Categorising sports injuries in epidemiological studies: the subsequent injury categorisation (SIC) model to address multiple, recurrent and exacerbation of injuries

Caroline F Finch,1 Jill Cook2

ABSTRACT
Objective Sports injuries are often recurrent and there is wide recognition that a subsequent injury (of either the same or a different type) can be strongly influenced by a previous injury. Correctly categorising subsequent injuries (multiple, recurrent, exacerbation or new) requires substantial clinical expertise, but there is also considerable value in combining this expertise with more objective statistical criteria. This paper presents a new model, the subsequent injury categorisation (SIC) model, for categorising subsequent sports injuries that takes into account the need to include both acute and overuse injuries and ten different dependency structures between injury types.

Methods The suitability of the SIC model was demonstrated with date ordered sports injury data from a large injury database from community Australian football players over one playing season. A subsequent injury was defined to have occurred in the subset of players with two or more reported injuries.

Results 282 players sustained 469 subsequent injuries of which 15.6% were coded to categories representing injuries that were directly related to previous index injuries. This demonstrates that players can sustain a number of injuries over one playing season. Many of these will be unrelated to previous injuries but subsequent injuries that are related to previous injury occurrences are not uncommon.

Conclusion The handling of subsequent sports injuries is a substantial challenge for the sports medicine field—both in terms of injury treatment and in epidemiological research to quantify them. Application of the SIC model allows for multiple different injury types and relationships within players, as well as different index injuries.

INTRODUCTION
Sports injuries are often recurrent and there is wide recognition that a subsequent injury (of either the same or a different type) can be strongly influenced by a previous injury.1–5 Reasons for this could include similar mechanisms or risk factors (especially intrinsic risk factors) involved in injury causation or limited tissue healing from an earlier injury.6 7

The terms ‘repeat’, ‘recurrent’ or ‘multiple’ injuries are used in the literature interchangeably to describe instances in which more than one injury occurrence is studied. This is problematic because ‘multiple/repeat’ injuries could still include possibly correlated injuries that are not exactly the same (eg, an athlete sustains an ankle injury, resulting in loss of range of movement and strength and leading to other injuries such as Achilles tendinopathy or an injury further up the kinetic chain). The common use of the term ‘recurrent/repeat’ implies more than one occurrence of exactly the same injury, whereas ‘multiple injuries’ implies several different and unrelated injuries. In some applications, it would be correct to talk about either multiple or recurrent injuries alone, but there are other instances in which both type of injuries are of interest and could occur to an athlete over a specified period. For this reason, like Hamilton et al,4 we prefer the term ‘subsequent’ injury to encompass both multiple and recurrent injuries, as well as to allow for ‘exacerbation’, which is reinjury before resolution of prior injury.

Acute injury can be defined as any physical complaint that is caused by the inability of the body’s tissues to maintain its structural and/or functional integrity following an instantaneous transfer of energy to the body (eg, from an impact or sudden movement). For example, a sudden twist of the knee may cause an anterior cruciate ligament tear that causes immediate pain and functional impairment. In contrast, injuries that are caused by an accumulated energy transfer, rather than a clearly identifiable single event, are often called overuse injuries.8 9 However, none of the published injury definitions to date adequately cover all forms of sports injury.

Correctly categorising subsequent injuries (multiple, recurrent, exacerbation or new) is important because if correct definitions are not used, there is the potential for either under-reporting or overestimation of a new injury incidence and/or injury recurrences.10 While correct classification requires considerable clinical expertise, there is also value in combining this with more objective statistical criteria. Existing sports injury classifications are inadequate because they are unable to capture true recurrence. Moreover, there is currently no existing holistic system that has both clinical application and statistical robustness.

The aims of this paper are therefore to present a new model for categorising subsequent injuries in the sports injury context. Its application to a large injury data collection demonstrates the key features and attributes of the model. It is hoped that the model will be applied to other injury datasets and evaluated further to examine its wider application by both clinicians and researchers.

