# Methods for Using Narrative Text from Injury Reports to Identify Factors Contributing to Construction Injury

**Jessica Bondy, MHA,[1*] Hester Lipscomb, PhD,[2]**
**Kenneth Guarini,[3] and Judith E. Glazner, MS[1]**

**Background** *Several methods exist for classifying injuries from written text, thereby identifying possible points of intervention. We describe an innovative method for such classification.*
**Methods** *Using Haddon's matrix as a framework, two independent reviewers coded text from over 4,000 injury reports into a qualitative software package to identify factors contributing to injuries sustained during construction of Denver International Airport (DIA). We compared our classification scheme with three others.*
**Results** *This process created a coded data set, an expanded version of Haddon's matrix adapted for construction injury, and coding rules for interpreting narrative text. Haddon's matrix provides a flexible theoretical framework for coding information about a spectrum of contributing factors.*
**Conclusions** *Narrative descriptions from injury reports can provide detail on circumstances surrounding injuries and identify factors contributing to injury. Forms guiding investigators to explicitly consider human, organizational, and environmental factors could foster more complete descriptions of factors contributing to construction injury.* Am. J. Ind. Med. 48:373–380, 2005. *© 2005 Wiley-Liss, Inc.*

***KEY WORDS: construction injury; injury etiology; text analysis; injury classification schemes***

## INTRODUCTION

Although work-related injury and fatality rates as well as Workers' Compensation payments indicate that construction is one of the most hazardous of all industries [Lipscomb et al., 1996; Glazner et al., 1998; BLS, 2003], epidemiological research has only recently been conducted to identify the etiology of non-fatal construction injuries. One reason is that measuring exposure to etiological factors is difficult and expensive on a construction site, in part because construction inherently involves a changing environment. Characteristics more easily measured—such as person-hours of work by trade, industry, or age—are not themselves exposures but rather mediators [Park, 2002]. Rates derived from such denominators provide information about high-risk groups but they give little information about how injuries might be prevented.

Conversely, exposures—such as working at height, poor housekeeping, or use of power tools—can theoretically be assessed with frequent on-site inspections by trained observers. But quantifying exposures remains an elusive goal. Sexias et al. [1998] proposed assessing hazards on a

[1]Department of Preventive Medicine and Biometrics, University of Colorado School of Medicine, Denver, Colorado
[2]Department of Community and Family Medicine, Duke University Medical Center, Durham, North Carolina
[3]OMI Safety Services, Inc., Denver, Colorado
Work was performed at the University of Colorado School of Medicine and Duke University Medical Center.
Contract grant sponsor: National Institute for Occupational Safety and Health; Contract grant number: 1R01 OH003613.
*Correspondence to: Jessica Bondy, Department of Preventive Medicine, Campus Box B119, University of Colorado Health Sciences Center, 4200 E. Ninth Avenue, Denver, CO 80262. E-mail: Jessica.Bondy@UCHSC.edu

Accepted 24 August 2005
DOI 10.1002/ajim.20228. Published online in Wiley InterScience (www.interscience.wiley.com)

construction site using an inspection checklist applied to a random sample of locations on the site. Three safety professionals assessed whether each hazard was present and how adequately it was safeguarded. The study found significant differences among the three inspectors with regard to the assessment of both the presence of hazards and the adequacy of protection, despite prior training and what Sexias termed ''a highly simplified observational task.''

An alternative approach to identifying factors contributing to injury is to analyze narratives provided in documents related to the injury. Because uninjured workers are not part of these analyses, this approach does not quantify risk. Nevertheless, it can provide important information about possible points of intervention. Researchers at NIOSH have successfully used narrative fields from both the National Traumatic Occupational Fatalities (NTOF) surveillance system and the National Electronic Injury Surveillance System (NEISS) to investigate the etiology of specific types of occupational injury and death, such as forklift rollovers and logging deaths [Stout and Jenkins, 1995; Stout, 1998]. In construction, this technique has been applied to falls from suspension scaffolds, trench cave-ins, falls through skylights and roof openings, roof bolting injuries, and ladder falls [Cohen and Lin, 1991; Althouse et al., 1997; Stout, 1998]. United States Army databases have been used to investigate factors contributing to military drownings [Bell et al., 2001] and injuries involving military transport operators [Lincoln et al., 2004]. Finally, text from insurance claims records has been used to study circumstances surrounding motor vehicle crashes in construction work zones [Sorock et al., 1996].

