Reviewer’s report

Title: PICO, PICOS and SPIDER: A comparison study of specificity and sensitivity in three search tools for qualitative literature

Version: 2 Date: 8 May 2014

Reviewer: Morwenna Rogers

Reviewer’s report:

Overall
The question is well defined and the topic is worth investigating (due to problems identifying qualitative literature, as identified by the authors). I think it is worth emphasizing however, that although PICO, PICOS and SPIDER are potentially useful tools for helping to formulate search strategies, they are not methods by which researchers working on reviews are beholden. The most useful resource for devising and running search strategies is the expertise of an information specialist, and I notice that although the advice of an IS and Librarian were sought in devising the detailed search strategy, they are not listed either as authors or acknowledged at the end of the paper, which is disappointing. However, I understand that many research teams do not have access to information specialists and therefore tools such as the three examined here are useful aids. Therefore I believe this study does add value to search methods work and the findings are worthy of publication.

Major compulsory revisions
1. PICO, PICOS and SPIDER are useful tools in helping to formulate a search strategy. However, I think in places the paper implies that these are more than simply tools that assist with a process. For example, in the background (paragraph 2), PICO is described as a ‘traditional model of searching’. Rather, it is a tool designed to aid in the structuring of clinical questions that can also be used to assist with keyword selection. I recommend that this be clarified.

2. The most useful resource for devising and running search strategies is the expertise of an information specialist, and I notice that although the advice of an IS and Librarian were sought in devising the detailed search strategies, they are not listed either as authors or acknowledged at the end, which would have added strength to the paper. It would perhaps be useful to highlight that the tools being studied can aid in the search process especially where there is an absence of an experienced information specialist in the review team.

3. There is limited discussion on the choice of search terms used and whether these had any bearing on the results. For instance, I noticed in the Cooke study that under the D component of the tool, the additional terms ‘questionnaire’, ‘survey’, and ‘focus group’ are used. Was there any reason why they were omitted here? Could this have had an effect on the results?
4. The tables are mixed up. Under method; search strategy, PICO and SPIDER are described in the text with instructions to insert tables 2 & 3, which are PICO and PICOS, rather than PICO and SPIDER (which is Table 4). Under Results: ‘see Table 3’ is not the correct table (I'm not sure which table is correct here – is there one missing?). Under ‘proportion of relevant articles’ where Table 4 should be inserted – I think this should be Table 5 (please check).

5. Were all 14 named articles of relevance checked for presence on all three databases at any point (CINAHL, in particular)? If so it should be mentioned in the methods. If not I would recommend this is done and stated. Otherwise this could be a reason why the papers weren't caught (rather than the failure of the strategy/tool)

6. At the end of the results section there appears to be a heading ‘Sensitivity and specificity’ with no text underneath. Is this a leftover heading or has part of the paper not yet been written?

7. Discussion: paragraph 2. It is stated that the PICOS model fits as a halfway point in both sensitivity and specificity between the other two tools. I think this is misleading – it was a lot more specific than the PICO search, but not much more sensitive than the SPIDER search (it only found 1 additional relevant hit). I would recommend that this is re-phrased.

8. Discussion: final paragraph (and also the last line of the conclusion). I cannot see how PICO or SPIDER relate to other aspects of the search process mentioned (citation searching and reference checking). This should be either explained or removed.

9. Conclusions: As above, I’m not sure PICOS is a compromise between the two others. I think the study has found that it results in a much more specific search than using PICO, for this particular topic, but it still has poor sensitivity; only a little better than the SPIDER model. I agree with the authors that more direct comparisons between PICOS and SPIDER are required in order to establish which one is the more appropriate tool in situations where time and resources are limited.

Minor essential revisions

1. PICO, PICOS and SPIDER are described initially as search tools but then later on as models. This should be made consistent throughout the paper.

2. Background: paragraph 2. Sentence starting ‘Due to its target…. ’ is confusing. Does this actually mean that due to its target literature base, some of the elements of PICO i.e. Comparison and Intervention are not relevant to qualitative literature? Or that due to its literature base search terms derived from the Comparison/Intervention elements of PICO (e.g. ‘control group’ or ‘intervention’) will not locate qualitative literature? Clarification is needed.

3. Background: paragraph 5. I recommend that the components of the acronym
SPIDER are stated here as they are for PICO and PICOS above.

4. Table 1 needs to be linked to information about the topic (multiple sclerosis) in the text before inserting the table otherwise it’s not clear why the MS search terms are in column 1.

