Reviewer's report

Title: Validation of triggers and development of a pediatric trigger tool to identify adverse events

Version: 2 Date: 1 September 2014

Reviewer: David Inwald

Reviewer's report:

Major compulsory revisions

1. The Methods section does not explain the results reported in the Results section adequately. Methods "Two stage review process" describes methodology in which notes are screened for triggers, then the notes are reviewed and a decision is made as to whether or not AEs are present. In Results, the initial three paragraphs reflect this methodology. There then follows the paragraph "Trigger outcome" which apparently reports a different set of results which the Methods section fails to explain. For example, in the initial part of the Results section we are told that 296 records contained 1066 potential AEs. From the Methods my assumption is that this means 296 records were identified by the RNs to contain triggers and on review a certain number of AEs were identified. In "Trigger outcome", however, we are told that triggers were identified in 417 records. These results are not apparently consistent with each other. Please clarify the Methods and/or Results.

2. Methods "Reliability and validity" is unclear. Usually kappa statistics are used to report consistency between reviewers. Why was this methodology not used here? Please clarify.

3. Results paragraph 3 which describes the agreement between reviewers is very unclear and difficult to follow. Please revise.

Minor essential revisions

1. It is not clear why the dataset could not be reanalysed using the final trigger list to produce an internal validation. Surely this would be a very useful exercise. Can this failure be explained in the Discussion please.

2. Recently in the UK a paediatric trigger tool has been developed and validated (see attached files). Some reference should be made to this and to some of the other paediatric trigger tools which have been developed around the world. This should include a rationale as to why there is a need to develop new tools such as the one the authors present in this MS rather than to validate previously developed tools. ie why do we need more and more tools which do the same job?

Discretionary revisions

1. The MS is very long; 4911 words. In particular the Discussion is very wordy. If possible it should be shortened.
Level of interest: An article of importance in its field

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.