantifying the costs of morbidity and mortality one must consider the nature of the injury and the body part involved.[37]

While levels 1 and 2 of the OIICS codes are useful for characterizing overall trends in agricultural and logging injury, emerging trends or the impact of targeted safety interventions may benefit greatly from the detail available at levels 3 and 4. The detail available for the type of event and source injury of injury at OIICS levels 3 and 4 was missing for one third to almost one half of the data for both categories. As seen in the results, this was primarily due to the inability to assign these categories based on only ICD-9-CM E-codes in the hospital data. For the remaining cases where these categories could be assigned at levels three and four, there does appear to be sufficient detail to inform interventions. Lastly, it is important to note that making small distinctions for the third and fourth level does increase the time to code each record.

Many categories at the third and fourth levels comprise only a small percentage of the total cases. These small strata would *not* be the focus of a current intervention. However, by coding to these levels of detail, the surveillance system takes on sensitivity for subtle shifts in injury patterns over time wherein a previously rare event suddenly reaches an incidence where intervention may be necessary.

For using the body part and nature of injury categories to estimate the cost the morbidity in this population, we see that the majority of the data are missing at level 4 for both categories. Further, nearly half of the data are missing at level 3 for the part of the body. This strongly calls into question the cost-effectiveness of coding these two categories to detail levels 3 and 4. This raises the question of whether these costs can be estimated using only levels 1 and 2. We would contend that levels 1 and 2 are in fact sufficient to make cost estimates and characterize the burden of disease. Therefore, if it is necessary to trim costs, it would seem that levels 3 and 4 of these two OIICS categories could be sacrificed.

Limitations

Unfortunately, the study does not have the ability to quantify to the extent to which cases were not identified by the system. While this issue is of particular concern in quantifying the costs and burdens of agricultural and logging morbidity, it does not prevent the system from detecting trends over time provided that the methods used remain constant.

While many cases are identified in the hospital dataset, this dataset may be limited in its ability to be coded using OIICS, especially for type of event and source of injury. Hospital data is a billing dataset, and the E-codes used for this study are not required for financial reasons. E-codes are known to be left unassigned, even when appropriate, and we would lose the ability to identify and code such cases.

CONCLUSIONS

Hospital and PCR data can be coded using the OIICS coding scheme to provide useful information for injury surveillance in agriculture and logging. A delicate balance of effectiveness and efficiency must be sought for the maximum utility of such a surveillance system. While simply coding to levels 1 and 2 would prove faster, and therefore save money, it would sacrifice the potentially important information gained from coding at these more detailed levels.

Exploring techniques that could enhance the coding process, such as Bayesian methods, are a promising step in bridging the gap between efficiency and detail. We also plan to explore the utility of machine coding as our next step, now that we have created a coded dataset from which the computer can 'learn'. Advancements in these methods have already been seen on other occupational injury datasets,[38-41] and we intend to apply them to our agricultural and logging injury surveillance system.

Based on these data, it is our recommendation that the type of event and source of injury be coded to level 4 wherever possible. For the nature of injury and body part, it may be most cost efficient to code only as far as level 2.

Acknowledgements

This research is funded by the Centers for Disease Control and Prevention National Institute for Occupational Safety and Health (CDC-NIOSH) Grant Number 2U54OH007542-11. We thank the Maine Emergency Medical Services Bureau, the New Hampshire Bureau of Emergency Medical Services, the Maine Health Data Organization, and the New Hampshire Department of Health and Human Services (DHHS) for the use of their data. We are grateful to Deb Dalton, Emma Garrison, Melissa Horsman, Alissa Kent, Pam Tinc, Rebecca Russell and Marybeth Vargha for their help with record verification and coding. Additionally, we appreciate the thoughtful review of Edward Fitzgerald, PhD, Kitty Gelberg, PhD, and Lloyd Wilson, PhD.

References

1. BLS. Fatal Work Injuries, 2014. 2015; http://www.bls.gov/iif/oshwc/cfoi/cfch0013.pdf. Accessed December 2, 2015.
2. CDC. Updated Guidelines for Evaluating Public Health Surveillance Systems. *MMWR* 2001; http://www.cdc.gov/mmwr/preview/mmwrhtml/rr5013a1.htm. Accessed August 3, 2015.
3. Day L, Voaklander D, Sim M, et al. Risk factors for work related injury among male farmers. *Occup Environ Med.* 2009;66(5):312-318.
4. DeRoo LA, Rautiainen RH. A systematic review of farm safety interventions. *American journal of preventive medicine.* 2000;18(4 Suppl):51-62.
5. Fox S, Ricketts M, Minton JE. Worker Injuries Involving the Interaction of Cattle, Cattle Handlers, and Farm Structures or Equipment. *J Agric Saf Health.* 2015;21(1):3-12.
6. Hawk C, Donham KJ, Gay J. Pediatric Exposure to Agricultrual Machinery:. *Journal of Agromedicine.* 2010;1(1):57-74.
7. Hoskin A. F, T. A. Miller, W. D. Hanford, S. R. Landes. Occupational injuries in agriculture: a 35-state summary. *NIOSHTIC-2 Publications Search - 20039768* 1988; http://www.cdc.gov/niosh/nioshtic-2/20039768.html. Accessed June 25, 2015.
8. Pratt DS, Marvel LH, Darrow D, Stallones L, May JJ, Jenkins P. The dangers of dairy farming: the injury experience of 600 workers followed for two years. *Am J Ind Med.* 1992;21(5):637-650.
9. Pedersen DH, Wilkins JR, 3rd, Bean TL, Mitchell GL, Crawford JM, Jones LA. Agricultural hazard data from a population-based survey of cash grain farms: Ohio observations. *Applied occupational and environmental hygiene.* 1999;14(5):299-305.
10. Forst L, Erskine T. Farm injuries in Ohio, 2003-2006: a report from the emergency medical services prehospital database. *J Agric Saf Health.* 2009;15(2):171-183.
11. Goldcamp EM. Work-related non-fatal injuries to adults on farms in the U.S., 2001 and 2004. *J Agric Saf Health.* 2010;16(1):41-51.
12. Layne LA, Goldcamp EM, Myers JR, Hendricks KJ. Youth living on Hispanic-operated farms: injuries and population estimates in the U.S., 2000. *J Agric Saf Health.* 2009;15(4):377-388.
13. Olorunnishola OA, Kidd-Taylor A, Byrd L. Occupational injuries and illnesses in the solid waste industry: a call for action. *New solutions : a journal of environmental and occupational health policy : NS.* 2010;20(2):211-223.
14. OSHA. Partial exemption for employers with 10 or fewer employees. - 1904.1. 2015; https://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=STANDARDS&p_id=9632. Accessed June 25, 2015.
15. Boden LI, Ozonoff A. Capture-recapture estimates of nonfatal workplace injuries and illnesses. *Ann Epidemiol.* 2008;18(6):500-506.
16. Azaroff LS, Levenstein C, Wegman DH. Occupational injury and illness surveillance: conceptual filters explain underreporting. *Am J Public Health.* 2002;92(9):1421-1429.
17. Costich J. Who pays for agricultural injury care? *J Agromedicine.* 2010;15(1):54-57.
18. Karttunen JP, Rautiainen RH. Characteristics of and risk factors for compensated occupational injury and disease claims in dairy farmers: a case-control study. *J Agric Saf Health.* 2013;19(3):191-206.
19. Fan ZJ, Bonauto DK, Foley MP, Silverstein BA. Underreporting of work-related injury or illness to workers' compensation: individual and industry factors. *J Occup Environ Med.* 2006;48(9):914-922.
20. BLS. Survey of Occupational Injuries and Illnesses (SOII) Respondents Home Page. 2015; http://www.bls.gov/respondents/iif/. Accessed June 25, 2015.
21. CDC-NIOSH. Extramural Research and Training Programs - State Surveillance Portfolio. 2015; http://www.cdc.gov/niosh/oep/statereports.html. Accessed November 25, 2015.
22. NEMSIS. History of NEMSIS. *National Emergency Medical Services Information System (NEMSIS)* http://www.nemsis.org/theProject/historyofNemsis.html. Accessed January 11, 2016.

23. Landman AB, Lee CH, Sasson C, Van Gelder CM, Curry LA. Prehospital Electronic Patient Care Report Systems: Early Experiences from Emergency Medical Services Agency Leaders. *PLoS One.* Vol 72012.
24. Landsteiner AM, McGovern PM, Alexander BH, Lindgren PG, Williams AN. Incidence Rates and Trend of Serious Farm-Related Injury in Minnesota, 2000-2011. *J Agromedicine.* 2015;20(4):419-426.
25. Johnson CM, Lagares-Garcia JA, Miller SL. When the bough breaks: a 10-year review of logging injuries treated at a rural trauma center in Pennsylvania. *Am Surg.* 2002;68(7):573-581.
26. Scott EE, Hirabayashi L, Krupa NL, Sorensen JA, Jenkins PL. Developing Surveillance Methodology for Agricultural and Logging Injury in New Hampshire Using Electronic Administrative Data Sets. *J Occup Environ Med.* 2015;57(8):866-872.
27. Scott EE, Krupa NL, Horsman M, Jenkins PL. Estimation of agricultural and logging injury incidence in Maine using electronic administrative data sets. *J Agromedicine.* 2015;20(2):195-204.
28. Earle-Richardson GB, Jenkins PL, Scott EE, May JJ. Improving agricultural injury surveillance: a comparison of incidence and type of injury event among three data sources. *Am J Ind Med.* 2011;54(8):586-596.
29. Scott EE, Earle-Richardson G, Krupa N, Jenkins P. A correction factor for estimating statewide agricultural injuries from ambulance reports. *Ann Epidemiol.* 2011;21(10):767-772.
30. BLS. Occupational Injury and Illness Classification System (OIICS). 2015; http://wwwn.cdc.gov/wisards/oiics/. Accessed June 25, 2015.
31. BLS. Census of Fatal Occupational Injuries (CFOI) - Current and Revised Data. 2015; http://www.bls.gov/iif/oshcfoi1.htm#2012. Accessed June 25, 2015.
32. G T, Windau JA, D D. Using the BLS Occupational Injury and Illness Classification System as a safety and health management tool. *Compensation and Working Conditions.* 1996:19-28.
33. CDC. Technical Information Work-RISQS. 2015; http://wwwn.cdc.gov/wisards/workrisqs/techinfo.aspx#caseIdentification. Accessed March 5, 2016.
34. Gorucu S, Murphy D, Kassab C. Occupational and Nonoccupational Farm Fatalities Among Youth for 2000 Through 2012 in Pennsylvania. *Journal of Agromedicine.* 2015;20(2):125-139.
35. Marsh SM, Jackson LL. A comparison of fatal occupational injury event characteristics from the Census of Fatal Occupational Injuries and the Vital Statistics Mortality System. *Journal of Safety Research.* 2013;46:119-125.
36. McHugh ML. Interrater reliability: the kappa statistic. *Biochemia medica.* 2012;22(3):276-282.
37. Kwon YJ, Lee SJ. Compensation for occupational injuries and diseases in special populations: Farmers and soldiers. *Journal of Korean Medical Science.* 2014;29:S24-S31.
38. Vallmuur K, Marucci-Wellman HR, Taylor JA, Lehto M, Corns HL, Smith GS. Harnessing information from injury narratives in the 'big data' era: understanding and applying machine learning for injury surveillance. *Inj Prev.* 2016.
39. Marucci-Wellman HR LM, Corns HL. A practical tool for public health surveillance: Semi-automated coding of short injury narratives from large administrative databases using Naïve Bayes algorithms. *Accident Analysis & Prevention.* 2015;84(November):165–176.
40. Marucci-Wellman H, Lehto M, Corns H. A combined Fuzzy and Naive Bayesian strategy can be used to assign event codes to injury narratives. *Inj Prev.* 2011;17(6):407-414.
41. Lehto M, Marucci-Wellman H, Corns H. Bayesian methods: A useful tool for classifying injury narratives into cause groups. *Injury Prevention.* 2009;15(4):259-265.

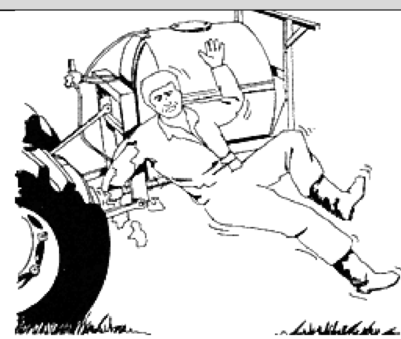
TABLES AND FIGURES

Table 1. Example of Assigning OIICS Codes to an Injury Narrative

OIICS Coding Scheme Definitions[30]

Nature	the principal **physical characteristic(s)** of the injury or illness
Part of Body	the part of the body **directly** affected by the previously identified nature of injury or illness
Source and Secondary Source	the objects, substances, equipment, and other factors that were **responsible for** the injury or illness incurred by the worker or that precipitated the event or exposure
Type of Event	the **manner** in which the injury or illness was produced or inflicted by the source of injury or illness

Patient X was injured while attempting to perform maintenance near the rear of his tractor, close to the power take-off (PTO) shaft. The shaft was running (spinning at 540 revolutions per minute) and the PTO shield was missing. The patient's right arm became entangled in the rotating shaft. The entanglement caused an avulsion injury to the forearm and hand, and the patient was transported to the nearest trauma center.



http://nasdonline.org/static_content/documents/1650/b0215016.gif

OIICS Levels	OIICS Level Title	Nature of Injury	Part of Body	Source of Injury	Type of Event
1	Division	1: Traumatic Injuries and Disorders	4: Upper Extremities	8: Vehicles	6: Contact with Objects and Equipment
2	Major Grouping	13: Open wounds	48: Multiple upper extremities locations	86: Off-road and industrial vehicles—powered	64: Caught in or compressed by equipment or objects
3	Detailed Level 3	131: Amputations, avulsions, enucleations	481: Hand(s) and arms(s)	863: Tractors, PTOs	641: Caught in running equipment or machinery
4	Detailed Level 4	1312: Avulsions, enucleations	4819 : Hand(s) and arm(s), n.e.c.	8634: Power take-off (PTO)	6411: Caught in running equipment or machinery during maintenance, cleaning

Table 2. Kappa Values for Double Coding Agricultural and Logging Injuries Using OIICS

