Journal of Agromedicine
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/wagr20

Estimation of Agricultural and Logging Injury Incidence in Maine Using Electronic Administrative Data Sets

Erika E. Scott[a], Nicole L. Krupa[b], Melissa Horsman[a] & Paul L. Jenkins[b]

[a] The New York Center for Agricultural Medicine and Health, Bassett Healthcare Network, Cooperstown, New York, USA

[b] Bassett Research Institute, Bassett Healthcare Network, Cooperstown, New York, USA

Published online: 23 Apr 2015.

Journal of Agromedicine, 20:195–204, 2015

ISSN: 1059-924X print/1545-0813 online
DOI: 10.1080/1059924X.2015.1009668

Taylor & Francis
Taylor & Francis Group

Estimation of Agricultural and Logging Injury Incidence in Maine Using Electronic Administrative Data Sets

Erika E. Scott,[1] Nicole L. Krupa,[2] Melissa Horsman,[1] and Paul L. Jenkins[2]
[1]The New York Center for Agricultural Medicine and Health, Bassett Healthcare Network, Cooperstown, New York, USA
[2]Bassett Research Institute, Bassett Healthcare Network, Cooperstown, New York, USA

ABSTRACT. Agriculture and forestry rank among industries with the highest rates of occupational fatality and injury. Establishing a nonfatal injury surveillance system is a top priority in the National Occupational Research Agenda. Recently, new sources of data such as Pre-Hospital Care Reports (PCRs) and hospitalization data have transitioned to electronic databases. Using narrative free text and location codes from Maine PCRs, along with International Classification of Diseases (ICD)-9 External Cause of Injury Codes (E-codes) in Maine hospital data, researchers are designing a surveillance system to track farm and forestry injury that utilizes electronic match-merging of the two data sources. For 2008, PCR records produced a total of 104 true agricultural cases. Of these, 66 (63%) were identified from the keyword/visual inspection process alone, 25 (24%) were identified by the farm checkbox only, and the remaining 13 (13%) by both methods. For the 150 unique injury events found in hospitalization data, 146 had the initial episode of care documented in only one of the three hospital files. The emergency department (ED) file had the largest number of these (123/146 = 84.2%), followed by the outpatient file (12/146 = 8.2%) and the inpatient file (11/146 = 7.5%). Of the 250 unique agricultural injuries identified (100 PCR only + 146 hospital only + 4 from both), 66 (26%) would not have been identified without free text review of PCR narrative. The false-positive rate (97.14%) keyword searches underscores that without visual inspection, it is not an effective strategy. Both sources of data (PCR and hospital data) need to be used in a continued surveillance system.

KEYWORDS. Agriculture, electronic databases, forestry, injury surveillance

PURPOSE

Agriculture and forestry rank among industries with the highest rates of occupational fatality and injury. According to data from the Bureau of Labor Statistics, logging workers had the highest fatality rate of any occupation for 2012, with a rate of 127.8 deaths per 100,000 full-time equivalent (FTE) workers. Farmers, ranchers, and agricultural managers also topped the list, with 216 fatalities in 2012, for a national fatal injury rate of 21.3/100,000 FTE.1 To compare, the overall worker fatality rate in 2012 was 3.4/100,000 FTE.[1]

Nearly 91% of America's farms are considered small farms, and the Occupational Safety and Health Administration (OSHA) does not have formal jurisdiction over the safety of these workplaces.[2] OSHA records are common data

Address correspondence to: Erika E. Scott, Northeast Center for Occupational Health, One Atwell Road, Cooperstown, NY 13326, USA (E-mail: erika.scott@bassett.org).

sources for major industry,[3] although they have limited use in agriculture, especially for small family-owned farms. 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), enforcing recordkeeping is difficult.[4] Although it is acknowledged that forestry workers experience elevated rates of injury and fatality, little work has been done in this arena. Moreover, many gaps still exist in injury data for this industry.[5]

Establishing a nonfatal injury surveillance system akin to the Census of Fatal Occupational Injuries remains a top priority for agricultural injury epidemiologists.[6] The first strategic goal of the National Occupational Research Agenda (NORA) is to "Improve surveillance within the Agriculture, Forestry, and Fishing Sector to describe: the nature, extent, and economic burden of occupational illnesses, injuries, and fatalities; occupational hazards; and worker populations at risk for adverse health outcomes."[7]

