Classifying sports medicine diagnoses: a comparison of the International classification of diseases 10-Australian modification (ICD-10-AM) and the Orchard sports injury classification system (OSICS-8)

K Rae, H Britt, J Orchard, C Finch

Background: The International classification of diseases 10-Australian modification (ICD-10-AM) and the Orchard sports injury classification system (OSICS-8) are two classifications currently being used in sports injury research.

Objectives: To compare these two systems to determine which was the more reliable and easier to apply in the classification of injury diagnoses of patients who presented to sports physicians in private sports medicine practice.

Methods: Ten sports physicians/sports physician registrars each coded one of 10 different lists of 30 sports medicine diagnoses according to both ICD-10-AM and OSICS-8 in random order. The coders noted the time taken to apply each classification system, and allocated an ease of fit score for individual diagnoses into the systems. The 300 diagnoses were each coded twice more by “expert” coders from each system, and these results compared with those of the 10 volunteers.

Results: Overall, there was a higher level of agreement between the different coders for OSICS-8 than for ICD-10-AM. On average, it was 23.5 minutes quicker to complete the task with OSICS-8 than with ICD-10-AM. Furthermore, there was also higher concordance between the three coders with OSICS-8.

Conclusions: Based on the results of this study, OSICS-8 appears to be the preferred system for use by inexperienced coders in sports medicine research. The agreement between coders was, however, lower than expected. It is recommended that changes be made to both OSICS-8 and ICD-10-AM to improve their reliability for use in sports medicine research.

A number of commentators have discussed the best way to collect sports injury data and to standardise the information collected so that different studies can be compared.1 2 A standardised classification system for sports medicine diagnoses is one critical component of this, and various classifications have been developed for this purpose.3 4

In injury surveillance research, the first step is defining what constitutes an injury, including inclusion and exclusion criteria. Then, when recording data a sound description of the injury in terms of its bodily location, pathology, and severity is required to characterise the injury. This information can then be expressed as a diagnostic code. Past studies have sometimes used non-specific diagnostic codes to analyse data—for example, coding all knee injuries (which could potentially include anterior cruciate ligament rupture, patellar tendinopathy, etc) as “knee injury”. This has been seen to limit the usefulness of the collected data and its ability to be used in subsequent research.5 7

In most data collections systems, once an injury has been assigned a diagnosis, it is then coded to enable analysis. There are only a few standardised classification systems used in sports injury research. The International classification of diseases (ICD) is one of the most common medical classification systems used internationally.4 Australia has made some local modifications to this system for the recording of morbidity, and the latest version of ICD-10-AM was published in 2002.2 The ICD-10-AM classifies diseases/conditions for every body system and is used in the hospital care setting. Its specific usefulness for classifying diagnoses in sports medicine practice has not been evaluated.

Another classification system widely used in sports injury research in Australia is the Orchard sports injury classification system (OSICS).4 The OSICS has been used for injury research undertaken through Cricket Australia12 13 and the Australian Football League (AFL).11 It has also been used in other sports both in Australia and overseas.12 13 Various injury surveillance systems such as Injury Tracker14 and the Australian Sports injury data dictionary15 also incorporate OSICS. Although OSICS has been applied in several contexts, its reliability has not been assessed to date.

Sports injury surveillance requires a sports injury classification system that allows classification of injuries and illnesses sustained across the broad spectrum of sports participation, but which can also provide both detailed and broad information to characterise the injuries. Such a system needs to be both accurate and detailed enough to inform clinical practice while being easy to use. The aim of this study was to compare the ease of application and the inter-coder reliability of ICD-10-AM and OSICS-8 to classifying sports medicine diagnoses encountered in common sports medicine practices and to make recommendations for improvements to these classification systems.

