

RESEARCH ARTICLES

Use of narrative analysis for comparisons of the causes of fatal accidents in three countries: New Zealand, Australia, and the United States

A Williamson, A-M Feyer, N Stout, T Driscoll, H Usher

Abstract

Objectives—To investigate the utility of narrative analysis of text information for describing the mechanism of injury and to compare the patterns of the mechanism of injury for work related fatalities in three countries.

Methods—Three national collections of data on work related fatalities were used in this study including those for New Zealand, 1985–94 (n=723), for Australia, 1989–92 (n=1220), and for the United States, 1989–92 (16 383). The New Zealand and Australian collections used the type of occurrence standard code for the mechanism of injury, however the United States collection did not. All three databases included a text description of the circumstances of the fatality so a text based analysis was developed to enable a comparison of the mechanisms of injury in each of the three countries. A test set of 200 cases from each country dataset was used to develop the narrative analysis and to allow comparison of the narrative and standard approaches to mechanism coding.

Results—The narrative coding was more useful for some types of injury than others. Differences in coding the narrative codes compared with the standard code were mainly due to lack of sensitivity in detecting cases for all three datasets, although specificity was always high. The pattern of causes was very similar between the two coding methods and between the countries. Hit by moving objects, falls, and rollovers were among the five most common mechanisms of workplace fatalities for all countries. More common mechanisms that distinguished the three countries were electrocutions for Australia, drowning for New Zealand, and gunshot for the United States.

Conclusion—Narrative analysis shows some promise as an alternative approach for investigating the causes of fatalities.

(*Injury Prevention* 2001;7(Suppl I):i15–20)

Keywords: narrative; text analysis; fatalities; occupational injury

Classification and coding systems have been developed to provide some systematic basis for collecting, aggregating, and comparing injury data in areas such as severity,¹ nature, and body region of injury,² in addition to comparing the *International Classification of Diseases* (ICD)9 and ICD10 classification systems.³ In theory, these coding systems allow for collection and comparison of data in the same form from multiple sites. In practice, however, there are considerable difficulties in using these systems.

An alternative approach is to use narrative or text descriptions of the circumstances of the injury.⁴ This approach allows data collectors and/or coders to describe the circumstances of the incident in as much detail as they feel is necessary. The main benefits of the approach are that the circumstances of the injury can be reported in the words of the data collector, coders are not frustrated in making decisions on how to code difficult cases, there is less possibility of misclassification errors, and less training of coders is needed to ensure compatibility of coding. Narratives provide the opportunity to reflect the broader circumstances of the incident by providing much more information than coded variables, and, where coding decisions are difficult to make and may be unreliable, a narrative or text based approach is likely to be more suitable.

Narrative approaches have some disadvantages, however, mainly in the difficulty in coding text descriptions systematically and reliably. This means that it may be impossible to use multiple coders and the reliability of descriptions from single coders may also vary over time. In addition, data collectors (for example, medical personnel), due to time constraints or disinterest, may collect insufficient text about why and how the injury occurred.

The usefulness of narrative information to describe the circumstances of work related fatalities was examined through a three country collaboration between the United States, New Zealand, and Australia. The project was part of the International Collaborative Effort on Injury (ICE) and includes comparison of national surveys of work related fatal accidents occurring over a similar period (1989–92 inclusive for Australia and the United States and 1985–94 for New Zealand, where the

NSW Injury Risk Management Research Centre and School of Psychology, University of New South Wales, Sydney, Australia
A Williamson
H Usher

New Zealand Environmental and Occupational Health Research Centre, University of Otago, Dunedin, New Zealand
A-M Feyer

National Institute for Occupational Safety and Health, Division of Safety Research, Morgantown, West Virginia, USA
N Stout

National Occupational Health and Safety Commission, Sydney, Australia
T Driscoll

Correspondence to:
Dr Ann Williamson, NSW Injury Risk Management Research Centre, Building B10, University of New South Wales, Sydney, NSW 2052, Australia
a.williamson@unsw.edu.au

number of cases is considerably smaller). To date the comparison has been very useful in comparing a range of demographic and job related factors⁵ and has identified some clear targets for further attention. This paper extends this analysis through a comparison of the mechanism of injury for the three countries.

