Reviewer's report

Title: Safety checklists for use by medical care teams in acute hospital settings

Version: 2 Date: 27 May 2011

Reviewer: Karena Hewson

Reviewer's report:

• General Comments:

The question posed by the authors is well defined. It covers a topic that has applicability to a wide range of acute hospital settings, contributes to the evaluation of patient safety initiatives that may be of use in acute care settings and defining areas for further exploration. Importantly, this topic has not yet been explored via a systematic review, and is timely given an increase in the amount of projects and studies that are looking at implementing safety checklists with the aim of improving care delivered to patients.

The methods are appropriate and reasonably well-described (see comments below). Appropriate databases were used for the literature review. Appropriate inclusion criteria, though exclusion criteria not explicit (see Minor Essential Revisions). Good description of quality assessment and the inclusion of the critical appraisal framework was useful.

The results were generally well-presented and described. Only a few questions and comments (see below) made for clarification purposes. Good description of study quality and summary of study findings.

The manuscript does adhere to the relevant standard for reporting for the most part, with the exception of there being no discussion of this systematic review’s limitations.

The discussion and conclusions are well-balanced and adequately supported by the data (although I have a few questions and comments relating to discussion section which are outlined below). Good discussion of issues arising from the systematic review, particularly the unknown impact of education and training on the effectiveness of checklist implementation, and whether checklists were used properly. Another good discussion point was the lack of validation of checklists prior to use. For those that have done some work to make the checklist a valid tool to use, the methods used to date are not robust. The cautionary note provided by the authors re: extrapolation of short-term outcomes from studies to longer term predictions about effectiveness is warranted – though long term evaluation and sustainability could also be incorporated under future research.

The authors clearly acknowledge the work upon which they are building, although it is not entirely clear how this came about i.e. was it part of a pre-determined research/project plan (see Discretionary Revisions). Although the
title of the paper is an adequate description of the study type & method, it doesn’t convey the findings i.e. limited evidence pertaining to the effectiveness of safety checklists (see Minor Essential Revisions). The abstract accurately conveys what has been found, however it is too long and requires revising (see Major Compulsory Revisions below).

Overall, the writing is acceptable – only a few areas for improvement have been suggested for revision.

• Major Compulsory Revisions:

Abstract:
1. At 433 words the abstract is too long (author guidelines state maximum is 350 words). Needs to be more succinct – particularly the results (i.e. stick to key findings) and conclusion (should only be 1-2 sentences max.). Examples of where words could be cut down include:
   - Consider deleting 5th sentence under ‘Results’ section i.e. ‘Measured outcomes were diverse... etc.’
   - Consider deleting 7th sentence under ‘Results’ section i.e. ‘This means the results of the studies...etc.’
   - 2nd sentence under ‘Conclusions’ section i.e. ‘The included studies suggest...’ – don’t need to repeat the level of bias, keep in results. A better sentence may read something like – ‘The included studies suggest some benefits of using medical checklists to improve protocol adherence and patient safety, but due to the risk of bias evidenced, they should be interpreted with caution.’

Discussion:
2. Need to state the limitations of this systematic review.

• Minor Essential Revisions:

Title:
3. Include the key finding into the title of the paper.

Method:
4. What are the exclusion criteria? It appears that multi-faceted interventions such as those used by DuBose et al (2008) and Wall et al (2005) were excluded. The exclusion criteria need to be clearly stated under the ‘Inclusion and exclusion criteria’ heading.

5. The process of selecting and appraising the studies is unclear. Did both reviewers independently select and appraise each of the studies? Or did one select and both did the appraisals? How were the two appraisals brought together? How were any conflicts dealt with?

6. The last sentence under the ‘Quality assessment’ heading is long and a bit unclear – would suggest that authors review for clarity.

Results:
7. Re: sentence ‘From all the found articles, 224 full text articles were retrieved for review’ - Please clarify whether articles that were available in full-text on the database were the only ones retrieved? Or did you obtain articles that fit the criteria regardless of whether they were available online or not e.g. through library services?

8. I strongly recommend noting the difference in how LOS was calculated & evaluated across the different studies – particularly in ICU settings where multiple studies all have a different way of looking at LOS. For example, Agarwal showed LOS based on a calculation of the number of days a patient had been in the ICU at midnight which would exclude those patients discharged during the day (i.e. a very crude and far from accurate measure). This could be done in the results section (i.e. text) and/or in Table 2.

