

# Emergency department coding of bicycle and pedestrian injuries during the transition from ICD-9 to ICD-10

M Karkhaneh,<sup>1</sup> B E Hagel,<sup>2,3</sup> A Couperthwaite,<sup>4</sup> L D Saunders,<sup>1</sup> D C Voaklander,<sup>1,5</sup> B H Rowe<sup>1,4</sup>

<sup>1</sup>School of Public Health, University of Alberta, Edmonton, Alberta, Canada

<sup>2</sup>Department of Pediatrics, University of Calgary, Calgary, Alberta, Canada

<sup>3</sup>Department of Community Health Sciences, University of Calgary, Calgary, Alberta, Canada

<sup>4</sup>Department of Emergency Medicine, University of Alberta, Edmonton, Alberta, Canada

<sup>5</sup>Alberta Centre for Injury Control and Research, Edmonton, Alberta, Canada

## Correspondence to

Brian H Rowe, Department of Emergency Medicine, University of Alberta, 1G1.43 Walter C. Mackenzie Centre, 8440-112th Street, Edmonton, Alberta T6L 2B7, Canada; brian.rowe@ualberta.ca

Accepted 25 May 2011

Published Online First

24 June 2011

## ABSTRACT

**Background** The international classification of diseases version 10 (ICD-10) uses alphanumeric expanded codes and external cause of injury codes (E-codes).

**Objective** To examine the reliability and validity of emergency department (ED) coders in applying E-codes in ICD-9 and -10.

**Methods** Bicycle and pedestrian injuries were identified from the ED information system from one period before and two periods after transition from ICD-9 to -10 coding. Overall, 180 randomly selected bicycle and pedestrian injury charts were reviewed as the reference standard (RS). Original E-codes assigned by ED coders (ICD-9 in 2001 and ICD-10 in 2004 and 2007) were compared with charts (validity) and also to ICD-9 and -10 codes assigned from RS chart review, to each case by an independent (IND) coder (reliability). Sensitivity, specificity, simple, and chance-corrected agreements ( $\kappa$  statistics) were calculated.

**Results** Sensitivity of E-coding bicycle injuries by the IND coder in comparison with the RS ranged from 95.1% (95% CI 86.3 to 99.0) to 100% (95% CI 94.0 to 100.0) for both ICD-9 and -10. Sensitivity of ED coders in E-coding bicycle injuries ranged from 90.2% (95% CI 79.8 to 96.3) to 96.7% (95% CI 88.5 to 99.6). The sensitivity estimates for the IND coder ranged from 25.0% (95% CI 14.7 to 37.9) to 45.0% (95% CI 32.1 to 58.4) for pedestrian injuries for both ICD-9 and -10.

**Conclusion** Bicycle injuries are coded in a reliable and valid manner; however, pedestrian injuries are often miscoded as falls. These results have important implications for injury surveillance research.

## BACKGROUND

Edmonton is one of the two largest cities in Alberta, with a population of 1 034 945 (city = 730 372; other metropolitan areas = 304 573) (reported by Statistics Canada, 2006). Four periods of a biannual Canadian community health survey (2001–07) demonstrated that the prevalence of recreational bicycle use in Alberta among youth (12–17 years of age) was in the range 58–65% and the mean number of times adolescents bicycled in the past 3 months was in the range 16–30%. The prevalence of recreational bicycling for adults (18+) was 24–28% and mean number of times adults bicycled was 17–19 in the past 3 months.<sup>1</sup> In the same study the prevalence of commuting bicycle use among youth in Alberta was 31–35% and among adults was 6–7%.<sup>1</sup>

A nationwide Canadian study demonstrated that 2% of all hospitalisations were due to bicycle-related injuries during a 10-year period (1994–2004).<sup>2</sup> A 5-year study (1991–95) in British Columbia demonstrated that 4% of visits of all children (1–19 years of age) to the emergency department (ED) resulted from bicycle-related injuries.<sup>3</sup>

In order to conduct surveillance studies, we require valid coding of patient's diseases and circumstances of the event leading to admission to EDs or hospitals. Inpatient coding has been established in hospitals for use in disease and mortality surveillance, epidemiological studies, billing and financial planning, and policy analyses.<sup>4 5</sup>

The International Classification of Diseases versions 9 (ICD-9) and 10 (ICD-10) have been used to code diseases and other health problems recorded on many types of health and vital records, including death certificates, hospitalisation and ED data. The ICD-10 classification is the latest in a series which has its origins in the 1850s.<sup>6</sup> ICD-10 was endorsed by the 43rd WHO assembly in May 1990 and began implementation in WHO member states in 1994.<sup>7</sup> The differences between ICD-9 and ICD-10 are substantial, not only in disease classification, but also in coding rules. As the ICD-9 system has been used by many hospitals and clinics for years and is still used in many US centres, this transition introduced some challenges for long-term and comparative studies.<sup>8</sup>

ICD-9 diagnosis codes consist of the following: 3-digit numeric characters (001–999), with two decimals representing illnesses and conditions; alphanumeric E-codes (E000–E999), describing external causes of injuries, poisonings, and adverse effects; and V codes (V01–V82), describing factors influencing health status and contact with health services. ICD-10 uses 3-digit alphanumeric codes (A00–Z99) with two decimals.<sup>7 9</sup> There are many other changes in ICD-10 that have been described in detail elsewhere.<sup>10</sup> Canada implemented ICD-10 for the classification of cause of death, beginning in 2000.<sup>8</sup> In an agreement with WHO, Canada adopted an enhanced version of ICD-10-CA by keeping the main structure of ICD-10, yet including more subgroup definitions using a third decimal and introducing the first Canadian classification of intervention.<sup>11</sup>