Accepted 13 February 2013

Centre for Healthy and Safe Sport (CHASS), University of Ballarat, Ballarat, Victoria, Australia

Department of Physiotherapy, School of Primary Health Care, Monash University, Melbourne, Victoria, Australia

Correspondence to Professor Caroline Finch, Centre for Healthy and Safe Sport (CHASS), University of Ballarat, SMB Campus, PO Box 668, Ballarat, VIC 3353, Australia; cfinch@ballarat.edu.au
DEFINING SUBSEQUENT INJURY RISK IN RELATION TO RECOVERY FROM AN INITIAL OR INDEX INJURY

To date, subsequent injuries have generally been considered from a clinical management and return-to-play (or time away from sport to recover from injury) perspective.1–3 10–12 However, there are many issues with defining and recording subsequent injuries in the sporting context. Sports injury surveillance guidelines and several conceptual papers describe the complex issues associated with properly classifying injuries as recurrent, reinjury, exacerbations or overuse.4 8 13 14 There has also been considerable attention given to consensus statements for sports injury surveillance, with the vast majority considering recurrent injury as one type of subsequent injury.13–16 Most of these follow the suggestions of Fuller et al.,7 though a new conceptualisation of overuse injury has been proposed by Bahr,4 and a refined definition of subsequent injury was given more recently by Hamilton et al.4 Reported definitions of subsequent injury have been both sport-specific and context-specific and whether or not the person has fully or partially recovered (or continues to play injured) from a previous injury has been considered to varying degrees.

The most recent taxonomy of subsequent sports injuries4 recommends that injuries be coded as: (1) new injury—different location, (2) local injury=same location but different type and (3) recurrent injury=same location and same type. In this taxonomy, the last two categories represent injuries that could be correlated with an initial injury. It has been accepted that the nature (at a tissue level), rather than just the general location of the injury (eg, knee or thigh), is also important when categorising repeat injuries.7 Others have argued that a recurrent injury should only be defined as one that is of the same type occurring at the same site but only after full recovery from an earlier injury.7 13 14 None of this prior work, however, discusses subsequent injuries from a statistical viewpoint, and therefore adequate recognition of the various dependencies (or possible relationships) between injuries within players is lacking in the sports medicine literature.

A NEW MODEL OF SUBSEQUENT INJURY CATEGORISATION (THE SIC MODEL)

The SIC model is one that takes into account the need to include both acute and overuse injuries and ten different dependency structures between injury types, and is outlined in table 1. Compared to the existing sports injury definitions, it adopts a more statistically oriented approach towards categorising injury types that makes provision for all potential within-person injury dependencies.

This approach is consistent with statistical modelling principles.17 The model was constructed using Bayesian data analysis principles that first require that a full probability model of all the possible outcomes be constructed.18 The starting point was therefore to construct a probabilistic model that included all combinations of the possible relationships between two injuries, irrespective of how likely they were. These categories were then checked to ensure that they had clinical meaning and that they were relevant to most sports injury data collections. In Bayesian statistical parlance, this is the prior predictive distribution of subsequent injury outcomes.

Table 1 also maps this new SIC model against the recent Hamilton et al7 subsequent injury definition and the Fuller et al7 recurrent/reinjury/exacerbation definitions. The overuse definitions are consistent with the approach advocated by Clarsen et al.19

It is recognised that for some real-world datasets some of the dependency categories within the SIC model may not apply (with the potential for some zero count cells), but it is still important to include them in an overall categorisation model, especially from a statistical model point of view. The categories within the SIC model are therefore holistic and intended to capture all types of subsequent injuries, while also being mutually exclusive. For example, to make a clinical case for category 5, a back-related hamstring strain could have an unclear clinical diagnosis; at times, it could present only like a hamstring strain and at other times, only like a low back issue. A player who presents twice with this injury will fit category 5. Category 6 is
easier to define and would include cases such as a player who sustains a fracture from a direct blow in one season and then fully recovers but sustains a similar fracture in the following season; this clearly fits category 6 rather than categories 7–10. Similarly, the category ‘no injury’ is included for model completeness from a statistical viewpoint, but is expected to have zero counts.