The sources of narratives used in these studies have included death certificates, emergency room reports, injury reports [Althouse et al., 1997; Wiehagen et al., 2001], and insurance claims [Sorock et al., 1996]. Even with short narratives researchers generally have found information deduced from narrative data to be more accurate and complete than data about the same events as recorded in administrative sources [Jenkins and Hard, 1992; Sorock et al., 1996; Bell et al., 2001]. Moreover, narratives often contain enough detail to identify information about etiological factors contributing to occupational injuries that is not otherwise available [Dement et al., 2003].

First Reports of injury (FRIs), required by law to file Workers' Compensation claims, contain narratives describing the circumstances of injury, but the information in these reports is rarely used in systematic analysis. On some projects, Accident Investigation Reports (AIRs) may be available for a subset of Workers' Compensation injuries. These reports may represent an important untapped resource for studying injury etiology.

Approaches to extracting information from narratives also differ across studies. Some researchers have attempted to autocode narratives, using keyword searches [Sorock et al., 1996; Williamson et al., 2001] and artificial intelligence [Stout, 1998]. More commonly, researchers use keyword searches to select narratives for review, but code the narrative manually, often employing a conceptual framework. One of the most common of these is Haddon's matrix [Runyan, 1998; Conroy and Fowler, 2000; Higgins et al., 2001], created to classify factors contributing to automobile crash injuries.

We used Haddon's matrix [Haddon, 1968, 1972] as the conceptual framework to code text from reports on injuries sustained in the construction of Denver International Airport (DIA) to elucidate factors contributing to injury. Our work departs from previous analyses of injury narratives in that we have included all injuries experienced on a specific construction site rather than focusing on fatalities or on a particular type of injury.

## MATERIALS AND METHODS

DIA was built between December 1990 and August 1994, employing over 32,000 individuals in all three major construction industry sectors (SIC codes 15–17) and virtually all construction trades. In total, over 31M person-hours were worked, generating 4,634 Workers' Compensation claims with payment for medical expenses, impairment, or lost wages. Data from these claims were accumulated in an administrative database developed for the project's Owner-Controlled Insurance Program (OCIP).

We had access to this administrative database as well as to injury reports—FRIs and AIRs—from DIA's risk management office. The injured worker or his supervisor filled out FRIs for most Workers' Compensation claims. AIRs were filled out by safety inspectors working for the OCIP or for the contractor and were available for many of the more serious injury events and some of the less serious ones. Overall, 4,143 FRIs and 1,808 AIRs corresponding to 4,483 events were available. Of these, 323 were for illnesses, leaving 4,160 injury events with one or more injury reports.

Although FRI forms changed over the course of the project, all versions collected substantially the same information. On these forms, short narratives were solicited to describe the injury itself, what happened to cause the injury and failure to use safety devices or to obey rules. There were at least 17 different AIR forms, used by different investigating entities. The AIR form used by OCIP safety personnel included narratives describing the accident; unsafe objects, acts, or conditions; causes; what personal protective equipment (PPE) was being used; and what could have been done to prevent the event. Where multiple sources of data were available, we combined all sources to interpret the injury event.

Our methods were designed around three goals. First, we wanted to use a structured review tool that provided consistency of review, while still allowing us to code as many contributing factors as were relevant to a particular injury incident. We needed a tool that was structured enough to elicit the full range of contributing factors yet not so structured that it would result in many empty cells, given the small amount of detail we anticipated finding in our short narratives. Second, we wanted to organize the review so that similar incidents were reviewed together, again to improve consistency of coding. Finally, to keep costs down we wanted people who were not trained safety personnel to perform the review.