All countries collected text information on injury circumstances. Australia and New Zealand used the same coding system for mechanism of injury, whereas the United States did not. This meant that it was possible to use the Australian and New Zealand datasets to develop a text search technique where its accuracy for coding of the mechanism of injury could be compared to coding using existing standard codes. This text search technique could then be applied to the text information from the United States to generate a code for mechanism of injury, with the additional advantage of allowing estimation of the accuracy of the text generated coding. The aim of this study, therefore, was to compare patterns of the mechanism of injury for work related fatalities in three countries and to investigate the utility of narrative analysis of text information on the circumstances of injury occurrence.

Method

DATASETS

Three national datasets of occupational fatalities were used which covered the period 1989–92 inclusive for Australia (n=1220) and the United States (n=16 383) and the years 1985–94 inclusive for New Zealand (n=723). The collections were made compatible by developing minimum criteria for all three datasets then selectively removing cases from each dataset that did not meet those criteria as discussed in depth in a previous paper.⁵ The main criteria for selection of cases were that the fatality was work related and did not include traffic fatalities.

All datasets contained similar variables (age, gender, industry, occupation, etc), although only the Australian and New Zealand datasets included a code for the mechanism of the injury. All three datasets, however, included a short text field, which provided a description of the circumstances of the injury. For the United States and New Zealand collections, the text field was brief, usually around one sentence long, whereas the Australian dataset included a more detailed text description.

ANALYSIS

The analysis involved a number of steps.

(1) First, the text search was developed based on mechanism codes within the Type of Occurrence Classification System (TOCS),⁶ a coding system used by Australia and New Zealand. The type of occurrence coding system was based on the ICD9 coding system.³ For this exercise, for the purposes of simplification, an abbreviated coding system was used in which the standard TOC mechanism was coded to the one digit or two digit level. The

New Zealand data were used as the text descriptions of injury occurrence were most similar to the style of the United States text in length and detail of description. The Microsoft Access programme was employed to generate a text search.

(2) The text search was then refined on a random sample of 200 New Zealand fatalities by comparing results with the standard TOCS codes. The refining process continued until modification of text terms did not result in further gains in accuracy. The sensitivity, specificity, and positive predictive value of the coding were calculated for each mechanism category. For this purpose, sensitivity and specificity were defined and calculated using conventional methods.⁷ Sensitivity describes the capacity to detect deaths that are really cases (true positives). It was defined in this study as the ratio of the number of deaths correctly identified as cases over the number of cases that should have been detected. Specificity describes the capacity of the coding system to detect deaths that should not be called cases and therefore to avoid false positives. This was defined as the ratio of the number of deaths identified as non-cases over the number of deaths that should not have been detected as cases. The positive predictive value of coding describes the capacity of the test (in this case the text search system) to detect cases correctly. For this exercise, positive predictive value was defined as a ratio of the number of deaths correctly identified as cases over the total number of deaths identified as cases. All of these measures were expressed as percentages.

(3) For the next step, a sample of 200 cases in the United States was coded for mechanism using the standard TOCS code. This coding was done independently by two coders and shown to be accurate ($\kappa=0.81$). The text search mechanism code was then used on these cases and the results compared with the standard TOC mechanism code for this dataset. The text search was refined further in the same way as before to maximise sensitivity, specificity, and positive predictive value of coding.

(4) The text search was then applied to a random sample of 200 cases from the Australian database to investigate its applicability to more comprehensive text material and again positive predicted value, sensitivity, and specificity were calculated for each mechanism category.

(5) Finally the text search was applied to all cases in the datasets for each country. It was possible to calculate sensitivity, specificity, and positive predictive value of mechanism coding for the entire Australian and New Zealand datasets and to estimate the likely error in text coding for the entire United States dataset. The distribution of cases in each mechanism category could then be calculated and compared between countries using the new narrative generated coding system for each country and using the standard TOC mechanism coding system for Australia and New Zealand.