**USING THE SIC MODEL TO CATEGORISE INJURIES: INCORPORATING MULTIPLE INDEX INJURIES**

Consistent with the Bayesian statistical modelling approach, the next stage was to calculate the posterior probability distribution by applying the model to an existing dataset and, in this case, to count the number of injury cases assigned to each category.

Comparison in a pairwise fashion to an earlier injury is required to assess the recurrent nature of an injury. Access to reliable information before the first injury recorded or onset of the problem is difficult, and so the first index injury needs to be taken as the first recorded injury for a given athlete in a case series or new study. Using data from existing prior sources could potentially bias the data in relation to the duration and quality of the information contained in those records. For overuse injuries, this first index injury can only be taken to be the first (recorded) time presentation of symptoms as there is no way to determine loss of structural and functional integrity of body tissue that remains asymptomatic, and asymptomatic pathology in the sporting population is high.

Depending on the nature of the first index injury, it may not be reasonable to assume that all subsequent injuries will be related to that one. If an athlete sustains several injuries, and these are recorded in sequence, then it is possible that there will be different index injuries. This is analogous to the competing risks issue that arises in survival analysis where people can be at different index injuries. This can be contrasted to the 12% shown in table 2, where injuries were categorised according to the Hamilton et al scheme.

**APPLICATION OF THE SIC MODEL TO SPORTS INJURY DATA**

The suitability of the SIC model was assessed with sports injury data from the Preventing Australian Football Injuries through eXercise (PAFIX) study. Full details of the study methods and data collection processes have been published elsewhere as has a specific description of the nature, body region and cause of these injuries. All injuries were assigned a diagnosis code according to the internationally recognised Orchard Sports Injury Coding System (OSICS-10).

The OSICS-10 codes assigned to each injury were used to categorise them, and decisions were supplemented with the body region and nature of injury codes, as well as text narrative descriptions of the injury events. All injuries for a given person were ordered by date of occurrence, to provide a temporal sequence of injuries for each player. The first injury was taken to be the first index injury against which the SIC model was applied, but it was also recognised that there could be more than one index injury within a given series of injuries in some athletes.

Over the 2007 and 2008 playing seasons, the PAFIX study prospectively recorded 1082 injuries sustained by 1564 community Australian football players. A total of 36% of all injuries sustained at least one injury (range 1–8).

All injuries were ordered sequentially according to the date of their occurrence to determine the first index injury and all other subsequent injuries following it (either another index injury or related injuries). A subsequent injury was defined to have occurred in the subset of players with two or more reported injuries. Overall, 469 subsequent injuries were sustained by 282 players. According to the Hamilton et al sports injury taxonony, 12% of these subsequent injuries were recurrent and hence likely to be dependent on previous injuries (table 2).

Table 3 classifies the 469 subsequent injuries in the PAFIX dataset according to the SIC model categories, relative to the first index injury recorded in the dataset (ie, the first index injury) and other identified subsequent index injuries. This table shows what the model looks like when applied to a sports injury dataset and provides the Bayesian posterior distribution of the model categories. The most likely reason for there being no cases coded to category 6 and this dataset recorded injuries to players over one playing season only. Collectively, 15.6% of all subsequent injuries in the PAFIX community Australian football injury dataset would be coded to categories representing injuries that are directly related to previous index injuries. This can be contrasted to the 12% shown in table 2, where injuries were categorised according to the Hamilton et al scheme.

