

# Accuracy of injury coding under ICD-9 for New Zealand public hospital discharges

J Langley, S Stephenson, C Thorpe, G Davie

*Injury Prevention* 2006;12:58–61. doi: 10.1136/ip.2005.010173

See end of article for authors' affiliations

Correspondence to: Professor J D Langley, Injury Prevention Research Unit, Department of Preventive and Social Medicine, University of Otago, PO Box 913, Dunedin, New Zealand; john.langley@ipru.otago.ac.nz

Accepted 12 December 2005

**Objective:** To determine the level of accuracy in coding for injury principal diagnosis and the first external cause code for public hospital discharges in New Zealand and determine how these levels vary by hospital size.

**Method:** A simple random sample of 1800 discharges was selected from the period 1996–98 inclusive. Records were obtained from hospitals and an accredited coder coded the discharge independently of the codes already recorded in the national database.

**Results:** Five percent of the principal diagnoses, 18% of the first four digits of the E-codes, and 8% of the location codes (5th digit of the E-code), were incorrect. There were no substantive differences in the level of incorrect coding between large and small hospitals.

**Conclusions:** Users of New Zealand public hospital discharge data can have a high degree of confidence in the injury diagnoses coded under ICD-9-CM-A. A similar degree of confidence is warranted for E-coding at the group level (for example, fall), but not, in general, at higher levels of specificity (for example, type of fall). For those countries continuing to use ICD-9 the study provides insight into potential problems of coding and thus guidance on where the focus of coder training should be placed. For those countries that have historical data coded according to ICD-9 it suggests that some specific injury and external cause incidence estimates may need to be treated with more caution.

Injury is increasingly being recognized as a major cause of morbidity worldwide. In response to this, injury prevention researchers and practitioners in several countries have argued for diagnostic and external cause coding of all injuries resulting in inpatient hospital treatment.<sup>1</sup> New Zealand is one of a small number of countries that have been coding the nature and external cause of injury for hospital inpatients, nationally, for many years. Details of the system are available from the New Zealand Health Information Service (<http://www.nzhis.govt.nz/documentation/nmds-dictionary>). The coded information has been used extensively in New Zealand to estimate the incidence of injury events, monitor trends, and evaluate interventions. Despite its practical importance, however, information on the reliability of the coding has, with one exception, been confined to unpublished audits by the New Zealand Health Information Service (NZHIS) of random samples of all public hospital discharges. These audits have not included sufficient injury discharges to allow for any meaningful analyses of errors by injury diagnostic or external cause groups. The only published study has the same shortcoming and in addition was based on a single hospital.<sup>2</sup>

Although published literature describing the coding accuracy of specific ICD diagnoses (for example, hip fracture)<sup>3</sup> and evaluating other cause of injury coding systems (for example, US Military use of STANAG)<sup>4</sup> exists, there are few published studies which might provide insight into the level of error to be expected for the full range of ICD-9 external cause codes and injury codes for large samples of hospital inpatients from the general population.

Two exceptions have examined external cause coding but the focus has been on high level external cause groupings.<sup>1 5</sup> Only one study has examined diagnostic coding and it too was restricted to high level groupings of codes.<sup>6</sup>

One issue, which has not been addressed to date, is the degree to which error is a function of number of discharges coded. In New Zealand, public hospitals are responsible for

the coding of their own discharges with the size of the community each hospital serves varying substantially. Consequentially, the large hospitals have teams of coders. In these situations coding errors may be lower due to the opportunity for peer review. On the other hand, the complexity of the discharges requiring coding may be higher.

The aims of this study were to determine the level of accuracy in coding for injury principal diagnosis and the first external cause code for public hospital discharges in New Zealand and determine how these levels vary by hospital size.

## METHODS

The National Minimum Data Set (NMDS) is collated by the NZHIS and includes records of all discharges from hospitals following injuries occurring in New Zealand which result in inpatient treatment.<sup>7</sup> The vast majority of injury discharges from hospitals in New Zealand are from public hospitals. The most recent year for which published data are available for private hospitals indicated that in 1995 there were 1296 injury and poisoning discharges.<sup>8</sup> The comparable figure for public hospitals was 66 054.<sup>9</sup>

Diagnoses and external causes (E-codes) are coded according to the various versions of the International Classification of Diseases. This study was confined to records coded under the Australian version of the International Classification of Disease, 9th Revision Clinical Modification (ICD-9-CM-A).<sup>10</sup> Coding personnel at the time ICD-9-CM-A was being used were administration clerks with mainly on-the-job training. Peer review took place only in the form of "discussing" the difficult events among peers, particularly in the larger hospitals. The Ministry of Health performed regular external coding audits.

