

QUT Digital Repository:
<http://eprints.qut.edu.au/>



McKenzie, Kirsten and Enraght-Moony, Emma L. and Walker, Susan M. and McClure, Roderick J. and Harrison, James E. (2008) Accuracy of External Cause of Injury Coding in Hospital Records: A Systematic Review. *Injury Prevention*.

© Copyright 2008 BMJ Publishing Group

Accuracy of External Cause of Injury Coding in Hospital Records: A Systematic Review

Kirsten McKenzie^{1*}, Emma Louise Enraght-Moony¹, Susan Mary Walker¹, Roderick John McClure², James Edward Harrison³

¹National Centre for Classification in Health, Queensland University of Technology, Brisbane, Queensland, Australia

²Monash University Accident Research Centre, Monash University, Melbourne, Victoria, Australia

³Research Centre for Injury Studies, Flinders University, Adelaide, South Australia, Australia

*Corresponding author:

National Centre for Classification in Health

School of Public Health, Queensland University of Technology,

KELVIN GROVE 4059, Queensland, Australia

Email: k.mckenzie@qut.edu.au

Phone: +61 7 3138 9753

Fax: +61 7 3138 5515

WORD COUNT (Full Text): 2992 words

WORD COUNT (Abstract): 236 words

NUMBER OF TABLES: 1

ABSTRACT

Objective: To appraise the published evidence regarding the accuracy of external cause of injury codes in hospital records.

Design: Systematic review

Data sources: Electronic databases searched included PubMed, PubMed Central, Medline, CINAHL, Academic Search Elite, Proquest Health and Medical Complete, and Google Scholar.

Snowballing strategies were employed by searching the bibliographies of retrieved references to identify relevant associated articles.

Selection criteria: Studies were included in the review if they assessed the accuracy of external cause of injury coding in hospital records via a recoding methodology.

Methods: The papers identified through the search were independently screened by two authors for inclusion. Due to heterogeneity between studies meta-analysis was not performed.

Results: There has been very limited research conducted examining the accuracy of external cause coding for injury-related hospitalization using medical record review and recoding methodologies, with only five studies matching the selection criteria. The accuracy of external cause coding using ICD-9-CM ranged from around 64% when examining exact code agreement, to around 85% when examining agreement for broader groups of codes.

Conclusions: While researchers may be able to use broad external cause groupings coded in ICD-9-CM with some confidence, for very specific codes researchers should exercise caution until further research is conducted to validate these data. As all previous studies have been conducted using ICD-9-CM, research is needed to quantify the accuracy of coding using ICD-10-AM, and validate the use of these data for injury surveillance purposes.

KEYWORDS

Injuries, Public health surveillance, Coding

Injuries are a significant cause of morbidity and mortality internationally, with the World Health Organization estimating that fatal injuries affect almost 6 million people worldwide, hospitalizations account for around 30 times as many deaths and emergency department presentations account for around 300 times as many deaths [1]. Hospital separations and mortality data are routinely used to monitor and assess injury causation and incidence, to inform injury research, policy and practice.

The ICD is the major system in use worldwide for the coding of morbidity data, and the ICD-10-AM is a modification of this classification used in all Australian hospitals [2]. The ICD-10-AM is used to assign alphanumeric codes to diagnoses, procedures, and external causes of injury recorded in patient medical records; to enable analysis and comparison of Australian morbidity data. In addition, the ICD-10-AM is used in eleven other countries worldwide and is being evaluated for use in an additional 16 countries.

Medical record reviews have a long history in the disease diagnosis area to validate and assess the accuracy of coding of different clinical diagnoses. There is considerable pressure for accurate diagnosis coding from those responsible for casemix funding and resource allocation, and from the clinical researchers using data for clinical categorization of diseases and epidemiological purposes. Hence, there has been an interest over many years and countries to assess the accuracy of diagnosis coding. A review by Williamson (2004), reported 129 published documents on accuracy in morbidity coding. Similarly, Campbell et al in a systematic review of diagnosis coding accuracy identified 30 studies in the United Kingdom alone [3, 4]. Despite the plethora of research on the accuracy of diagnosis coding, there has been very limited research conducted examining the accuracy of external cause of injury coding in hospital data [5-9]. Currently there is a lack of knowledge, understanding, and familiarity of researchers with using hospital data for injury surveillance, and very few injury researchers driving a program of quality assurance of these data.