Recently, several data sources have become more widely available to researchers. Emergency medical services (EMS) data have attracted new attention as a source for surveillance. Traditionally used for quality assurance and quality control by the Bureau of Emergency Medical Services, the narrative free text contains an abundance of useful data, not only for injury surveillance, but for many other outcomes.[8] There are two methods in which a farm or logging injury can be identified from the ambulance report: having the "farm" location box checked off, or through the free text account of the event. Free text fields have been used for agricultural- and logging-related traumatic injury in the past.[9] Farm injuries can be identified in hospital data using International Classification of Diseases (ICD)-9 external cause of injury codes, also known as "E-codes." In the past few decades, E-code use has increased substantially and has proven valuable in injury prevention research.[10] Other research has been beneficial in linking workers compensation claims with hospital discharge data,[11] and this general concept can be applied to EMS Pre-hospital Care Reports (PCRs) and hospital records. Data linkage allows the maximum case ascertainment,[12] which is very important when studying injury events that are difficult to identify in administrative databases. Agricultural cases are well known for posing a challenge for injury surveillance researchers, as (1) there are limited useful terms and variables for capturing such cases, and (2) many patients do not present at the hospital,[13] even for injury events.

METHODS

Subjects (Population at Risk)

The subjects are all persons who come into contact with agricultural or logging activities in the state of Maine and are therefore at risk for agricultural or logging injury. Because of the inherent difficulties with establishing the exact size of this population, it will be taken as the number of agricultural and logging workers as defined by the Census of Agriculture and US Census. This convention creates a situation where small numbers of individuals, for example, visitors to a farm, could contribute to the numerator for rate calculations without being counted in the denominator. This phenomenon would be expected to slightly inflate the injury rates. However, this overestimation is counterbalanced by missed cases, with the net effect of these two opposing forces difficult to quantify.

Data Acquisition

Administrative clearance to receive the data was obtained from the Maine Health Data Organization and the Maine Bureau of Emergency Medical Services to receive a data set that included personal identifiers. For the EMS data, the data manager had access to the state EMS portal from which the data were downloaded. The PCR contains variables related to a patient's treatment at the scene and during transport. In addition, narrative fields contain a written description of the injury or illness event.[14] The hospital data (inpatient, outpatient, and emergency department) were received on CD from the Maine Health Data Organization. These data, used for hospital billing, contain codes related to the injury and illness diagnosis, treatment, and discharge of patients. EMS data, which were also received from the New Hampshire Bureau of Emergency Medical Services, were utilized for certain comparative purposes. All data were read into permanent SAS (SAS Institute, Cary, NC) data sets for analysis.

Definition of Agricultural and Logging Events

The case definition describes exactly what the surveillance system is intended to identify and quantify. A traumatic agricultural injury is defined as when energy is transferred to an individual (a) from an agricultural source (e.g., tractor, bull), (b) while in an agricultural location, or (c) while doing an agricultural activity that results in physical harm severe enough to require medical attention. Excluded are injury events that occur on the farm involving sources not associated with farm activity at any time (e.g., skateboard, barbecue grill, etc.). Likewise, a logging-related injury is the occurrence of an event where energy is transferred to an individual from a forestry or logging source (e.g., tree, chainsaw), or while doing a logging activity (e.g., felling, limbing, splitting logs, and lumber) that results in physical harm for which immediate medical attention is sought. It is important to note that both definitions exclude injuries for which immediate medical attention is not sought. This exclusion is due to the design of the injury surveillance system. However, there is justification for limiting surveillance to emergent injury events, as these injuries are the most severe, and are more often to be identified in the data sources as work-related. Many other studies use this definition.[15–18]

In the hospital data, agricultural events were identified using ICD-9 external cause of injury codes. Specifically, ICD-9 codes E849.1 (farm location [codes exist that negate the inclusion of E849.1—available from researchers]) and E919.0 (contact with agricultural machinery [provided the following location E-codes are not included: E849.3, E849.4, E849.6, E849.7]) denoted agricultural injuries. There are currently no E-codes for logging in the ICD-9 system.

In the PCR data, if the farm location checkbox was marked, this was taken as a preliminary indication of agricultural relatedness. In addition, the free text string for the narrative was also searched for specific keywords. Keywords used for the agricultural and logging cases were derived from the National Institute for Occupational Safety and Health (NIOSH) NORA Dictionary of AgFF,[7] in conjunction with input from two agricultural researchers. The initial set of keywords used is available from the researchers. All keywords that returned any records are listed in Table 1. To maximize the number of records available for examination of the keywords, data from New Hampshire were also utilized, as will be explained below.