	PCR	Hospital
Nature of Injury		
Division (Level 1)	0.4944	0.3215
Major Grouping (Level 2)	0.5011	0.7634
Detail 3 (Level 3)	0.5246	0.6356
Detail 4 (Level 4)	0.4779	0.5816
Part of Body		
Division (Level 1)	0.7536	0.8933
Major Grouping (Level 2)	0.6310	0.7894
Detail 3 (Level 3)	0.5054	0.5834
Detail 4 (Level 4)	0.4019	0.3936
Source of Injury		
Division (Level 1)	0.8073	0.7418
Major Grouping (Level 2)	0.8244	0.6938
Detail 3 (Level 3)	0.8042	0.6465
Detail 4 (Level 4)	0.7223	0.6234
Type of Event		
Division (Level 1)	0.8405	0.8504
Major Grouping (Level 2)	0.7221	0.6501
Detail 3 (Level 3)	0.6534	0.6318
Detail 4 (Level 4)	0.5550	0.5734

Table 3. Percentage of Classifiable Cases Using OIICS Definition By Level

	PCR (N=1,041) Percent	**Hospital (N=570)** Percent	**Combined (N=1,611)** Percent
Nature of Injury			
Level 1 (Division)	95.0	99.8	96.7
Level 2 (Maj. Grouping)	89.5	97.7	92.4
Level 3 (Detailed)	82.1	76.3	80.1
Level 4 (Detailed)	46.4	39.8	44.1
Part of Body			
Level 1 (Division)	100.0	98.8	99.6
Level 2 (Maj. Grouping)	85.2	94.4	88.5
Level 3 (Detailed)	53.4	59.0	55.4
Level 4 (Detailed)	3.4	7.2	4.7
Type of Event			
Level 1 (Division)	95.3	88.8	93.0
Level 2 (Maj. Grouping)	93.7	30.2	71.2
Level 3 (Detailed)	90.0	12.5	62.6
Level 4 (Detailed)	77.1	3.5	51.1
Source of Injury			
Level 1 (Division)	94.8	68.4	85.5
Level 2 (Maj. Grouping)	94.4	65.4	84.2
Level 3 (Detailed)	92.6	9.5	63.2
Level 4 (Detailed)	85.0	1.9	55.6

Table 4. Percent Estimates of Nature of Injury using Different Denominators, PCR, Hospital, & Combined

Note: N/A indicates code not includable in numerator or denominator

			PCR		Hospital		Combined	
		OIICS Code Description	Denominator - Number of Records (N=1,041) %	Denominator - Number with code %	Denominator - Number of Records (N=570) %	Denominator - Number with code %	Denominator - Number of Records (N=1,611) %	Denominator - Number with code %
OIICS Level 1	1	Traumatic injuries and disorders	91.16	95.96	97.37	97.54	93.36	96.53
	2	Diseases and disorders of body systems	0.48	0.51	0.88	0.88	0.62	0.66
	5	Symptoms, signs, and ill-defined conditions	2.11	2.22	1.58	1.58	1.92	1.99
	7	Exposure to disease- no illness incurred	1.25	1.31	0	0	0.81	0.83
	9	Nonclassifiable	5.00	N/A	0.18	N/A	3.29	N/A
OIICS Level 2	N/A	Nonclassifiable	10.47	N/A	0	0	6.77	N/A
	11	Traumatic injuries to bones, nerves, spinal cord	10.18	11.37	11.40	11.67	10.61	11.48
	12	Traumatic injuries to muscles, tendons, ligaments, joints, etc.	0	0	14.74	15.08	5.21	5.64
	13	Open wounds	17.68	19.74	24.56	25.13	20.11	21.76
	14	Surface wounds and bruises	0	0	17.02	17.41	6.02	6.51
	18	Multiple traumatic injuries and disorders	7.30	8.15	16.67	17.06	10.61	11.48
	19	Other traumatic injuries and disorders	39.96	44.64	8.07	8.26	28.68	31.03
	Other	Other Categories (under 5 % each)	14.41	16.10	7.54	5.39	11.98	12.09
OIICS Level 3	N/A	Nonclassifiable	17.87	N/A	23.68	N/A	19.93	N/A
	111	Fractures	8.45	10.29	11.23	14.71	9.44	11.78
	123	Sprains, strains, tears			13.16	17.24	4.66	5.81
	131	Amputations, avulsions, enucleations	2.21	2.69	2.63	3.45	2.36	2.95
	132	Cuts, lacerations	14.02	17.08	0	0	9.06	11.32
	141	Abrasions, scratches	0	0	2.11	2.76	0.74	0.93
	143	Bruises, contusions	2.21	2.69	13.33	17.47	6.15	7.67
	183	Fractures and other injuries	0	0	11.93	15.63	4.22	5.27
	189	Multiple traumatic injuries and disorders, n.e.c.	3.84	4.68	3.51	4.60	3.72	4.65
	197	Nonspecified injuries and disorders	38.42	46.78	5.96	7.82	26.94	33.64
	Other	Other Categories (under 2% each)	12.98	15.79	12.46	16.32	12.79	15.97
OIICS Level 4	N/A	Nonclassifiable	53.60	N/A	60.18	N/A	55.93	N/A
	1212	Dislocation of joints	1.06	2.28	0	0	0.68	1.55
	1232	Sprains	0	0	13.16	33.04	4.66	10.56
	1311	Amputations	1.44	3.11	2.63	6.61	1.86	4.23
	1832	Fractures (except rib, trunk fractures) and internal injuries	0	0	1.05	2.64	0.37	0.85
	1839	Fractures and other injuries, n.e.c.	0	0	10.00	25.11	3.54	8.03
	1971	Crushing injuries	1.06	2.28	4.04	10.13	2.11	4.79
	1972	Soreness, pain, hurt—nonspecified injury	35.54	76.60	1.58	3.96	23.53	53.38
	Other	Other Categories (under 1% each)	7.30	15.73	7.36	18.51	7.32	16.62

Table 5. Percent Estimates of Part of Body using Different Denominators, PCR, Hospital & Combined

Note: N/A indicates code not includable in numerator or denominator

Level	Code	OIICS Code Description	PCR: Denominator - Number of Records (N=1,041) %	PCR: Denominator - Number with code %	Hospital: Denominator - Number of Records (N=570) %	Hospital: Denominator - Number with code %	Combined: Denominator - Number of Records (N=1,611) %	Combined: Denominator - Number with code %
OIICS Level 1	1	Head	13.74	14.82	10.88	11.01	12.73	13.42
	2	Neck, Including throat	0.58	0.62	0.18	0.18	0.43	0.46
	3	Trunk	22.38	24.15	15.09	15.28	19.80	20.88
	4	Upper Extremities	14.7	15.85	38.07	38.54	22.97	24.21
	5	Lower Extremities	19.79	21.35	22.98	23.27	20.92	22.05
	6	Body Systems	2.59	2.8	1.4	1.42	2.17	2.29
	8	Multiple Body Parts	18.92	20.41	10.18	10.3	15.83	16.69
	9	Other Body Parts	7.3	N/A	1.23	N/A	5.15	N/A
OIICS Level 2	N/A	Nonclassifiable	14.79	N/A	5.61	N/A	11.55	N/A
	11	Cranial region, including skull	5.38	6.31	0	0	3.48	4.05
	32	Back, including spine, spinal cord	7.78	9.13	5.44	5.76	6.95	8.09
	34	Pelvic region	5.09	5.98	0	0	3.29	3.83
	44	Hand(s)	0	0	23.51	24.91	8.32	9.68
	51	Leg(s)	12.97	15.22	9.47	10.04	11.73	13.66
	53	Foot (feet)	0	0	6.67	7.06	2.36	2.75
	89	Other multiple body parts	13.64	16.01	5.61	5.95	10.80	12.57
	Other	Other Categories (under 5 % each)	40.35	47.35	36.5	38.66	38.98	45.38
OIICS Level 3	N/A	Nonclassifiable	46.59	N/A	41.05	N/A	44.63	N/A
	111	Brain	2.88	5.4			1.86	3.36
	132	Eye(s)	0	0	2.98	5.06	1.06	1.91
	322	Lumbar region	4.42	8.27	0	0	2.86	5.16
	341	Hip(s)	3.17	5.94	0	0	2.05	3.70
	423	Forearm(s)	0	0	3.16	5.36	1.12	2.02
	442	Finger(s), fingernail(s)	2.21	4.14	12.46	21.13	5.83	10.54
	448	Hand(s) and finger(s)	0	0	4.21	7.14	1.49	2.69
	512	Knee(s)	3.46	6.47	0	0	2.23	4.04
	513	Lower leg(s)	0	0	2.46	4.17	0.87	1.57
	518	Multiple leg locations	0	0	3.68	6.25	1.30	2.35
	899	Multiple body parts, n.e.c.	13.16	24.64	5.44	9.23	10.43	18.83
	Other	Other Categories (under 2% each)	24.11	45.14	24.56	41.66	24.27	43.83
OIICS Level 4	N/A	Nonclassifiable	96.64	N/A	92.81	N/A	95.28	N/A
	5181	Knee(s) and leg(s)	0	0	2.63	36.59	0.93	19.74
	5189	Multiple leg locations, n.e.c.	0	0	1.05	14.63	0.37	7.89
	Other	Other Categories (under 1 % each)	3.36	100	3.51	48.78	3.41	72.37

Table 6. Percent Estimates of Source of Injury using Different Denominators, PCR, Hospital & Combined

Note: N/A indicates code not includable in numerator or denominator

			PCR		Hospital		Combined	
			Denominator - Number of Records (N=1,041)	Denominator - Number with code	Denominator - Number of Records (N=570)	Denominator - Number with code	Denominator - Number of Records (N=1,611)	Denominator - Number with code
		OIICS Code Description	%	%	%	%	%	%
OIICS Level 1	1	Chemicals and Chemical Products	0.48	0.51	1.4	2.05	0.81	0.94
	2	Containers, Furniture, and Fixtures	0.48	0.51	0	0	0.31	0.36
	3	Machinery	4.32	4.56	48.07	70.26	19.80	23.17
	4	Parts and Materials	1.83	1.93	0	0	1.18	1.38
	5	Persons, Plants, Animals, and Minerals	63.88	67.38	10.88	15.9	45.13	52.80
	6	Structures and Surfaces	3.84	4.05	1.58	2.31	3.04	3.56
	7	Tools, Instruments, and Equipment	9.41	9.93	5.61	8.21	8.07	9.44
	8	Vehicles	8.93	9.42	0.35	0.51	5.90	6.90
	9	Other Sources	1.63	1.72	0.53	0.77	1.24	1.45
	N/A	Nonclassifiable	5.19	N/A	31.58	N/A	14.53	N/A
OIICS Level 2	N/A	Nonclassifiable	5.57	N/A	34.56	N/A	15.83	N/A
	31	Agricultural and garden machinery	0	0	46.49	71.05	16.45	19.54
	51	Animals	45.05	47.71	9.47	14.48	32.46	38.57
	72	Handtools—powered	7.59	8.04	0	0	4.90	5.83
	86	Off-road and industrial vehicles—powered	5.28	5.6	0	0	3.41	4.06
	Other	Other Categories (under 5% each)	36.51	38.65	9.48	14.47	26.94	32.01
OIICS Level 3	N/A	Nonclassifiable	7.4	N/A	90.53	N/A	36.81	N/A
	515	Mammals, except humans	44.86	48.44	0	0	28.99	45.87
	587	Trees, logs, limbs	16.81	18.15	0	0	10.86	17.19
	722	Cutting handtools—powered	7.59	8.2	0	0	4.90	7.76
	732	Cutting handtools—power not determined	0	0	2.46	25.93	0.87	1.38
	863	Tractors, PTOs	4.51	4.88	0	0	2.92	4.62
	Other	Other Categories (under 2% each)	18.83	20.33	7.01	74.07	14.65	23.18
OIICS Level 4	N/A	Nonclassifiable	14.99	N/A	98.07	N/A	44.38	N/A
	5154	Horses and other equines	43.52	51.19	0	0	28.12	50.56
	5871	Trees	9.03	10.62	0	0	5.83	10.49
	5872	Logs	2.11	2.49	0	0	1.37	2.46
	5873	Limbs, branches—unattached	4.9	5.76	0	0	3.17	5.69
	7221	Chainsaws—powered	7.59	8.93	0	0	4.90	8.82
	8631	Farm tractor	2.88	3.39	0	0	1.86	3.35
	Other	Other Categories (under 1% each)	14.98	17.62	1.93	100	10.37	18.64

Table 7. Percent Estimates of Type of Event using Different Denominators, PCR, Hospital & Combined

Note: N/A indicates code not includable in numerator or denominator

			PCR		Hospital		Combined	
			Denominator - Number of Records (N=1,041)	Denominator - Number with code	Denominator - Number of Records (N=570)	Denominator - Number with code	Denominator - Number of Records (N=1,611)	Denominator - Number with code
		OIICS Code Description	%	%	%	%	%	%
OIICS Level 1	1	Violence and other injuries by persons or animals	12.39	13	7.72	8.7	10.74	11.55
	2	Transportation Incidents	39.77	41.73	2.63	2.96	26.63	28.64
	3	Fires and Explosions	0.48	0.5	0	0	0.31	0.33
	4	Falls, Slips, Trips	9.99	10.48	10	11.26	9.99	10.75
	5	Exposure to Harmful Substances or Environments	1.92	2.02	2.46	2.77	2.11	2.27
	6	Contact with Objects and Equipment	28.91	30.34	60.18	67.79	39.98	42.99
	7	Overexertion and Bodily reaction	1.83	1.92	5.79	6.52	3.23	3.47
	N/A	Nonclassifiable	4.71	N/A	11.23	N/A	7.01	N/A
OIICS Level 2	N/A	Nonclassifiable	6.34	N/A	69.82	N/A	28.80	N/A
	13	Animal and insect related incidents	12.1	12.92	7.72	25.58	10.55	14.82
	23	Animal & other non-motorized vehicle transp. incidents	33.33	35.59	0	0	21.54	30.25
	43	Falls to lower level	6.53	6.97	0	0	4.22	5.93
	62	Struck by object or equipment	23.25	24.82	5.79	19.19	17.07	23.98
	Other	Other Categories (under 5% each)	18.45	19.7	16.67	55.23	17.82	25.02
OIICS Level 3	N/A	Nonclassifiable	9.99	N/A	87.54	N/A	37.43	N/A
	132	Struck by animal	10.18	11.31	0	0	6.58	10.52
	231	Animal transportation incident	33.33	37.03	0	0	21.54	34.42
	273	Nonroadway noncollision incident	2.5	2.77	0	0	1.61	2.58
	433	Other fall to lower level	5.09	5.66	0	0	3.29	5.26
	623	Struck by falling object or equipment other than power vehicle	11.05	12.27	0	0	7.14	11.41
	625	Injured by handheld obj.or equipment	6.63	7.36	0	0	4.28	6.85
	632	Struck against stationary object or equipment	0	0	3.16	25.35	1.12	1.79
	Other	Other Categories (under 2% each)	21.23	23.6	9.3	74.65	17.01	27.18
OIICS Level 4	N/A	Nonclassifiable	22.86	N/A	96.49	N/A	48.91	N/A
	1321	Trampled by or stepped on by animal	2.59	3.36	0	0	1.68	3.28
	1322	Kicked by animal	5	6.48	0	0	3.23	6.32
	1329	Struck by animal, n.e.c.	1.25	1.62	0	0	0.81	1.58
	2313	Thrown, fell, or jumped from animal being ridden	32.56	42.22	1.58	45	21.60	42.28
	2631	Jack-knifed or overturned, roadway	1.44	1.87	0	0	0.93	1.82
	2731	Jack-knifed or overturned, nonroadway	1.83	2.37	0	0	1.18	2.31
	4332	Other fall to lower level 6 to 10 feet	1.15	1.49	0	0	0.74	1.46
	Other	Other Categories (under 1% each)	31.32	40.59	1.93	55	20.92	40.95