Abbreviations: ICD-10-AM, International classification of diseases 10-Australian modification; OSICS-8, Orchard sports injury classification system; SMIS, Sports medicine injury surveillance.
METHODS

To provide a realistic and representative sample set of injury diagnoses treated by sports medicine physicians, a random sample of diagnoses was obtained from the Sports medicine injury surveillance (SMIS) database. The SMIS study was conducted over a 12 month period. Of these, 64.3% sought initial attention from the sports medicine clinics over the 12 month period. Of these, 64.3% sought initial treatment from a sports physician, who recorded a sports medicine diagnosis on the injury surveillance record forms in free text.

The diagnoses of all SMIS cases treated by a sports physician were reviewed, and cases of duplicate injury diagnoses removed. This yielded a list of about 3000 individual injury diagnoses. A 10% random sample of these diagnoses was selected and used to generate 10 alphabetical lists of 30 diagnoses to be given to the coders. These lists were then randomly reordered twice to form two new lists for the coders to use.

Each list was then coded by:

1. An expert with each classification system
   a. a trained ICD-10 coder, with a degree in health information management and currently working as a researcher in general practice
   b. the developer of OSICS, author J.O., a sports physician
2. The primary author (K.R.) self trained in both systems and a senior sports medicine registrar
3. One of 10 volunteer sports medicine practitioners with no formal training in, or significant experience with, either classification system. These 10 sports medicine practitioners were either experienced sports medicine registrars or sports physicians—that is, doctors currently doing advanced training in or having completed their fellowship with the Australasian College of Sports Physicians (ACSP), recruited from the ACSP list of registrars and fellows living in Sydney.

The two experts each coded all 300 cases according to the particular classification system with which they were experienced, the primary author (K.R.) coded all 300 cases with both classifications, and the 10 volunteer sports medicine practitioners each coded their list of 30 diagnoses with first one and then the other classification system in random order.

The ICD-10-AM is available in paper and electronic format. The paper format consists of five books in total, two of which were supplied to the volunteer coders for the completion of this task. The two books supplied were volumes 1 and 2, the Tabular list of diseases (590 pages of codes) and the Alphabetic index of diseases (582 pages) respectively. The Tabular list of diseases is set up in chapters according to body systems. Most sports medicine diagnoses are found in chapter XIII (Diseases of the musculoskeletal system and connective tissue; codes M00–M99; divided into subsections based on pathology) and chapter XIX (Injury, poisoning and certain other consequences of external causes; the injury sections are divided into subsections based on anatomy and then pathology).

For the paper format, the coder could either use the index by looking up a keyword, or try to find the code by looking up the relevant section of the coding book. The electronic format requires that the coder types in keywords and selects from the codes that appear in a pick list. All coders used the paper format.

The OSICS is also available in both paper and electronic format. The paper format consists of seven pages of codes organised into anatomical and then functional groupings. The electronic format, similar to ICD-10-AM, relies on the coder typing in a keyword and selecting a code from those located by the program. Again the volunteer coders and the primary author used the paper format, and the expert coder used the computer format to complete the task.

Each of the 10 volunteer coders was given brief training on the layout of both classifications and instructions on how to complete the task. They were asked to note down the time at which they both started and completed the task. They were asked to provide a single code for each diagnosis, unless the injury was actually two or three separate injuries that should not be catered for by one code—for example, “medial and lateral epicondylitis”. In this case they were allowed to use multiple codes. They were asked to complete the list of diagnoses in the order on the sheet and not to proceed to the next diagnosis until they had completed the current one. They were also asked to allocate a number from 1 to 4 based on how well they thought the code fitted that particular diagnosis: 1, a perfect fit; 2, a reasonable fit but there should be a better code; 3, the code allocated was a last resort; 4, unable to code a particular diagnosis.

Results were analysed in four ways

1. Average pairwise agreement. The coders were grouped into three pairs: (a) expert and author; (b) author and volunteer; (c) volunteer and expert. Agreement (% of all codes) for these three pairs was determined and then used to determine the average pairwise agreement for each classification system.
2. Total agreement and total disagreement. The proportion of diagnoses where all three coders who were coding the diagnosis agreed on the code given was calculated and expressed as a percentage, as was the proportion where all three coders selected disparate codes.
(3) Time to complete the task. The time taken from start to finish for each volunteer coder with each classification system was also recorded. The average was compared using the paired t test.
(4) Subjective analysis of unsatisfactory codes. All diagnoses that generated total disagreement between the coders or those codes that had been allocated a 3 or 4 on the above described four point goodness of fit scale by the volunteer coder were deemed unsatisfactory and were subjectively reviewed by one of the authors (KR). The one, two, or three different code(s) chosen by the coders for each diagnosis were reviewed to determine which was the most correct and the reasons for dissatisfaction or disagreement with coding that particular diagnosis.