E-codes have been widely used in surveillance system for mortality in traffic-related injuries<sup>12</sup> and morbidity of bicyclists, pedestrians, and those with sport and recreational related injuries<sup>13–15</sup>;



This paper is freely available online under the BMJ Journals unlocked scheme, see <http://injuryprevention.bmj.com/site/about/unlocked.xhtml>

however, coding issues might have led to errors in the interpretation of research findings.<sup>16</sup> Appropriate coding by ICD-9 and ICD-10 has always been an important issue for health surveillance and health services research. Many studies have evaluated the validity and/or reliability of ICD-9 coding for external causes of injuries (from now on called E-coding for both ICD-9 and ICD-10),<sup>5 17–21</sup> or focused on principal diagnosis.<sup>22</sup> Other studies evaluated the validity/reliability of ICD-10 for E-coding<sup>23 24</sup> or only principal diagnosis in ICD-10.<sup>25</sup>

For those countries that implemented ICD-10, the transition from ICD-9 to ICD-10 may have had an impact on the trends of causes of injuries. Bridge coding studies, to date, have evaluated the impact of a coding change by focusing on principal causes of mortality and not external cause of injury.<sup>8 10 26–30</sup> One study examined the usefulness of ICD-10-CM in capturing public health diseases (reportable diseases, leading cause of death and morbidity/mortality related to terrorism) and reported agreement levels of coders when coding such diseases in ICD-9-CM and ICD-10-CM. They found that ICD-10-CM was more specific and fully captured more diseases than ICD-9-CM; however, coders were more consistent in coding ICD-9-CM than ICD-10-CM.<sup>31</sup>

A long-term surveillance study in Alberta, Canada has shown that transition from ICD-9-CM to ICD-10-CA appeared to cause a decrease in the number of motor-vehicle-related deaths/hospital admissions, with a smaller impact on motor-vehicle ED visits.<sup>32</sup> Similar studies have demonstrated that transition from ICD-9 to ICD-10 can affect ranking of causes of death,<sup>29</sup> possibly resulting in a decrease in diseases such as pneumonia or an increase in cerebrovascular diseases.<sup>30</sup>

In Alberta, ICD-10-CA codes were implemented on 1 January 2000 for deaths and 1 April 2002 for morbidity data (hospitalisation and ED records).<sup>32</sup> Concurrently, the Alberta government implemented a law mandating bicyclists <18 years of age to wear helmets, effective 1 May 2002.<sup>33</sup> Given the timing of the bicycle helmet legislation and the coding change, it was essential to investigate whether the coding transition may have influenced the overall incidence of cycling-related injuries independent of the legislation. Therefore, the aim of this study was to evaluate the reliability and validity of ED coders in applying ICD-9-CM and ICD-10-CA external cause of injury codes for bicyclists.

In our study we used pedestrian injuries to establish how coding changes affected injury trends in another vulnerable road user group not affected by bicycle helmet legislation.

## METHODS

### Case selection

We identified all cycling- and pedestrian-related injuries from the Emergency Department Information System (EDIS) software.<sup>34</sup> This system captures data on all patients presenting to the ED, including patient demographics, illness and circumstances of injury, times of arrival and care, injury descriptions, symptoms, consultations, and triage/vital signs assessment. Using the patient's complete paper chart, medical record nosologists (henceforth referred to as ED coders) assign ICD-9-CM before or ICD-10-CA, after 1 April 2002, after reviewing physician-assigned diagnoses at the time of ED discharge (home or hospital). EDIS review and case selection was performed for the four busiest cycling months of the year (May to August) in three separate years (2001=pre-transition to ICD-10; 2004; 2007=post-transition to ICD-10).

On the basis of ICD-9 and ICD-10 E-code descriptions,<sup>7</sup> bicycle and pedestrian injuries were defined (see appendix) and

used by investigators to identify all cases from the EDIS database. The keywords from these definitions were used for searching cases of bicycle and pedestrian injuries admitted to the EDs. Keywords for bicycle injuries included: bike, biking, cycle, bicycle, bicycling, bike injuries, cycle injuries, bicycle injuries, biking, and tricycle. Keywords for pedestrian injuries included: pedestrian, walking, jogging, car-ped, side walk, curb, cross walk, hit by (bicycle, motorcycle, car, or bus), ran over, parking lot. A variety of misspellings of bicycle (eg, bik, bicycl) and pedestrian terms (e.g., wlk, jogin) were also used to make sure we have not missed any cases due to typing mistakes.

After retrieving all relevant cases, two separate pools of bicycle and pedestrian injuries were prepared from adjudication with senior nursing staff (making sure they were valid bicycle and pedestrian injuries); research assistants randomly selected and reviewed 180 bicyclist and 180 pedestrian presentations (360 in total) from three hospital EDs in Edmonton (University of Alberta Hospital, Stollery Children's Hospital, and North East Community Health Center). Our sample included 60 injured cyclists and 60 injured pedestrians in each year.