CHAPTER 5
INJURY TRENDS

Trends in Non-Fatal Agricultural and Logging Injury in Maine and New Hampshire: Results from a Low-Cost Passive Surveillance System

Erika Scott, Erin Bell, Liane Hirabayashi, Nicole Krupa, and Paul Jenkins

ABSTRACT

Agriculture and logging are dangerous industries, and although data on fatal injuries exists, less is known about non-fatal injuries. The purpose of this study is to describe trends in agricultural and logging morbidity in Maine and New Hampshire from 2008-2010 using a newly established passive surveillance system. Demographics and specifics of the event were recorded for each incident case. The average age of injured people in Maine and New Hampshire was 41.7 for agriculture and 45.4 for logging. Women comprised 43.8 percent of agricultural injuries, whereas almost all logging-related injuries occur to men. Machinery (n=303) and animal-related (n=523) injuries accounted for most agricultural incidents. Of all injured women, over 60% sustained injuries due to animal related causes. Agricultural injuries were spread across the two states with clustering in southern New Hampshire and south central Maine, with additional injuries in the Aroostook County area, which is located in the northeast part of the state. Seasonal variation in agricultural injuries was seen, with peaks in the summer months. Logging injuries (n=68) were primarily due to trees, heavy equipment, and chainsaws. Overlap between the two industries was seen for tree-related work (cutting wood on farm property). It was possible to correlate rises in certain morbidity rates with coincident economic trends. Our methods are able to capture traumatic injury in agriculture and logging, in sufficient detail to prioritize interventions and to evaluate outcomes. The system is low-cost and has the potential to be sustained over a long period. Differences in rates of animal and machinery-related injuries suggest the need for state-specific safety prioritization within agriculture and logging.

BACKGROUND

Agricultural and logging workers are at high risk of injury compared to workers in other industries. Elevated rates of fatal injury among agricultural workers and loggers (26.0 deaths per 100,000 full-time equivalent workers (FTE) and 109.5/100,000 FTE, respectively) contrast sharply with the all-worker fatal injury rate of 3.3/100,000 FTE.[1] While federal and state programs provide data on fatal injury in these industries, challenges exist in adequately capturing non-fatal injury events.[2] This is a well-documented limitation of many existing surveillance systems.[2-4]

While efforts have been made to characterize non-fatal logging injury in the Pacific Northwest,[5-8] the Deep South,[9,10] and Appalachia,[11-16] less is known of how Northeastern logging differs from these other regions. As for agricultural injury, in the 2006 growing season, Brower and colleagues found an injury rate of 7.9 per 100 full time equivalents (FTE) for Maine migrant and seasonal farm workers, but only based on one study.[17] In New Hampshire, recent estimates (average of 50 'sure to be' farm cases per year) were considered undercounts because only hospital data was used as a single source for farm injury surveillance.[18]

Maine and New Hampshire are small states in terms of agriculture, however the industry is still important to the economies of both states.[19,20] The average market value of agricultural products sold in Maine and New Hampshire in 2007 was $617 million and $199 million respectively.[19,20] Similar to agriculture, the Northeast is diverse in terms of forest operations, ranging from larger mechanized businesses to smaller hand-harvest operations.[21] Forest products are a primary driver of the Maine economy, encompassing nearly 30 percent, or $885 million, of the state's total exports, and account for four percent of New Hampshire's economy.[22,23]

Traumatic injuries in both industries require substantial medical care and lead to lost work-time, often creating economic hardship for the patient, the patient's family, and the employer.[24-27]

Part of the challenge in adequately capturing work-related injury events in the agriculture, forestry, and fishing (AFF) sector is the ever-changing work-environment. Furthermore, for agriculture, the work-site is often the place of residence.[28-30] This is especially true in the Northeast, where small family-owned farms are the norm.[31] Therefore, traditional means of capturing occupational injury, such as workers' compensation data, are not adequate to estimate agricultural injury rates because many of the small family farms fall outside the scope of workers' compensation coverage.[27,32]

Previous methods of quantifying injuries in these industries have mostly proved cost prohibitive. [33] We have developed a methodology that combines existing variables and free-text narrative to identify agricultural and logging injuries in hospital data and pre-hospital care reports (PCR). These data are typically available for minimal cost and while they are primarily collected for other purposes, they are likely to be in electronic form, or transitioning to this format. Research has shown that combining coded and narrative data provide enough detail to reconstruct injury events, which can aid in the prioritization of safety interventions.[34] This research will fill a gap in the understanding of farm and logging injuries, and has the potential to increase awareness of these data sources for other public health purposes.

Using data from our newly developed surveillance system, we will describe trends in agricultural and logging morbidity in Maine and New Hampshire from 2008-2010 and suggest priorities for safety interventions. We hypothesize that trends in agricultural and logging morbidity will be similar between Maine and New Hampshire, and therefore interventions can be regionally focused, instead of state specific. A secondary goal is to quantify the relative cost of the surveillance system overall, in context of the information it provides.

METHODS

SURVEILLANCE SYSTEM AND DATASET CREATION

Injury details were obtained by sub-setting data from the New Hampshire Bureau of Public Health Statistics & Informatics, the Maine Health Data Organization and the Bureau of Emergency Medical Services (EMS) of both states. EMS data were downloaded directly from each state's EMS portal. The hospital data (inpatient, outpatient, and emergency department) were received via secure file transfer.

The process of identifying agriculture and logging-related injuries in these administrative data sets has been described in detail in previous publications.[35,36] Each injury event identified was coded using the Occupational Injury and Illness Classification System (OIICS)[37] by two independent coders using the specific methodology provided by the National Institute of Occupational Safety and Health (NIOSH).[37] The events were also classified as either agricultural (FAIC code 1, 3-10) or logging-related (FAIC code 2) based on the American Society of Agricultural and Biological Engineers (ASABE) Farm and Agricultural Injury Classification Codes (FAIC).[38]

A convention was adopted such that an injury was only assigned to the logging industry in cases where this could be definitively established. This required that the narrative of the PCR specifically state that the injured person: 1) was a "logger", 2) was working at a "logging" site, or 3) was using *heavy* logging equipment (feller-buncher, skidder, etc.). This convention resulted in a large number of tree-related incidents (incidents where the narrative contained words such as chainsaw, tree, log, etc) being assigned as agriculturally, as opposed to logging, related. Although it was believed that the majority of these incidents were in fact agriculturally related, this decision is still somewhat arbitrary. Therefore, these incidents, although discussed in the results under the agricultural heading, are also discussed as their own category.

Population at Risk

The population at risk over this brief time period (2008 to 2010) covered by the data was a constant as measured by the 2007 Census of Agriculture and the American Community Survey (Table 1). Rather than dived the injury counts by this constant value, the time trends for the data are presented as simple counts.

Data Analysis

This is a new surveillance database containing only three years of data; therefore, these data are presented as simple line graphs, for both states combined. The gender distribution of the injured subjects was compared to the population distribution using the Z test of a hypothesized population proportion. The mean age of the injured was compared to the mean age of the population using the Z test of a hypothesized population mean. These comparisons were carried out separately for agriculture and logging. Comparisons between those injured in New Hampshire and Maine agriculture or logging were conducted using chi-square (for categorical variables). Probability values less than 0.05 were considered statistically significant. All data were analyzed using SAS 9.3.1 (Cary, NC).

Spatial analysis was conducted by linking the frequency of the injury location (or secondarily, the patient's home address) by zip code for the three-year period, to the zip code level shape file for Maine and New Hampshire. Results were displayed using proportional symbols. Maps were created in ESRI ArcMap10 (Redlands, CA). Only a map for agricultural injury was generated, as these data were too sparse to warrant a map for logging-related injury.

Evaluating the influence of economics on morbidity was done using the General Autoregressive Conditional Heteroscedasticity model (GARCH model), specifically comparing the price of home heating fuels with woodcutting related injury. GARCH modeling was chosen to account for autocorrelation in seasonal data, and heteroscedasticity among varying seasons. The additional data used in GARCH modeling was obtained from the U.S. Energy Information Administration.[39] This protocol was approved by the Institutional Review Board of the Mary Imogene Bassett Hospital and University at Albany.

Cost Estimates

The cost of the surveillance system was calculated using staffing costs at the Northeast Center (NEC), which encompassed time to 1) apply and receive these data, 2) clean and match-merge data, 3) subset data to agricultural and logging cases, 4) code data using FAIC and OIICS coding scheme, 5) analyze data, and 6) publish and disseminate data. Additional costs include the data handling fees. The personnel time was calculated using full-time equivalents (FTE) and was added together with the data fees.

Results

Agriculture (total of 1,517 cases)

The average age of those who sustained an agricultural injury in Maine and New Hampshire was 41.7 (Maine 41.9, New Hampshire 41.6) (Figure 1). This is significantly younger ($p<.0001$) than the average age of all farmers in Maine and New Hampshire from the 2007 Census of Agriculture (average age 56.3).[19,20]

Women made up 47.8% of those injured in New Hampshire versus 39.5% of those injured in Maine (p=0.0002). For the two states combined, almost half (43.8%) of those sustaining an agricultural injury were women.. This percentage is significantly higher than the overall percentage of women (27.4%) in the agricultural workforce in these two states (p<.0001). Of all female injuries occurring in agricultural across the two states combined, 65.1 percent involved animals. In contrast, for male victims of agricultural injury, only 9.5 percent could be attributed to animal-related causes. When horse-related injuries were removed from the agricultural dataset, the average age of an injured person rose to 43.7, and the percentage of women dropped to 25.2% in the non-horse-related agricultural injury dataset. Looking only at horse-related injuries, the average age of the victim was 36.9, and over 88% were women (Table 2).

Agricultural injuries were widely distributed across the two states, with some clustering in southern New Hampshire and south central Maine. An additional clustering of injuries was seen in the Aroostook County area, which is located in the northeast part of the state (Figure 2). A greater proportion of agricultural injuries occurred within the same zip code where the patient lived (51.4%), versus logging injuries (33.8%).

The two largest categories of agricultural injuries were for machinery (n=303) and animals (n=523) (Figures 4&5). Of all animal-related injuries, 87% were caused by horses (n=454). Machinery-related cases occurred with a greater frequency in Maine (21.8% of all agricultural injuries) than in New Hampshire (16.6%). Conversely, animal-related injuries occurred with a higher frequency in New Hampshire (47.6%) versus Maine (24.8%). Both of these findings were statistically significant.

While the time-trend for machinery incidents appeared to be relatively stable, there was some indication of an upward trend for animal-related incidents over the three years. Seasonal patterns appear in the agricultural data, with clear peaks in the summer months for each of the three years under surveillance (Figure 3).

Logging (total of 68 cases)

The average age of those who sustained a logging injury in Maine and New Hampshire was 45.4 (Maine 47.0, New Hampshire 43.5) (Figure 6). There was no statistically significant difference in the age of those injured versus the mean age of all logging workers in the Maine and New Hampshire logging industry. Logging injuries were sustained almost exclusively (97.0%) by men, who comprise almost the entire logging workforce. The logging injuries were too sparse to display geographically due to confidentiality concerns.

Seasonal variation in injury incidence was not seen for logging (Figure 7). Some indication of an upward trend that peaked in the summer and fall of 2008 can be seen. The majority of logging-related injuries were caused by trees (n=43), followed by heavy equipment and vehicles (n=13) and chainsaws (n=8).

Injuries Attributable to Both Industries

Chainsaw (n=79) and tree related (n=175) injuries made up a significant proportion of the injury events that could not be definitively attributed to agriculture versus logging (Figures 8&9). For chainsaw incidents, there were two noteworthy peaks observed in the summers of 2008 and 2009. The 2008 peak appears to be consistent with an upward trend in tree injuries observed during the course of 2008. Cases unclassifiable by source of injury accounted for twenty percent (20%) of Maine injury records but only four percent (4%) of New Hampshire records.

Summary of the cost of the system

Personnel time

The majority of the cost for the surveillance system is contained within personnel time. Staffing for the surveillance system is currently at 1.5 full time equivalents (FTE). The most time intensive part of the surveillance process is coding hospital records and PCRs, which took approximately two working weeks (at 1.0 FTE).

Other Costs

The only data cost incurred was for hospitalization data from the Maine Health Data Organization, which totaled $9,700.00 for three years of data (outpatient, inpatient, and ED). There was no cost for either PCR dataset, or the New Hampshire hospitalization data.

Predicting Morbidity using Economic Trends

There were no significant findings for propane or heating oil, however natural gas prices could be associated with woodcutting injuries. A negative correlation was found for natural gas and woodcutting injuries in the winter (-0.93) and spring (-0.99), along with a positive correlation in the summer (0.72) and fall (0.69) (Table 3). Though these data are preliminary, interaction is seen between the season, natural gas prices and woodcutting injuries ($p=0.02$) (Table 4).

Discussion

Consistent with past research,[40-43] machinery and animal-related injuries are the most common types of agricultural injuries identified in Maine and New Hampshire from 2008 to 2010. Machinery injuries were consistent with occupationally related tasks, such as tractor and skidsteer operation, harvesting, and processing crops. Livestock was a major source of injuries, but the majority (87%) of these were caused by horses, which implies that these may have been recreational. Women experienced a larger proportion of animal-related injuries than men, which may be explained by two factors: 1) women's roles on the farm have traditionally been focused on animal care, as opposed to machinery operation,[44] and 2) women make up a significant portion of recreational horse riders.[45]

Horse related injuries were an important source of the morbidity. These events, while agricultural in nature, may be either recreational or work-related. Currently, our surveillance system does not have the ability to make this distinction. A significantly higher proportion of animal and horse-related injury events were identified in New Hampshire as compared to Maine. New Hampshire has more horse farms than Maine, which is in congruence with this finding.[19,20] Despite the inability to distinguish which of these cases are occupationally related, the frequency of horse-related injuries implies the need for a safety intervention. There appears to be an increase in animal-related injuries over the three-year period, but this will have to be followed over a longer period to verify a statistically significant change.