**Abbreviations:** NMDS, National Minimum Data Set; NZHIS, New Zealand Health Information Service; VIMD, Victorian Inpatient Minimum Database.

**Table 1** Correct and incorrect coding summary

	Principal diagnosis	% of total coded (95% CI)	E-code	% of total coded (95% CI)
Correct	1579	95 (93–96)	1238	74 (72–76)
Correct to 4 digit level, 5th digit incorrect	9	0 (0–1)	131	8 (7–9)
Correct to 3 digit level, 4th digit incorrect	20	1 (1–2)	57	3 (3–4)
Correct to group level, 2nd or 3rd digit incorrect	26	2 (1–2)	168	10 (9–12)
Incorrect at group level	36	2 (2–3)	76	5 (4–6)
Total coded	1670		1670	

A simple random sample of 1800 discharges recorded in the NMDS was selected from the period 1996–98 inclusive. This allowed for the 95% confidence interval around the overall estimate of coding error to be no wider than plus or minus 5%. This interval represented a balance between precision and the cost associated setting a narrower interval. 1998 was the last year the majority of hospital discharges in New Zealand were coded in ICD-9-CM-A.

The sample was limited to those discharges:

- where the principle diagnosis was injury (injury code 800–999);
- which were the first discharge for this injury; and
- which were from a public hospital.

The first restriction reflects the criteria NZHIS uses in its statistical summaries of cases with an external cause code. The second restriction was to avoid sampling those injury events which occurred as multiple records due to readmissions for the treatment of the same injury. The methods for this have been described elsewhere.<sup>11</sup> The last restriction was for pragmatic reasons. The majority of private hospitals in New Zealand do not code the nature and circumstances of injury of their discharges. Irrespective as illustrated above 98% of inpatient treatment in the acute phase is managed by public hospitals.

A list of the medical records of each discharge selected in the random sample was provided to the hospitals before arrival. If retrieved, they were reviewed either at the relevant hospital or at the nearest major hospital in the region. The latter procedure was put in place to avoid the high travel costs of reviewing a small number of records at any one hospital. An accredited expert (not one of the authors) on ICD-9-CM-A independently coded the principal diagnosis and the E-code.

The accredited coder was instructed to code the discharge independently of the codes already recorded in the NMDS.

However, in some cases there was access to the previously selected codes as some hospitals have them listed in the summary information in the medical record.

By comparing the codes thus generated and those of the NMDS, the selected NMDS discharges were separated into five mutually exclusive categories:

1. Correct
2. Correct to the 4 digit level, 5th digit incorrect
3. Correct to the 3 digit level, 4th digit incorrect
4. Correct to the group level, 2nd or 3rd digit incorrect
5. Incorrect at the group level.

The rate of incorrect coding of the E-code and the principal diagnosis was then estimated as the proportion of the sample falling into the 2nd, 3rd, 4th, or 5th categories above. Results include proportions correct and 95% confidence intervals calculated using Wilson's method following the recommendation of Newcombe.<sup>12</sup> For the diagnosis codes, all five digits are used to specify the diagnosis. For the E-code, the first four digits specify the mechanisms while the 5th digit indicates the location of injury event.

It was hypothesized that small hospitals would have a higher error rate than large hospitals on the grounds that for difficult cases coding clerks in small hospitals would have fewer peers with experience they could consult with. We defined large hospitals as those with a mean annual throughput for the study period of 3000 or more injury discharges. Hence, the large hospitals were those of Auckland, Starship, Middlemore, Waikato, Wellington, and Christchurch. The remaining hospitals were defined as small hospitals. A logistic regression was used to test this hypothesis with an outcome variable of correct/incorrect (categories 2–5 above). Principal diagnosis, grouped into six categories (Skull/Neck/Trunk Fracture, Extremity Fracture, Intracranial Injury, Open Wound, Other Effects/Complications, and Other), was added as an explanatory variable to control for differences in case mix.