5. Methods: paragraph 1. I’m not sure what is meant by ‘no time limit was imposed on searches’. Should this be no date restriction? Please clarify.

6. Methods: paragraph 2. The formatting of the Sinfield reference is not consistent with the rest of the paper.

7. Methods: paragraph 3. Was there a particular reason to exclude conference abstracts from the study? These could be a source of unpublished qualitative literature. If they wouldn’t have come up anyway (e.g. if they are not routinely found on CINAHL, MEDLINE or EMBASE) then there is no need to list it as an exclusion criteria for the purposes of this study.

8. Methods; search strategy. In the first sentence, mesh should be MeSH. However, as MeSH is specific to MEDLINE I would suggest just stating ‘…search terms used a mixture of medical subject headings and keywords’

9. Results: paragraph 2. The sentence starting ‘the time spent searching’ needs clarification. Is this the time to run the searches or the time spent running the searches and carrying out the screening process?

10. Results: paragraph 3. The sentence starting ‘two authors reviewed...’ should be in the methods section.

11. Results: paragraph 4. The sentences starting ‘For the PICO model...’ it should read ‘title and abstract stage’ throughout in order to be consistent (rather than ‘at the title stage’).

12. Results: paragraph 4. The grammar is wrong in the sentence starting ‘For the PICO model...’. the comma after ‘these’ and before 34 should not be there as it changes the meaning (34/34 instead of 12/34)

13. Results: paragraph 9. First sentence – 3articles rather than 3 articles

14. Results: paragraph 10. Sentence starting ‘Only two articles....’ It would sound better if it was ‘found on both MEDLINE and EMBASE’ rather than ‘across MEDLINE and EMBASE’

15. Results: paragraph 10. Sentence starting ‘Differences were found...’ is confusing. I have translated this as there were differences in the number of relevant papers identified both on each database and by using different models? Could this be rewritten more clearly?

16. Discussion: paragraph 1. In the sentence starting ‘As previously identified...’ it is not clear if the eleven studies missed were in the Cooke study or this study.
17. Discussion: paragraph 2. Sentence starting ‘This might be applicable….; the formatting of the reference [30; 3] is inconsistent with the rest of the document

18. Discussion: paragraph 3. Sentence starting with “This issue…." doesn’t flow well from the preceding paragraph – it is unclear what the issue is. Can this be clarified? The following sentence starting ‘The low overlap…’also doesn’t flow well and is therefore hard to understand what is meant. I think it can be stated more clearly here that the PICO model identified all the papers and therefore should be the preferred choice where time is not a factor.

19. Discussion: paragraph 6. Paragraph starting ‘The word qualitative..’. Please clarify whether this means all databases and models in this part of the discussion, or just PICO

20. There are multiple formatting and consistency errors in the references list that need sorting out. Some but not all: Ref 3, Volume number is in italic (also in ref. 6); Ref 8, the last author is followed by a full stop, not a colon like the others; Ref 10, the year is in italic etc. etc.

21. Please check the spelling of Cooke throughout – it is misspelt ‘Cook’ in places.

Discretionary revisions

1. Background paragraph 6 starting ‘Cooke et al.…’, ‘tested’ sounds clearer than ‘replicated’. I suggest amending to either ‘the SPIDER work was replicated’ or ‘the SPIDER tool was tested’

2. Methods paragraph 2: I would drop ‘In this article’ from the start of sentence 3 as this is explicit.

3. Results: Why are the results from CINAHL listed first for the 3 tools? I would consider this is the least relevant database to the study (and therefore results given last) despite the lack of hits from it being an interesting additional finding.

4. Results: paragraph 8. The sentence starting ‘The SPIDER tool identified….’is a little clumsy. This could be more clearly written e.g. ‘The SPIDER tool identified 6 relevant articles out of 94 articles across all three databases (6.38%) compared to…’

5. Results: paragraph 10. Sentence starting ‘CINAHL Plus…’ looks like repetition here from previous paragraphs and could be left out. If not, the phrase ‘in this article’ in the middle of the sentence is unnecessary.

6. Results: paragraph 10. Final sentence starting ‘Studies not identified…’ would sit better in the discussion rather than results section.

7. Discussion: paragraph 8: Interesting that Google Scholar is suggested here. I am not aware that this can be comprehensively searched in the same manner as other databases so personally wouldn’t make this suggestion for further research.
Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Needs some language corrections before being published

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