Over the three years for which data are available, some patterns have become apparent. As expected, a cyclical seasonal pattern emerged in agricultural injury, with most injuries occurring in the summer months, when farmers are most active. The same pattern was not identified in logging, nor was it expected. Logging work can be done almost year around, with a slowing in the spring during 'mud season'. Fall and winter remain popular times for logging, as the trees have dropped their leaves, and frozen ground allows trees to be more easily removed from the forest.

The zip codes for the incident location were only available for injuries identified through the pre-hospital care reports (ambulance runs), whereas zip codes of the patients residence were available for both pre-hospital care reports and hospitalization data. Logging injuries were more likely to occur in a zip code that differed from residence zip code, than agricultural injuries (p=0.005). Since the farmstead is often the home and the workplace, it is expected that a greater proportion of agricultural injuries occurred at or near home property. Notably, greater concentrations of injuries appear in the northern tip of Maine, which has the highest productivity for crops in the state, along with southern New Hampshire, and south central Maine.[46]

Interestingly, tree and chainsaw injuries appeared to increase in summer and fall 2008, which is the same time when home heating fuel prices reached their peak during the Great Recession.[39] We hypothesize that homeowners increasingly relied on firewood during this time, which in turn, led to an increase in chainsaw and tree-related injury events. As prices returned to a lower level, these types of injuries declined. Preliminary results seem to support these conclusions Woodcutting injuries are a crossover event, as they are sustained by professional loggers, farmers, as well as rural homeowners. Our ability to distinguish between these occupational and lifestyle categories are a current limitation of the surveillance system. In addition, it becomes difficult to estimate a denominator population due to the pervasiveness of woodcutting activity among these groups. None the less, these injures are serious in nature, and similar interventions and safety messaging are likely to be appropriate for all the groups.

This surveillance system will gain in value by continuing over time. Once data have been accumulated over a sufficient number of years, control charts will be employed to track significant changes in each injury category. One of the central tenets of statistical analysis is that it is easier to measure a change in something than the thing itself. The notion that every agricultural or logging injury could be captured by a cost-effective monitoring system is simply not realistic, nor is it necessary provided the system remains in place long enough for changes over time to be measured. Monitoring morbidity and mortality over time has historically failed because the systems that were developed were too expensive to be maintained in the long term.[47]

These data, which were collected at relatively low-cost in an ongoing manner, provide enough detail (often to the type of machinery or animal, for instance) for injury epidemiologists to understand common themes in agriculture and logging injuries, as well as track relatively rare events over time. Since the basic steps of the passive system have been established, adding additional years of data, or additional states, should be easier than the initial set-up of the system. Further, we are currently exploring ways to enhance the system by using advanced statistical techniques to reduce or eliminate the need for extensive visual inspection of records.[48-52]

Limitations

The inability to firmly establish occupational relatedness for all of the injury events is a limitation of the system. This is a particularly acute issue as related to horse injuries. Administrative data are not maintained specifically for surveillance; therefore, there are challenges when using these sources to track injury. Some records are missing key variables that would inform us of the source of injury or type of event, and this can vary from state to state. In New Hampshire, only a small portion of agricultural injury records could not be assigned meaningful OIICS codes, whereas this proportion reached twenty percent in Maine. While this passive system is relatively low cost when compared to surveys or active systems, these data may not be as current as data obtained using another more active system. We anticipate this time lag in data acquisition to become shorter as electronic reporting becomes the industry standard for not only hospitalization data, but for pre-hospital care reports, as well.

Currently, the system has data for only three years, which limits the ability to perform most formal statistical analyses, such as control charts and time series analysis. These methods will be employed increasingly as data become available for additional states and additional years. It is anticipated that at least five years of data will be necessary for these purposes.

CONCLUSIONS

Machinery and animal-related events were the most frequent cause of injuries in the agricultural industry in Maine and New Hampshire from 2008-2010. Within animal injuries, horse-riding activities are a primary concern. The geographic distribution of agricultural injuries appeared consistent with agricultural areas within Maine and New Hampshire. Manual felling was a cause of both logging and agricultural injury where the injury was caused 1) by the chainsaw itself, and 2) by the tree or limb that was cut. Our methods are able to capture traumatic injury in agriculture and logging with sufficient detail to prioritize interventions and to evaluate outcomes. The system is low-cost, relative to surveys or active surveillance, and has the potential for continued use in years to come. Differences in the frequency of animal and machinery-related injuries suggest the need for state specific-safety prioritization.

ACKNOWLEDGEMENTS

This research is funded by the Centers for Disease Control and Prevention National Institute for Occupational Safety and Health (CDC-NIOSH) Grant Number 2U54OH007542-11. We thank the Maine Emergency Medical Services Bureau, the New Hampshire Bureau of Emergency Medical Services, the Maine Health Data Organization, and the New Hampshire Department of Health and Human Services (DHHS) for the use of their data. We are grateful to Deb Dalton, Emma Garrison, Melissa Horsman, Alissa Kent, Pam Tinc, Rebecca Russell and

Marybeth Vargha for their help with record verification and coding. Additionally, we appreciate the thoughtful review of Edward Fitzgerald, PhD, Kitty Gelberg, PhD, and Lloyd Wilson, PhD.

References

1. BLS. Fatal Work Injuries, 2014. 2015; http://www.bls.gov/iif/oshwc/cfoi/cfch0013.pdf. Accessed December 2, 2015.
2. Rosenman KD, Kalush A, Reilly MJ, Gardiner JC, Reeves M, Luo Z. How much work-related injury and illness is missed by the current national surveillance system? *J Occup Environ Med.* 2006;48(4):357-365.
3. Boden LI. Capture-recapture estimates of the undercount of workplace injuries and illnesses: sensitivity analysis. *Am J Ind Med.* 2014;57(10):1090-1099.
4. Leigh JP, Marcin JP, Miller TR. An Estimate of the U.S. Government's Undercount of Nonfatal Occupational Injuries. 2004;46(1):10-18.
5. Holman RG, Olszewski A, Maier RV. The epidemiology of logging injuries in the northwest. *J Trauma.* 1987;27(9):1044-1050.
6. Stover B, Wickizer TM, Zimmerman F, Fulton-Kehoe D, Franklin G. Prognostic factors of long-term disability in a workers' compensation system. *J Occup Environ Med.* 2007;49(1):31-40.
7. Franklin GM, Haug J, Heyer N, Checkoway H, Peck N. Occupational carpal tunnel syndrome in Washington State, 1984-1988. *Am J Public Health.* 1991;81(6):741-746.
8. Paulozzi LJ. Fatal logging injuries in Washington state, 1977 to 1983. *Journal of occupational medicine. : official publication of the Industrial Medical Association.* 1987;29(2):103-108.
9. Lynch SM, Smidt MF, Merrill PD, Sesek RF. Incidence of MSDs and neck and back pain among logging machine operators in the southern U.S. *J Agric Saf Health.* 2014;20(3):211-218.
10. Bordas RM, Davis GA, Hopkins BL, Thomas RE, Rummer RB. Documentation of hazards and safety perceptions for mechanized logging operations in East Central Alabama. *J Agric Saf Health.* 2001;7(2):113-123.
11. Bell JL, Grushecky ST. Evaluating the effectiveness of a logger safety training program. *J Safety Res.* 2006;37(1):53-61.
12. Bell JL, Helmkamp JC. Non-fatal injuries in the West Virginia logging industry: using workers' compensation claims to assess risk from 1995 through 2001. *Am J Ind Med.* 2003;44(5):502-509.
13. Johnson CM, Lagares-Garcia JA, Miller SL. When the bough breaks: a 10-year review of logging injuries treated at a rural trauma center in Pennsylvania. *Am Surg.* 2002;68(7):573-581.
14. Helmkamp JC, Lundstrom WJ. Work-related deaths in West Virginia from July 1996 through June 1999: surveillance, investigation, and prevention. *J Occup Environ Med.* 2000;42(2):156-162.
15. Helmkamp JC, Derk SJ. Nonfatal logging-related injuries in West Virginia. *J Occup Environ Med.* 1999;41(11):967-972.
16. Wolf C DG. Logging Work Injuries in Appalachia. *U.S. Department of Agriculture, Forest Service, Northeastern Forest Experiment Station.* 1978.
17. Brower MA, Earle-Richardson GB, May JJ, Jenkins PL. Occupational injury and treatment patterns of migrant and seasonal farmworkers. *J Agromedicine.* 2009;14(2):172-178.
18. Weiss D AK, Scott E. *Identifying the gaps in the methodology of New Hampshire Farm Injury Surveillance Using Hospital Discharge Data.* NH Division of Public Health Service; December 2015 2015.
19. USDA. State Profile Maine. In: NASS, ed2007.
20. USDA. State Profile New Hampshire. In: NASS, ed2007.
21. Bureau of Labor Statistics. Occupational Employment Statistics Logging Equipment Operators 45-4022. 2015; http://www.bls.gov/oes/current/oes454022.htm#st, 2016.
22. Council MFP. Maine Exports. 2014; http://maineforest.org/wp-content/uploads/2015/03/Maine-Forest-Products-Council-Trade-Statistics-World.pdf. Accessed July 14, 2015.
23. (NEFA) NSFA. *The Economic Importance of New Hampshire's Forest-Based Economy 2011.* 2011.

24. Leigh JP. Economic burden of occupational injury and illness in the United States. *Milbank Q.* 2011;89(4):728-772.
25. Santana VS, Araújo-Filho JB, Albuquerque-Oliveira PR, Barbosa-Branco A. Occupational accidents: Social insurance costs and work days lost. *Revista de Saude Publica.* 2006;40(6):1004-1012.
26. Leigh JP, Waehrer G, Miller TR, Keenan C. Costs of occupational injury and illness across industries. *Scand J Work Environ Health.* 2004;30(3):199-205.
27. Leigh JP, McCurdy SA, Schenker MB. Costs of occupational injuries in agriculture. *Public Health Rep.* 2001;116(3):235-248.
28. Gorucu S, Murphy D, Kassab C. Occupational and Nonoccupational Farm Fatalities Among Youth for 2000 Through 2012 in Pennsylvania. *Journal of Agromedicine.* 2015;20(2):125-139.
29. Meeker BJ, Carruth A, Holland CB. Health hazards and preventive measures of farm women. Emerging issues. *AAOHN journal : official journal of the American Association of Occupational Health Nurses.* 2002;50(7):307-314.
30. Penn State University. The Work Environment. *National Safe Tractor and Machinery Operation Program* 2004; https://www.sdstate.edu/sdces/resources/crops/HOSTA/upload/1-1-The-Work-Environment-C.pdf. Accessed March 14, 2016.
31. USDA. *2012 Census of Agriculture Full Report.* 2015.
32. Costich J. Who pays for agricultural injury care? *J Agromedicine.* 2010;15(1):54-57.
33. CDC - Looking to the Future for Agriculture Injury Surveillance at NIOSH - Agruiculture, Forestry and Fishing Resources. 2015; http://www.cdc.gov/niosh/agforfish/aginjurysurv.html. Accessed June 29, 2015.
34. Lincoln AE SG, Courtney TK, Wellman HM, Smith GS, Amoroso PJ. Using narrative text and coded data to develop hazard scenarios for occupational injury interventions. *Inj Prev.* 2004;10:249-254.
35. Scott EE, Krupa NL, Horsman M, Jenkins PL. Estimation of agricultural and logging injury incidence in Maine using electronic administrative data sets. *J Agromedicine.* 2015;20(2):195-204.
36. Scott EE, Hirabayashi L, Krupa NL, Sorensen JA, Jenkins PL. Developing Surveillance Methodology for Agricultural and Logging Injury in New Hampshire Using Electronic Administrative Data Sets. *J Occup Environ Med.* 2015;57(8):866-872.
37. BLS. Occupational Injury and Illness Classification System (OIICS). 2015; http://wwwn.cdc.gov/wisards/oiics/. Accessed June 25, 2015.
38. ASABE. Farm and Agricultural Injury Classification (FAIC) Code. *ASAE S575.1*: ASABE; 2002.
39. Petroleum & Other Liquids. EIA; 2016. http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=M_EPD2F_PRS_NUS_DPG&f=M. Accessed May 12, 2016.
40. Landsteiner AM, McGovern PM, Alexander BH, Lindgren PG, Williams AN. Incidence Rates and Trend of Serious Farm-Related Injury in Minnesota, 2000-2011. *J Agromedicine.* 2015;20(4):419-426.
41. Gorucu S, Murphy DJ, Kassab C. A Multi-Year Analysis of Fatal Farm and Agricultural Injuries in Pennsylvania. *J Agric Saf Health.* 2015;21(4):281-298.
42. Forst L, Erskine T. Farm injuries in Ohio, 2003-2006: a report from the emergency medical services prehospital database. *J Agric Saf Health.* 2009;15(2):171-183.
43. Hawk C, Donham KJ, Gay J. Pediatric Exposure to Agricultrual Machinery:. *Journal of Agromedicine.* 2010;1(1):57-74.
44. Missikpode C, Peek-Asa C, Young T, Swanton A, Leinenkugel K, Torner J. Trends in non-fatal agricultural injuries requiring trauma care. *Injury epidemiology.* 2015;2(1):30.
45. Lin CY, Wright J, Bushnik T, Shem K. Traumatic spinal cord injuries in horseback riding: a 35-year review. *The American journal of sports medicine.* 2011;39(11):2441-2446.

46. USDA. County Profile - Aroostook County Maine. 2007; https://www.agcensus.usda.gov/Publications/2012/Online_Resources/County_Profiles/Maine/cp23003.pdf. Accessed May 19, 2016.
47. CDC. Future Directions for the Surveillance of Agricultural Injuries; Public Meeting; Request for Comments. 2015; https://www.federalregister.gov/articles/2015/02/26/2015-03949/future-directions-for-the-surveillance-of-agricultural-injuries-public-meeting-request-for-comments. Accessed June 26, 2015.
48. Marucci-Wellman HR LM, Corns HL. A practical tool for public health surveillance: Semi-automated coding of short injury narratives from large administrative databases using Naïve Bayes algorithms. *Accident Analysis & Prevention.* 2015;84(November):165–176.
49. Bertke SJ, Meyers AR, Wurzelbacher SJ, Measure A, Lampl MP, Robins D. Comparison of methods for auto-coding causation of injury narratives. *Accid Anal Prev.* 2015;88:117-123.
50. Bertke SJ, Meyers AR, Wurzelbacher SJ, Bell J, Lampl ML, Robins D. Development and evaluation of a Naïve Bayesian model for coding causation of workers' compensation claims. *J Safety Res.* 2012;43(5-6):327-332.
51. Marucci-Wellman H, Lehto M, Corns H. A combined Fuzzy and Naive Bayesian strategy can be used to assign event codes to injury narratives. *Inj Prev.* 2011;17(6):407-414.
52. Lehto M, Marucci-Wellman H, Corns H. Bayesian methods: A useful tool for classifying injury narratives into cause groups. *Injury Prevention.* 2009;15(4):259-265.