**Table 2** Incorrect coding by diagnosis group

E-code	Diagnosis group	Total	% correct (95% CI)	% Correct to 4 digit level, 5th digit incorrect	% Correct to 3 digit level, 4th digit incorrect	% Correct to group level, 2nd or 3rd digit incorrect	% Incorrect at group level
820–829	Lower limb fracture	263	97 (95–99)	0	1	1	0
810–819	Upper limb fracture	228	96 (93–98)	2	1	0	0
996–999	Complications	222	94 (90–96)	0	1	2	3
850–854	Intracranial	175	93 (88–96)	0	0	6	2
960–979	Poisoning	108	92 (85–96)	0	2	4	3
870–897	Open wounds	204	93 (88–95)	1	3	2	1
800–805	Skull, neck/trunk fractures	144	98 (94–99)	0	1	0	1
	Other diagnoses	326	93 (90–96)	1	1	0	5
	Total	1670	95 (93–96)	1	1	2	2

**Table 3** Incorrect coding by E-code group

E-code group		Records	% Correct (95% CI)	% Error in 1st 4 digits (95% CI)	% Error in location code (5th digit) only (95% CI)
E880-888	Unintentional Falls	507	73 (69-77)	18 (15-22)	9 (7-11)
E916-928	Other accidents	370	75 (71-80)	19 (15-23)	6 (4-9)
E870-879	Medical injury	242	68 (62-74)	17 (13-22)	15 (11-20)
E930-949					
E810-819	Motor vehicle traffic crash	157	81 (74-86)	17 (12-23)	3 (1-6)
E800-807	Other transport	91	70 (60-79)	22 (15-32)	8 (4-15)
E820-848	Various	126	72 (64-79)	22 (16-30)	6 (3-11)
E960-969	Intentional Homicide/Assault	86	78 (68-85)	20 (13-29)	2 (1-8)
E950-959	Suicide/Self-inflicted	83	88 (79-93)	1 (0-7)	11 (6-19)

## RESULTS

Of the 1800 cases from 58 hospitals in the sample, 99 records could not be retrieved. Of the remaining 1701 a further 31 were discarded after coding either because they did not meet study criteria or the case was not recorded according to the study protocols. This left 1670 cases from 52 hospitals for analysis.

Five percent of the principal diagnoses, 18% of the first four digits of the E-codes, and 8% of the location codes (5th digit of the E-code), were incorrect (table 1).

Table 2 shows the level of incorrect coding by diagnosis group. The percentage of discharges coded correctly was greater than 90% for all diagnosis groups.

Table 3 shows the level of incorrect coding by E-code group. Discharges classified as intentional were more likely to be correct than discharges for unintentional injuries. Only Suicide/Self-inflicted injury had a substantially lower than average level of incorrect coding in the first four digits of the E-code. Medical injury had the highest level of incorrect coding of the location code.

There were no substantive differences in the level of incorrect coding (in any category) of principal diagnoses and E-codes between large and small hospitals. In large hospitals 44 (5.4%) of 821 principal diagnoses were incorrectly coded compared to 47 (5.5%) in 849 principal diagnoses from small hospitals, (OR = 0.97, 95% CI 0.63 to 1.48). Of the 821 E-codes from large hospitals, 217 (26.4%) were incorrectly coded whereas in small hospitals 215 (25.3%) of 849 E-codes were incorrectly coded, (OR = 1.06, 95% CI 0.85 to 1.32). Case mix appeared not to confound these relationships as the estimates were relatively unchanged when groupings of principal diagnoses were included in the models (principal diagnoses: OR = 0.99, 95% CI 0.65 to 1.52; E-codes: OR = 1.06, 95% CI 0.85 to 1.32).

## DISCUSSION

From an initial sample of 1800 there were 1670 available to be recoded. The major loss (n = 99) was attributable to the failure to locate the medical record. There are several possible explanations for this including: the record being in use, off-site (that is, it had been requested by another hospital), or incorrectly filed.

Our finding of 5% incorrect coding in principal diagnoses compares very favorably with that reported for the Victorian Inpatient Minimum Database (VIMD) in Australia at 19%.<sup>6</sup> The level of incorrect coding for E-codes, 18%, is similar to that reported for the VIMD, 16%<sup>6</sup> and the Washington State based study, 13%.<sup>5</sup> We have been unable to identify any comparable study that would enable us to compare the rate for location coding (8%).

Contrary to our hypothesis we did not find significant differences in the level of incorrect coding of E-codes between large and small hospitals.

There were two potential limitations of this study. First, some of the "incorrect" codes could have been due to mistakes by the accredited coder. It was our intention to have mismatches plus a sample of other cases reviewed by a second accredited coder. Unfortunately this was not possible as although there were many other experienced coders, there was only one other accredited ICD-9-CM-A coder in New Zealand and that person was not available to assist with the project. Moreover, at the time this study was undertaken, there were very few coders with formal training in ICD coding. Irrespective it would have been impractical, for cost reasons, to send a second coder out to the various hospitals to recode the mismatched codes, given the numbers involved.

A second limitation was that although our intention was to have the medical records reviewed blind to the original diagnosis and E-codes, this was not practical. It would have required a third party examining each record and removing any evidence as to how the discharge was coded (for example, the discharge summary sheet) before it was reviewed by the accredited coder. We are unable to estimate the proportion of records that would have had such evidence in them, but NZHIS advises that this would be very few (personal communication, T Vanderburg, 2004).