A specific data extraction form was designed to capture necessary information from patients' paper charts. Using the extracted information, an independent expert coder (IND coder) was employed to assign both ICD-9-CM and ICD-10-CA codes. The IND coder was not aware of any previous coding associated with a bicyclist or pedestrian injury, nor the study hypothesis. After providing both ICD-9-CM and ICD-10-CA codes for each case, we merged these data with administrative data from the ambulatory care classification system (ACCS), a central electronic database for diagnosis, procedure, healthcare utilisation, and follow-up of emergency department patients in Alberta, Canada which was originally produced by ED coders. Therefore, each case had an ICD-9-CM code (before 1 April 2002) or an ICD-10-CA code (after 1 April 2002) assigned by ED coders as usual practice, forming part of the electronic administrative health record, with ICD-9-CM and ICD-10-CA codes assigned by the IND coder.

### Analysis

Data were analysed using Stata IC V.11.<sup>35</sup> Examining validity, we calculated sensitivity, with 95% CIs, as the proportion of all cycling injuries identified through our chart review (reference standard) that were similarly coded as bicycle injuries by the ED and IND coders. Specificity was calculated as the proportion of all no bicycle injuries in our charts that were not coded as bicycle injuries by ED or IND coders. Similar sensitivity and specificity estimates and 95% CIs were produced for ED (ACCS data) and IND coders for pedestrians.

Simple percentage agreement between the two coders was calculated. Since simple percentage agreement does not account for agreement by chance, we used Cohen's  $\kappa$  statistic, a measure of chance-corrected proportional agreement.<sup>36</sup>  $\kappa$  agreement was defined a priori as almost perfect (0.81–1.0), substantial (0.61–0.8), moderate (0.41–0.60), fair (0.21–0.40), slight (0.0–0.20), or poor (<0.0).<sup>37</sup>

We constructed separate 2×2 tables for ICD-9-CM and ICD-10-CA by year. We calculated percentage agreement and  $\kappa$  for coding between the ED coders and the IND coder. For sensitivity and agreement analysis, pedestrians were used as negative cases for bicyclists and vice versa.

After we finished our analyses and on the basis of our reference standard medical chart reviews, many of the pedestrian injuries were not E-coded accurately; therefore, we decided to

perform a post-hoc investigation for those misclassified E-codes among pedestrian injuries.

### Sample size

We based our sample size on sensitivity, or the proportion of all EDIS identified cycling injuries in Edmonton transferred to ACCS. Our focus was on estimation (CIs) rather than statistical testing. For a CI width of 15% ( $\pm 7.5\%$ ), assuming a worst case of 50% sensitivity, we would require 171 subjects. Therefore, with 180 subjects, the 95% CI around the estimate of sensitivity was expected to be less than 10%.

### Ethics

We obtained ethical approval from the University of Alberta Health Research Ethics Board. Patients were not contacted during this study.

## RESULTS

### Validity of E-coding by ED and IND coder

Sensitivity of E-coding bicycle injuries by ED coders in comparison to the reference standard (RS) ranged from 90.2% (95% CI 79.8 to 96.3) in 2007 to 96.7% (95% CI 88.5 to 99.6) in both 2001 and 2004 (table 1, bicycle injuries). Sensitivity of E-coding bicycle injuries by the IND coder in comparison to the RS ranged from 95.1% (95% CI 86.3 to 99.0) in 2007 to 100% (95% CI 94.0 to 100) in 2001 (table 1, bicycle injuries).

Sensitivity of E-coding pedestrian injuries by ED coders in comparison to the RS ranged from 25.0% (95% CI 14.7 to 37.9) in 2004 to 38.3% (95% CI 26.1 to 51.8) in 2001. The sensitivity estimates for the IND coder in coding pedestrian injuries compared with the RS ranged from 30.0% (95% CI 18.8 to 43.2) in 2004 to 43.3% (95% CI 30.6 to 56.8) in 2001 (table 1, pedestrian injuries).

Specificities for bicycle injuries were 98.3–100% and for pedestrian injuries were all 100% (not presented in table 1).

### Validity of E-codes in ICD-10 and ICD-9

The results of the validity analysis showed that sensitivity of E-coding bicycle injuries by the IND coder using ICD-10 for the pre-transition year of 2001 was 98.3% (95% CI 91.1 to 100); sensitivity for ICD-9 for post-transition was 98.3% (95% CI 91.1 to 100) in 2004 and 96.7% (95% CI 88.7 to 99.6) in 2007 (table 1, shaded rows bicycle injuries). Sensitivity of E-coding for pedestrian injuries by the IND coder using ICD-10 for the pre-transition year (2001) was 45% (95% CI 32.1 to 58.4); retesting results of ICD-9 for post-transition were 25.0% (95% CI 14.7 to 37.9) in 2004 and 37.3% (95% CI 25.0 to 59.0) in 2007 (table 1, shaded rows pedestrian injuries).

### Reliability of E-coding between ED and IND coders

Examining chance-corrected agreement ( $\kappa$ ) and applying Landis and Koch's<sup>37</sup> ranking of  $\kappa$ , agreement between ED coders and the IND coder for bicycle injuries was almost perfect, ranging between 0.88 and 0.97 ( $\kappa_{\text{pooled}}=0.94$ ; 95% CI 0.91 to 0.98). Similarly, almost perfect agreement was seen in the comparison of ED coders to the IND coder for pedestrian injuries ranging between 0.90 and 0.97 ( $\kappa_{\text{pooled}}=0.92$ ; 95% CI 0.87 to 0.98) (table 2).