For the keyword “field” a random sample of 500 records was visually inspected to determine if the injury was actually agriculture- or logging-related. This sampling was necessary due to the large number of records (18,217) containing this keyword. For all other keywords, all records were inspected. The proportion that were actually related to agriculture or logging were then summarized for each keyword. 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.

Treatment of Duplicates and Merging of Data Sets

Prior to merging, records with duplicate data for all five variables in the hospital file (emergency department [ED] and inpatient) were consolidated into a single record that incorporated all of the data from both records and the frequency of missing values on the matching variables was assessed using PROCFREQ in SAS. The merging of the PCR and hospital data occurred in three phases. In phase 1, the two data sets were match-merged using patient’s gender, admission date, ZIP code, date of birth, and the facility code receiving the patient. An exact match on all five of these variables was required. Multiple PCR records that were duplicates on

TABLE 1. Text String Search for Keywords Returning at Least One True Agricultural or Logging Case

Keyword	Returned	True hits	Hit rate	Keyword	Returned	True hits	Hit rate
Pruning	1	1	1	Crop	12	0	0
Three point hitch	1	1	1	Sheep	12	0	0
Tie Down	2	1	0.500	Buck	11	0	0
Chainsaw	8	3	0.375	Wedge	11	0	0
Barn	15	4	0.267	Silo	11	0	0
Chain	63	10	0.159	Bull	10	0	0
Horse	64	9	0.141	Hitch	9	0	0
Hay	8	1	0.125	Udder	9	0	0
Tractor	128	11	0.086	Chute	8	0	0
Skidd	54	3	0.056	Pasture	7	0	0
Log	262	11	0.042	Loader	6	0	0
Farm	135	5	0.037	Bunker	5	0	0
Bale	27	1	0.037	Goat	5	0	0
Tree	310	11	0.035	PTO	5	0	0
Wagon	39	1	0.026	Cow	4	0	0
Chicken	68	1	0.015	Pens	4	0	0
Feed	68	1	0.015	Pig	4	0	0
Woods	229	3	0.013	Ram	4	0	0
Blade	116	1	0.009	Winch	3	0	0
Pen	286	0	0	Amish	2	0	0
Cart	139	0	0	Buggy	2	0	0
Animal	53	0	0	Entanglement	2	0	0
Mower	50	0	0	Pesticide	2	0	0
Plow	49	0	0	Pipeline	2	0	0
Plant	48	0	0	Sanitizer	2	0	0
Straw	46	0	0	Spreader	2	0	0
Calv	34	0	0	Vacuum pump	2	0	0
Implement	34	0	0	Beaters	1	0	0
Fence	32	0	0	Bobcat	1	0	0
Auger	29	0	0	Breeding	1	0	0
Timber	23	0	0	Cleanser	1	0	0
Trough	23	0	0	Cultivat	1	0	0
Hog	22	0	0	Digger	1	0	0
Chopp	21	0	0	Fertilizer	1	0	0
Bind	20	0	0	Kickback	1	0	0
Combine	20	0	0	Kicker	1	0	0
Irrigation	19	0	0	Manure	1	0	0
Arch	18	0	0	Methane	1	0	0
Stall	16	0	0	Rake	1	0	0
Cable	12	0	0	Scraper	1	0	0
				Sheave	1	0	0

five variables had one record deleted. These deletions were done at random after it was shown that deleting based on preset criteria did not produce different results.

In phase 2, for any subject who did not match between the PCR and hospital files using the five variables, an attempt was made to match with the same variables excluding facility code. An exact match was required on all four variables (patient's gender, admission date, ZIP code, date of birth). Duplicates were handled in this phase prior to matching in an analogous manner to that described above, but using only the four matching variables.

In phase 3, this entire process was repeated for both the four- and five-variable match except that 1 day was added to the date of ambulance transport. This was done in an attempt to match records where the hospital admission occurred early the following morning of the ambulance run. This protocol was approved by the Institutional Review Board of the Mary Imogene Bassett Hospital.

Coding of Confirmed Cases as to Agricultural or Logging Origin

Two independent researchers reviewed each narrative that was retained in the PCR. From this information, true cases were classified as either "forestry or logging" (FAIC-2) or "agriculturally related" (FAIC-1, FAIC-4, FAIC-5, FAIC-6, FAIC-7, FAIC-8, FAIC-9) according to the definitions found in the American Society of Agricultural and Biological Engineers (ASABE) Farm and Agricultural Classification (FAIC) Code.[19] The results were compared and any discrepancies in coding between the two researchers were resolved by a third researcher.