A systematic review of the literature was conducted to appraise the available evidence on the accuracy of external cause coding in hospital records in Australia. This review provides evidence to evaluate the validity of national injury estimates which are based upon these morbidity data.

METHODS

Study Question

What is the accuracy of ICD external cause of injury coding in hospital admissions records in Australia?

Search Strategy

The following search phrase was used to search a range of databases and results were collated by two reviewers: ("external cause" OR e-code OR "e code") AND injury AND (quality OR validity OR reliability OR accuracy OR concordance OR consistency OR completeness OR documentation) AND (coding OR ICD) AND hospital AND (recod* OR abstract* OR review*). The text search was conducted in association with the following MeSH terms strategy: ((Medical Records/*classification) OR (International Classification of Diseases+)) AND (Wounds and Injuries/*classification OR Wounds and Injuries/*etiology) AND Documentation/standards AND Hospitals.

The databases which were searched included: PubMed, PubMed Central, Medline, CINAHL, Academic Search Elite, Proquest Health and Medical Complete, and Google Scholar. No time restrictions were included when searching these databases to ensure all articles indexed within each database were retrieved. In addition to the systematic keyword search approach snowballing strategies (i.e. following up on citations that emerge from other citations) were employed by searching the bibliographies, and citation links, of retrieved references to identify relevant associated articles. Grey literature searches were conducted to identify locally published reports and presentations. In addition, handsearching of the following key journals for articles on external cause data was conducted: Journal of Trauma, Injury and Infection Control, American Journal of Public Health, Australia & New Zealand Journal of Public Health, Injury Prevention.

Inclusion/exclusion criteria

The papers identified through the search were independently screened by two authors (KM and EEM) for inclusion. Studies were included in the review if they assessed the accuracy of external cause of injury coding in hospital records via a recoding methodology (n=5). Seventy-nine studies were excluded which were either: a) not recoding studies (i.e. epidemiological studies, data/policy recommendation reports (n=73) etc); b) studies not specifically focusing on community injuries (i.e. recoding studies on other clinical diagnoses, recoding studies on adverse events) (n=2).; and, c) studies where data was collected from emergency department records only (n=4). (Note: a large

number of irrelevant papers which were not recoding studies were returned in PubMed Central using this search phrase).

Synthesis of Study Results

Papers were reviewed and summarized in tabular and text form. Due to heterogeneity between studies meta-analysis was not performed.

RESULTS

There has been very limited research conducted examining the accuracy of external cause coding for injury-related hospitalization using medical record review and recoding methodologies. Only five studies were found that matched the selection criteria [5-8, 10]. The details of these five studies are summarized in Table 1.

Study Setting, Population and Study Design

All published studies of this nature have been conducted using hospital data coded using ICD-9-CM. No studies have been conducted on ICD-10-AM, which has been used in Australia since 1998. Three of the studies were undertaken in the United States [6, 7, 10], one in New Zealand [5], and one of these studies was undertaken in Australia [8].

The number of case records reviewed ranged from 323 cases to 1670 cases, with the range of data obtained from hospitalisations occurring in the years 1985 through to 1998. All except one of the studies [10] selected cases based on a principal diagnosis of an injury (ICD-9-CM Code range 800–999), whilst the remaining study selected cases based on the presence of an external cause code. Within the studies, a mixture of simple random sampling and stratified random sampling was used to select cases for review.

All of the studies utilized an independent coder to review and recode the selected medical records, with three of the studies specifically stated that attempts were made to blind the reviewer to the original codes [5, 7, 10]. Only one of the studies stated that additional information was abstracted from the medical record in addition to the recoding task, with a narrative description of the cause of injury and place of occurrence recorded separately for each form from the medical record [6].

‘Accuracy’ Measures and Statistical Analysis

Accuracy of coding was largely operationalised as the concordance/agreement between the original codes and the recoded data. Each of these studies examined accuracy in terms of levels of agreement, including complete external cause code agreement, agreement to the 4th digit ICD code,

agreement to the 3rd digit ICD code, agreement to the group level and disagreement being the main ‘accuracy’ categories used. Differences in the assignment of intent and/or mechanism were explored in four out of five of these studies to further explore where the differences in coding patterns could be identified [5-7, 10].

Statistical analysis was largely descriptive, showing percentage agreement of coding. Two studies attempted to identify correlates of coding accuracy using logistic regression, to identify whether certain characteristics (such as hospital size, length of stay, patient age etc) correlated with a higher likelihood of being assigned a different code to the original code [5, 8].