## Structure for Data Abstraction

For our review tool, we employed an adaptation of Haddon's matrix [Haddon, 1968, 1972]. This framework originally involved classifying factors that contributed to motor vehicle injuries into three categories—human, vehicle/equipment, and environment—and three time periods: before, during, or after the crash. We adapted this matrix to construction by changing the categories to human, objects, environment, and organization. We found that establishing agreement on when an incident began, and thus what factors were in place *before* the incident began, was highly problematic, a difficulty that has been noted by other researchers [Runyan, 1998]. We consequently collapsed the pre-event and event categories. Post-event factors included items that contributed to severity of injury after the fact, such as delayed evacuation.

## Classification of Injuries

One innovation of this work was to group injuries by what we called mechanism of injury event (MOIE), as opposed to the American National Standards Institute [ANSI, 1963] variable ''mechanism of injury'' (MOI). The essence of the difference between these coding systems is that MOIE attempts to reflect the initial event happening to the victim in the chain of events that caused the injury. In contrast, MOI usually captures the final energy exchange in that chain of events. So, for example, under MOIEs, falls starting with a slip or trip are distinguished from falls starting when a surface collapses or moves; the MOI taxonomy does not distinguish these events.

Because our categories concern the initial energy exchange *involving the injured person*, they do not always encompass the first energy exchange in the chain of events that led to the injury. For example, ''foreign body eye'' does not embody the energy exchange that brought the foreign body to the eye, distinguishing, for instance, debris blown by wind from debris falling from a face mask. Instead, this information was coded in the contributing factors. In describing the Merseyside Accident Information Model (MAIM), Davies et al. [1998] illustrated the distinction between the initial event in the ''accident'' and the initial event involving the injured person with the example of a workman who drops a hammer while on a scaffold. The hammer strikes a pedestrian walking under the scaffold. In this case, the ''accident'' sequence starts with the dropping of the hammer, while the first event involving the injured person is being struck by the hammer.

We ended up with 13 major categories of MOIE accounting for 3,784 injuries (90.1% of all injuries with reports): burn; caught; cumulative trauma; electrical shock; fall/no slip; foreign body in eye; jumped; motor vehicle; slips/trips; straining/overexertion; struck against; struck by; twisted by/pulled by.

## Methods of Abstraction and Quality Assurance

To find similar kinds of injuries, a single reviewer examined all safety reports to code basic information about the injury: body part, nature of injury, and ''mechanism of injury event'' (MOIE). Each injury incident was assigned to a single MOIE, and incidents were then grouped together for a second review, focused on contributing factors.

Because it was deemed more complex and open to interpretation, coding contributing factors into Haddon's matrix was performed by *two* research assistants, working independently. These assistants were not safety experts but applied coding rules developed by the research team. Reviewers were instructed to code all contributing factors into the revised Haddon matrix, without regard to whether the factor was clearly causal. An item that clearly was not a contributor was not coded, but when there was some question about its contribution to injury, we erred on the side of inclusion as illustrated in Table I. Reviewers were not asked to assign weights to the factors—to identify which might be the primary or secondary factor—nor were they to code information about the sequence of events except to distinguish post-event factors from others.

After a set of injuries had been coded by each reviewer independently, the reviewers compared their results and tried to resolve any discrepancies. Independent coding by the two reviewers was in complete agreement—that is, all codes matched—for approximately two thirds of the injury events. After discussion among themselves, the reviewers were able to reach total agreement on approximately 83% of the total number of injury events. When the reviewers could not resolve a discrepancy, the report was referred to investigators (JG, JB) for interpretation. Discrepancies that still could not be resolved (9% of all events) were referred to a member of the research team who had been Manager of Safety on the DIA construction project (KG) for questions of fact, terminology, or, occasionally, interpretation. This adjudication process sometimes resulted in new coding rules and subcategories of Haddon's matrix, which were then applied by the reviewers in their subsequent reviews. Occasionally, a new rule would require the reviewers to revisit injury reports