Data Analysis

The proportion of the agricultural population with at least one injury requiring medical care was estimated as all agricultural cases identified divided by the total number of agricultural workers in the state. Similar logic was used to estimate the proportion of loggers with at least one injury requiring medical attention.

Table 1 was created to show the actual "hit rate" produced for each keyword that produced at least one true agricultural case. Any keyword found in at least 100 records, and which did not return at least one true farm case on visual inspection, was then examined in the New Hampshire data. If this keyword also did not result in the identification of a single true case in the New Hampshire data, it was deemed to be unnecessary for future free text review. A complete list of all 134 keywords searched is available from the authors.

RESULTS

PCRs

A total of 115,001 PCR records were received for 2008. Of these, 0.88% of PCR records had nothing entered in the free text field. Ultimately, as will be described below, a total of 104 of these (104/115,001 = 0.09%) contained a true agricultural (92) or logging (12) case.

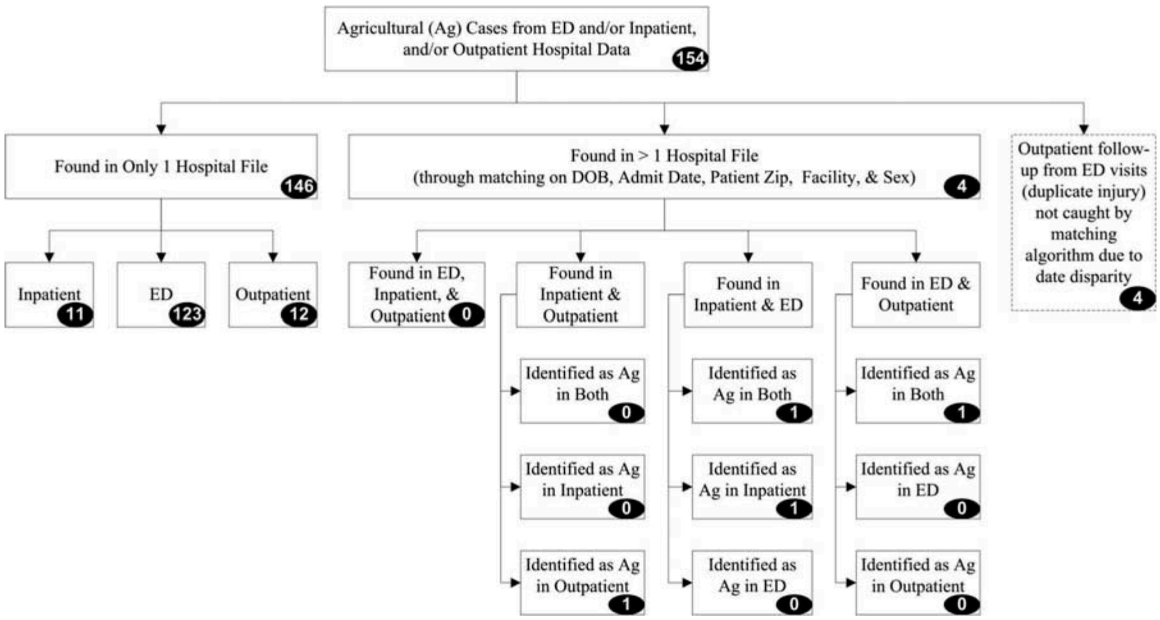
Of the 134 keywords searched in the PCR records, 45 were not found in any text strings. A total of 22,832 of the 115,001 PCR records contained at least 1 of the remaining 89 keywords. Eight of these keywords did not return a single true agricultural or logging case in either the Maine or New Hampshire PCR data. These eight were therefore considered to be of no use in keyword searches. Table 1 shows the results of the text string search and visual inspection for the remaining 81 that returned at least one true injury. If the 45 keywords that were not found in any text string and the other 8 keywords that produced no true cases were eliminated, the number of keywords would be reduced to 81 and the number of records would be reduces from 22,832 to 2,766. Visual review of these 2,766 records could be accomplished by an experienced technician within approximately 24 working hours.

Of the 104 true agricultural or logging cases, 66 (63%) were identified from the keyword/visual inspection process alone, 25 (24%) were identified by the farm checkbox only, and the remaining 13 (13%) by both methods. Therefore, the keyword/visual inspection process identified a total of 79 true cases (a true case may have contained more than one keyword). A system that employs these 81 keywords without visual inspection would therefore be expected to have a false-positive rate of (2,766 − 79)/2,766 = 97.14%. In addition to the 38 true agricultural cases for which the farm box was checked, 33 additional records with the farm box checked were found not to be true agricultural cases. Thus, the false-positive rate for the farm check box was 33/71 = 46.5%.