RESULTS
Figure 1 shows the pairwise agreement between groups of coders. The average pairwise agreement between the groups of two coders was 57.2% for OSICS-8 and 35.3% for ICD-10-AM. Figure 2 shows a comparison of the average total agreement (where all three coders selected the same code), agreement (where two coders selected the same code), and disagreement (where all the three different selected codes) for OSICS-8 and ICD-10-AM. Average total agreement was higher for OSICS-8 than ICD-10-AM, and average total disagreement was higher for ICD-10-AM. Across the 10 volunteer coders, the time taken using OSICS-8 to code the diagnoses ranged from 23 to 60 minutes, with a mean (SD) of 36.5 (9.9) minutes compared with 39–90 minutes with a mean (SD) of 60 (18.9) minutes for ICD-10-AM. This difference was significant ($t_9 = -5.26$, p<0.001).

Reasons for total disagreement about which code to assign, or in cases where the code selected was deemed unsatisfactory by the volunteer coder as outlined in the methods section appeared to be due to both deficiencies in the classification systems and limitations in the way the diagnoses were written by the clinicians in the SMIS system (table 1).

DISCUSSION
To our knowledge, this is the first study comparing the application of specific classification systems with the coding of sports medicine diagnoses. Langley and Chalmers have analysed the place of occurrence codes in ICD-9 in a sports

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Subjective assessment of the limitations of the two classification systems for sports medicine diagnoses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reason for disagreement</td>
<td>ICD-10-AM</td>
</tr>
<tr>
<td>(1) Human error*</td>
<td>Complexity of musculoskeletal system and organisation based on pathology caused fatigue and boredom, as stated by coders</td>
</tr>
<tr>
<td>(a) Poor organisation of system affected navigation across the documentation</td>
<td></td>
</tr>
<tr>
<td>(b) Relationship of codes to diagnosis</td>
<td>Codes in chapter XIII do not relate to the diagnosis in any recognisable pattern, potentially causing translational errors</td>
</tr>
<tr>
<td>(c) Lack of formal training in classification system</td>
<td>Coders chose “unspecified” rather than “not otherwise specified” or “other” codes</td>
</tr>
<tr>
<td>(i) Errors use of “unspecified” codes</td>
<td></td>
</tr>
<tr>
<td>(ii) Omission of part of code</td>
<td>Chapter XIII—4th letter specifying body location for “other” codes often omitted</td>
</tr>
<tr>
<td>(d) Lack of understanding of sports medicine diagnosis</td>
<td>Occurred with expert coder as a result of having no previous experience in the practice of sports medicine</td>
</tr>
<tr>
<td>(2) Deficiencies in the codes available</td>
<td>e.g. tib. Post tendon tear—required localising to ankle or foot</td>
</tr>
<tr>
<td>(a) Location of injury</td>
<td>e.g. haematoma Achilles does not exist—either injury to Achilles (anatomy) or contusion of ankle (pathology)</td>
</tr>
<tr>
<td>(i) Codes too specific regarding location of injury</td>
<td></td>
</tr>
<tr>
<td>(ii) Codes not specific enough for location of injury</td>
<td></td>
</tr>
<tr>
<td>(b) Codes do not encompass all details of the diagnosis causing coder to choose:</td>
<td></td>
</tr>
<tr>
<td>(i) Pathology v anatomy</td>
<td></td>
</tr>
<tr>
<td>(ii) 1 of 2 different pathologies in one injury</td>
<td>Combination injuries, e.g. fracture dislocation, do not have a code in either system</td>
</tr>
<tr>
<td>(iii) 1 of 2 structures injured in one body part</td>
<td>e.g. ankle sprains—ICD only allows one injury to be coded. Not possible to code for a sprain of more than one ligament</td>
</tr>
<tr>
<td>(c) Post-surgical conditions</td>
<td>Dilemma of coding pre-surgical v post-surgical condition</td>
</tr>
<tr>
<td>(3) Deficiencies in diagnosis given</td>
<td>Problem—too many “other” codes</td>
</tr>
<tr>
<td>(a) Location e.g. “lateral ligament injury”</td>
<td></td>
</tr>
<tr>
<td>(b) Diagnosis, e.g. acceleration injury</td>
<td>Problem—too many “other” codes</td>
</tr>
<tr>
<td>(c) Terminology, e.g. hamstring impingement syndrome</td>
<td>Problem—too many “other” codes</td>
</tr>
<tr>
<td>(4) Classification system omits or under-represents diagnoses</td>
<td>Stress fractures, osteitis pubis, osteochondral injuries, specific tendinopathies</td>
</tr>
<tr>
<td>(5) Chronicity of injury uncertain</td>
<td>Choice of chapter XIII v chapter XIX, especially for knee and back</td>
</tr>
<tr>
<td>(6) Overuse injuries</td>
<td>Confusion whether overuse injuries can be coded in chapter XIX</td>
</tr>
</tbody>
</table>

