Use of Narrative Text Fields in Occupational Injury Data

by Nancy Stout, Ed.D. and E. Lynn Jenkins, M.A.

Surveillance data on occupational injuries and deaths are frequently analyzed in order to identify high–risk worker populations, characterize injury circumstances, and determine potential risk factors. Results drive injury prevention efforts at the national, state and local levels, and impact legislation and regulatory policy. The inclusion of narrative fields in injury surveillance data allow for identification of specific hazards and injury incidents. NIOSH maintains several surveillance systems that include narrative fields.

Analyses of these narrative entries, through computerized key–word searches and manual review, has allowed us to go beyond the limits of coded data to better understand specific circumstances and risks.
NIOSH's primary system for surveillance of fatal occupational injuries is the National Traumatic Occupational Fatalities System, or NTOF. NTOF is comprised of information from death certificates for people who die from injuries at work. Death certificates are provided by all 50 states for cases that meet these criteria: age 16 years or older, external cause of death (E800–E999), and a positive response to the "injury at work?" item. The NTOF database contains data on fatal occupational injuries since 1980.

We recently released a publication: "Fatal Injuries to Workers in the United States, 1980–1989: A Decade of Surveillance", that Dr. Satcher described, which provides an overview of the NTOF surveillance system and contains both national and state–specific analyses of worker deaths for the decade of the 1980s.

We frequently use NTOF data to respond to requests for information, from within CDC as well as from other federal organizations, states, and the public. In the U.S. there are two agencies that share responsibility for occupational safety and health: NIOSH is a research organization that conducts research, assesses risk, and develops prevention efforts, while the Occupational Safety and Health Administration, or OSHA, is a regulatory agency that develops and enforces occupational safety and health regulations. NIOSH frequently provides data and testimony to OSHA when they are proposing new rulemaking for occupational safety. This is important to us because this is one of the most effective ways we can implement our findings to shape national policy to prevent occupational injuries.

One reason NTOF is so useful is that it contains narrative data that allow us to examine detail that is not typically available, because most databases consist solely of coded data. NTOF contains narrative entries for industry, occupation, an injury description, and immediate, underlying and contributory causes of death. To demonstrate the value of these narrative data, I would like to describe some examples of analyses of narrative injury data that have impacted national programs or policies.

First, OSHA requested data pertaining to their proposed rule for confined space entry. Deaths in confined spaces are not limited to an industry or occupation group, nor are they specific to one cause of death, so they are difficult to identify in any surveillance data, and impossible if the data are all coded.

With NTOF, we were able to first select E–Codes that could potentially be confined space–related deaths (e.g., suffocation, poisoning). This narrowed the data to about 2000 cases for a 6–year period. Then we reviewed the injury description narratives to confirm deaths in confined spaces (758 cases for 1980–1985). The narrative information is not always detailed or specific enough to identify all such deaths, but we were able to determine that at least 126 deaths a year occurred in confined spaces, and were able to further analyze this subset to identify characteristics of these victims. For example, they were most frequently employed in construction or mining, and the leading causes of death were mechanical suffocation, poisoning by gas, explosion, and drowning. OSHA used these data to help justify the need for and determine the specifications for the recently enacted OSHA confined space safe–entry procedures.

In developing the Agricultural Initiative in 1989, the U.S. Congress requested data illustrating the major occupational safety hazards related to agricultural machinery. We knew from previous narrative data analysis that tractors were the leading cause of death, and we knew from literature and from working with the data that tractor roll–overs were

a major problem, but that had not been documented or quantified at the national level. Again we turned to the narrative data.

We first selected cases with E–codes for agricultural machines, then we did a keyword search for the word "tractor" (which is not a unique E–code). Finally, we manually reviewed the injury descriptions. We were able to document that tractors accounted for 69 percent of all agricultural machinery–related deaths (1523 cases), and that 52 percent of these resulted from roll–overs and 16 percent were run–overs. Funding was allocated for several injury prevention programs as a result of these findings.

Incidentally, in another analysis of narrative data we discovered that many tractor incidents are missed by limiting analysis to the E–code for agricultural machines, because according to ICD–9 rules, tractors are correctly coded as motor vehicles if they are on a public roadway under their own power.