TABLES AND FIGURES

Table 1. Denominator Data in Agriculture (2007) and Logging(2010)

	Maine	New Hampshire	Total
Agricultural Workers	28,697	12,042	40,739
Logging Workers	2,380	975	3,355

Table 2. Characteristics of Horse-Related Injuries and All Other Agricultural Injuries

	Average Age	Age Mode	Min, Max Age	Percent Female
Horse -Related Injuries	36.9	15	0, 82	88.5
All Other Agricultural Injuries	43.7	49	0, 91	25.2

Table 3. GARCH Model, Woodcutting Injuries and Home Heating Fuel Costs

Variable	Winter	Spring	Summer	Fall
Natural Gas	-0.93 (ns)	-0.99 (p=.06)	0.72 (ns)	0.69 (ns)
Propane	-0.73 (ns)	-0.98 (ns)	0.68 (ns)	-0.56 (ns)
Heating Oil	-0.54 (ns)	-0.99 (p=.06)	0.72 (ns)	-0.16 (ns)

Table 4. GARCH Model, Interaction between Season and Natural Gas

Variable	DF	Standard Estimate	Error	t Value	Approx Pr > |t|
Intercept	1	34.37	9.12	3.77	0.01
Natural Gas	1	-3.97	1.31	-3.02	0.02
Season	1	-4.79	3.48	-1.38	0.21
Interaction	1	1.54	0.54	2.83	0.02

Figure 1. Age at the Time of Agricultural Injury, Maine and New Hampshire, 2008-2010

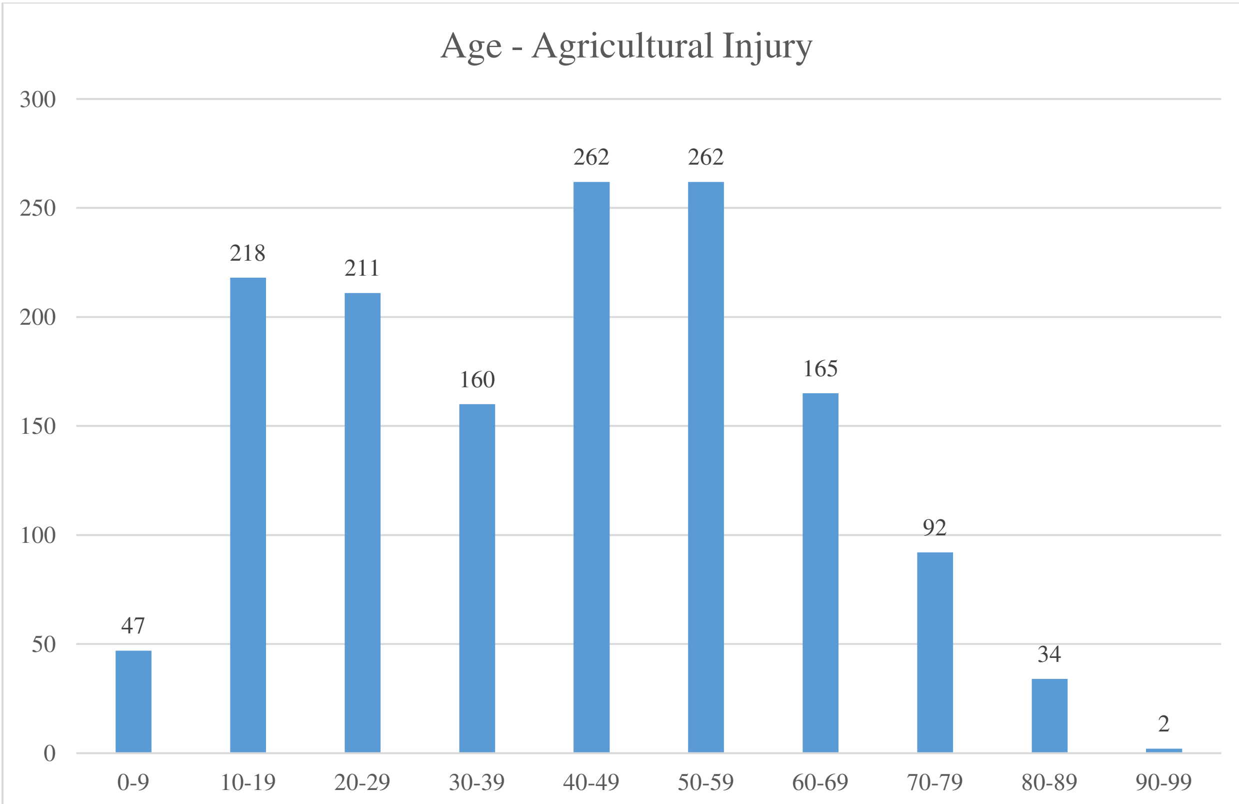


Figure 2. Location of Agricultural Injuries, Maine and New Hampshire, 2008-2010

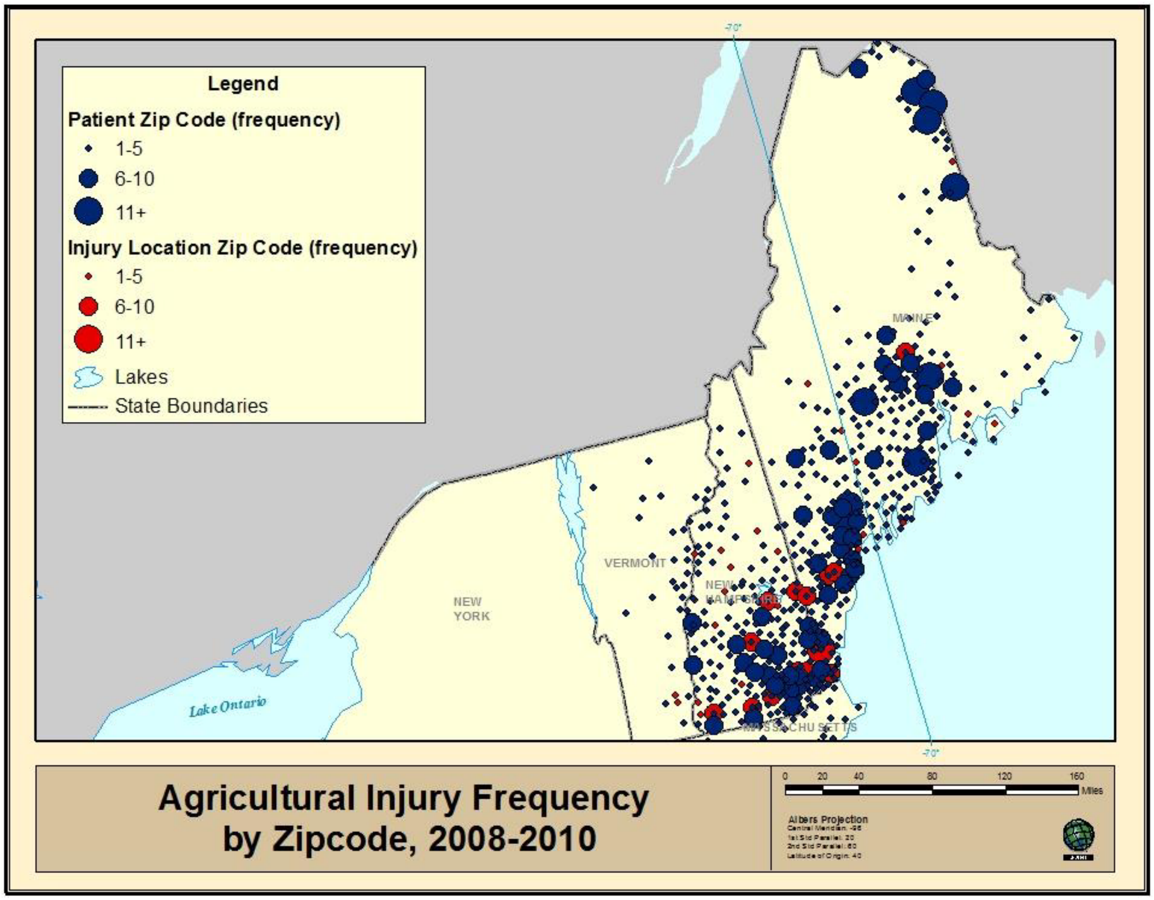


Figure 3. Agricultural Injury, by Season, Maine and New Hampshire, 2008-2010

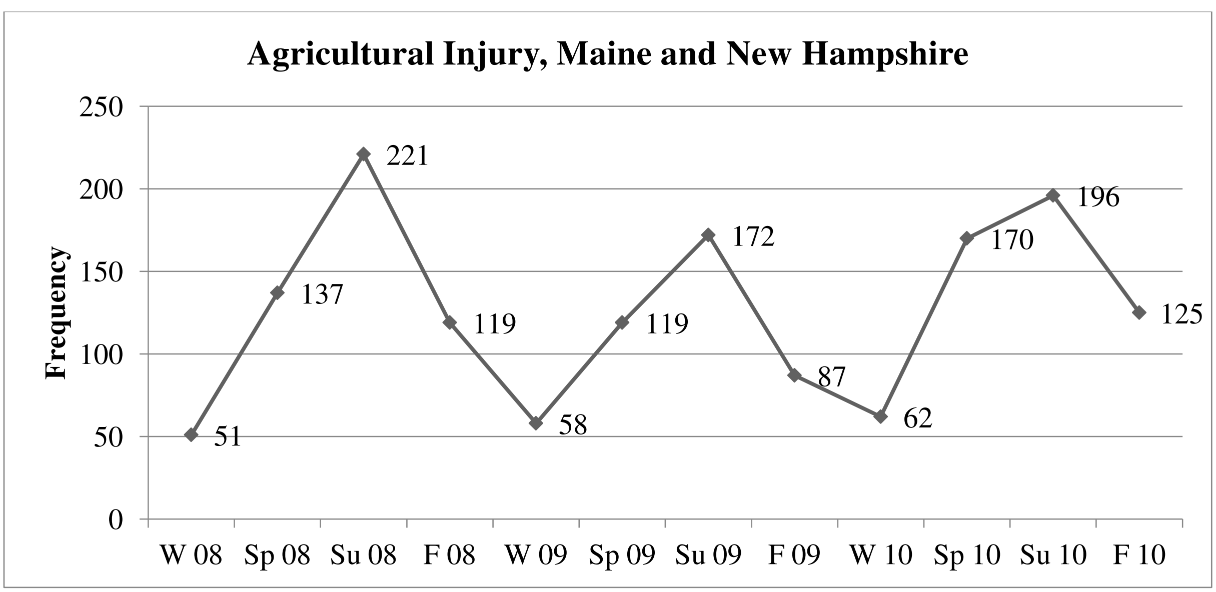


Figure 4. Agricultural Machinery Injuries, by Season, Maine and New Hampshire, 2008-2010

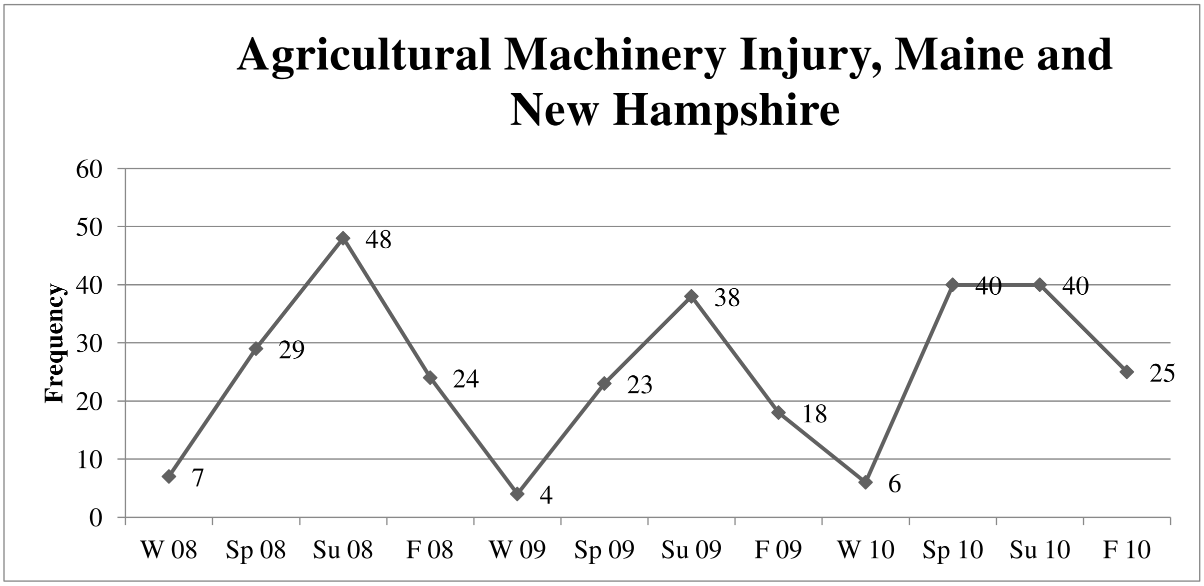


Figure 5. Agricultural Animal-related Injuries, by Season, Maine and New Hampshire, 2008-2010