**Multiple index injuries**

One of the challenges with coding injury data from sequences of this type collected over a period of time (eg, a playing season) is identifying index injuries as the ‘subsequent’ injury status always needs to be determined in relation to an earlier injury in a series. When players only sustain two injuries, the first chronological injury is taken as the index injury and the

<table>
<thead>
<tr>
<th>Number of injuries per player</th>
<th>Number of players</th>
<th>Total number of subsequent injuries</th>
<th>Number (%) of subsequent injuries according to injury category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>New</td>
</tr>
<tr>
<td>2</td>
<td>161</td>
<td>161</td>
<td>128</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(80%)</td>
</tr>
<tr>
<td>3</td>
<td>86</td>
<td>172</td>
<td>133</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(77%)</td>
</tr>
<tr>
<td>4</td>
<td>18</td>
<td>54</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(67%)</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>32</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(71%)</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>25</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(80%)</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>18</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(55%)</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(71%)</td>
</tr>
<tr>
<td>Total</td>
<td>282</td>
<td>469</td>
<td>355</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(76%)</td>
</tr>
</tbody>
</table>

Cases in which the players sustained only one injury are not included because they were not ‘subsequent’ to another injury. Local injuries were classified when subsequent injuries had an Orchard Sports Injury Coding System (OSICS-10 code) to the same body region but a different pathology. Recurrent injuries had exactly the same four-digit OSICS-10 code.
subsequent status of the second injury is determined in relation to that index injury. However, when players sustain more than two injuries, it is possible that the first injury is no longer the appropriate index injury for subsequent injuries. For example, one player sustained the following four injuries sequentially by date order: fracture of finger (OSICS-10: WFFX), fracture of finger (OSICS-10: WFFX), lower leg muscle injury (OSICS-10: QMXZ), lower leg muscle injury (OSICS-10: QMXZ). It was considered that the second injury was subsequent and related to the first one but that the third injury, while subsequent to the first two, was not related to either of them. The fourth injury was unrelated to the first two but related to the third injury, and so the third injury was taken to be a second index injury for the purposes of assigned clarification codes.

Table 4 summarises the number of different index injuries found in the PAFIX data, according to the cases with specific overall numbers of injury. For example, of those players with three injuries, 50 (58.1%) had three unique/index injuries, 34 (39.5%) had two index injuries and 2 (2.3%) had one index injury. Similarly, of the players with two injuries, just over one in seven players (15.5%) had only one index injury, with the second injury being related to the first one.

### Table 4 Distribution of the number of unique index injuries in the sets of injuries sustained by community football players over one playing season

<table>
<thead>
<tr>
<th>Overall number of injuries</th>
<th>Number of players with this number of index injuries</th>
<th>Number of players with this number of index injuries</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>1</td>
<td>0 0 0 0 0 1</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>0 2 1 0 2 0</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>0 0 2 0 1 2</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>0 2 0 3 3</td>
</tr>
<tr>
<td>4</td>
<td>18</td>
<td>0 4 9 5</td>
</tr>
<tr>
<td>3</td>
<td>86</td>
<td>2 34 50</td>
</tr>
<tr>
<td>2</td>
<td>161</td>
<td>25 136</td>
</tr>
</tbody>
</table>
considered in many studies. The fact that there are so many index injuries casts further doubt on studies which only report the first injury or time to first injury, as these are clearly ignoring valuable data on many injury cases. Coding of data according to the SIC model shows that more injuries (15.6%) were directly related to previous index injuries than would be identified by using the Hamilton et al\(^4\) injury classification schemes. Identifying all related injuries is also critical to better management of injuries.

Any modelling of subsequent injury is dependent on the data collected in sports injury studies and two key issues require attention. Robust data collection techniques require attention if this model is to work adequately. Ensuring that there are clear definitions of recurrence, aggravation and exacerbation, as are provided by the SIC model, would help in identifying if the injuries were new index injuries. This paper used a Bayesian statistical modelling approach\(^5\) to underpin the SIC model, and its ability to accurately model sports injury data depends on both of these factors. As with any Bayesian statistical model, it will be important that the SIC model is further tested with other prospectively collected sports injury data by other researchers.