*Where the correct code had been identified by one coder but was missed by the other two coders.
†Chapter XIII, Diseases of the musculoskeletal system and connective tissue.
‡Chapter XIX, Injury, poisoning and certain other consequences of external causes.
injury application, but did not analyse its diagnostic codes. Unfortunately, there is no gold standard for the classification of sports medicine diagnoses. We therefore chose to compare two different systems, OSICS-8, developed specifically for sports medicine injury research, and ICD-10-AM, the classification published by the World Health Organisation, which was developed to classify all possible medical injuries and illnesses. Although we could have compared OSICS with other specific sports injury classifications such as NAIRS5 (National athletic injury/illness reporting system) widely used in America, we were unable to find any literature supporting their reliability. The comparison we performed has highlighted some important problems with both classifications in relation to their application to typical sports medicine diagnoses encountered in clinical practice.

Overall, we found that the level of agreement for both classifications fell below what would be considered acceptable for the accurate assignment of codes. Britt et al in their study of the use of the International classification of primary care in general practice adopted the level of 70%, set by Bensinger and at the 10th WONCA Conference on Family Medicine, as an acceptable level of positive agreement as a measure of validity. If this is used as a marker, the above results show that, although agreement for OSICS-8 (57.2%) was better than for ICD-10-AM (35.3%), it was still not acceptable by this measure.

It is possible that the nature of our coders could have biased these comparisons. For OSICS-8, all coders were sports physicians or sports physician registrars, whereas for ICD-10-AM coding, the professional coder had no knowledge of clinical sports medicine. This lack of specific sports medicine clinical knowledge may have affected this coder’s ability to accurately assign codes. At the time of coding, the professional ICD coder was therefore permitted to ask questions to help her understand the diagnoses she was coding. These questions were answered by email with no referral to ICD-10-AM coding, the professional coder had no knowledge of the written diagnosis, and subsequent application of the most appropriate code was not clearly identified, which was one of the problems identified in secondarily coding with OSICS-8.

This study was not able to determine whether any improvements could be made in the classification accuracy with better education before using the classifications and/or increased experience in using the classifications. Coders had less than five minutes interactive education on each classification system, although they could familiarise themselves with both systems for as long as they wanted before starting to code. They also had only 30 non-repeated diagnoses to code. Therefore there is theoretically much potential for improvement in accuracy and reliability with learning in both systems. Future studies need to be undertaken to evaluate the level of training and or experience required for a coder to plateau in the number of errors they make when coding.