Hospital Data

A total of 4,859,235 hospital records were received for 2008. Of these, 652,704 came from the emergency department, 4,055,853 were outpatient records, and 150,678 were inpatient records. Through the process described below, a total of 150 (150/4,859,235 = 0.003%) true agricultural cases were identified from these

FIGURE 1. Distribution of agricultural cases found in hospital data.



records. Given the current ICD-9 system, which does not have E-codes for logging, it was not possible to identify logging cases from hospital data.

The process of eliminating duplicates and match-merging the ED, inpatient, and outpatient files produced a preliminary total of 154 unique agricultural cases (Figure 1). Upon further inspection, it was determined that 4 of these 154 did not represent a unique injury, but were follow-up visits for treatment of the same injury (because these follow-up visits did not occur on the same day as the initial visit, they were not immediately recognized as duplicates during the match-merge). As these 4 are not considered further, the results discussed below relate only to 150 initial episodes of care.

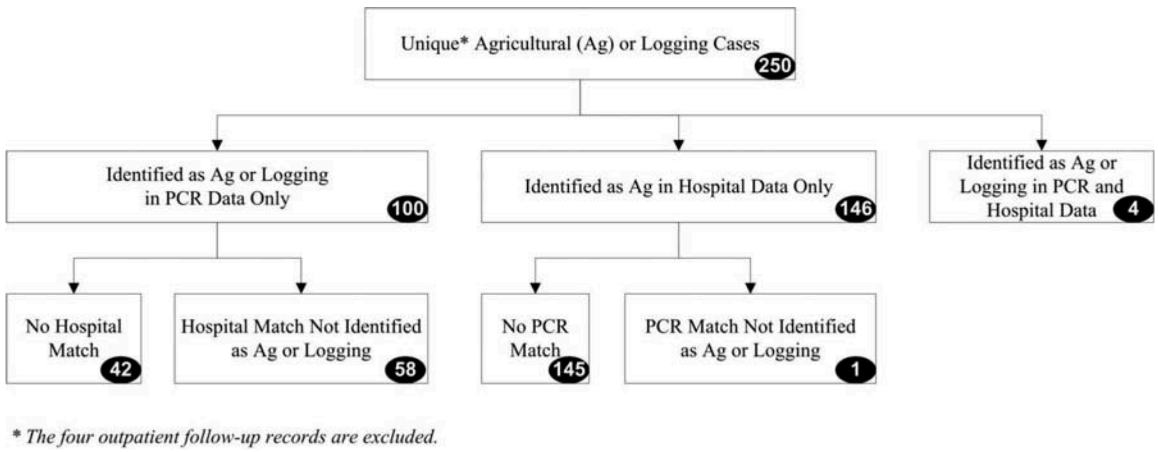
For the remaining 150 unique injury events, 146 (97.3%) had the initial episode of care documented in only one of the three hospital files. The ED file had the largest number of these (123/146 = 84.2%), followed by the outpatient file (12/146 = 8.2%) and the inpatient file (11/146 = 7.5%).

Four of the injury events had the initial episode of care documented in two of the three hospital files. In two of these four cases, both files noted agricultural relatedness and would therefore have inflated the total injury count by 2. In the other two cases, only one of the two files noted agricultural relatedness. No event was documented in all three files.

Overlap Between PCR and Hospital Data

Four of the five matching variables had some missing values in the PCR data set, with 34% of the records missing at least one of the five. As the data were downloaded based on the date of the incident, this identifier was present for all the records. Of the remaining four, missing values were seen for gender (6.0%), date of birth (DOB; 6.2%), ZIP code (6.5%), and permanent facility identifier (31.5%). In the hospital data, less than 1/100 of a percent are missing, as these data go through extensive editing prior to release. Of the 150 hospital records indicating an agricultural injury, 5 (5/150 = 3.3%) could be linked to a PCR record, and 3 of these 5 PCRs also indicated agricultural origin. In 1 of the other 2 cases, the PCR file indicated logging origin. Conversely, of the 104 true agricultural or logging cases found in the PCR, 62 (62/104 = 59.6%) could be linked to a hospital record. Among those 62 hospital

FIGURE 2. Distribution of unique agricultural or logging cases.



records, only 4 (4/62 = 6.4%) indicated agricultural origin. Figure 2 shows the distribution of unique agricultural and logging cases.