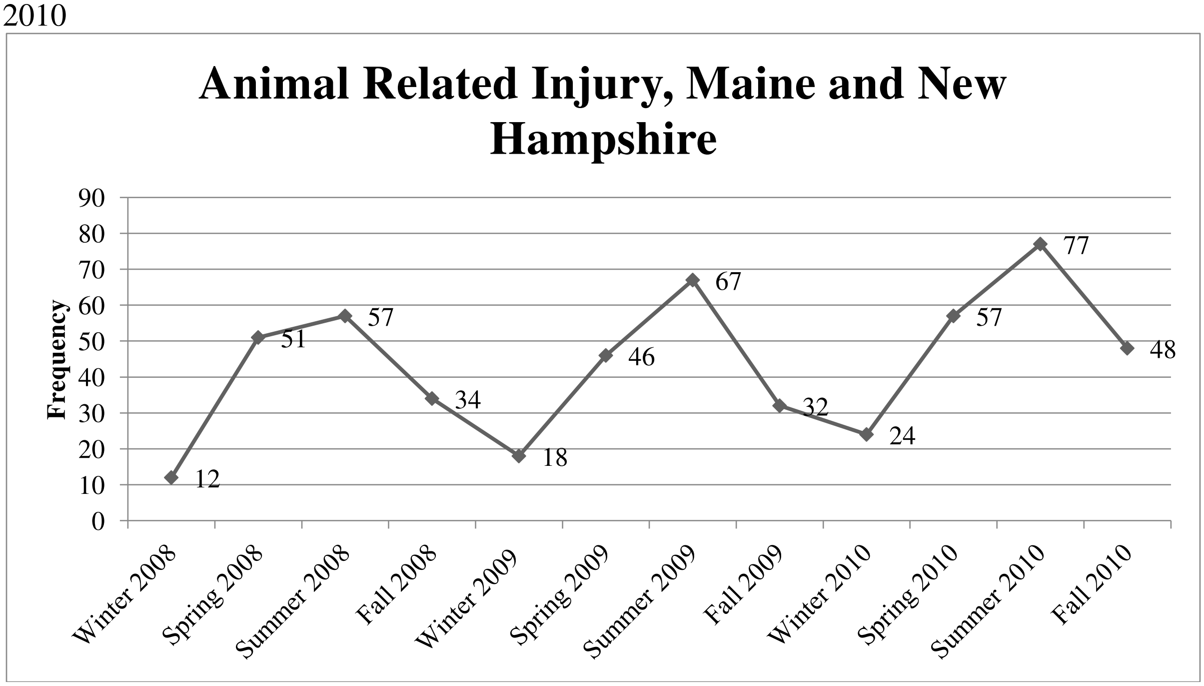


Figure 6. Age at Time of Injury in the Logging Industry, Maine and New Hampshire, 2008-2010

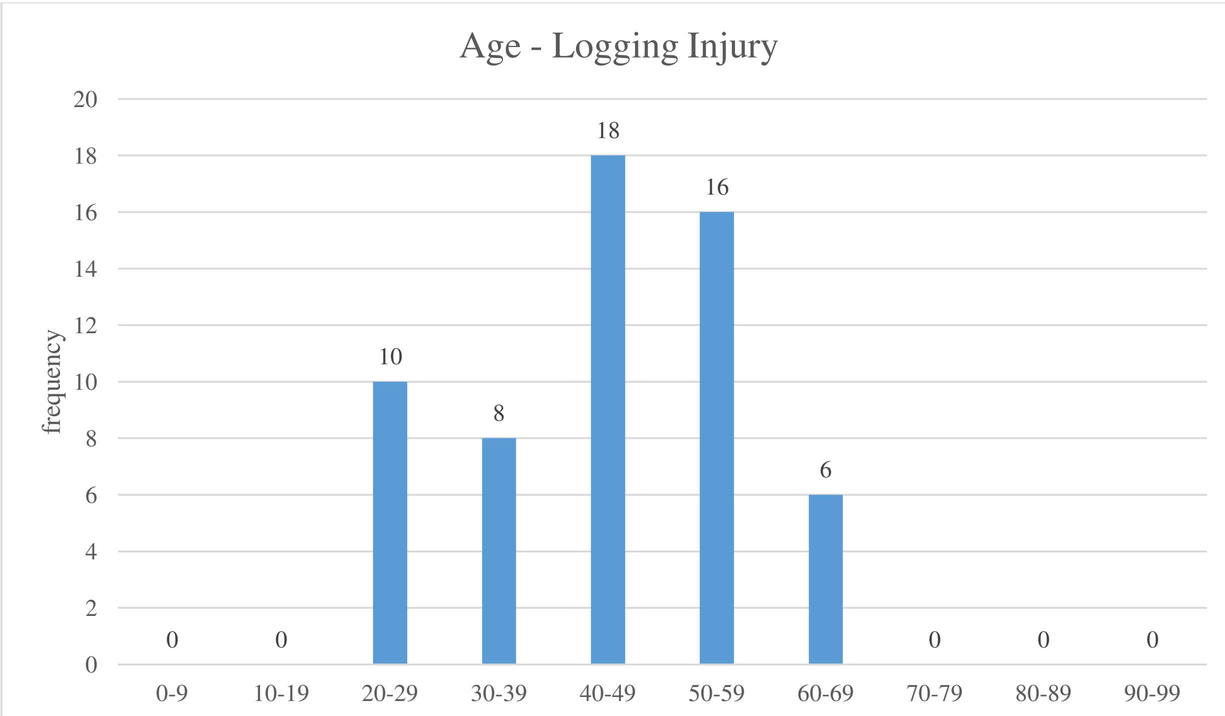


Figure 7. Logging Injuries by Season, Maine and New Hampshire, 2008-2010

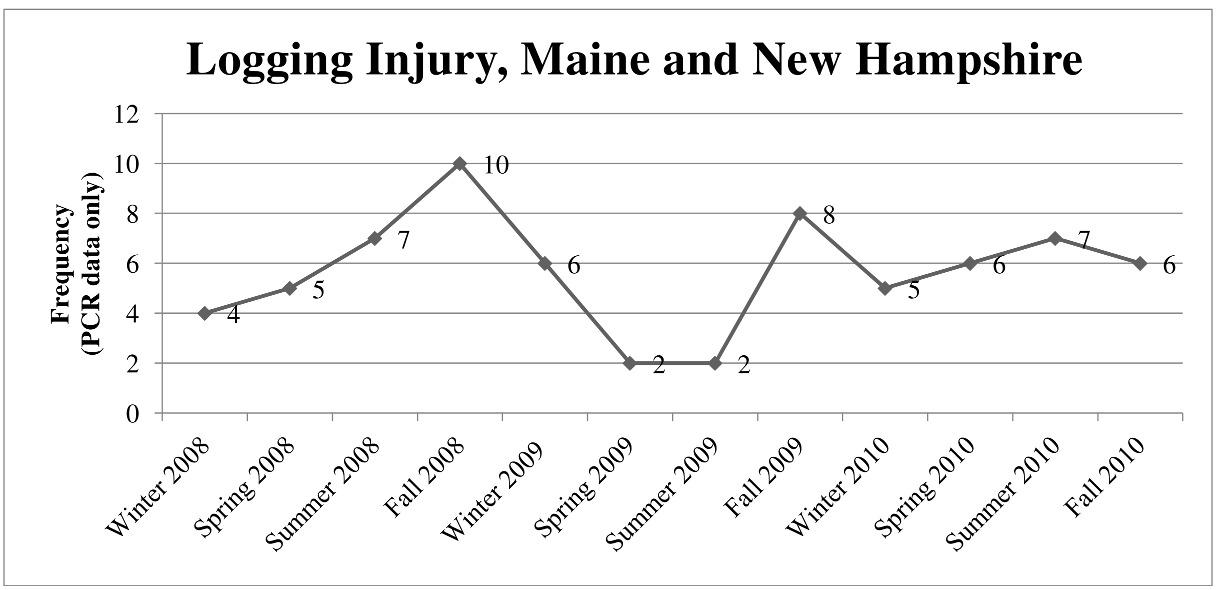


Figure 8. Chainsaw -related Injuries, Maine and New Hampshire, 2008-2010

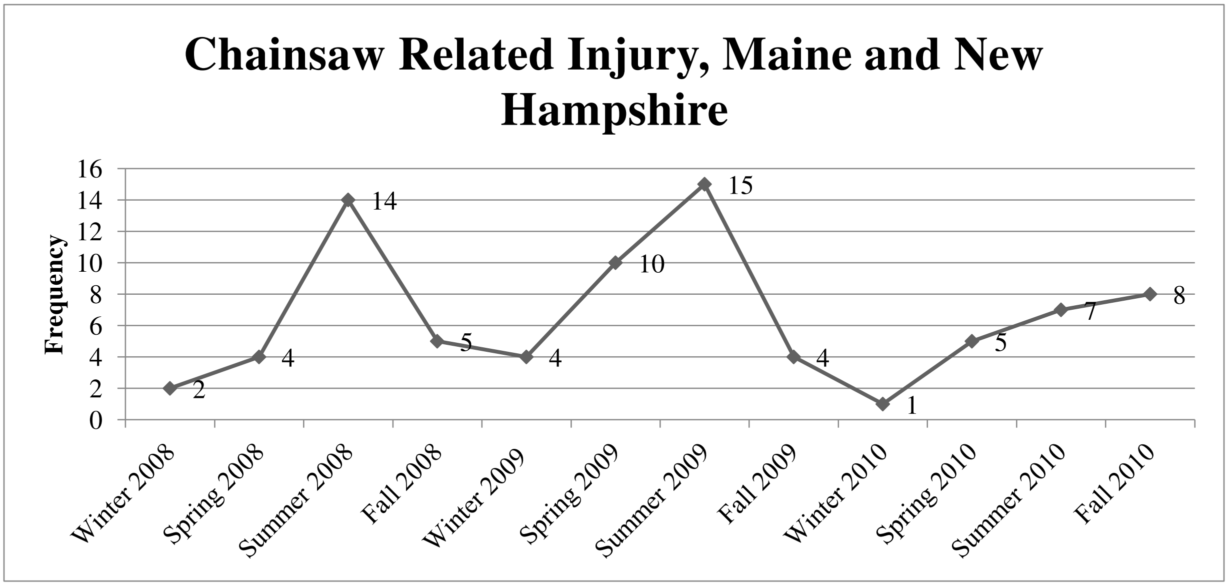
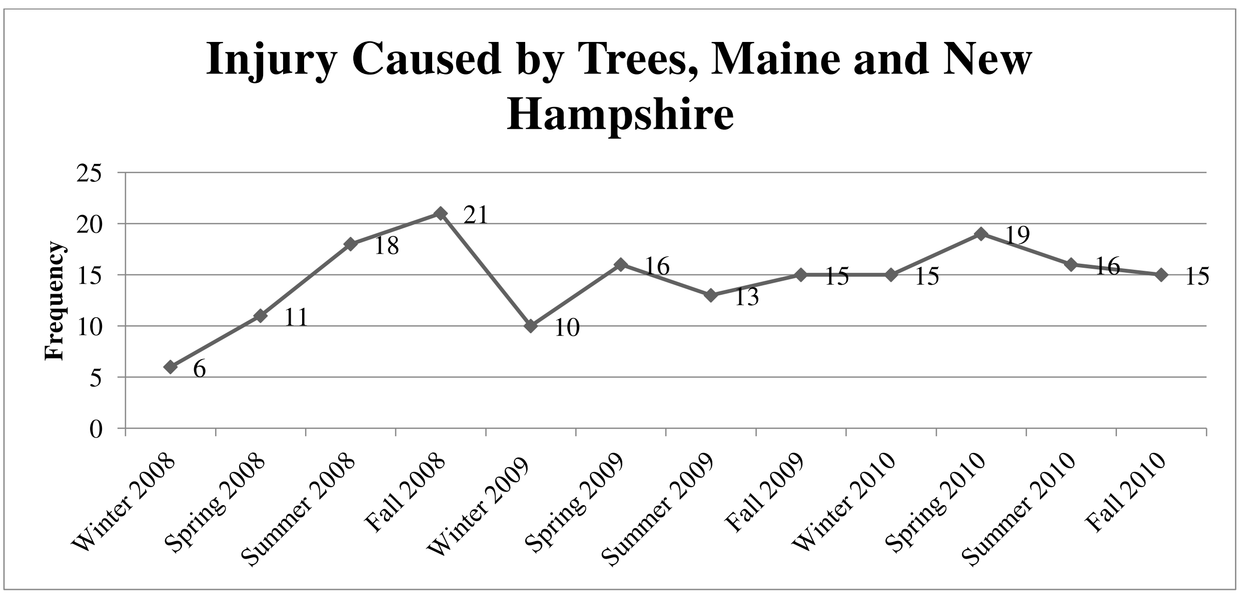


Figure 9. Tree related Injuries, Agriculture and Logging, Maine and New Hampshire, 2008-2010



CHAPTER 6
CONSIDERATIONS AND FUTURE STEPS

SUMMARY

For 2008-2010, 1,585 agricultural and logging-related injuries were identified in Maine and New Hampshire using our surveillance methods. Despite the restructuring of the New Hampshire hospital data file in 2010, which more than tripled the number of records received from this source, our methodology was able to account for such a change and identify a relatively consistent number of agricultural injuries over these three years. Administrative databases undergo transformations in content and management and this underscores the ability of the system to account for large changes in its source data and still track changes over time. The cumulative incidence of injuries in Maine and New Hampshire was found to be 12.4/1,000 for agricultural workers, compared to 10.4-12.2/1,000 for logging workers. These findings are consistent with other recent agricultural morbidity estimates. The surveillance system described in this dissertation will subsequently be referred to as the Surveillance System for Traumatic AFF Injuries at NEC (SUSTAIN).

Hospital and PCR data could be coded using the Occupational Injury and Illness Classification System (OIICS)[1] coding scheme to provide useful information for injury surveillance in agriculture and logging. While simply coding to levels one and two would prove faster, and therefore save money, it would sacrifice the potentially important information gained from coding at these more detailed levels for some categories. Based on these data, it is our recommendation that the type of event and source of injury be coded to level four wherever possible because these elements create the foundation for safety interventions. For the nature of injury and body part, it may be most cost efficient to code only as far as level two.

Machinery and animal-related events were the most frequent cause of injury in the agricultural industry in Maine and New Hampshire from 2008-2010. Within animal-related injuries, horseback riding activities are a primary concern, especially among women riders. The geographic distribution of agricultural injuries appeared consistent with agricultural areas within Maine and New Hampshire. Manual felling was a cause of both logging and agricultural injury when the injury was caused 1) by the chainsaw itself, and/or 2) by the tree or limb that was cut. Our methods are able to capture traumatic injury in agriculture and logging, with enough detail to prioritize interventions and to evaluate outcomes, using a system that is low-cost and has the potential for continued use in years to come. Differences in rates of animal and machinery-related injuries suggest the need for state specific safety prioritization.

Methodological Considerations

The Surveillance System for Traumatic AFF Injuries at NEC (SUSTAIN) is dynamic, and it will continue to evolve over time, adapting to changing administrative data. In the process of system development, we were confronted with changes to state-based record keeping that influenced the number, and quality, of records that we received. We anticipate these events as part of maintaining a robust, passive surveillance system, and we have identified several methodological issues to consider, discussed in further detail below.

Calculating Denominators

Using the proper denominator is an important aspect of calculating injury rates. While the numerator accounts for people injured in agricultural and logging events captured by the surveillance system for a particular time span, the denominator is more of a snapshot of the population at a given point in time. It is ideal to match the numerator and denominator as closely as possible (same year), but for this study and many others, it is not always feasible.