We conclude that users of the NMDS can have a high degree of confidence in the injury diagnoses coded under ICD-9-CM-A. A similar degree of confidence is warranted for E-coding at the group level (for example, fall), but not, in general, at higher levels of specificity (for example, type of fall). Starting in 1998 ICD-10 (ICD-10-AM)<sup>13</sup> was progressively introduced. There is thus no point in making recommendations based on the findings presented here for improving E-coding in New Zealand. Such a recommendation, if it is warranted, awaits the findings of our proposed replication of this study using ICD-10-AM coded discharges as its introduction represented a major change to the ICD coding structure, especially for external causes.<sup>14</sup>

Although these findings are particularly relevant to those using New Zealand data, they, together with the findings from comparable studies, are relevant to other countries. For those countries that are continuing to use ICD-9 they provide insight into potential problems of coding and thus guidance on where the focus of coder training should be placed. For those countries with historical data coded according to ICD-9 these findings suggest that some specific injury and external cause incidence estimates may need to be treated with more caution.

### Key points

- Diagnostic and external cause coding of all injuries resulting in inpatient hospital treatment is important for injury prevention.
- There are few published studies which might provide insight into the level of error to be expected for the full range of ICD-9 external cause codes and injury codes for large samples of hospital inpatients from the general population.
- For those countries that are continuing to use ICD-9 the findings provide guidance on potential problem areas.
- For those countries with historical data coded according to ICD-9 they suggest that some specific injury and external cause incidence estimates may need to be treated with more caution.

### ACKNOWLEDGEMENTS

The Health Research Council of New Zealand funded this research. The authors wish to thank Marie Gregan, Accredited NZHIS coder, Medical Record Staff throughout New Zealand, and the New Zealand Health Information Service for their assistance with this project. The helpful review of an earlier version of this paper by Colin Cryer is appreciated.

### Authors' affiliations

J Langley, S Stephenson, G Davie, Injury Prevention Research Unit, Department of Preventive and Social Medicine, University of Otago,

Dunedin, New Zealand

C Thorpe, Health Pac, Ministry of Health, Wellington, New Zealand

### REFERENCES

- 1 Langlois JA, Buecher JS, O'Connor EA, *et al*. Improving E coding of hospitalizations for injury: do hospital records contain adequate documentation? *Am J Public Health* 1995;**85**:1261-5.
- 2 Smith MW. Hospital discharge diagnoses: how accurate are they and their international classification of diseases (ICD) codes? *NZ Med J* 1989;**102**:507-8.
- 3 Fox KM, Reuland M, Hawkes WG, *et al*. Accuracy of medical records in hip fracture. *J Am Geriatr Soc Jun*, 1998;**46**:745-50.
- 4 Amoroso PJ, Smith GS, Bell NS. Qualitative assessment of cause-of-injury coding in US military hospitals: NATO Standardization Agreement (STANAG) 2050. *Am J Prev Med* 2000;**18**(Suppl 3):174-87.
- 5 LeMier M, Cummings P, West TA. Accuracy of external cause of injury codes reported in Washington State Hospital discharge records. *Inj Prev* 2001;**7**:334-8.
- 6 MacIntyre CR. Accuracy of ICD-9-CM codes in hospital morbidity data, Victoria: implications for public health research. *Aust N Z J Public Health* 1997;**21**:477-82.
- 7 Ministry of Health. National Minimum Dataset (Hospital Events). Wellington: New Zealand Health Information Service, 4 March, 2003.
- 8 New Zealand Health Information Service. Selected Morbidity Data for Private Hospitals 1995. Wellington: Ministry of Health, 1998.
- 9 New Zealand Health Information Service. Selected Morbidity Data for Publicly Funded Hospitals 1995/96. Wellington: Ministry of Health, 1997.
- 10 National Coding Centre. Australian Coding Standards for ICD-9-CM. In: *Faculty of Health Sciences*. Sydney: National Coding Centre, 1996.
- 11 Langley JD, Stephenson SCR, Cryer C, *et al*. Traps for the unwary in estimating person based injury incidence. *Inj Prev* 2002;**8**:332-7.
- 12 Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med* 1998;**17**:857-72.
- 13 National Centre for Classification in Health. ICD-10-AM Tabular List of Diseases. 1st edition. Australia: National Centre for Classification in Health, 1998.
- 14 Langley JD, Chalmers DJ. Coding the circumstances of injury: ICD-10 a step forward or backwards? *Inj Prev* 1999;**5**:247-53.