### Post-hoc results for pedestrian E-coding

Approximately 3.4% of pedestrian injuries that we identified and confirmed through EDIS and chart review were not assigned an external cause of injury by the ED coder. ED coders also misclassified between 59% (ICD-9-CM) and 69% (ICD-10-CA)

**Table 1** Sensitivity of emergency department (ED) and independent (IND) coder in comparison to reference standard (medical charts) in coding external cause of bicycle and pedestrian injuries in the three EDs in Edmonton, Alberta, Canada

Activity	Reference standard (Medical charts)				Sensitivity* % (95% CI)
	a†	b	c	d	
<b>Bicycle injuries</b>					
ED coder 2001 (ICD-9)	58	0	2	60	96.7 (88.5 to 99.6)
ED coder 2004 (ICD-10)	58	0	2	60	96.7 (88.5 to 99.6)
ED coder 2007 (ICD-10)	55	0	6	59	90.2 (79.8 to 96.3)
<b>Overall</b>	<b>171</b>	<b>0</b>	<b>10</b>	<b>179</b>	<b>94.5 (90.1 to 97.3)</b>
IND coder 2001 (ICD-9)	60	0	0	60	100.0 (94.0 to 100)
IND coder 2001 (ICD-10)‡	59	0	1	60	98.3 (91.1 to 100)
IND coder 2004 (ICD-10)	59	1	1	59	98.3 (91.1 to 100)
IND coder 2004 (ICD-9)‡	59	1	1	59	98.3 (91.1 to 100)
IND coder 2007 (ICD-10)	58	0	3	59	95.1 (86.3 to 99.0)
IND coder 2007 (ICD-9)‡	59	0	2	59	96.7 (88.7 to 99.6)
<b>Overall</b>	<b>180</b>	<b>1</b>	<b>1</b>	<b>178</b>	<b>99.4 (97.0 to 100)</b>
<b>Pedestrian injuries</b>					
ED coder 2001 (ICD-9)	23	0	37	60	38.3 (26.1 to 51.8)
ED coder 2004 (ICD-10)	15	0	45	60	25.0 (14.7 to 37.9)
ED coder 2007 (ICD-10)	21	0	38	61	35.6 (23.6 to 49.1)
<b>Overall</b>	<b>59</b>	<b>0</b>	<b>120</b>	<b>181</b>	<b>33.0 (26.1 to 40.4)</b>
IND coder 2001 (ICD-9)	26	0	34	60	43.3 (30.6 to 56.8)
IND coder 2001 (ICD-10)‡	27	0	33	60	45.0 (32.1 to 58.4)
IND coder 2004 (ICD-10)	18	0	42	60	30.0 (18.8 to 43.2)
IND coder 2004 (ICD-9)‡	15	0	45	60	25.0 (14.7 to 37.9)
IND coder 2007 (ICD-10)	22	0	37	61	37.3 (25.0 to 50.9)
IND coder 2007 (ICD-9)‡	22	0	37	61	37.3 (25.0 to 50.9)
<b>Overall</b>	<b>67</b>	<b>0</b>	<b>112</b>	<b>181</b>	<b>37.4 (30.3 to 45.0)</b>

\* Sensitivity= $a/(a+c)$

† a: chart (+) coder (+) or true positive; b: chart (-) coder (+) or false positive; c: chart (+) coder (-) or false negative; d: chart (-) coder (-) or true negative

‡ Double coding

ED, emergency department; IND, independent

of pedestrian injuries. Of the 59% misclassified pedestrian injuries in ICD-9-CM, 70% were miscoded as falls, 18% as unspecified, and 3% as unspecified vehicle collision; 3% overexertion, and 6% other. Of the 69% misclassified pedestrian injuries in ICD-10-CA, 74% were miscoded as falls, 24% as overexertion, 1% as bitten dog, and 1% as striking stationary object; 5% had no E-code (not shown in figure 1). Missing or misclassified bicycling injuries did not exceed 4% (figure 1).

For the IND coder (who independently coded all bike and pedestrian injuries by ICD-9-CM and ICD-10-CA) there were no missing E-codes for pedestrian injuries; however, approximately 63% of the records were misclassified. Of these, 58% were misclassified as falls for both ICD-9-CM and ICD-10-CA codes. The IND coder had 1.1% missing codes (in ICD-10 CA) for bicycle injuries and 2.2% and 1.7% misclassification for ICD-9-CM and ICD-10-CA, respectively (figure 2).

## DISCUSSION

This study evaluated the reliability of ED coders in E-coding bicycle injuries in three Canadian EDs. We also studied the

**Table 2** Simple and chance-corrected agreements between emergency department (ED) and independent (IND) coders (head to head) in coding external cause of bicycle and pedestrian injuries in the three EDs in Edmonton, Alberta, Canada

Activity	IND coder				Agreement * %	κ statistics† Point estimate (95% CI)
	a‡	b	c	d		
Bicycle injuries						
ED coder 2001 (ICD-9)	58	0	2	60	98.3	0.97 (0.92 to 1.00)
ED coder 2004 (ICD-10)	57	1	3	59	96.7	0.93 (0.87 to 0.99)
ED coder 2007 (ICD-10)	53	2	5	60	94.2	0.88 (0.80 to 0.97)
Total and pooled κ	171	0	10	179	97.2	0.94 (0.91 to 0.98)
Pedestrian injuries						
ED coder 2001 (ICD-9)	23	0	3	94	97.5	0.92 (0.84 to 1.00)
ED coder 2004 (ICD-10)	15	0	3	102	97.5	0.90 (0.78 to 1.00)
ED coder 2007 (ICD-10)	21	0	1	98	99.2	0.97 (0.92 to 1.00)
Total and pooled κ	59	0	8	293	97.8	0.92 (0.87 to 0.98)

\*Agreement=(a+d)/(a+b+c+d).