For agricultural injury cases, the denominator is derived from the Census of Agriculture, conducted by the United States Department of Agriculture through the National Agricultural Statistics Survey.[2] This census is conducted every five years, in years ending in the digit '2' or '7'. Census dates are not always congruent with numerator dates; therefore, a rule was established to make reporting consistent. Once released, that particular census data will be used for the stated year, plus four years after (e.g., for the 2007 census, 2007-2011). Likewise, when 2012 numerator data are available, data from the 2012 Census of Agriculture will serve as the denominator. The same logic is applied to logging-related injury, however these population estimates are derived from the American Community Survey.[3]

Changes in Administrative coding

On October 1, 2015, the United States adopted the International Classification of Diseases Clinical Modification, 10th edition (ICD-10-CM), which supplanted the 9th edition as the standard for coding hospital data.[4,5] With this transition, the choices for external cause of injury codes (E-codes) have dramatically increased. E-codes are the method by which we identify agricultural cases in outpatient, inpatient and emergency department data. Using the ICD-9 system, we are limited to identifying agricultural cases with hospital data. However, the availability of ICD-10-CM E-codes in newer data (October 2015 on) will allow us to explore the feasibility of including supplementary agricultural cases, in addition to forestry and fishing injury events in our surveillance system.

It will be several years before data from 2015 are introduced into the surveillance system. The surveillance system relies on transforming narrative text and administrative codes into standardized categories via the Occupational Injury and Illness Classification System (OIICS).[1] At that point, the crosswalk between E-codes and OIICS will be re-assessed for comparability.

Data Acquisition

The continued success of SUSTAIN rests upon the availability of administrative data well into the future. Availability can be defined in two ways: 1) these data are collected by the appropriate entity [state health departments, cost-containment boards] and 2) they remain available to researchers through data use agreements (DUA) or memorandums of understanding (MOU). As new electronic administrative data sources become available, we intend to explore their utility in SUSTAIN, in an effort to reduce the undercount of injury events. Trauma registries are a new dataset that may be valuable to the system. Table 1 shows the twelve states under the Northeast Center's jurisdiction, and the status of these data in each state.

Table 1. Potential Sources of data for SUSTAIN

	Existing Administrative Data		
	PCR with Narrative	Hospital	Trauma Registry
State			
Maine	**X**	**X**	N/A
New Hampshire	**X**	**X**	Y
Vermont	X	X	X
Massachusetts	X	X	W
Connecticut	**X**	**X**	W
New York	N/A	**X**	X*
Pennsylvania	N/A	Y	Y
New Jersey	W	W	N/A
Delaware	Y	W	Y
Maryland	Y	W	Y
West Virginia	N/A	Y	Y

X	Data Applications Complete
X	Data Applications Underway
W	Waiting – State Collects Data, unable to release
N/A	Appropriate Data Not Collected by State
*	Contains Industry and Occupation Variables
Y	Data available, applications not completed

Validating the SUSTAIN System

We are able to measure injury rates using our system and quantify if they increase, decrease, or remain the same; however, it is difficult to quantify the undercount that is inherent in such surveillance data. Validating the results of this passive surveillance system is an exercise well spent, and we intend to do just that by conducting a survey in conjunction with the National Agricultural Statistics Service (NASS). Results from the forthcoming survey will be compared to the results obtained from the electronic system (for those states where electronic data are available). In addition to comparing the overall estimate of incidence between the two methods, it will also be possible to compare the rank ordering of the four categories (nature of injury, part of body, type of event, and sources of injury) of OIICS codes. These comparisons will be helpful in the ongoing attempt to refine and improve the passive surveillance system.

The survey design has already been implemented by the Central States Center for Agricultural Safety and Health (CS-CASH) in the Midwest,; therefore, the results in the Northeast can be compared to those from the Midwest. Any differences in injury rates that are identified may produce information to use in future multi-region interventions. These may clarify regional differences in exposures that will help in understanding contributing factors as well as enhance collaborative efforts between the two centers. The mailed survey will be administered by the National Agricultural Statistics Service (NASS) in New York, Maine, New Hampshire, Vermont, Pennsylvania, Maryland, Delaware, Massachusetts, Rhode Island, Connecticut, New Jersey, and West Virginia using similar protocols to the Midwest survey. The survey is designed to assess demographics, injury outcomes and contributing factors for all persons (adults and children) who either live or work on the farm. Thus, the data on individual persons at risk will be clustered by farm.

Future Directions

Occupational Surveillance at the National Level

Last year saw the closing of a some long term surveillance systems that tracked agricultural injury in the United States.[6] More broadly, attention has been focused on improving occupational health surveillance, in general. The National Academies of Sciences, Engineering, and Medicine have recently begun a study aimed at 'Developing a Smarter National Surveillance System for Occupational Safety and Health in the 21st Century'.[7] This process will bring together experts in the field to guide the next wave of surveillance research. Close attention will be paid to the outcome of this study, and we plan to incorporate their best practices into our system.

Surveillance Working Group

Our organization, the Northeast Center for Occupational Health and Safety in Agriculture, Forestry, and Fishing (NEC), will be participating in the newly proposed NIOSH Agricultural Safety and Health Centers Surveillance Work Group (SWG), a collaborative effort to advance and improve surveillance of fatalities, injuries and illnesses in the agriculture, forestry and fishing (AFF) sectors. For the first time, organizations are partnering to coordinate surveillance activities between the NIOSH-funded agricultural safety and health centers, the NORA surveillance working group, government agencies, insurance companies and other entities conducting surveillance related to the agriculture, forestry, and (where indicated) fishing sectors. The cross-center collaboration of the Surveillance Working Group (SWG) is a pioneering step to standardize agricultural, forestry and fishing surveillance in the US.

The goals of the group will include standardizing and increasing the cost-effectiveness of existing surveillance methods, identifying and addressing surveillance gaps, coordinating with existing, external surveillance data sources, such as Bureau of Labor Statistics (BLS) or NASS and increasing access to and visibility of reliable AFF data.

The Next Five Years

The NEC functions on five-year funding cycles from the Centers for Disease Control and Prevention National Institute of Occupational Safety and Health (CDC-NIOSH). The grant, which supported this dissertation, provided funding from 2011 through 2016. We recently applied for additional funds to continue this important surveillance research. If awarded, we will complete several aims, which are outlined in the following section.

The proposed research builds upon this dissertation and previous studies[8-10] that used a variety of data sources to identify injuries for the agricultural and logging industries.[8,11] The long-term goal of the study will be to improve the low-cost injury surveillance methods for the agricultural, forestry and commercial fishing industries in the Northeast. This will enhance understanding of the causes of traumatic injuries, identify high-risk groups, and allow for ongoing program evaluation. If successful, similar methods can be employed to capture data in other areas of the country and for other occupational groups. To achieve the overarching aims of this proposal, several data sources will be combined. Short-term goals include optimizing narrative keyword searches, investigating state and regional trauma databanks, and exploring the utility of ICD-10-CM E-Codes for AFF injury identification. Lastly, the study will collect injury and illness data using a survey designed by the Central States Center for Agricultural Safety and Health (CS-CASH), which will permit regional comparisons of injury data and validation of the passive surveillance system. The proposed research is innovative as it takes a multi-faceted, low-cost approach to identifying and classifying occupational injuries.

Expanding Geography

The geographic network of the Northeast Center (NEC) includes: Maine, New Hampshire, Massachusetts, Vermont, Rhode Island, New York, Connecticut, New Jersey, Pennsylvania, Maryland, Delaware, and West Virginia. We currently have data from Maine, New Hampshire, and New York and are actively seeking data from Connecticut, Vermont, Maryland, and New Jersey. Every effort will be made to work with each state in question to facilitate the data requests. In the past, NEC has worked with states to develop data use agreements (DUA) and memorandums of understanding (MOU) if no such agreements were in place, and will continue to support states in this manner. Involvement will range from a state's farmers participating in the proposed injury survey, to the delivery of state based administrative datasets to be included in the surveillance system.

Expanding into New Industries

Fishing

While efforts have been made to quantify fishing related fatalities, historically, non-fatal traumatic injury surveillance in this field has been difficult.[6,12] Expansion of the surveillance to the commercial fishing industry will involve both text string searches and the use of ICD-10-CM E-codes. First, a list of fishing keywords (gantry, winch, overboard, PFD, lifejacket, fishing boat, etc.) will be used in the electronic text search algorithms that will be described in detail in the methods section of specific aim three. This list will be developed using the NIOSH NORA Dictionary of AFF[7] in consultation with a panel of commercial fishermen, Fishing Partnership Support Services (FPSS) and NIOSH fishing researchers. The keyword list will be refined using the methodology of Scott et al.[10]

Industries in New Hampshire (Beyond Agriculture, Forestry, and Fishing)

The Northeast Center (NEC) will be collaborating with the New Hampshire Occupational Health Surveillance Program (NH OHSP) to identify work-related (non-AFF) traumatic injuries from pre-hospital care reports (PCR) and hospitalization data, using techniques similar to those developed to identify agricultural injuries. The New Hampshire Trauma and Emergency Medical Services Information System (TEMSIS) mandates the collection of a 'work-related' checkbox in the PCR. This field indicates that the injury or illness occurred while working. When the work-related checkbox is marked, a free text field appears, prompting the emergency medical technician to fill in the patient's industry and job title. In addition, the narrative is retained electronically, and can be visually inspected to confirm or deny work-relatedness. For the hospital data, workers' compensation as payer will be the proxy for a work-related injury or illness, as a specific 'work-related' variable does not currently exist in the hospitalization data.

Lowering the Cost of SUSTAIN

In the next grant cycle, much energy will be invested in improving the search and coding algorithms to reduce the cost of the passive surveillance system, while not affecting data quality. One innovative approach to improve the search algorithm for the electronic text strings is the use of Bayesian methodology.[13-18] These analyses will be conducted separately for commercial fishing, agriculture and logging. The key to the success of the Bayesian approach is the identification of words that occur with high frequency in true cases and relatively low frequency in non-AFF cases. These words will best discriminate and move the posterior in the correct direction. The proposed study intends to significantly reduce the burden of visual inspection of records by increasing the specificity of the methods over and above that which currently exists. It is hoped that this increase in specificity can be achieved without degrading the sensitivity of the method. The Bayesian methodology is intended to reduce the number of injury records that require visual inspection down to a reasonable number such that this process can be a part of an ongoing cost-effective system.

Public Health Significance

Public health surveillance is defined as "the ongoing systematic collection, analysis, and interpretation of health-related data essential to the planning, implementation, and evaluation of public health practice, closely integrated with the timely dissemination of these data to those who need to know. The final link in the surveillance chain is the application of these data to prevention and control".[19] The idea, that public health surveillance is as much about collecting, analyzing and interpreting data, as it is disseminating the final product to be used to enhance the greater health of the population, is a very important one.[19] While refining the collection, analysis, and interpretation of these data are still important goals, what is essential now, is the dissemination of these data internally at NEC, and externally, to research partners.

Investing in strategies to improve the surveillance of traumatic injury helps to plan a course of action to reduce the occurrence of these events. Using data from SUSTAIN, we see a need for focused injury reduction programming for horseback riding, and we plan to explore this in the future. Results from SUSTAIN also validate the need for continuing safety interventions such as the ROPS rebate program, which is in the process of launching into a national initiative. The ROPS rebate program, which is currently available in New Hampshire, provides a financial incentive of up to seventy percent of the cost of a roll bar and installation, to a maximum of $865. Additionally, the bevy of machinery-related injuries in SUSTAIN warrants the continued expansion of power-take off (PTO) shielding intervention, where we are using social marketing techniques to promote the use of PTO shields, and actively distributing PTO shields throughout the United States. The PTO shield we sell fits ninety percent of power-take off shafts and protects the operator from coming into contact with the spinning shaft, thus avoiding near certain injury or death.

Information about logging injury from SUSTAIN has aided our decision to apply for funding which will focus on the health and safety of logging workers in Maine. We have organized a network of partners who include the Professional Logging Contractors of Maine, Certified Logging Professionals of Maine, University of Maine Cooperative Forestry Research Unit, Insurance companies, and the Maine Department of Labor, among others, to explore the root causes of logging-related injury. In addition to obtaining detailed information on injury events, we intend to collect health related data that assess the injury and disease burden of the population.

Overall, surveillance is the first of many steps in the public health model. We (at NEC) are fortunate to have the ability to follow each step of the public health model, and see it come full circle. Interventions are informed by quality data, and from that point, research to practice (R2P) is implemented. Changes in workplace safety habits are understood by the surveillance mechanism through a reduction of injury rates. Then, the process is repeated depending on the priorities of the surveillance system and the available resources.

References

1. BLS. Occupational Injury and Illness Classification System (OIICS). 2015; http://wwwn.cdc.gov/wisards/oiics/. Accessed June 25, 2015.
2. USDA. Census of Agriculture Methodology. 2015; http://www.agcensus.usda.gov/Publications/2012/Full_Report/Volume_1,_Chapter_1_US/usappxa.pdf. Accessed July 20, 2015, 2015.
3. Census U. American Community Survey. 2010; http://www.census.gov/acs/www/#. Accessed June 25, 2015.
4. CDC. ICD - ICD-9 - International Classification of Diseases, Ninth Revision. 2015; http://www.cdc.gov/nchs/icd/icd9.htm. Accessed August 26, 2015.
5. CDC. Public Health Transition to ICD-10-CM/PCS. 2015; http://www.cdc.gov/nchs/icd/icd10cm_pcs.htm. Accessed May 27, 2016.
6. CDC - Looking to the Future for Agriculture Injury Surveillance at NIOSH - Agruiculture, Forestry and Fishing Resources. 2015; http://www.cdc.gov/niosh/agforfish/aginjurysurv.html. Accessed June 29, 2015.
7. National Academies of Sciences EaM. Developing a Smarter National Surveillance System for Occupational Safety and Health in the 21st Century. 2016; http://dels.nas.edu/Study-In-Progress/Developing-Smarter-National-Surveillance-System/DELS-BANR-14-01. Accessed May 31, 2016.
8. Scott EE E-RG, Krupa N, Jenkins PL. A Correction Factor for Estimating Statewide Agricultural Injuries from. *Annals of Epidemiology.* 2011;21:767–772.
9. Earle-Richardson GB JP, Scott EE, May JJ. Improving Agricultural Injury Surveillance: A Comparison of Incidence and Type of Injury Event Among Three Data Sources. *American Journal of Industrial Medicine.* 2011;54:586-596.
10. Scott EE, Krupa NL, Horsman M, Jenkins PL. Estimation of agricultural and logging injury incidence in Maine using electronic administrative data sets. *J Agromedicine.* 2015;20(2):195-204.
11. Earle-Richardson GB, Jenkins PL, Scott EE, May JJ. Improving agricultural injury surveillance: a comparison of incidence and type of injury event among three data sources. *American Journal of Industrial Medicine.* 2011;54(8).
12. Mode NA, Hackett EJ, Conway GA. Unique occupational hazards of Alaska: Animal-related injuries. *Wilderness and Environmental Medicine.* 2005;16(4):185-191.
13. Bertke SJ, Meyers AR, Wurzelbacher SJ, Measure A, Lampl MP, Robins D. Comparison of methods for auto-coding causation of injury narratives. *Accid Anal Prev.* 2015;88:117-123.
14. Marucci-Wellman HR LM, Corns HL. A practical tool for public health surveillance: Semi-automated coding of short injury narratives from large administrative databases using Naïve Bayes algorithms. *Accident Analysis & Prevention.* 2015;84(November):165–176.
15. Bertke SJ, Meyers AR, Wurzelbacher SJ, Bell J, Lampl ML, Robins D. Development and evaluation of a Naïve Bayesian model for coding causation of workers' compensation claims. *J Safety Res.* 2012;43(5-6):327-332.
16. Marucci-Wellman H, Lehto M, Corns H. A combined Fuzzy and Naive Bayesian strategy can be used to assign event codes to injury narratives. *Inj Prev.* 2011;17(6):407-414.
17. Lehto M, Marucci-Wellman H, Corns H. Bayesian methods: A useful tool for classifying injury narratives into cause groups. *Injury Prevention.* 2009;15(4):259-265.
18. Nanda G, Grattan KM, Chu MT, Davis LK, Lehto MR. Bayesian decision support for coding occupational injury data. *J Safety Res.* 2016;57:71-82.
19. CDC. *Comprehensive Plan for Epidemiologic Surveillance.* Atlanta1986.