Of the 250 unique agricultural (238) and logging (12) injuries identified (100 PCR only + 146 hospital only + 4 from both), 66 (26%) would not have been identified without free text review of the PCR narrative. Of the 250 cases, 208 (83%) were treated at a hospital facility. Of these 208, 63 (31%) had an associated PCR, which implies that the patient was transported via ambulance.

Estimates of Injury Rates

Because there are currently no ICD-9 E-codes for logging injuries, it is not possible to identify them in the hospital data. Therefore, the 12 logging cases identified by case review of PCR records must be considered only a partial count.

The 2007 Census of Agriculture[20] estimates the number of workers in agriculture in Maine (operators + hired workers) to be 28,697. As the census is conducted only every 5 years, this estimate was considered to be the most relevant for 2008. From this, the cumulative incidence of agricultural injury requiring medical attention in Maine for 2008 is estimated to be 238/28,697 = .0083 or 8.3/1000 (95% confidence interval [CI] = 0.0073–0.0093) workers.

Estimation of the cumulative incidence in the logging industry requires an assumption. Specifically, it is assumed that if logging E-codes were available for hospital data, the ratio of cases identified between the PCR versus hospital records would be the same as that observed for agricultural cases. Although the validity of this assumption cannot yet be proven, it is believed to be logical for two reasons. First, the location of the work is rural in both cases. Thus, the distance from health care facilities is in all likelihood similar. Second, the types of injuries are probably similar in that they result from vigorous outdoor activities. For the 238 agricultural cases, 89 were identified from the PCR alone, 146 from the hospital data alone, and 3 from both sources. Therefore, using only the PCR data, the study would have identified 92 (89 + 3) agricultural cases.

From this, it is estimated that combining these sources produces a count for agricultural cases that is 2.59 (238/92) times higher than the count obtained from the PCR data alone. Applying this same multiplier to the logging count from the PCR would yield an estimate of 31.1 logging injuries (12 × 2.59). Using data from the American Community Survey,[21] it is estimated that there are 2,380 logging workers in the state of Maine (averaged over 2006–2010). Therefore, the cumulative incidence of injuries requiring medical attention in the Maine logging industry is 31.1/2,380 = .0131 = 13.1/1,000 (95% CI = 0.0085–0.0177) workers.

CONCLUSIONS

These data indicate that the cumulative incidence of injuries requiring medical attention is 12.9/8.2 = 1.57 times higher in the logging versus agricultural industry in Maine. This number must be considered tentative due to the assumption that was required for the logging incidence estimate. However, rates of fatal injury in logging consistently surpass those of agriculture1; therefore, this is not out of the realm of possibility. In fact, Fosbroke and colleagues[22] estimated that over a lifetime of working in logging, 62.7/1000 workers were estimated to die from an occupational injury.

The large percentage of agricultural cases in the PCR that could only be identified by visually searching text strings (63%) is important from several perspectives. The estimated 24 manhours required to perform these searches annually would not be prohibitive for a cost-effective surveillance system. However, it is certainly not as fast as electronic detection of a farm checkbox. A far larger problem exists in that many states do not have this free text in their electronic PCR files. In these cases, it will be necessary to apply a large correction factor to accurately estimate the total number of cases.

The extremely high false-positive rate ([2,766 − 79]/2,766 = 97.14) for our electronic searches of text strings underscores that this method, without subsequent visual inspection, is not an effective search algorithm. Thus, our surveillance system currently requires this visual inspection. The possibility exists that more sophisticated computer search algorithm for text strings might remove or reduce this need. It can readily be seen that a surveillance system that utilized only PCR records would have only captured 104/248 = 41.9% of the agricultural and logging morbidity. Use of only hospital records would have not fared much better, capturing only 148/248 = 59.7%. As explained previously, 4 of the 248 episodes of care were captured by both files, which is why these numbers sum to slightly greater than 100%. These results underscore the necessity of including additional data sources to better capture all injury data. Further, the lack of detail in the hospital E-codes reduces the ability of the system to identify likely causes of many of the injuries and thus limits its current usefulness. It is hoped that future iterations of E-codes, including those contained in ICD-10, will contain more specific information.

The proportion of PCR agricultural cases that could be matched to a hospital record (62/104 = 59.6%) indicates that the majority of ambulance runs for agricultural cases result in treatment at a hospital facility, although the hospital rarely identifies it as an agricultural case. This proportion is in all likelihood higher if one considers that some of these cases were not linked due to electronic merging problems, missing data, etc. As shown in the results, 48% of the PCRs were missing at least one of the five matching variables. Unfortunately, this is the only effect that is possible to quantify. Only 5 of the 148 (3.4%) agricultural cases identified in the hospital data were found to have an ambulance run (PCR). This is an indication that a large proportion of agricultural injuries arrive at the hospital by means other than an ambulance.