†κ statistics=(probability of agreement) – (probability of by-chance agreement)/1 – (probability of by-chance agreement).

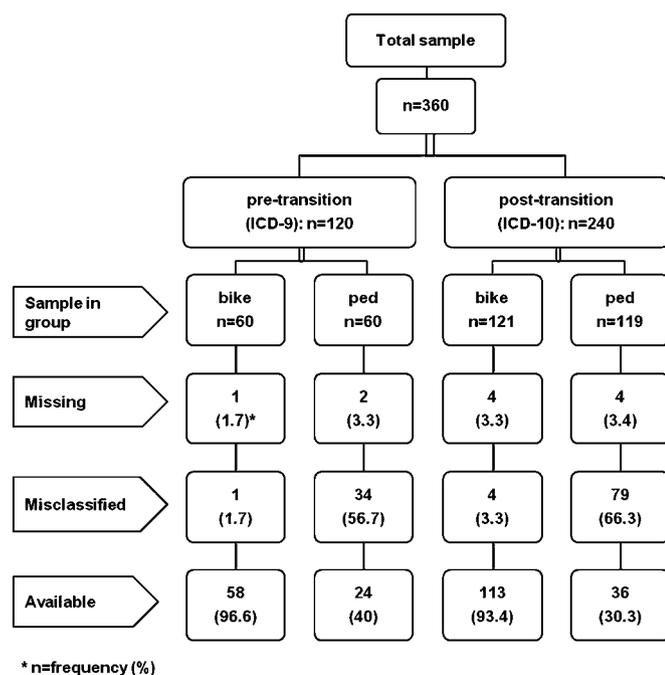
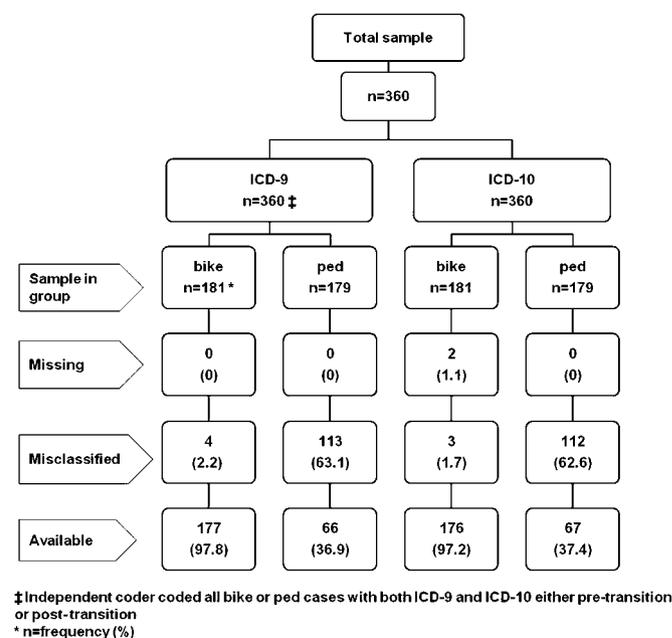
‡a: chart (+) coder (+) or true positive; b: chart (-) coder (+) or false positive; c: chart (+) coder (-) or false negative; d: chart (-) coder (-) or true negative.

validity of ICD-9-CM and ICD-10-CA in E-coding bicycle injuries, using pedestrian E-coding as a comparison group. The results showed that agreement for E-coding bicycle injuries was consistently high before and after transition from ICD-9-CM to ICD-10-CA between IND/ED coders and medical charts. Reviewing documented information, the IND coder was able to assign relevant E-codes for bicycle injuries. The limited differences between ED coders and the IND coder have demonstrated the high quality of bicycle injury E-coding, supporting conclusions drawn from bicycle studies using hospital and ED administrative databases.

The difference between ICD-9 and ICD-10 for external causes of injuries among pedestrians and bicyclists is mostly related to the method of defining each code. In ICD-9 the letter 'E' at the beginning of codes is an indication of external cause of injury followed by three digits that specify both external cause of

injury and the circumstances of the injury (eg, E801=railway involving collision with other object); decimals will specify whether the injured person is a pedestrian or bicyclist. In ICD-10 the letter 'V' and the first digit will specify the external cause of injury and the injured person (V0 for pedestrian and V1 for bicyclist) and the second digit will specify the external cause of injury (eg, 5=pedestrian injured in collision with railway or railway vehicle). The decimal in ICD-10 is used for specifying circumstances of the injury event (e.g., traffic or non-traffic related).<sup>7-9</sup> We have shown in validity analysis by the IND coder that ICD-10 is a valid classification for E-coding all bicycle and pedestrian injuries as is ICD-9.