Another example is an OSHA request for data on fatalities to line clearance tree trimmers for testimony pertaining to standards for protective equipment for the electrical power industry. Tree trimmers can be coded into a number of different occupation categories (linemen, gardeners, laborer, etc.) so they cannot be identified by occupation codes. To identify tree trimmers in NTOF, we did a keyword search of the occupation, industry, and injury description fields for terms such as tree trimmer, arborist, and tree surgeon. We found 127 cases for 1980–1985. We were able to document that at least 21 tree trimmers were killed on the job each year, and that 41 percent were electrocutions, and 33 percent were fatal falls. Also, all but one were males and two–thirds were less than 35 years old. Again, these findings helped shape national occupational safety regulatory policy that recently went into effect. NIOSH also published an Alert on this topic to warn workers of the magnitude of this problem and the hazards they face from electrocution as well as from falls.

A few years ago, several deaths were brought to our attention of farmers who had entered their manure pits and been overcome and died from methane gas asphyxia. Tragically, other family members and co–workers also died in rescue attempts. We realized that we needed to alert the public of deaths due to methane asphyxia in manure pits on farms. Although we had investigated a few cases, we wanted to try to determine the national magnitude of this problem. So we subset potential cases from the NTOF data by appropriate E–codes, such as suffocation and poisoning, which narrowed us to 2000 cases, then we manually reviewed the injury descriptions to confirm cases that occurred in manure pits.

Although death certificates do not always contain enough detail to identify all cases, we did find that at least 16 people died from asphyxia in manure pits, and that 5 episodes resulted in multiple deaths from rescue attempts. We published an MMWR article on this and issued an Alert. Incidentally, we recently got a phone call from a farmer who was getting ready to go into his manure pit to fix something and he remembered seeing our publication. He called us to ask what he should do to stay alive when he went into his pit. We were able to appropriately advise him, and were pleased that our information is reaching those at risk.

Another OSHA request for information on hazards in the logging industry resulted in this analysis of narrative data. We selected cases coded as logging industry and examined the distribution of cause of death. After finding that a large proportion of cases were due to machines, we selected cases coded to the machinery E–code and reviewed the injury descriptions to better understand the circumstances of these deaths. We found that almost half of the machinery–related deaths in logging were the result of roll–overs. We had previously determined that the leading cause of death in this industry—struck by falling objects—was largely due to falling trees and logs. While that is not so surprising, we also found that of ALL worker deaths E–coded as being struck by falling objects, in all industries, 30 percent are from trees alone. Again, we could not have learned any of this from coded data. Without the narrative data, we would not have a clue as to what falling objects were killing workers or how we could attempt to prevent these deaths.

We also used narrative data to identify and understand worker deaths caused by falls from suspension scaffolds, to examine deaths from trench cave–ins, from falls through skylights and roof openings, deaths from forklift trucks, from electrocutions during work with scaffolds near overhead power lines, from entanglements with hay bailers that resulted in scalpings, and from homicides of convenience store workers. We published Alerts on many of these

topics, using these NTOF analyses in support of other data and information, to request assistance in preventing these deaths.

These are just a few examples of the uses and value of narrative data. None of these analyses would have been possible with coded data, as none of these cases are identifiable through E-codes or other coded data alone. The narrative information, particularly injury description, allow us to go beyond the limits of coded data to drive NIOSH research and initiatives as well as national regulatory policy, to prevent worker deaths.

We also want to emphasize the value of E-codes in addition to narrative data. In many of these examples the E-codes allowed us to subset a large database to a manageable number of cases for manual review or to determine which words to use in keyword searches. Without this ability, many cases may be missed. Also, to determine overall distributions of cases, such as identifying the leading causes of death or conducting international comparisons, coded data are essential. Unfortunately, this may be why databases that lack E-codes are under-utilized. Both coded and narrative data are valuable in injury surveillance efforts.

Another value of narrative data is the ability to code or recode variables to alternative coding schemes. For example, there are three US standard coding schemes for occupation, two for industry, and three for injury circumstances, as well as numerous schemes unique to agencies or other nations. Most schemes are not comparable, and data coded to one cannot be directly converted to another. This prohibits comparisons between databases, particularly international comparisons. It also often prevents the computation of rates, if denominator and numerator databases are coded to different schemes (which is frequently the case in the U.S.).

Coding schemes also change periodically. U.S. employment codes generally change every decade, and even internationally standardized codes such as the ICD are periodically modified. When surveillance systems convert to a new code structure, the ability to monitor trends over time is often lost. Narrative data can be recoded to provide comparability of data between systems, years, and countries.