Although emergency department data produced the greatest number of true cases (119) in the hospital data, inpatient (10) and outpatient (15) data contributed additional unique cases not captured using other methods. If available, inpatient and outpatient data should be included in such a surveillance system.

Limitations

Several sources of bias are known to influence the data and its analysis. First, previous research[23] showed that active surveillance can capture injury events that are not identified in PCR or hospital records therefore would be missed by a fully electronic system. However, active surveillance is a costly and time-consuming endeavor. Secondly, visitors to the farm or woodlot that become injured would not be included in the denominator; however, their cases could be included in the numerator. Currently this is unavoidable, as variables for industry or occupation do not exist in the electronic medical record, nor are found in the PCR. It is believed that current estimates of farm and logging injury still represent an undercount.

Using administrative databases for purposes other than originally intended is not without issue. It is known that significant errors can be made in assigning E-codes in hospital data.[24] The accuracy of ICD-9 coding has been questioned[25]; however, for hospital records, it remains the best method of identifying agriculturally related events. Other research has echoed the need for better use of E-codes and for expanded variables related to industry and occupation.[26,27] We further validated injury cases where on the farm location keyword was present, to ensure other diagnostic codes related to an injury event, and not to a chronic or acute illness. Overall injury patterns, measured using the same methodology year to year, can still be used to inform public health researchers. Considering the quality of information in light of the cost and effort to acquire such data, administrative databases remain a viable option for ongoing injury surveillance.

The transition to electronic reporting of administrative data has had several effects. Positively, data are easier to manage and transmit to data recipients, there is a higher level of consistency in reporting (e.g., drop down menus, automated checkboxes, and once methodology is firmly in place, the time lag for reporting is dramatically decreased. On the other hand, the transition process creates many problems in and of itself. New data management systems need to be established and tested, and users must be trained on new techniques of data entry. Serious delays may arise from this process and data sets can remain incomplete as this transition occurs. However, health data is continually generated as these transition processes take place.

In conclusion, multiple data sources capture a better picture of injury than using a single source.

ACKNOWLEDGMENTS

We thank the Maine Emergency Medical Services Bureau and the Maine Health Data Organization for the use of their data. We are grateful to Kaye Freling and Liane Hirabayashi for their help with record identification and manuscript review, respectively.

FUNDING

Funding for this study was provided by the National Institute for Occupational Safety and Health (2U54OH007542-11).

REFERENCES

1. Bureau of Labor Statistics. National Census of Fatal Occupational Injuries. Available at: http://www.bls.gov/iif/oshwc/cfoi/cfoi_revised12.pdf. Published 2012. Updated April 2014. Accessed May 1, 2014.

2. Occupational Health and Safety Administration. Agricultural Exemption, Enforcement Exemptions and Limitations under the Appropriations Act CPL 02-00-051. Available at: https://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=DIRECTIVES&p_id=1519. Published May 28, 1998. Accessed March 9, 2015.

3. Olorunnishola OA, Kidd-Taylor A, Byrd L. Occupational injuries and illnesses in the solid waste industry: a call for action. *New Solut*. 2010;20:211–223.

4. Occupational Safety and Health Administration. Injury and Illness Recordkeeping Forms. Available at: https://www.osha.gov/recordkeeping/RKforms.html. Updated 2013. Accessed November 20, 2013.

5. Smidt M. A description of forest industries and occupations with focus on forestry workers' jobs and injury and illness surveillance. Presented at the 34th Council on Forest Engineering; June 2011; Quebec, Canada. Available at: https://www.cirrelt.ca/COFE2011/proceedings/cofe2011proceedings.pdf. Accessed October 8, 2014.

6. Jenkins EL, Hard DL. Implications for the use of E codes of International Classification of Diseases and narrative data in identifying tractor-related deaths in agriculture, United States, 1980–1986. *Scand J Work Environ Health*. 1992;18:49–50.

7. National Institute for Occupational Safety and Health. National Occupational Research Agenda Agriculture, Fishing and Forestry. Available at: http://www.cdc.gov/niosh/nora/comment/agendas/AgForFish/pdfs/AgForFishDec2008.pdf. Published 2008. Accessed October 8, 2014.

8. National Emergency Medical Services Information System. Needs and Benefit Available at: http://www.nemsis.org/theProject/whatIsNEMSIS/needsAndBenefits.html. Published 2013. Accessed October 8, 2014.