9. It may also be worth mentioning the problem with calculating LOS as a mean e.g. Narasimhan assumes LOS is normally distributed, though this is often not the case – particularly for ICUs (i.e. positive skew due to short stays).

10. Pronovost only evaluated LOS from implementation of the daily goals sheet onwards, so didn’t really even have a historical control for comparison. I think this could be noted somewhere.

11. In Table 2, under the ‘Intervention’ column for the Weingarten (2004) study, authors write that ‘it was unclear if the interventions were the same across hospitals’. I conversely think it was clear that the interventions were NOT the same across hospitals. On page 161 of the Weingarten paper there are two tables summarising the number of hospitals by number of interventions implemented and the top 10 interventions and intervention combinations. The intervention pertaining to checklists also included other types of forms and reminders. The authors acknowledge that they were not prescriptive on content, method or implementation strategies, and that hospitals chose the intervention strategy that was most appropriate for their institution. I suggest altering the wording to more accurately reflect the nature of the interventions used in this study.

References:

12. Reference no. 4 is the incorrect reference – I believe it is supposed to be ref. 13 instead.

13. Reference no. 12 appears only on p. 10 in the discussion section – it has not been included elsewhere in the paper e.g. results of review, so not sure if it is meant to be there.

14. Reference list and in-text referencing needs to be fixed in light of the above revisions.

• Discretionary Revisions:

Method:

15. I think ‘All EBM’ should be ‘All EBM Reviews’

16. In the first sentence under ‘Inclusion and exclusion criteria’, suggest using
‘included all’ instead of ‘was’.

17. Second sentence under ‘Inclusion and exclusion criteria’ needs improving. Suggest ‘The intervention was care given with the use of safety checklists that addressed safety concerns, which were applied to patients by medical care teams, including junior medical staff.’

18. Third sentence under ‘Inclusion and exclusion criteria’ – suggest use ‘provided’ instead of ‘given’.

19. I’m not sure if the statement in ‘Missing data’ is relevant, or whether it needs to be made clearer as to what sort of data you sought from the authors of included studies.

Results:

20. Try to avoid repeating the phrase ‘study types’ twice within the same sentence (see first sentence under ‘search results’).

21. Suggest following changes to 7th sentence under ‘search results’ i.e. ‘Four clinical settings were covered, including five studies in the intensive care unit (ICU).’

Discussion:

22. I wonder whether the point around validation of checklists requires a little more discussion, particularly on the weaknesses of the studies that had attempted some validation, and what is required of further studies (and mention this more clearly under ‘implications for further research’)?

23. As per my comments re: LOS in results section, I think LOS requires further discussion – particularly the impact of checklists on patient outcomes and whether this is an adequate measure of quality improvement. There is some mention of this on p.10, but I think it is pretty clear that outcomes were NOT uniformly defined across the studies. There are arguably better markers of patient safety than LOS.

24. Re: sentence ‘More RCTs are needed in this area to increase the level of evidence’. Are the authors suggesting that RCT study designs are the only acceptable way of studying this topic, or are other robust study designs/methods also called for? See Jean Louis Vincent’s commentary in Crit Care Med 2010, v.38(10) for example.

25. Last complete sentence on p.8 ‘This may suggest that the setting...’ etc – is unclear, suggest that authors revise.

26. Re: 5th sentence in the second paragraph on p.9 ‘It is uncertain if other factors (e.g. new policy directives) could have influenced the behaviour of staff in caring for patients.’ – perhaps also consider the impact of the unit’s safety culture (see Bryan Sexton & Peter Pronovost’s work in ICUs).

27. The 6th sentence in the second paragraph on p.9 is unclear, consider revising.
28. Thought the ‘Implications for further research’ section could be stronger. Try to address the limitations in the studies reviewed e.g. outcome measurement, link between process and outcomes evaluated, ensuring adequate controls, patient selection, study designs, evaluation of checklist use, implementation model used, studies that include adequate validation work and evaluate the impact of extraneous variables such as safety culture of the unit and policy changes. Also consider incorporating long term evaluation and sustainability under this section.

29. Not sure of the value add re: refs 16 and 17 to implications for further research. The Wieser (2010) study appears to be a sub-study of the Haynes (2009) one. And the descriptive paper by Lyons does little to strengthen the argument for better study designs (retrospective or prospective). Consider deleting and replace with directions for further research (as per previous comment).

30. Consider writing a lead-in statement pertaining to the use of technology in the clinical setting prior to the future research directions of Southern Health, as it currently reads as an after-thought rather than a part of the research plan or process.

**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