There have been a number of efforts in recent years to develop software that automatically codes narrative data into numeric codes. Dr. Rosenberg described NCHS efforts in developing and improving software that codes cause of death. There is currently a collaborative effort underway in the U.S., between NIOSH, NCHS, and several other organizations, to develop intelligent software that will automatically code occupation and industry narratives into standard numeric codes. Similar efforts have also taken place in other countries and have been applied to other variables. Information about various coding software has not been assimilated, to my knowledge, and this may be another area for this International Collaborative Effort to address. It would certainly be valuable to coordinate and integrate these coding programs. Automated coding software makes it easier, less expensive, and more reliable to automate narrative data than to code variables prior to automation. Narrative data not only provide valuable detail, but also provide the flexibility to adapt existing data to changes and future needs.

References

1. Castillo, DN, Landen, DD, Layne, LA. 1994. Occupational Injury Deaths of 16- and 17-year-olds in the United States. American Journal of Public Health, 1994:84.

2. Etherton, JR, Myers, JR. 1990. The use of rollover protection on farm tractors in West Virginia. B.Das ed., Advances in Industrial Ergonomics and Safety II, Taylor and Francis, Philadelphia, PA. pp. 819-825.

3. Etherton, JR, Myers, JR, Jensen, RC, Russell. JC, Braddee, RW. 1991. Agricultural machine-related deaths. Amer. Jour. Pub. Hlth. 81(6):766-768.

4. Fosbroke, DE, Myers, JR. 1994. Logging fatalities in the United States by region, cause of death, and other factors—1980 through 1988. Journal of Safety Research, 25(2).

5. Jenkins, EL, Hard, DL. 1992. Implications for the use of E-codes of the International Classification of Diseases and narrative data in identifying tractor-related deaths in agriculture, United States, 1980—1986. Scand J Work Environ Health, 92(18) Suppl 2:49—50.

6. Myers, JR, Casini, VJ. 1989. Fatalities attributed to methane asphyxia in manure pits—Ohio,Michigan, 1989. Centers for Disease Control. Mortality and Morbidity Weekly, 38(33):583-586.

7. Myers, JR, Fosbroke, DE. Logging fatalities in the United States by region, cause of death, and other factors—1980 through 1988. Accepted for publication by the Journal of Safety Research, November 2, 1993.

8. Roerig, S, Melius, J, Casini, VJ, Myers, JR, Snyder, KA. 1992. Scalping incidents involving hale balers - New York. Centers for Disease Control. Mortality and Morbidity Weekly Report, 41:27:489-491.

9. Snyder, KA, Bobick, TG, Hanz, JL, and Myers, JR. 1992. Grain-handling fatalities in production agriculture, 1985-1989. In: American Society of Agricultural Engineers 1992 International Winter Meeting. December 15-18, 1992, Nashville, TN. Paper No. 92-5509.

10. Stout, N, Frommer, MS, Harrison, J. 1990. Comparison of work-related fatality surveillance in the U.S.A. and Australia. Journal of Occupational Accidents, 90(13):195-211.

11. Stout-Wiegand, N. 1987. Characteristics of Work-related Injuries Involving Forklift Trucks. Journal of Safety Research, 87(18):179-190.

12. U.S. Department of Health and Human Services. 1989. NIOSH Alert: Preventing Worker Deaths and Injuries from Falls through Skylights and Roof Openings. DHHS (NIOSH) Publication No. 90-100.

13. U.S. Department of Health and Human Services. 1990. NIOSH Alert: Request for Assistance in Preventing Deaths of Farm Workers in Manure Pits. DHHS (NIOSH) Publication No. 90-103.

14. U.S. Department of Health and Human Services. 1992. NIOSH Alert: Request for Assistance in Preventing Falls and Electrocutions During Tree Trimming. DHHS (NIOSH) Publication No. 92-106.

15. U.S. Department of Health and Human Services. 1992. NIOSH Alert: Request for Assistance in Preventing Worker Injuries and Deaths Caused by Falls From Suspension Scaffolds. DHHS (NIOSH) publication No. 92-108.

16. U.S. Department of Health and Human Services. 1994. Worker Deaths in Confined Spaces: A Summary of NIOSH Surveillance and Investigative Findings. DHHS (NIOSH) Publication No. 94-103.

17. U.S. Department of Health and Human Services. 1993. NIOSH Alert: Request for Assistance in Preventing Homicide in the Workplace. DHHS (NIOSH) Publication No. 93-109.