This study only looked at the situation where the diagnoses are determined and recorded by a clinician and secondarily coded by another. An alternative method would be for the injury to be diagnosed and classified by the same person, thus removing the possible confounder of interpretation by the coder of the written diagnosis. In this alternative (primary coding), there would be no time when the body part was not clearly identified, which was one of the problems identified in secondarily coding with OSICS-8.

To enable comparison of codes selected by coders for the same injury, all coders would have had to diagnose the injury from the patient’s clinical presentation and may or may not have selected the same diagnosis. This would have introduced a new reason for the use of different codes (disagreement on the clinical diagnosis). As our interest was not in diagnostic accuracy but purely in the application of the classifications, the original scenario was chosen.

In summary, this study has highlighted some of the problems inherent in diagnostic classifications applied to sports medicine diagnoses. Both ICD-10-AM and OSICS-8 classifications would benefit from revision and reorganisation to correct their deficiencies in the areas of organisational structure, directional advice, terminology, omitted codes, and either a lack of or overly expansive “other” categories.

When determining how to best code injuries/illnesses for sports medicine studies, we formulated a list of recommendations that should be considered by researchers to maximise accuracy in allocating codes (table 2).

### Table 2  Recommendations when using classification systems in injury research

| (1) The classification system should be one that: | (i) cater for all possible diagnoses  
| (ii) have clear rules regarding what codes to use in particular circumstances  
| (iii) is easy to use and navigate  
| (iv) have a high intraobserver and interobserver reliability |
| (2) The research group should: | (i) provide clear instructions to those clinicians assigning the diagnoses so that they provide written diagnoses that are able to be coded  
| (ii) consider having the coding done by the sports practitioner assigning the diagnosis so that there is no chance for misinterpretation of the diagnosis when being coded  
| (iii) have all coders knowledgeable in sports medicine diagnoses and/or able to communicate with the diagnosing practitioner if a query arises  
| (iv) have all coders well trained in the coding system or under close supervision by an experienced coder |
| (v) have established rules and guidelines about what aspect of the diagnosis to prioritise if the exact code is unavailable, e.g. location, pathology  
| (vi) have a system whereby group consensus can occur with any code where the coder is unsure regarding the best code to use |

CONCLUSION

There are many aspects to accumulating accurate injury surveillance data, with coding and classification of diagnoses being just one of these. Error can occur at many levels when trying to accurately code an injury including accurate diagnosis of the problem, accurate interpretation of the written diagnosis, and subsequent application of the most
appropriate classification and selection of the most appropriate code. This research has only assessed interpretation of the written diagnosis and assignment of the most appropriate code. As seen above, even with the diagnosis stated, the accuracy of application of the most appropriate code varies depending on the coder and the coding system. With all the above borne in mind, OSICS-8 would appear the preferred coding system for use in sports medicine research. It was the faster and easier system to use and provided the greater agreement between coders. It is freely available over the internet and is continually being modified, with new codes being added as feedback from users is received. However, OSICS-8 does have limitations in its organisational structure, specifically its hierarchical system of codes, and deficiencies in the areas of injuries to body parts unnamed and injuries not otherwise specified. These deficiencies need to be addressed in future versions of the coding system.

The ICD-10-AM has many potential shortcomings with regard to its use in the coding of sports medicine diagnoses. Its size and complexity means it is harder to navigate and takes longer to use. The diagnoses used by sports medicine practitioners are not well organised, nor easy to find. Several important groups of sports medicine diagnoses either omitted or inadequately covered by ICD-10-AM include stress fractures, degenerative tendinopathies, and osteitis pubis. Developments in sports medicine research have allowed refinement of the diagnostic terms used as pathology becomes better understood, but there do not seem to have been similar changes in ICD-10-AM to reflect the level to which diagnosis of sports injuries has progressed today.

This research, by highlighting deficiencies in both systems, will be useful in providing information for the ongoing improvement of both OSICS-8 and ICD-10-AM. Once changes are implemented, further research should be undertaken to determine whether they have improved speed and agreement. It will also hopefully allow more consideration be put into the coding of injuries for research studies to allow the recorded information to be more accurate, more retrievable, and more useable in the future.

REFERENCES