The final text search was developed after trialing on first New Zealand, then United States,

Table 1 Text search and standard mechanism coding for a sample of work related fatalities in New Zealand, Australia, and the United States

Mechanism	New Zealand (n=200)				Australia (n=200)				United States (n=200)			
	n=Mech code	% Positive predictive value	% Sensitivity	% Specificity	n=Mech code	% Positive predictive value	% Sensitivity	% Specificity	n=Mech code	% Positive predictive value	% Sensitivity	% Specificity
Drown	18	75	100	97	10	60	80	99	6	100	67	100
Hypoxia	5	38	100	96	4	50	75	99	6	100	33	100
Explosion	8	83	63	99	3	100	100	100	9	100	100	100
Electrocution	1	100	100	100	20	79	95	97	25	94	68	99
Chemical	4	67	50	99	2	33	50	99	0	—	—	99
Gunshot	3	100	100	100	10	91	100	99	35	100	97	100
Stabbing	3	100	100	100	2	50	100	99	0	—	—	100
Assault	4	75	75	100	3	100	100	100	3	100	67	100
Slide/cave-in	0	0	0	98	1	0	0	93	5	60	60	99
Rollover	9	50	88.9	96	10	50	90	95	21	85	81	99
Bite/sting	0	—	—	100	1	100	100	100	0	—	—	99
Contact with hot	4	50	75	99	5	60	60	99	3	0	0	99
Contact with cold	1	100	100	100	0	—	—	100	0	—	—	100
Hit by moving object (combined category)	69	72	67	86	73	65	62	81	30	61	73	99
Bitten by animal	0	—	—	100	1	100	100	100	0	—	—	99
Being trapped	28	67	21	98	13	25	15	97	10	50	10	99
Falls	47	64	96	99	35	40	74	95	29	57	86	99

and then Australian samples. The text search always began with the term most like the type of category (for example, *drown*, *hypox*, etc) then expanded to include synonyms and related words (for example, *blew up*, *inhale*, etc). For some categories exclusion search terms were used (for example, for *falls*, *Not like felled*, *felling*). Some codes required considerably more search terms in one country than others to achieve accurate coding. For example, all of the drowning cases in the dataset from the United States were obtained by using three search words, whereas the same three search words only obtained 60% of the drowning cases for the Australian database. Similarly, the searches varied in the specificity of picking up target cases. For example, a number of cases involving the words *roll*, *overturn*, or *flip* were detected, but only around half of them had been coded in the *rollover* category using the standard mechanism code.

Results

The results of final text searches on the 200 case samples from each of the datasets are shown in table 1. The results show fairly similar patterns of sensitivity, specificity, and positive predictive value between each of the

datasets. The main errors of coding for all datasets were in sensitivity rather than specificity of detecting cases. Coding for the New Zealand dataset showed the best sensitivity and specificity overall, followed by Australia, then the United States. The most troublesome codes were similar for each dataset. All datasets showed poor sensitivity of coding for *being trapped*.

Applying the text based coding system to the larger United States and Australian datasets resulted in some loss of predictive value and to a lesser extent sensitivity, but specificity of the coding system remained very high for both countries. The results of coding by text search for the entire dataset for New Zealand and Australia are shown in table 2 together with the results for the same 200 cases for the United States coding. The most notable change was for the Australian dataset, with predictive value decreasing for 10 codes by amounts ranging from 3% for *chemical exposure* to 67% for *assaults*. For New Zealand, decreases in positive predictive value were seen for only four codes and then only by a maximum of 30%. Predictive value actually improved for four codes in the New Zealand dataset, and for three codes for the Australian dataset. Sensitivity showed slightly better results when applied

Table 2 Text search coding of mechanism for all work related fatalities in New Zealand and Australia and for a sample of United States fatalities

Mechanism	New Zealand (n=723)				Australia (n=1220)				United States (n=200)			
	n=Mech code	% Positive predictive value	% Sensitivity	% Specificity	n=Mech code	% Positive predictive value	% Sensitivity	% Specificity	n=Mech code	% Positive predictive value	% Sensitivity	% Specificity
Drown	95	88	99	98	102	100	68	100	6	100	67	100
Hypoxia	10	40	80	98	5	22	80	98	6	100	33	100
Explosion	13	68	100	99	18	45	83	97	9	100	100	100
Electrocution	29	100	97	100	117	73	94	93	25	94	68	99
Chemical	8	80	50	100	19	30	16	99	0	—	—	99
Gunshot	15	89	79	100	41	76	85	98	35	100	97	100
Stabbing	6	100	83	100	11	31	100	97	0	—	—	100
Assault	6	45	83	99	7	33	71	99	3	100	67	100
Slide/cave-in	10	36	90	98	17	14	59	92	5	60	60	99
Rollover	87	59	82	92	60	29	75	84	21	85	81	99
Bite/sting	4	0	—	—	3	50	100	100	0	—	—	99
Contact with hot	4	0	—	—	15	38	53	98	3	0	0	99
Contact with cold	1	100	100	100	1	0	—	—	0	—	—	100
Hit by moving object (combined category)	243	79	88	89	431	83	75	77	30	61	73	99
Bitten by animal	0	—	—	—	2	100	100	100	0	—	—	99
Being trapped	96	55	44	94	77	27	38	88	10	50	10	99
Falls	95	44	93	82	174	44	79	68	29	57	86	99