**TABLE I.** Examples of Coding, Illustrating When Items Were and Were Not Coded as Contributing Factors

| | | |
|---|---|---|
| Narrative text | Working on runway—carrying bucket of concrete. Company paving truck drove by and accidentally sprayed tac coat on three employees. Employee was the only one injured. | Employee was pushing a tub of grout up a ramp, tub stopped and employee twisted right elbow trying to force the tub forward. |
| Haddon codes | Co-worker safety infraction<br>Motor vehicle<br>Material (tac) | Victim inappropriate action<br>Victim motion/pushing<br>Object size-weight (grout)<br>Material (grout)<br>Structure (ramp) |
| Comment | Concrete was not coded in this example because it clearly did not contribute to the event. | The grout was coded in this example (material, object-size weight) because the fact that the object being moved was heavy may have contributed to the injury. The ramp (structure) was included for the same reason. |

they had already coded. As the coding proceeded through the various MOIEs, we noted that each new MOIE had its own characteristic new ambiguities, so we coded only 30 injury incidents, gathered the resulting questions, and adjudicated them before proceeding to the remaining injury reports.

Narrative text on injury reports was very short, often one sentence or less, and sometimes included jargon. Nevertheless, most of the narratives contained some information about the factors contributing to the injury. Of the 4,146 injury events reviewed, 3,912 had sufficient information to be coded for MOIE. Of these, 3,670 had sufficient information to be coded for contributing factors. The average number of factors coded per injury incident was 2.7 with a range of 1–12.

We used a qualitative software package [QSR N5™, 2000] to store the narrative documents and the coding assigned to text without altering the original text in any manner. This software allowed us to revise our coding scheme as the review evolved and provided a way to view the part of the original report text that had caused us to apply a particular code. Linkage allowed us to review text for each report or group of reports. For instance, we could examine text from all reports for injuries involving tools or co-worker action. After reports were coded, coded data were integrated with data from the administrative database by transferring N5 coding into a SAS dataset and vice versa.

## RESULTS

The review process created two significant by-products, in addition to a coded data set. First, it created an expanded version of Haddon's matrix adapted for construction injury. As we reviewed the injury reports, we created a taxonomy of contributing factors, adding subcategories within each top-level category when we encountered common contributing factors. For example, the ''environment'' category ultimately included the sub-categories shown in Table II. Injury events were coded at the most specific level possible, but sub-categories were not developed where there were few injury events. The final version of Haddon's matrix (available upon request from the authors) had a total of 110 categories. The qualitative software described above proved indispensable in managing the taxonomy, since it was designed to permit adding codes as needed and viewing associated text, should decisions need to be revisited.

The second by-product of our process was a set of coding rules to help our reviewers interpret narrative text. Some were simple, for example, ''Vehicles that usually have license plates are coded as 'motor vehicle,' not 'heavy equipment'.'' Others depended on knowing the policies of this construction site, such as ''If working over 5 feet, personal protective equipment (PPE) is required, so a safety infraction should be coded if absence of PPE is documented.'' Still others were quite complex: ''When carrying something heavy, generally the object contributing to the weight of the load is coded. For instance, if a worker is carrying wet concrete in a bucket, the wet concrete is coded and not the bucket. The exception is if

**TABLE II.** Taxonomy of Environmental Contributing Factors in the Final Version of Haddon's Matrix Developed by Coding Text from Injury Reports

| Contributing factor |
|---|
| Wind |
| Weather conditions, excluding wind |
| Steep terrain. "Terrain" refers to outdoor, natural surfaces |
| Rough terrain (e.g., ruts, bumps, rocks) |
| Slippery terrain |
| Other terrain conditions |
| Slippery non-terrain walking surface (e.g., floor, steps, truck bed) |
| Other non-terrain walking surface conditions |
| Victim's stepping up or down (change in elevation) |
| Working overhead (e.g., drilling in ceiling) |
| Working at height (on scaffold, on stilts, etc.) |
| Working in a tight spot (e.g., confined area, awkward position) |
| Poor indoor lighting |
| Darkness (i.e., no lighting) indoors or outdoors |
| Other environmental conditions |

there is reasonable evidence that the object helping convey the load was a contributing factor, such as in many injuries with wheelbarrows.''