Five studies in the USA<sup>5, 17, 19-21</sup> and one in New Zealand<sup>18</sup> reported reliability of ICD-9-CM E-coding for injured patients. Exact code agreements between an IND coder and hospital nosologists were reported to be 55.6–82%. One study in Australia<sup>23</sup> and one in New Zealand<sup>24</sup> reported 67.6% and 71%,

**Figure 1** Frequency of missed/misclassified E-codes by emergency department (ED) coders for bicycle (bike) and pedestrian (ped) injuries in three EDs in Edmonton, Alberta, Canada.**Figure 2** Frequency of missed/misclassified E-codes by independent coder for bicycle (bike) and pedestrian (ped) injuries in three emergency departments in Edmonton, Alberta, Canada.

respectively, correct E-coding in ICD-10. A systematic review (including five studies) also demonstrated that the range of accurate E-coding in hospital records was between 65% (exact code agreement) and 85% (agreement for broader groups of codes).<sup>38</sup>

Our study found different results for E-coding of a comparison population of pedestrian injuries. Despite the low reliability of pedestrian injury coding E-coding between the IND and ED coders (table 2), both demonstrated many cases (over 50%) with incorrect E-codes (table 1, pedestrian injuries). The differences between all coders and the reference standard (medical charts) demonstrate the poor quality of pedestrian injury coding, and call into question conclusions drawn from any hospital and ED administrative databases examining pedestrian injuries.

Studying the accuracy of E-coding for work-related and non-work-related injuries in Massachusetts ED data, Hunt *et al* demonstrated that machinery injuries were misclassified in many cases (65%) to other external cause of injuries such as cut/pierce, struck by/against, falls, overexertion, and missing.<sup>17</sup> In the same study it was shown that there was misclassification for coded as not-specified (54%), not-elsewhere classified (31%), other specified (19%), natural/environmental (17%), fall (11%), fire/burn (11%), poisoning (9%), struck by/against (9%), cut/pierce (3%), transportation (1%), and overexertion (1%). Overall, all causes were misclassified 14% of the time to other groups of injuries.<sup>17</sup>

In another study, the percentage error in the 5th digit location for E-coding was between 2% (for homicide/assault) and 15% (for medical injury).<sup>5</sup> Another source of inconsistency between original and independent auditor codes appeared to be due to missing E-codes making up between 14%<sup>21</sup> and 20%<sup>20</sup> of injury cases. We investigated the many missing E-codes for pedestrian injuries and found that in only 3.3% (pre-transition to ICD-10) and 3.4% (post-transition to ICD-10) of the time did ED coders forget to E-code for pedestrian injuries; however, between 59% and 69% of pedestrian injuries were misclassified, mostly as falls. The IND coder had also miscoded a substantial proportion of pedestrian injuries as falls.

As suggested by other studies,<sup>5 17–19 21 24</sup> we also realise that E-coding for pedestrian injuries needs to be emphasised in the nosologist training programmes in order to reduce the number of misclassified cases. As fall was the main source of misclassification, we would suggest that a detailed search for location of the falls must be considered an important piece of information for pedestrian E-coding. It is quite likely that other mechanisms of injury would be subject to the same level of misclassification as our pedestrian injuries (e.g., struck by or against, falls in non-pedestrian settings) and further work in this area is required. Sensitivity of E-coding for bicycle injuries was high for both ED and IND coders, and due to misclassification, was lower for E-coding of pedestrian injuries. ED coders and our IND coder demonstrated a high degree of agreement regardless of E-coding for bicycle or pedestrian injuries (table 2).

### Limitations and strengths

Our study was not without limitations. Since we only conducted our study in three hospital EDs in Edmonton, our results may not be generalisable to other locations. In our study, the reference standard was developed through chart review by research staff and confirmed by clinical nurses. The decisions were made long after the discharge occurred and could not be validated further; however, we believe that we selected an unbiased group of both cyclists and pedestrians for coder review. Although it seems very unlikely, we may have missed some patients in the EDIS if ED research assistant failed to use appropriate keywords to identify bicycle or pedestrian injuries;

### What is already known on the subject

- ▶ In both ICD-9 and ICD-10, external cause of injury codes (E-codes) have been widely used for morbidity and mortality surveillance.
- ▶ Current studies suggest that reliability of E-coding injured patients for ICD-9 ranged from 56 to 82% and for ICD-10 ranged from 68 to 71%.
- ▶ Misclassification does occur in E-coding of specific groups of injuries, such as work-related injuries in approximately 14% of cases.

### What this study adds

- ▶ This study demonstrated that sensitivity of both emergency department (ED) and independent coders was >90% for bicycle injuries but lower than 50% for pedestrian injuries.
- ▶ The majority of misclassification of E-coding pedestrian injuries was recorded as fall.
- ▶ Although bicycle injuries were E-coded with high accuracy, unexpected high misclassification in pedestrian injuries for both emergency department and independent coders highlights the importance of validity and reliability examination of E-coding before drawing valid conclusions in surveillance studies.

however, missing cases does not affect the validity and reliability of our study. Since in our analysis we have only used pedestrians as the comparison group for bicyclists, it may be argued that we would overestimate levels of agreement because of the limited range of other non-cyclist choices. However, we suspect our choice of pedestrians as a comparator group will have led to a conservative estimate in that it would be hard to distinguish this group from cyclists. If we had chosen less similar mechanisms (e.g., farm injuries or motor vehicle injuries) as our comparison group, it is quite likely that the agreement would have been higher.