Study Findings

Studies examining external cause coding accuracy found that percentage agreement between coders ranged between 59% when examining very specific code assignment to 95% when examining broad category assignment (See Table 1).

The studies that evaluated the accuracy of the complete external cause code, reported an average percentage agreement of 64% (59% [10], 66% [7], and 67% [6]). Where accuracy was examined to the 4th digit ICD code level (with errors in the 5th digit); the percentage agreement of coders was reported as 82% agreement by both Langley et al and Langlois et al[5, 6]. Additionally, Langley et al’s data showed 85% agreement to the 3rd digit ICD code.

The studies which examined percentage agreement of coding by code block found variable results across the different code blocks. LeMeier et al examined the accuracy of coding by external cause mechanism (i.e. the degree to which coders agreed on the way in which the injury was sustained), and found 87% agreement in mechanism of injury [7]. Both LeMeier et al and Smith et al examined the agreement of coders in terms of the intent (i.e. unintentional, intentional self harm, assault) and found the percentage agreement in coding of intent was 95% and 86% respectively [7, 10]. In terms of the accuracy of codes for unintentional falls, code agreement was on average 70% (66% [7], 73% [5]). Motor vehicle traffic crashes were reported as having an agreement in code assignment of between 63% and 81% [5, 7]. Finally, Langley et al examined the extent of coder agreement within the intent blocks of intentional self-harm and assault, and found percentage agreement within these code blocks of 83% and 86% respectively [5].

MacIntyre et al examined the types of errors in external cause code assignment and identified three categories of errors: errors of omission (i.e. missing external causes); superfluous external cause

codes (i.e. unnecessary codes); and discrepant external cause codes (i.e. those were coders did not agree on code assignment as traditionally examined in recoding studies) [8]. They found that errors of omission accounted for 21% of errors identified; superfluous external cause codes accounted for 11% of errors identified; and discrepant external cause codes comprised 68% of errors identified.

Two studies examined correlates of coding accuracy using logistic regression. The first study examined the size of the hospital as a correlate of external cause coding accuracy, controlling for the principal injury type [5]. The second study examined several correlates of external cause coding accuracy, including: whether the admission was an emergency admission;, length of stay; number of diagnoses and procedures; type of injury; patients age; hospital; and mortality outcomes [8]. Both studies found none of the factors which were examined showed any significant correlation with coding accuracy [5, 8].

Table 1: Summary of Previous Research Using Medical Record Recoding Methods to Examine the Accuracy of External Cause Coding in Injury-Related Hospital Records

Study	Aim	Setting and Sample Period	Sample size and selection	Method	Accuracy Measures	Statistical analysis	Results
Langley et al, 2006	Determine level of coding accuracy for injury PDx and external cause code in hospital discharges.	New Zealand public hospitals from 1996-1998	1670 cases from 52 hospitals; Simple random sample of cases with injury as PDx.	Independent expert coder recoded medical records mostly blinded to original codes.	Accuracy categories: 1. Correct 2. Correct to 4 digit 3. Correct to 3 rd digit 4. Correct to group level 5. Incorrect group level	% correct; Logistic regression: IV - size of hospital Controlled - Injury PDx DV – correct/incorrect.	82% external cause coded data overall correct to 4 digit level (18% error). Unintentional % correct: 73% Falls 75% Other Injuries 81% MVTC Intentional % correct 86% Assault 83% Suicide Log reg: No difference in level of coding accuracy by size of hospital. Correct ranking of leading causes of injury. Incident rate ratios: 20% of other/unspec could have had mechanism coded; 48% underreporting of undetermined intent miscoded as unintentional; Agreement: 87% for mechanism 95% for intent 66% for complete e-code (34% error). Sensitivity and PPV: 95% and 93% for falls 97% and 88% for MVTC 99% and 94% for poisonings 92% and 94% for firearms.
LeMeier et al, 2001	Evaluate accuracy of external cause of injury codes in hospital discharges.	Washington State (USA) civilian hospitals from 1996	1260 cases from 32 hospitals; 32 hospitals selected which accounted for 80% of injury hospitalisations. Simple random sample of cases with injury as PDx.	Independent expert coder recoded medical records blinded to original codes.	Accuracy categories: 1. Correct 2. Correct to 4 digit 3. Correct to 3 rd digit 4. Correct to group level 5. Incorrect group level	Estimated counts and incidence ratios by mechanism; Sensitivity and PPV; Interrater agreement and kappa. Calculations used sample weight and population corrections where appropriate.	84% consistent external cause code assignment (16% error). Types of errors: 21% error of omission 11% superfluous e-codes 68% discrepant e-codes. Log reg: No significant predictor of error in e-codes.
MacIntyre et al, 1997	Ascertain the reliability of injury data in hospital discharges.	Victoria (Australia) public hospitals from 1994-1995	546 cases from 4 hospitals; 4 hospitals selected which accounted for 25% of injury hospitalisations. Random sample of cases with injury as PDx and an external cause coded.	Medical records reviewed and recoded by physician with knowledge of ICD-9-CM coding, with input from coder.	‘Consistency’ or ‘discrepancy’ between original codes and medical record content. Categories of ‘error’: 1. Errors of omission 2. Superfluous coding 3. Discrepancy.	Overall discrepancy rates; Diagnosis specific discrepancy rates; Logistic regression: IV – emergency admission, LOS, No. of diagnoses and procedures, type of injury, pt age, hospital, mortality. DV – error in e-code assigned.	84% consistent external cause code assignment (16% error). Types of errors: 21% error of omission 11% superfluous e-codes 68% discrepant e-codes. Log reg: No significant predictor of error in e-codes.