9. Stout N, Jenkins EL. Use of narrative text fields in occupational injury data. In: Proceedings of the International Collaborative Efforts on Injury Statistics; March 1995; Maryland. Available at: http://198.246.124.22/nchs/data/ice/ice95v1/c24.pdf. Updated 1995. Accessed April 1, 2014.

10. Lawrence BA, Miller TR, Weiss HB, Spicer RS. Issues in using state hospital discharge data in injury

control research and surveillance. *Accid Anal Prev*. 2007;39:319–325.

11. Bunn TL, Slavova S, Bathke A. Data linkage of inpatient hospitalization and workers' claims data sets to characterize occupational falls. *J Ky Med Assoc*. 2007;105:313–320.

12. Freidman L, Krupczak C, Brandt-Rauf S, Forst L. Occupational amputations in Illinois 200–2007: BLS vs. data linkage of trauma registry, hospital discharge, workers compensation databases and OSHA citations. *Injury*. 2013;44:667–673.

13. Schappert SM, Rechtsteiner EA. Ambulatory Medical Care Utilization Estimates for 2006. National Health Statistics Reports. Available at: http://www.cdc.gov/nchs/data/nhsr/nhsr008.pdf. Updated August 2008. Accessed April 1, 2014.

14. Maine Emergency Medical Services Data Dictionary. Available at: http://www.nemsis.org/support/stateProgressReports/documents/ME-121205DataDictionaryE1.0.pdf. Published 2006. Accessed October 8, 2014.

15. Day L, Voaklander D, Sim M, et al. Risk factors for work related injury among male farmers. *Occup Environ Med*. 2009;66:312–318.

16. Pickett W, Hartling L, Dimich-Ward H, et al. Surveillance of hospitalized farm injuries in Canada. *Inj Prev*. 2001;7:123–128.

17. Nogalski A, Lubek T, Sompor J, Karski J. Agricultural and forestry work-related injuries among farmers admitted to an emergency department. *Ann Agric Environ Med*. 2007;14:253–258.

18. Mehler LN, Schenker MB, Romano PS, Samuels SJ. California surveillance for pesticide-related illness and injury: coverage, bias, and limitations. *J Agromedicine*. 2006;11:67–79.

19. American Society of Agricultural and Biological Engineers. Farm and Agricultural Classification Code ASAE S575.1. Available at: http://elibrary.asabe.org/abstract.asp?aid=41297&redir=[confid=s2000]&redirType=standards.asp&dabs=Y. Updated 2007. Accessed April 1, 2014.

20. United States Department of Agriculture. Census of Agriculture. Table 7 Hired Farm Labor—Workers and Payroll: 2007, 46 Selected Operation and Operator Characteristics. Available at: http://www.agcensus.usda.gov/Publications/2007/Full_Report/Volume_1,_Chapter_1_State_Level/Maine/. Updated March 29, 2012. Accessed March 15, 2014.

21. United States Census Bureau. American Fact Finder EEO 1w. Detailed Census Occupation by Sex and Race/Ethnicity. Available at: http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=EEO_10_5YR_EEOALL1W&prodType=table. Updated 2006–2010 (5-year ACS data). Accessed March 15, 2014.

22. Fosbroke DE, Kisner SM, Myers JR. Working lifetime risk occupational fatal injury. *Am J Ind Med*. 1997;31:459–467.

23. Earle-Richardson G, Jenkins P, Scott E, May J. Improving agricultural injury surveillance: a comparison of injury incidence rates among three data sources. *Am J Ind Med*. 2011;54:586–596.

24. Schwartz RJ, Nightingale BS, Boisoneau D, Jacobs LM. Accuracy of E-codes assigned to emergency department records. *Acad Emerg Med*. 1995;2:615–620.

25. McKenzie K, Enraght-Moony EL, Walker SM, McClure RJ, Harrison JE. Accuracy of external cause-of-injury coding in hospital data. *Inj Prev*. 2009;15:60–64.

26. Hunt PR, Hackman H, Berenholz G, McKeown L, Davis L, Ozonoff V. Completeness and accuracy of International Classification of Disease (ICD) external cause of injury codes in emergency department electronic data. *Inj Prev*. 2007;13:422–425.

27. Davis LK, Hunt PR, Hackman HH, McKeown LN, Ozonoff VV. Use of statewide electronic emergency department data for occupational injury surveillance: a feasibility study in Massachusetts. *Am J Ind Med*. 2012;55:344–352.