We selected cases from one year coded by ICD-9-CM (2001) and two years coded by ICD-10-CA. Unlike other coding studies, we focused only on one external cause of injury (bicycle) and we examined a similarly vulnerable road user group (pedestrian). This is very helpful to make sure that using administrative data to study all bike-related injuries is reliable. Concurrently, validity of ICD-9-CM and ICD-10-CA was evaluated for E-coding of these two types of traffic-related injuries. Finding more than 50% misclassified E-codes for pedestrian injuries initiated a post-hoc investigation, showing that pedestrian injuries were often miscoded as falls. In our analysis we emphasised κ rather simple percentage agreement to test reliability of ED coder. This is also the first study in Canada evaluating the reliability of coders and the validity of the ICD-9-CM and ICD-10-CA systems for bicycle and pedestrian injuries. We selected our cases from the months of summer that included more cases of bicycle or pedestrian injuries, and used a random selection method from a pool of bicycle and pedestrian injuries.

### CONCLUSION

This study shows that ED coders are reliable in E-coding bicycle injuries using ICD-9-CM and ICD-10-CA systems. ICD-10-CA

and ICD-9-CM are valid classification tools in capturing bicycle injuries presenting to the ED. Pedestrian injuries, however, may be miscoded as falls, and this needs to be considered when examining ICD coded administrative data on these vulnerable road users. These results have important implications for injury surveillance research.

**Acknowledgements** The authors greatly appreciate the help of Anne Wiersma for her assistance in coding charts, and Debbie Boyko for her assistance in data extraction from charts.

**Funding** BEH holds the Alberta Children's Hospital Foundation Professorship in Child Health and Wellness, funded through the support of an anonymous donor and the Canadian National Railway Company, as well as the Alberta Heritage Foundation for Medical Research Population Health Investigator and Canadian Institutes of Health Research New Investigator Awards, Alberta, Canada. BHR is supported by the Government of Canada as a 21st Century Canada Research Chair in Emergency Medicine.

**Competing interests** None declared.

**Ethics approval** This study was conducted with the approval of the University of Alberta Health Research Ethics Board.

**Contributors** BEH (Lead) and BHR obtained study funding. BEH, MK, BHR, LDS, and DCV contributed to the conception and design of the study. BEH, MK, and BHR contributed to the acquisition of the primary data, analysis, and drafting of the article. All of the authors contributed to the interpretation of the data and revision of the article for important intellectual content and gave final approval of the manuscript prior to submission.