Table 3 Narrative text analysis and search terms for mechanism of injury of work related fatalities in New Zealand, Australia, and the United States

Mechanism	New Zealand % (n=723)	Australia % (n=1220)	United States % n=12 809	Search term
Drowning	14.8	5.7	3.1	Like “*drown*” or like “*overboard*” or like “*boat*”
Hypoxic atmosphere	2.8	1.5	1.1	(Like “*hypox*” or like “*inhal*” or like “*smoke*” or like “*overcome*” or like “*suffoc*” and not like (“*crush*”))
Explosion	2.6	2.7	4.3	Like “*explo*” or like “*blew up*”
Contact with electricity	3.9	12.3	5.7	Like “*elec*” or like “*volt*”
Contact with chemical	0.7	0.8	0.67	Like “*chem*” or like “*toxic*” or like “*poison*” or like “*ingest*”
Shot by gun	0.03	4.0	15.0	Like “*shot*” or like “*shoot*” or like “*gun*”
Stabbed by knife	0.7	2.9	1.7	Like “*stab*” or like “*knife*”
Assault with other weapon	1.5	1.2	1.0	Not like “*heart*” and not like “*knife*” and not like “*shot*” and (like “*assault*” or like “*attack*” or like “*fought*” or like “*beat*” or like “*murder*”)
Slide or cave-in	3.5	5.7	2.4	Like “*cave*” or like “*slide*” or like “*avalanche*” or like “*collapse*”
Rollover	16.7	12.6	8.9	Like “*roll*” or like “*rolled*” or like “*overturn*” or like “*O/T*” or like “*flip*”
Insect and spider bites/stings	0.0	0.5	0.05	(Like “*sting*” or like “*bit*” or like “*stung*” or like “*bitten*” or like “*bite”) and not like “*snake*” and not like “*animal*”
Contact with hot objects/exposure to environmental heat	0.1	1.7	0.4	(Like “*hot*” or like “*burn*” or like “*flame*” or like “*fire*” or like “*heat*” or like “*steam*”) and not like (“*steam train*”) and not like (“*elec*”)
Contact with cold objects/exposure to environmental cold	0.1	0.7	0.03	Like “*hypothermia*” or like “*cold*”
Being hit by moving objects (combined category)	36.4	31.8	31.3	(Like “*run over*” or like “*ran over*” or like “*struck*” or like “*pin*” or like “*crush*” or like “*hit by*” or like “*trap*” or like “*coll*” or like “*crash*” or like “*impact*”) and not like (“*roll*”)
Being trapped by moving machinery or between stationary and moving objects	10.7	8.9	3.4	Like “*trap*” or like “*caught*”
Falls, trips, and slips of a person	27.4	25.6	22.3	(Like “*fall*” or like “*falls*” or like “*fell*” or like “*trip*” or like “*slip*”) and not like (“*felling*”) and not like (“*felled*”)

to the larger dataset for Australia with only seven codes showing lower sensitivity and to a smaller extent (around 20%). Similar results were seen for New Zealand with loss of sensitivity seen only for four codes and then only around 15%. Specificity remained very high for both New Zealand and Australian datasets, only dropping by a few per cent.

The final text based code was applied to all three total datasets (see table 3) and the results showed the same top two causes of workplace fatalities in each country. The most common cause was being *hit by moving objects* (including motor vehicles), followed by *falls, slips, and trips of the person*. The patterns then differed between the countries. For New Zealand, the third and fourth most common mechanisms were *rollover* and *drowning*, whereas for Australia it was *rollover* and *contact with electricity* and for the United States, *shot by gun*, and *rollover*.