## Comparison of Existing Injury Coding Schemes

A variety of classification schemes has been used in occupational injury databases. The BLS system for classifying occupational injury is widely used but was not created to identify avenues for injury prevention. In contrast, other models have focused on injury prevention; Haddon's matrix [Conroy and Fowler, 2000; Higgins et al., 2001], Davies and Manning's MAIM system [Davies et al., 1998], and the Lincoln et al. [2004] reconstruction template are three such models. Each has strengths and weaknesses.

The BLS system requires that three items about the process of injury be recorded: the event or exposure, the immediate source of the injury, and the secondary source, that is, the object, substance, or person that generated the immediate source. By focusing the event codes on the manner in which the injury was produced by the immediate source, the BLS system may fail to identify the initial energy exchange. Since at most two items may be coded as sources of injury, some contributing factors may not be identified. In addition, BLS's taxonomy of events is very detailed, comprising 20 pages of codes and instructions, such that misclassification may be a problem. For example, unless a coder was very familiar with BLS event codes, he might not properly distinguish ''rubbed or abraded by friction or pressure'' from ''rubbed, abraded, or jarred by vibration.''

Lincoln et al. [2004] simplified an earlier injury model, the MAIM developed by Manning and Davies, which attempts to reconstruct the series of events leading to injury. MAIM represents an ''accident'' as a series of events, starting with the first unforeseen or unintended event and proceeding through an indeterminate set of events to the outcome. Each event is described in terms of a subject, a verb, and optionally a direct object. This method depends on extensive interviews with accident victims, which provide very complete information about details of the injury event. Lincoln restructured this model into a template which attempts to reconstruct the series of events leading to injury using nine key elements: general activity, specific task, contributing factor, precipitating mechanism, primary source, secondary source, injury event/exposure, and outcome. This system's structure elicits more complete information than the does the BLS system. In particular, the precipitating mechanism category identifies the initial energy exchange, and activity and task are included; these are potentially important items that can direct prevention and intervention activities. This template gives different weight to different factors: source, secondary source, contributing factor. On the other hand, it requires sophisticated coders and may not be appropriate for short injury reports such as FRIs, where complete information is not typically available. Also it shares the limitation of the BLS system that contributing factors in complex injury events may be missed since only one may be coded.

In the system we used, the categories are less detailed and more flexible than those of the BLS system. There are two major categories: MOIE and factors contributing to injury. The MOIE identifies the original energy exchange leading to injury, where that information was available. The contributing factors are designed to capture all human, object, environmental, and organizational factors involved in the injury. This system allows for the inclusion of multiple factors that could have contributed to injury, but assigns no relative weights to those factors. Also, it does not describe the complete sequence of events leading to injury.

To understand how these three systems would differ in the information they capture, consider the text from an example DIA injury report: ''I was back filling in a very [tight] area, upon coming out I was pushing up the wacker,[1] as Earl and Pat [were] pulling up the wacker with a rope, then they [slipped] down and I was trying to hold the wacker from falling on me, then the pressure of the wacker bent my back.'' The codes resulting from the three systems are shown in Table III.

In this example, the BLS method collects relatively little information. The source of the injury was the wacker, but no appropriate category for this class of equipment could be found in BLS source codes. BLS codes were very specific with respect to hand tools but included relatively few categories for construction equipment. The Lincoln reconstruction captures the sequence of events, but misses that co-workers dropped an object, that the size or weight of the object contributed to the injury, and that the victim was working overhead and was lifting improperly. The MOIE/Haddon approach portrays contributing factors more completely than does the Lincoln system, but to fully understand the sequence of events using the MOIE/Haddon approach, one would have to search for the original report using the qualitative software. In this instance, the precipitating mechanism is not captured in the MOIE, because the first event in the chain of events (co-workers slipping) did not happen to the injured party.