Study	Aim	Setting and Sample Period	Sample size and selection	Method	Accuracy Measures	Statistical analysis	Results
Langlois et al, 1995	Evaluate the quality and availability of cause of injury information in medical records and extent to which inadequate documentation contributes to incomplete e-coding.	Rhode Island (USA) hospitals from 1988-1990	1440 cases from unstated number of hospitals; Stratified random sample (based on specificity of e-code) of cases with injury as PDx from all hospitals.	Medical records reviewed and recoded by two independent expert coders. Two researchers reviewed codes and assigned final e-code and cases of disagreement were arbitrated by the first author. Narrative description of cause of injury, place of occurrence, and details regarding which health professional recorded the info (eg physician, nurse) recorded separately for each form in record (eg ambulance form, ED form, history).	Level of agreement of original codes and recoded data. Agreement levels: 1. Complete agreement to the 4 th digit 2. Agreement to the third digit 3. Agreement to the section 4. Disagreement. Level of documentation from specific forms assessed by comparing codes assigned from each form to codes assigned for complete record.	% agreement.	82% agreement for external cause codes where a specific external cause code assigned originally. Of 18% cases where disagreement present, 64% due to difference in intent. 70% of cases originally coded with a vague/unspecified e-code and 66% of cases where no e-code assigned originally were assigned a specific e-code on recode. Medical record documentation was sufficient to assign a specific e-code in 78% of cases and adequate for a vague e-code assignment in an additional 10% of cases. % of cases with specific cause info in different forms: 51% History form 42% ED record 40% Discharge summary.
Smith et al, 1990	Evaluate the usefulness of ICD external cause codes for injury surveillance.	Indian Health Service (USA) hospitals from 1985	323 cases from 2 hospitals	Medical records reviewed, cause of injury information abstracted, and external cause recoded by an independent coder who was blinded to codes originally assigned.	Agreement levels: 1. Complete agreement to the 3 rd and 4 th digit 2. Agreement to external cause group level 3. Agreement by intent.	% agreement.	63% agreement to the 3 digit ICD code level. 81% agreement to the external cause code group level. 86% agreement a the level of intent.

Discussion

One of the most notable findings of this systematic review was the considerable lack of research examining the accuracy of external cause of injury coding in hospital records; only five papers meeting the inclusion criteria for this review. Whilst these data are used routinely to monitor and assess injury causation and incidence, to develop burden of disease estimates, and to inform injury research, policy and practice, there is currently a limited empirical basis to validate the quality of these data.

This review showed that the accuracy of external cause coding using ICD-9-CM ranges from around 64% when examining exact code agreement, to around 85% when examining agreement to the three digit level. Differences in coding accuracy were evident when examining different external cause axes and code blocks. That is, agreement levels differed depending on whether the intent was deemed to be intentional or unintentional, and depending on what the mechanism causing the injury was (i.e. motor vehicle crash, fall etc). Thus, while researchers examining data coded from countries using ICD-9-CM may be able to use broad external cause code blocks with some level of confidence, for very specific code blocks researchers should exercise caution until further research is conducted to validate these data.