As a comparison, it was also possible to examine the patterns of injury causation that emerge for Australia and New Zealand when the standard mechanism coding system was used. Table 4 shows the results from applying

the standard mechanism codes and for comparison, the narrative derived codes for the United States. The same top cause, *being hit by moving vehicles* was seen, but then the pattern differs from that found using the text coding method. The next most common causes for New Zealand were *being trapped*, *drowning* and *falls, trips and slip* then *rollover*, and *falls, trips, and slips* were much less common. For Australia, the top two causes were the same as found using text, but *drowning* was much more common. *Rollovers* were much less common using the standard coding system. The finding of higher percentages in Australia of *slide or cave in* and *insect/spider bites* were also not found using the standard coding system.

Discussion

From these results, it seems the same general mechanisms cause most fatal accidents in all three countries. The narrative search results showed that in all three countries *being hit by moving objects* was the most common mechanism for the fatality, followed by *falls, trips, and slips*.

Table 4 Mechanism of injury for work related fatalities using standard mechanism codes for New Zealand, Australia and for comparison, narrative text search for the United States is shown in italics

Mechanism	Code	New Zealand % (n=723)	Australia % (n=1220)	United States % (n=12 809)
Drowning	58	13.1	8.4	3.1
Hypoxic atmosphere	64	1.4	0.4	1.1
Explosion	59	1.8	1.5	4.3
Contact with electricity	57	4.0	9.6	5.7
Contact with chemical	61	1.1	1.6	0.67
Shot by gun	82	2.1	3.4	15.0
Stabbed by knife	83	0.8	0.9	1.7
Assault with other weapon	84	0.8	0.6	1.0
Slide or cave-in	91	1.4	1.4	2.4
Rollover	93	12.0	4.9	8.9
Insect and spider bites/stings	63	0.6	0.2	0.05
Contact with hot objects/exposure to environmental heat	51, 53	0.6	1.2	0.4
Contact with cold objects/exposure to environmental cold	52	0.1	0.1	0.03
Being hit by moving objects (combined category)	20, 21, 22, 23, 24, 25, 26, 27, 28, 92	44.1	50.2	31.3
Being trapped by moving machinery or between stationary and moving objects	25, 26	13.3	6.3	3.4
Falls, trips, and slips of a person	1, 2, 3, 4, 5	13.1	14.3	22.3

The exceptions to this conclusion are that a larger percentage of fatalities in New Zealand can be attributed to drowning, a higher percentage in Australia to electrocutions, and in the United States the percentage of workplace deaths due to gunshot was much higher than seen in Australia or New Zealand. This result is consistent with previous analysis of this dataset which showed higher death rates for fishing related occupations in New Zealand compared with Australia and the United States.⁵ The higher percentage of gunshot fatalities in the United States compared with the other countries also might be expected due to the higher rates of gun ownership in the United States. The reason for the higher percentage of electrocutions in Australia is not readily apparent. Although a higher electricity voltage is standard in Australia compared with the United States (240 volts compared with 110 volts), New Zealand also uses the same higher voltage as Australia. There is also no indication that more workers in Australia are likely to be exposed to electrical hazards, for example, the percentage of Australians working in trades related areas is smaller than in the United States or New Zealand (21.1%, 26.2%, and 30.9% for Australia, United States, and New Zealand respectively). Consequently, this finding needs further investigation.

Compared to the standard method, the narrative coding method was more successful for some codes than others. Errors in the narrative coding occurred due to lack of sensitivity in detecting cases for all three datasets. Sensitivity ranged from 100% for some codes in each country dataset to as low as 16% for fatalities involving exposure to chemicals in the Australian dataset. Even the addition of a wide range of search terms did not improve sensitivity a great deal for some mechanism codes, such as hypoxia. Sensitivity was also poor for the United States dataset, in particular for hypoxia, although specificity was high.

Some text based codes were fairly successful for all three country datasets. For example, over three quarters of cases involving drowning, explosion, gunshot, rollover, and falls were detected in each country dataset. In contrast, sensitivity was poor for all three datasets for chemical exposure and being trapped as more than half of the cases that would have been picked up by the standard coding system were missed.