APPENDICES

Appendix A

Surveillance Characteristics

Table 1. Considerations for Surveillance Systems[1]

Justification	frequency/severity of problem direct/indirect costs isolated events or communicable disease public interest? interventions possible?
Objectives	measuring scope of problem underlying population monitoring changes feasible?
Case Definition	specific reproducible
Operational Considerations	staff how often are data analyzed? frequency of reporting?
Data Acquisition	data use agreements
Data Volume	seasonal or constant data? what variables are needed? maintaining timelines

Justification

First, there must be justification for establishing the system.[1] The disease or health related event must affect a significant portion of the population, or be severe enough to need excessive public health resources or expenditures and to warrant being under surveillance.

Objectives

Objectives must be established for any given surveillance system, as they are part of a thoughtful plan to effect public health change. Perhaps the most important, is measuring the scope of the public health problem. A well designed and administered surveillance system should address the points raised in the table above.[1]

Case Definition

To properly identify cases, there must be a case definition. The definition describes the epidemiology of the factor, or factors, that are chosen to detect the health related event and may include symptoms, quantitative medical results (laboratory results, diagnostic imaging, etc), behavioral patterns or demographic information .[1] Case definitions must be very specific in order to be reproducible, which is important for comparison across time and space (i.e. different locations or underlying populations). Planning a surveillance system requires detailed logistics. While it may seem straightforward to identify what templates are needed for data collection or if personal identifiers are needed (for example), these factors are integral to the success and feasibility of a good surveillance study.

Operational Considerations

Operational considerations should be at the forefront of planning a system, as these dictate what can be under study and how accurately. A surveillance system requires many inputs ranging from adequate staffing, data collection protocols, data use agreements (when gathering secondary data), security measures (both physical and computer), among others.

Data Acquisition

To facilitate the acquisition of data, cooperative agreements must be established between public health agencies and other data clearinghouses. It is of utmost importance to obtain acceptance among all participating parties. Data use agreements (DUA) and memorandum of understanding (MOU), along with institutional review board (IRB) documentation must be finalized before any surveillance activities may take place. Agencies involved should be kept up to date on study findings to ensure continued participation.

Data Volume

Other considerations to contemplate include expected data volume, seasonality in which data are collected and released, identifying exactly what variables are needed to answer the surveillance question, among others. It is important to mobilize the implementation of the surveillance system as soon as interested parties have completed necessary paperwork – this ensures that agencies do not lose interest or focus in the public health issue at hand.

Beyond the logistics of establishing a surveillance system, several attributes are important to the success of a surveillance system. Once the system is established, one must determine its performance. Both the structure and operation of the system should be relatively simple.[2]

Well-designed surveillance systems contain qualities that are standard, regardless of the type of surveillance. These attributes ensure the system runs smoothly and efficiently (see table 2).

Table. 2 Attributes of Surveillance Systems (adapted from Lee et al.)[2,3]

Attribute	**Definition**
Simplicity	Structure and ease of operation
Flexibility	Ability to adapt to changing needs, both technical and information
Data quality	Completeness and validity of data
Acceptability	Willingness of stakeholders to participate in system
Sensitivity	Proportion of cases detected, or the ability to detect outbreaks over time
Predictive value positive	Proportion of reported cases that have the event under surveillance
Representativeness	Ability to accurately describe the event over time and the distribution in the population
Timeliness	Speed between steps in the system, including receiving data
Stability	Reliability and availability

Evaluating the surveillance system is necessary to ensure the effective tracking of public health issues. The Centers for Disease Control and Prevention updated evaluation guidelines in 2001,[2] and these are often the standard to which assessments are conducted. This ensures quality data are being included in the system and that it remains effective. Important components of an evaluation include direct and indirect costs to maintain the system, the purpose of the system,

Simplicity

Items to evaluate include determining what data are needed to meet the case definition, the volume of extraneous data, does it add to the information for the health-related event, how many agencies receive this data, and is it integrated with other systems. The methods used to store and analyze data and the staff requirements all have a bearing on the simplicity of design. Ideally, the system should be built to require minimal maintenance, though reality often dictates greater operational involvement, especially at the inception of a new system (to identify and solve technical issues).

Flexibility

New demands on a surveillance system, for instance - using the same model for another disease, becomes a test of its flexibility. Entire surveillance systems may be adapted for use in identifying emerging trends, or new health events can be added to an existing case definition. Surveillance administrators may change the system, as needed, however these changes must be documented, as they will affect what data is captured. Simple systems are often more flexible, meaning that fewer components must be changed to look at an expanded case definition or additional data sources.[2]

Data quality

Data quality varies dependent on the initial source of information. Systems relying on secondary and administrative data inherently have data quality issues that cannot be resolved by the surveillance administrator, as these data are entered into the surveillance system, as is. However, it is imperative that the surveillance administrator understand the limitations of their data, and can account for such when reporting findings. Data completeness refers to the extent in which the template or reporting form is filled out, and data validity refers to the correctness of the information that is present on the form.

Acceptability

Participating organizations must find the system acceptable, to continue either reporting cases, or providing the administrative datasets. Other factors that influence acceptability of a system include the perceived importance of the disease or event, how much time and effort are involved to report cases, what safe guards are in place for confidentiality, the level of trust in those safe guards, whether data produced by the final surveillance system is useful to the agency, and voluntary versus mandatory reporting.[2]

Sensitivity

The sensitivity of a surveillance system refers to both the proportion of cases (injury/disease) identified by the system, and its ability to detect changes in case numbers over time.[2] Both measures are affected by what events take place in the population under surveillance and the ability of diagnostic tests (or the recording of proper information) to determine a true case. To measure the sensitivity of the system, outside data needs to be sampled. Occasionally medical chart review or registries aid in this process. Additionally, the capture-recapture technique can be used to assess sensitivity, as well.[2]

The quality of a surveillance system relates to the completeness of data, and how timely those data are. If data for a surveillance system are gathered from administrative datasets, researchers lack control over data completeness. Quality assurance and quality control lie with the originating agency. Frequently, these types of data (administrative datasets) are the best available and the most cost effective. However, limitations of these sources exist. For example, if detailed data are missing, the power of the surveillance study can be questioned, and results cannot be generalized to specific outcomes.

Predictive value positive

The positive predictive value (PVP), which refers to the proportion of reported cases that are truly a disease (or injury) in the surveillance system, is an important measure related to sensitivity, which refers to the number of diseased (or injured) that are correctly identified as having the disease (or injury). A high positive predictive value will indicate there are not many false positive cases.[2] The same could be said for specificity, or the probability a reported case will be reported will not be a true case. Lastly, one must consider the representativeness of the surveillance system. Does this system capture cases equally across differing demographics? Does it underreport in some areas, leading to assumptions of increased prevalence in others? Ideally, the system will well represent the burden of disease or injury in the underlying population.[2]

Representativeness

Representativeness refers to the ability to use surveillance data to describe a health related event within a certain population. The data's denominator is used to describe the population from which the data is selected, therefore can be informative to whom the results of the surveillance system can be generalized to. Representativeness also relates to who is included in the numerator data of a surveillance system. For instance, if ambulance reports are a primary source of data, then the results may not generalize to patients in the hospital that were not transported by an ambulance.

Timeliness

How timely data can be introduced into the surveillance system is one factor; however, another important factor is the processing, cleaning and analyzing of the data. Timeliness may mean different things for different types of health events. The sense of urgency is much greater for an infectious disease that spreads quickly among people, versus an injury surveillance system, where even data a few years old may still be relevant.

Stability

Surveillance systems should be stable over time, providing reliable estimates on the health event under watch. This is accomplished by the convergence of the elements outlined in the section above. The surveillance system's ability to provide data in a routine fashion lends stability to the system. The availability of such data to stakeholders is valuable in guiding program planning and targeting new interventions.

APPENDIX B

INDUSTRY AND OCCUPATION CLASSIFICATION SCHEMES

Industry sectors refer to creation of goods or services within a particular area of the economy. Occupation refers to specific job title with certain job tasks within an industry. Classification schemes exist for both industry and occupation, and are important in assigning titles and groupings for particular workers. The North American Industrial Classification System (NAICS) is used by federal agencies and researchers to standardize industry classifications.[4] It contains twenty major sectors, which are shown in Table 3.

Table 3. NAICS Classification System[4]

Sector	Description
11	Agriculture, Forestry, Fishing and Hunting
21	Mining, Quarrying, and Oil and Gas Extraction
22	Utilities
23	Construction
31-33	Manufacturing
42	Wholesale Trade
44-45	Retail Trade
48-49	Transportation and Warehousing
51	Information
52	Finance and Insurance
53	Real Estate and Rental and Leasing
54	Professional, Scientific, and Technical Services
55	Management of Companies and Enterprises
56	Administrative and Support and Waste Management and Remediation Services
61	Educational Services
62	Health Care and Social Assistance
71	Arts, Entertainment, and Recreation
72	Accommodation and Food Services
81	Other Services (except Public Administration)
92	Public Administration

The Standard Occupational Classification (SOC) is used by the Bureau of Labor Statistics (BLS) and other federal agencies to subset all occupations into 840 detailed definitions.[5] There are 23 major groups and 97 minor groups within the scheme. This allows for detailed data analysis at the occupational level.

Several coding systems have been developed to address occupational injury and illness. The BLS developed the Occupational Injury and Illness Classification System (OIICS)[6] in 1992 to provide consistency in coding for occupational injury and illnesses. Since then, they have collaborated with NIOSH to provide updates to this coding scheme. The scheme has coding trees for nature of injury, part of body, type of event and source of injury. The categories may be expanded or collapsed to provide greater or less detail, depending on the available injury information.

In addition to overarching coding schemes such as OIICS, other special codes have been developed to aid the injury researcher in classifying cases. One example used frequently for the agriculture, forestry and fishing industry is the American Society of Agricultural and Biological Engineers (ASABE) Farm and Agricultural Injury Classification (FAIC)[7] scheme. FAIC codes separate workers and bystander injuries on the farm, which can be very useful information.

APPENDIX C

TABLE 4. KEYWORDS USED FOR FREE-TEXT SEARCH

AGRICULTURE	COW	FREE STALL	MANURE	SLACKLINE	UDDER
AMISH	CHOKER	GRAIN BIN	METHANE	SLURRY STORE	UNCAPPING
ANIMAL	CHOPP	CORN BIN	POST HOLE DIGGER	SPREADER	UNHITCH
ARCH	FORAGE	OAT BIN	FENCE POST	SPRING POLE	WEDGE
AUGER	COMBINE	GRAVITY BIN	POWER TAKE OFF	STALL	WINCH
BALE	CROP	APPLE BIN	POWER TAKE-OFF	STRAW	YARDER
BALING	HAY	GUYLINE	PTO	THREE POINT HITCH	YARDING
BARN	DEBARK	GUYWIRE	PRUNING	3 POINT HITCH	CHAIN
TREE	DEFACER	HARROW	FEED	3 PT HITCH	CLEANSER
BEATERS	IMPLEMENT	HITCH	RAM	TIE DOWN	FENCE
BIND	DRIVELINE	HOG	ROPS	TIE STALL	SANITIZER
BLADE	DRIVE LINE	CHICKEN	MOWER	ENTANGLEMENT	RAKE
BREEDING	HORSE	IRRIGATION	SHEAVE	PEN	TEDDER
BUCK	FARM	KICKBACK	WAGON	PENS	POULTRY
BUGGY	CALV	KICKER	SHEEP	PESTICIDE	DIGGER
BULL	FERTILIZER	CHUTE	SILAGE	PIG	CART
BUNKER	SPRAYER	LIMBING	SKIDSTEER	PIPELINE	GOAT
CABLE	PASTURE	LIVESTOCK	BOBCAT	PLOW	SILO
CHAIN SAW	FOPS	LOADER	SKIDD	TIMBER	SKID STEER
CHAINSAW	WOODS	LOG	SLABB	TRACTOR	VACUUM PUMP
CULTIVAT	FREESTALL	SCRAPER	SLACK LINE	TROUGH	PLANT

References

1. Nash DEZ. Public Health Surveillance. In: EPIC, ed. *Epidemiology and Population Health Summer Institute*. New York, New York: Columbia University; 2013.
2. CDC. Updated Guidelines for Evaluating Public Health Surveillance Systems. *MMWR* 2001; http://www.cdc.gov/mmwr/preview/mmwrhtml/rr5013a1.htm. Accessed August 3, 2015.
3. Lee LM, Teutsch SM, Thacker SB, Louis MES. *Principles and Practice of Public Health Surveillance.* 3rd Edition ed. New York, New York: Oxford University Press; 2010.
4. Special Projects Staff SSSD. North American Industry Classification System (NAICS) Main Page. 2010.
5. BLS. Standard Occupational Classification (SOC) System. 2015; http://www.bls.gov/soc/. Accessed June 2016, 2015.
6. BLS. Occupational Injury and Illness Classification System (OIICS). 2015; http://wwwn.cdc.gov/wisards/oiics/. Accessed June 25, 2015.
7. ASABE. Farm and Agricultural Injury Classification (FAIC) Code. *ASAE S575.1*: ASABE; 2002.