Haddon's matrix is less scenario-oriented than are some other structured review tools, such as MAIM [Davies et al., 1998]. With written narratives, no follow-up questions are possible, so such a highly structured approach would result in many variables with missing values. The Lincoln et al. [2004] data model template, while simpler, still attempts to depict an injury scenario. Although the Lincoln team was working with reports rather than interviews, it had the advantage of

[1] A wacker is a powered hand tool used for tamping earth. There are several kinds of wackers. The one involved in this event was likely three times the size of a lawn mower and weighed approximately 200 pounds.

**TABLE III.** Coding of Work-Related Injury Data Under Three Injury Classification Systems

| Category | BLS | Lincoln[a] | MOIE/Haddon |
|---|---|---|---|
| Event, type of accident | Strain/Sprain | Bodily reaction | Straining motion |
| Source/agency | No appropriate code | Wacker | — |
| Secondary source | Co-workers | Co-workers | — |
| Contributing factor(s) | — | Slippery terrain | Victim motion/pushing-pulling<br>Victim improper lifting<br>Co-worker dropped object<br>Equipment<br>Object size-weight<br>Environment/terrain/slippery<br>Environment/working overhead |
| Precipitating mechanism | — | Coworker slip/trip | None |
| Activity | — | Backfilling | — |
| Task | — | Pushing up wacker | — |

[a]Lincoln reconstruction: During backfilling, when pushing up wacker, slippery terrain contributed to a coworker slip/trip involving a wacker and co-workers that caused a bodily reaction.

narratives that were relatively long and complete compared with those available to us.

That Haddon did not try to capture the chain of events leading to an injury, but rather included all factors in three general time categories, reflects his belief that countermeasure priorities should not be pursued in the same order as the sequence of events leading to the injury [Haddon, 1968, 1972]. He found the chain of events less relevant than the totality of contributing factors in identifying targets for prevention. Even a very detailed description of the sequence of events leading to an injury does not translate directly into a set of potential points of intervention; the importance of training, housekeeping, or organizational responsibility is not contained in a list of objects, activities, and body positions but must be deduced from them, and this task requires knowledge of the field in which the event occurs. For example, if a worker uses a ladder that fails, does this reflect a failure on the part of the organization (providing inadequate equipment) or on the part of the worker (not inspecting the equipment before using it)? Only knowledge of the expectations in the field can resolve whether either or both factors are involved.

## DISCUSSION

### Strengths

The use of Haddon's matrix provides a theoretical framework around which to code information about a wide array of contributing factors. We believe that, by requiring our coders to consider such a spectrum of factors, this framework elicited more complete coding than would otherwise have occurred. When combined with MOIE it allows a focused selection of events for more detailed analysis, where intervention points can be more readily identified.

Coding contributing factors was made more accurate by using two independent coders and involving a safety expert. It was expedited by coding injuries of the same type together and by adjudicating the disagreements between the coders after a small number of injury events of a given type had been coded. Finally, the process was facilitated by using a qualitative software package that allowed us to store links from our codes to the original, supporting text. This allowed us to revisit decisions and, during analysis, to explore factors not originally identified as important when coding was done. Injuries on this project resulting from falls from height and from slips/trips have been so analyzed [Lipscomb et al., 2004, 2005].

Analyzing all injuries experienced on a specific construction site rather than focusing on fatalities or on a particular type of injury, permitted us to identify the role of contributing factors across many types of injuries. This approach makes it possible to identify contributing factors which, while perhaps not *most* important to any specific type of injury, play an important cumulative role across all injury types. For example, motor vehicles and heavy equipment were involved in every MOIE and nearly every type of work, not just those coded specifically as motor vehicle accidents, as would be expected. In addition, analyzing the whole site allowed us to tie our coding to administrative information, permitting us to explore the number of injuries or amount of Workers' Compensation payments involving particular contributing factors per payroll and hours worked [Glazner et al., 2005].

### Limitations

Although we were able to code the great majority of injury events, we were hampered by limited information on the injury forms available to us. Even when complete data were available and coders agreed on the facts of the event, they could still disagree about what aspects of the event contributed to its occurrence. In this situation, we opted to be inclusive in our coding.