Despite the finding that the text based approach tends to underestimate the number of cases involving many of the mechanism codes, coding is very specific, even using a very simple set of text terms. Sensitivity and specificity are often a trade-off. In this study, a conservative approach to case detection was employed to maximise specificity, as misclassification of cases was of greater disadvantage than assigning codes to more cases, but with known error. Text searching with high specificity, even if sensitivity is less than desired, has useful advantages. Knowing the sensitivity of individual codes can allow accurate analysis of cases within those categories. For example we can now electronically identify cases of electrocution from the datasets

for further analysis of that subset. Moreover, classification through text searching that results in acceptable accuracy of codes can be a tremendous savings of time and skilled effort, particularly on large datasets, even if remaining cases must be coded by hand.

In this analysis it was only possible to estimate the errors of text search coding for the United States dataset by using the sample of 200 cases that were coded using the standard method specifically for this exercise. It is, therefore, not possible to know exactly how much error there is in the text search result for the entire United States dataset. It would be expected, however, that the change for the United States dataset would be very similar to those for the other two datasets. Interestingly, for Australia and New Zealand datasets, predictive value and sensitivity improved for some codes (drowning, electrocution, and being hit by moving vehicles) by as much as 60% when the text search was applied to the large dataset indicating the success of the particular text searches for these codes. For most codes, the specificity of coding hardly changed when applied to the large dataset, but sensitivity and predictive value fell. Clearly this is the result of the particular strategy used in this study whereby false positives were minimised, but at a cost of a higher risk of missing true cases. This necessarily led to underestimation of the number of cases.

The accuracy and sensitivity of narrative coding could be improved by developing a text search dictionary that contains a set list of terms to describe particular types of injury causes. Based on the experiences from this study, it seems that this would certainly be feasible, although any improvements in coding would need to be balanced by the additional need for coder training. As more sophisticated automatic text search methods become available, it may be possible to overcome some of the problems encountered in this study.

This study revealed that there are relatively few differences in causes of occupational fatalities between the three countries. The differences that were identified are prime targets for further study to examine international differences in risk factors and prevention measures for these causes. The study also revealed that a text based approach to coding mechanism of injury can reflect the same general patterns of injury causation as a much simpler, standard method. The results showed that the two approaches differ in the types of errors they produce. The specificity of the text search was very high for most mechanism codes. The positive predictive value of the text search method, however, was variable, being high for some mechanisms, but moderate for some of the more common mechanisms such as falls, being trapped, and rollovers. Similarly, the sensitivity was variable, but was 75% or more for most major mechanisms in all three countries. Overall, the narrative text search approach is a promising alternative, or addition, to manual coding, particularly with some knowledge of the form that errors are likely to take, as provided by this study.

Funding for this project was provided by the National Institute for Occupational Safety and Health, New Zealand Environmental and Occupational Health Research Centre, and New Zealand Occupational Safety and Health Service.

The authors wish to acknowledge the International Collaborative Effort (ICE) on Injury Statistics for contributions to this research. The ICE is sponsored by the National Center for Health Statistics, US Centers for Disease Control and Prevention with funding from the National Institute of Child Health and Development, National Institutes of Health.

The views expressed in this paper are those of the authors and do not necessarily reflect those of the National Occupational Health and Safety Commission.

- 1 Association for the Advancement of Automotive Medicine. Committee on Injury Scaling. *The abbreviated injury scale—1990 revision (AIS-90)*. Des Plaines, IL: AAAM, 1990.
- 2 MacKenzie E, Champion H. Development of a matrix for classifying injuries according to their nature and body region. *Proceedings of the international collaborative effort on injury statistics*. 1999, Vol 111; 17.1–17.3.
- 3 World Health Organization. *International classification of diseases*. 9th Revision. Geneva: World Health Organization, 1977.
- 4 Stout N. Analysis of narrative text fields in occupational injury data. In: Feyer A-M, Williamson A, eds. *Occupational injury. Risk, prevention and intervention*. 1998: 15–20.
- 5 Feyer A-M, Williamson AM, Stout N, et al. Comparison of work related fatal injuries in the United States, Australia, and New Zealand: method and overall findings. *Inj Prev* 2001;7:22–8.
- 6 National Occupational Health and Safety Commission. *Type of occurrence classification system*. Canberra: Australian Government Publishing Service, 1990.
- 7 Hennekens CH, Buring JE. *Epidemiology in medicine*. Boston, US: Little Brown and Company, 1987.