**Provenance and peer review** Not commissioned; externally peer reviewed.

## REFERENCES

- Dennis J, Potter B, Ramsay T, et al. The effects of provincial bicycle helmet legislation on helmet use and bicycle ridership in Canada. *Inj Prev* 2010;**16**:219–24.
- Canadian Institute for Health Information. *Head Injuries in Canada: A Decade of Change (1994-1995 to 2003-2004)*. Ottawa, Ontario Canada, 2006. <https://secure.cihi.ca/estore/productbrowse.htm?locale=en&Lang=en-US> (accessed Sep 2009).
- Linn S, Smith D, Sheps S. Epidemiology of bicycle injury, head injury, and helmet use among children in British Columbia: a five year descriptive study. *Inj Prev* 1998;**4**:122–5.
- Alexander S, Conner T, Slaughter T. Overview of inpatient coding. *Am J Health Syst Pharm* 2003;**60**(21 Suppl 6):S11–14.
- Smith SM, Colwell LS Jr, 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**:279–81.
- World Health Organization. *History of International Classification of Diseases*. <http://www.who.int/classifications/icd/en> (accessed Feb 2010).
- World Health Organization. *International Classification of Diseases-10th revision (ICD-10)*. 2003. <http://www.who.int/classifications/icd/en/> (accessed Sep 2009).
- Geran L, Tully P, Wood P, et al. *Compatibility of ICD-10 and ICD-9 for Mortality Statistics in Canada*. Statistics Canada, 2005. <http://www.statcan.ca> (accessed Sep 2009).
- World Health Organization. *ICD-9-CM, Millenium Edition*. Los Angeles, California, USA: Practice Management Information Corporation, 2001.
- Wang FL, Morrison K, Schopflocher DP, et al. *ICD-9 to ICD-10 Coding with Reference to Causes of Death Grouping in Alberta (working document version 1.0)* 2006.
- Canadian Institute for Health Information. *The Canadian Enhancement of ICD-10 (final report)*. Ottawa, Ontario Canada, 2001. [http://secure.cihi.ca/cihiweb/en/downloads/codingclass\\_icd10enhan\\_e.pdf](http://secure.cihi.ca/cihiweb/en/downloads/codingclass_icd10enhan_e.pdf) (accessed Sep 2009).
- Pan SY, Ugnat AM, Semenciw R, et al. Trends in childhood injury mortality in Canada, 1979-2002. *Inj Prev* 2006;**12**:155–60.
- Dempsey RL, Layde PM, Laud PW, et al. Incidence of sports and recreation related injuries resulting in hospitalization in Wisconsin in 2000. *Inj Prev* 2005;**11**:91–6.
- Cook A, Sheikh A. Trends in serious head injuries among English cyclists and pedestrians. *Inj Prev* 2003;**9**:266–7.
- Kypri K, Chalmers DJ, Langley JD, et al. Adolescent injury morbidity in New Zealand, 1987-96. *Inj Prev* 2002;**8**:32–7.
- Leonard PA, Beattie TF, Gorman DR. Under representation of morbidity from paediatric bicycle accidents by official statistics—a need for data collection in the accident and emergency department. *Inj Prev* 1999;**5**:303–4.
- Hunt PR, Hackman H, Berenholz G, et al. Completeness and accuracy of International Classification of Disease (ICD) external cause of injury codes in emergency department electronic data. *Inj Prev* 2007;**13**:422–5.
- Langley J, Stephenson S, Thorpe C, et al. Accuracy of injury coding under ICD-9 for New Zealand public hospital discharges. *Inj Prev* 2006;**12**:58–61.
- 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.
- Langlois JA, Buechner JS, O'Connor EA, et al. Improving the E coding of hospitalizations for injury: do hospital records contain adequate documentation? *Am J Public Health* 1995;**85**:1261–5.
- Schwartz RJ, Nightingale BS, Boisoneau D, et al. Accuracy of e-codes assigned to emergency department records. *Acad Emerg Med* 1995;**2**:615–20.
- MacIntyre CR, Ackland MJ, Chandraraj EJ, et al. 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.
- McKenzie K, Enraght-Moony EL, Waller G, et al. Causes of injuries resulting in hospitalisation in Australia: assessing coder agreement on external causes. *Inj Prev* 2009;**15**:188–96.
- Davie G, Langley J, Samaranyaka A, et al. Accuracy of injury coding under ICD-10-AM for New Zealand public hospital discharges. *Inj Prev* 2008;**14**:319–23.
- Henderson T, Shephard J, Sundararajan V. Quality of diagnosis and procedure coding in ICD-10 administrative data. *Med Care* 2006;**44**:1011–19.
- Salmerón D, Cirera L, Saez M, et al. Influence of the introduction of the ICD-10 on tendencies of mortality by causes (1980-2004). *Gac Sanit* 2009;**23**:144–6.
- Cano-Serral G, Perez G, Borrell C, et al. Comparability between ICD-9 and ICD-10 for the leading causes of death in Spain. *Rev Epidemiol Sante Publique* 2006;**54**:355–65.
- Richardson DB. The impact on relative risk estimates of inconsistencies between ICD-9 and ICD-10. *Occup Environ Med* 2006;**63**:734–40.
- Anderson RN, Rosenberg HM. Disease classification: measuring the effect of the Tenth Revision of the International Classification of Diseases on cause-of-death data in the United States. *Stat Med* 2003;**22**:1551–70.
- Rooney C, Griffiths C, Cook L. The implementation of ICD-10 for cause of death coding—some preliminary results from bridge coding study. *Health Stat Q* 2002;**13**:31–41.
- Watzlaf VJ, Garvin JH, Moieni S, et al. The effectiveness of ICD-10-CM in capturing public health diseases. *Perspect Health Inf Manag* 2007;**4**:6.
- Alberta Centre for Injury Control & Research. *Alberta Injuries: Comparison of Injuries in Alberta's Health Region, 2006*. Edmonton, Alberta Canada, 2009. <http://acirc.ca/data/injury-data-reports> (accessed Sep 2010).
- Alberta Transportation. *Traffic Safety in Alberta (Mandatory Bicycle Helmets)*. 2004. <http://www.saferoads.com/safety/educators/efaqa.html> (accessed Sep 2006).
- Innes G, Murray M, Grafstein E. A consensus-based process to define standard national data elements for a Canadian emergency department information system. *CJEM* 2001;**3**:277–84.
- STATA. *Statistical Software*. 11 edn. College Station, TX: Statacorp LP, 2009.
- Cohen J. A Coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960;**20**:37–46.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;**33**:159–74.
- McKenzie K, Enraght-Moony EL, Walker SM, et al. Accuracy of external cause-of-injury coding in hospital records. *Inj Prev* 2009;**15**:60–4.

## APPENDIX 1 Definition of bicycle and pedestrian injuries used by investigators for identifying cases from Emergency Department Information System (EDIS)

### Bicycle injury

The bicycle injury should have happened in a public area (e.g. streets, highways, parks, bicycle pathways, commuter route). They included:

- ▶ Bicycle rider or passenger hit by a motor-vehicle (including motorcycle, moped and other motorized vehicle)
- ▶ Bicycle had a collision with another bicycle
- ▶ Bicycle hit a stationary object
- ▶ Bicycle hit a moving object or being hit by that moving object (e.g. train, animal)
- ▶ Bicyclist fell off the bike
- ▶ Rider on a unicycle
- ▶ Bicyclists on a reclined bicycle or reclined tricycle
- ▶ Bicyclists on a tandem bicycle

### Pedestrian injury

Pedestrian is a person who shares the road or commuting route with other road users (motorized or non-motorized vehicles) or commuters in public places (e.g. parks streets, bicycle pathways, residential pathways). This is a public road way use, but non-motorized and non-wheeled transportation. They included:

- ▶ A person hit or run over by a motorized vehicle (including motorcycle, moped and others) in roadways
- ▶ A person hit or run over by a non-motorized vehicle (e.g. bicycle, tricycle, scooter and other) in roadways
- ▶ A person injured on a roadway while walking (e.g. tripped over curb or fell over tree root)
- ▶ A person walking on the roadway and injured due to hitting light poles, trapping in a hole, and hit by a loose object of traffic control device.
- ▶ A person injured while walking on one side of the bicycle (not riding it) in a roadway