Most of the subsequently published sports injury studies have been restricted to one playing season because players are not exposed to injury risk when they are not participating in their sport. This is an issue in aetiological studies because a higher risk of subsequent injury could manifest in the next playing season, and therefore restricting studies to only one playing season probably underestimates the risk of subsequent injury.\(^6\) The few studies that have followed players over consecutive seasons have then treated each season as an independent observation period, and accumulated injury counts over all seasons. Moreover, injuries that do not occur during a playing season are often ignored, even though it could well be expected that injured players would have their treatment and conditioning regimens maintained or even increased in the off-season.\(^7\) Injuries that occur in the ‘off-season’ or ‘pre-season’ are not reported in incidence studies that only report new injury occurrences during a playing season, even if they could influence subsequent injury during that season. There is a clear need for sports injury studies to incorporate injury experiences over more than one season, as well as during the off-season periods, if they are to describe the incidence of subsequent injury accurately over time.

A better understanding of the nature, causes and outcomes of subsequent injuries will have major implications for sports medicine clinical practice because returning players to preinjury function, ensuring recovery of their physical functioning and preventing future injuries are the major goals of treatment. From the preventive health, clinical treatment and injury management perspectives, subsequent injuries are problematic for players, as well as for the coaches who need to manage their training and participation workloads and the clinicians who manage the rehabilitation and recovery from those injuries.\(^8\) Importantly, just because an injured player has no symptoms does not mean that they have had full tissue recovery.\(^9\) \(^26\)

In conclusion, the recognition and handling of subsequent sports injuries pose a substantial challenge for the sports medicine field—both in terms of injury treatment and in epidemiological research to quantify them. This study has found that about one in six (or 15.6%) of all subsequent sports injuries should be coded to categories representing injuries that are directly related to previous index injuries. Moreover, injured sports participants can sustain both recurrent injuries and the same injury again, and either can be dependent upon earlier index injuries. Importantly, this study has also shown that it is likely to be inappropriate to use the same index (eg, only the first injury in a chronological sequence) for all subsequent injuries. Together, this justifies the applicability of the SIC model through its ability to allow for multiple different injury types within players that recognises that the within-player dependencies are likely to be stronger for recurrent (exact same) injuries or injury exacerbations, rather than multiple (different but new) injury outcomes.

In summary, this paper sets out a holistic way of reviewing subsequent injuries that has strong clinical application and relevance. The SIC model needs further evaluation through application to other datasets, but the only way to do that is to put it out for evaluation by other epidemiologists. There will still be challenges in knowing how to best analyse the data classified according to the SIC model, but the frailty survival model approach recently advocated by Ullah and colleagues\(^5\) \(^27\) and multistate process survival models show promise.

**What are the new findings?**

- This paper presents a new model to categorise subsequent (multiple, recurrent, exacerbation or new) sports injuries, derived from statistical considerations of dependencies between different injuries and sports medicine clinical considerations.
- Greater capacity to classify subsequent injuries will allow improved understanding of the effect of initial injury on re-injury and injury exacerbations for both clinical management and injury prevention.

**Acknowledgements** Kathy Diamantopoulou (ACRISP, MIRI) undertook the data extraction and some data coding. Tom Ranger (Department of Physiotherapy) also assisted with the injury coding. The categorised data were extracted from the PAFIX database, which was generated through a separate NHMRC project grant (ID 400937, chief investigators CF Finch, DG Lloyd and BC Elliott; project coordinators D Twomey and T Doyle). Rory Wolfe and John Orchard are thanked for their comments on a draft version of this model. This work was conducted as part of the research programme of the Australian Centre for Research into Injury in Sports and its Prevention (ACRISP), one of the International Research Centres for Prevention of Injury and Protection of Athlete Health supported by the International Olympic Committee (IOC).

**Contributors** CFF conceived the study, led the development of the new model and had major responsibility for the writing of the paper. IC contributed to the development of the model, coordinated the injury diagnosis classification and contributed to the writing of the paper.

**Funding** CFF was supported by an NHMRC Principal Research Fellowship (ID 565900). The Australian Centre for Research into Injury in Sport and its Prevention (ACRISP) is one of the International Research Centres for Prevention of Injury and Protection of Athlete Health supported by the International Olympic Committee (IOC).

**Competing interests** None.

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

**Open Access** This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/3.0/.

**REFERENCES**

et al