Coding all contributing factors, without regard to whether they were *clearly* causal, may have added some ‘‘noise’’ to our final dataset. Specifically, because we had no separate category for ‘‘source of injury,’’ we included as factors items that were not true points of intervention. For example, if a worker fell to the floor, the floor was coded as an object contributing to the injury, yet it would be absurd to propose that injuries could be reduced by eliminating floors. Adding a category for source of injury and following the Lincoln et al. [2004] emphasis on contributing factors as elements that *increased* the risk of injury might help distinguish true points of potential intervention. Thus, for example, the floor might be coded as a contributing factor if it was covered with rebar but otherwise would be coded as ‘‘source’’ and not as a contributing factor.

Employing reviewers who were not safety experts required development of extensive coding rules. While this exercise might create cost savings in the long run if the rules are reapplied in future projects, it was an arduous and time-consuming effort, and at least some of the resultant rules were specific to the DIA work site. Inexperienced reviewers need explicit instruction about how to handle complex and sometimes ambiguous narratives. The advantage of making the decision rules explicit is that they may be examined and should result in more consistent decisions than might be expected from implicit coding, even by experts, but exposition of implicit knowledge into decision rules is a difficult task.

The cost of coding injury reports ($14.60 per report) was considerable and may limit the applicability of this method. As noted above, this cost might be somewhat reduced in future projects, to the extent that the taxonomy of contributing factors and the coding rules could be re-used for other construction injury analyses.

## CONCLUSION

Classification schemes vary in terms of how granular they are, how structured, how much emphasis they place on the sequence of events, how much source information is needed to use them, and how skilled the coders must be. A highly structured system may not provide a place to capture all relevant information. A classification scheme with a complicated taxonomy may lead to misclassification unless used by highly skilled coders. A system with many variables risks creating a database with much missing data unless it is being filled out from detailed reports or from interviews. Researchers must constantly make trade-offs among these issues in choosing a classification scheme appropriate to their needs and resources.

Nevertheless, prompting the people who fill out injury reports to be more complete in their reporting would seem to be an indisputable improvement. Some reporting forms included check boxes about important items (e.g., use of PPE), but our impression was that these boxes were checked without much thought and the information contained in them could not be trusted. An alternative that the authors have been investigating is the use of AIRs that use conditional branching so that important questions are asked based on the kind of injury reported. Forms that guide investigators to consider human, organizational, and environmental factors explicitly could foster more complete descriptions of factors contributing to construction injury. In addition, the inclusion of text allows the post-hoc investigation of issues that are not always evident at the time of initial data collection even when conditional branching is used for broad mechanisms of injury [Lipscomb et al., 2003].

With this experiment in using a common source of short narrative injury descriptions, we believe we have added to the spectrum of tools available to researchers seeking to describe the etiology of construction injury. This approach will become more powerful if it is possible to revise standard reporting forms to use conditional branching so that important questions are asked based on the kind of injury reported.

## ACKNOWLEDGMENTS

The authors thank Kathleen Kreiss, Stacy Pocrass, and Harry Shannon, who assisted in development of the injury report coding process and use of the taxonomy underlying it. We are grateful to Jennifer Paddock and Eric Lupo, who expertly coded the thousands of injury reports used in this study. For provision of the DIA administrative data and injury reports, we thank Sue Christenson, DIA Risk Manager.

## REFERENCES

Althouse RC, Klishis MJ, Grayson RL. 1997. Microanalysis of roof bolter injuries. Appl Occup Environ Hyg 12:851–857.

American National Standards Institute, Inc (ANSI). 1963. American national standard method of recording basic facts relating to the nature and occurrence of work injuries. New York: ANSI, Inc.

Bell NS, Amoroso PJ, Yore MM, Senier L, Williams JO, Smith GS, Theriault A. 2001. Alcohol and other risk factors for drowning among male active duty U.S. Army soldiers. Aviat Space Environ Med 72: 1086–95.

BLS (Bureau of Labor Statistics) 2003. http://www.bls.gov/iif/oshwc/osh/os/ostb1236.pdf

Cohen HH, Lin L. 1991. A scenario analysis of ladder fall accidents. J Saf Res 22:31–39.

