Reviewer's report

Title: The development of an Evaluation Framework for Injury Surveillance Systems

Version: 1 Date: 22 December 2008

Reviewer: Ronan Lyons

Reviewer's report:

Major Compulsory Revisions:

1. The topic of this paper interests me enormously. However, I found it quite a difficult paper to read and to follow. I think this paper needs to be substantially edited to allow the reader to follow the logic more easily. It was not until the very end that I realized that the answer to which variables were included in EFISS was to be found in Tables 4 and 5.

There are several aspects of the paper where I struggled to follow the arguments presented. On page 6 there is mention of calculating means, medians, standard deviations and interquartile ranges. Later on there is a discussion which shows that means and SDs were used. There was no mention of why means had been chosen over medians or whether subsequent analysis required the use of either.

Minor Essential Revisions:

1. It appears that a cut off of 5.9 was chosen to identify a ‘high’ score but no discussion or justification for this decision is given.

2. The same could also be said for the decision to base ‘consistency’ on <1, 1-2 and > 2 SD. These decisions appear to have considerable implications, given that the sample size was n=7 and a single different answer could mean a change of category.

3. I also found it odd that the ‘Type of data collected is adequate for injury surveillance’ was regarded by the panel as being of ‘low importance’ whereas concepts of ‘positive predictive value’ and usefulness’ were regarded as being of ‘high importance’. It is difficult to envisage what is meant by PPV in injury surveillance. How does this relate to concepts of sensitivity and specificity in this particular use? Does sensitivity refer to the ability to detect all injuries in samples attending emergency departments or admitted to hospital or to all injuries in the community, whether brought for hospital treatment or not? What about ‘representativeness’ – does this refer to the sample being representative of all injuries treated in the hospital or to all occurring in the community served by the hospital or system? ‘Usefulness’ appears to be judged (Table 7) on the proportion of variables in unspecified core and optional data sets. These issues should be discussed in the paper.
4. Page 8 mentions that the aim of the Delphi round was to ‘review the proposed definitions for the eleven characteristics which had not been previously defined’ and states that these are in table 2 but Table 2 does not include the definitions – they are later found in Table 5. Presumably Table 5 was what was shown to the panel. I wonder whether the panelists were provided with additional information not presented to the readers.

In conclusion, I struggled somewhat to understand the flow of this paper. I suspect that the basic idea is sound but it is difficult to follow the logic at times.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

'I declare that I have no competing interests'