As all previous studies have been conducted using ICD-9-CM, urgent research is needed to quantify the accuracy of external cause coding using ICD-10 (and the clinical variants of the ICD-10 such as ICD-10-AM, ICD-10-CM, and ICD-10-CA), and validate the use of these data for injury surveillance purposes. ICD-10 external cause codes are considerably different in terms of structure to ICD-9-CM codes, and the clinical variations of ICD-10 provide additional codes for place of injury and activity at the time of the injury as well as increased levels of specificity across code blocks. As a consequence, it is to be expected that the accuracy of coding under the ICD-10 classification system would vary from that which was coded under the ICD-9-CM classification system.

Key Messages

What is already known on this subject?

- Injuries are a significant cause of morbidity and mortality in the Australian population.
- Hospital separations and mortality data are routinely used to monitor and assess injury causation and incidence, to inform injury research, policy and practice.

What this study adds?

- There is currently a limited empirical basis to validate these data with only five studies identified internationally which examine the accuracy of external cause data in hospital separation data (using ICD-9-CM which has been superseded in many countries by ICD-10).
- The accuracy of external cause coding using ICD-9-CM ranges from around 64% when examining exact code agreement, to around 85% when examining agreement for broader groups of codes.
- While researchers may be able to use broad external cause code blocks with some level of confidence, for very specific external cause codes researchers should exercise caution until further research is conducted to validate these data using the current version of ICD.

Competing Interests

No competing interests.

Funding

This research is funded by an Australian Research Council Linkage Project grant, Injury Prevention and Control Australia, the Victorian Department of Human Services, and the Queensland Health - Health Information Centre.

CONTRIBUTORSHIP

Author 1 (McKenzie) contributed to the conceptual design of the manuscript, and was responsible for conducting the systematic literature review, writing the first draft of the manuscript, compiling all authors responses, and preparing the final version of the manuscript. The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence (or non exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd and its Licensees to permit this article (if accepted) to be published in IP and any other BMJPGGL products to exploit all subsidiary rights, as set out in our licence

<http://ip.bmj.com/ifora/licence.pdf>.

Author 2 (Enraght-Moony) contributed to the conceptual design of the manuscript, assisted with the systematic literature review, and reviewed and commented on each draft of the manuscript.

Author 3 (Walker) contributed to the conceptual design of the manuscript, provided context to the manuscript in terms of clinical coding processes, and reviewed and commented on each draft of the manuscript.

Author 4 (McClure) contributed to the conceptual design of the manuscript, provided context to the manuscript in terms of injury prevention implications, and reviewed and commented on each draft of the manuscript.

Author 5 (Harrison) contributed to the conceptual design of the manuscript, and provided context to the manuscript in terms of injury surveillance implications.

References

- [1]Holder Y, Peden M, Krug E, Lund J, Gururaj G, Kobusingye O, eds. Injury Surveillance Guidelines. Geneva: World Health Organisation 2001.
- [2]Walker S, McEvoy S. Injury classification systems. In: McClure R, Stevenson M, McEvoy S, eds. *The Scientific Basis of Injury Prevention Control*. Melbourne: IP Communications 2004:51-61.
- [3]Campbell SE, Campbell MK, Grimshaw JM, Walker AE. A systematic review of discharge coding accuracy. *Journal Of Public Health Medicine*. 2001 2001/09//;23(3):205-11.
- [4]Williamson D. Clinical coding audits: an annotated bibliography. *Health Information Management*. 2004;33(1):21-7.
- [5]Langley J, Stephenson S, Thorpe C, Davie G. Accuracy of injury coding under ICD-9 for New Zealand public hospital discharges
10.1136/ip.2005.010173. *Inj Prev*. 2006 February 1, 2006;12(1):58-61.
- [6]Langlois JA, Buechner JS, O'Connor EA, Nacar EQ, Smith GS. Improving the E coding of hospitalizations for injury: do hospital records contain adequate documentation? *Am J Public Health*. 1995 Sep;85(9):1261-5.
- [7]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.
- [8]MacIntyre CR, Ackland MJ, Chandraraj EJ. Accuracy of injury coding in Victorian hospital morbidity data. *Aust N Z J Public Health*. 1997;27(7):779-83.
- [9]McKenzie K, Harding L, Walker S, Harrison J, Enraght-Moony E, Waller G. The quality of cause-of-injury data: Where hospital records fall down. *Australian and New Zealand Journal of Public Health*. 2006;30(6):509-13.
- [10] Smith SM, Colwell LS, Sniezek JE. An evaluation of external cause-of-injury codes using hospital records from the Indian Health Service, 1985. *Am J Public Health*. 1990;80(3)(279-281).