Conroy C, Fowler J. 2000. The Haddon Matrix: Applying an epidemiological research tool as a framework for death investigation. Am J Forensic Med Pathol 21:339–342.

Davies JC, Stevens G, Manning DP. 1998. Understanding accident mechanisms: On analysis of the components of 2516 accidents in a MAIM database. Saf Sci 29:25–58.

Dement JM, Lipscomb H, Li L, Epling C, Desai T. 2003. Nail gun injuries among construction workers. Appl Occup Environ Hyg 18: 374–83.

Glazner JE, Borgerding JA, Lowery JT, Bondy J, Mueller KL, Kreiss K. 1998. Construction injury rates may exceed national estimates: Evidence from the construction of Denver International Airport. Am J Ind Med 34:105–112.

Glazner J, Bondy J, Lezotte DC, Lipscomb H, Guarini K. 2005. Factors contributing to construction injury at Denver International Airport. Am J Ind Med 47:27–36.

Haddon W. 1968. The changing approach to the epidemiology, prevention, and amelioration of trauma: The transition to approaches etiologically rather than descriptively based. Am J Public Health 58: 1431–1438.

Haddon W. 1972. A logical framework for categorizing highway safety phenomena and activity. J Trauma 12:193–207.

Higgins DN, Casini VJ, Bost P, Johnson W, Rautiainen R. 2001. The fatality assessment and control evaluation program's role in the prevention of occupational fatalities. Inj Prev 7(suppl 1):127–133.

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 18:49–50.

Lincoln AE, Sorock GS, Courtney TK, Wellman HM, Smith GS, Amoroso PJ. 2004. Using narrative text and coded data to develop hazard scenarios for occupational injury interventions. Inj Prev 10: 249–254.

Lipscomb HJ, Kalat J, Dement JM. 1996. Workers' compensation claims of union carpenters 1989–1992: Washington state. Appl Occup Environ Hyg 11:56–63.

Lipscomb HJ, Dement JM, Nolan J, Patterson D, Li L. 2003. Nail gun injuries in residential carpentry: Lessons from active injury surveillance. Inj Prev 9:20–24.

Lipscomb HJ, Glazner J, Bondy J, Lezotte D, Guarini K. 2004. Analysis of text from injury reports improves understanding of construction falls. J Occup Environ Med 46 :1166–1173.

Lipscomb HJ, Glazner J, Bondy J, Lezotte D, Guarini K. 2005. Injuries resulting from slips and trips in construction. Appld Ergon (in press) DOI 10.1016/j.apergo.2005.07.008

N5, QSR International Pty. Ltd., Bundoora, Victoria, Australia. 2000.

Park RM. 2002. Hazard identification in occupational injury: Reflections on standard epidemiologic methods. Int J Occup Environ Health 8:354–362.

Runyan CW. 1998. Using the Haddon matrix: Introducing the third dimensions. Inj Prev 4:302–307.

Sexias NS, Sanders J, Sheppard L, Yost MG. 1998. Exposure assessment for acute injuries on construction sites: Conceptual development and pilot test. Appl Occup Environ Hyg 13:304–312.

Sorock GS, Ranney TA, Lehto MR. 1996. Motor vehicle crashes in roadway construction workzones: An analysis using narrative text from insurance claims. Accid Anal Prev 28:131–138.

Stout N. 1998. Analysis of narrative text fields in occupational injury data. In: Feyer A-M, Williamson A, editors. Occupational injury . Risk, prevention and intervention. New York: Taylor and Francis. p 15–20.

Stout N, Jenkins EL. 1995. Narrative text fields in occupational injury data. Proceedings of the International Collaborative Effort on Injury Statistics, March:1–24.

Wiehagen WJ, Mayton AG, Jsapal JS, Turin FC. 2001. An analysis of serious injuries to dozer operators in the U.S. mining industry. NIOSH April:1–23.

Williamson A, Feyer AM, Stout N, Driscoll T, Usher H. 2001. Use of narrative analysis for comparisons of the causes of fatal accidents in three countries: New Zealand, Australia, and the United States. Inj Prev 7(suppl I):i15–20.