Reviewer's report

**Title:** Diagnostic accuracy of the STRATIFY clinical prediction rule for falls: A systematic review and meta-analysis

**Version:** 2 **Date:** 28 February 2012

**Reviewer:** Geeske Peeters

**Reviewer's report:**

**General comments:**
This paper summarizes the literature regarding diagnostic accuracy of STRATIFY, a clinical prediction rule (CPR) to predict people at high risk of falls. Falls prediction is an important but complex topic. The methodology is adequate and in line with current guidelines. Although this paper is generally well written, I have some suggestions for improvement. My main concern is that the authors do not explain why this particular instrument was chosen for this paper, while many CPRs have been described in the literature. In addition, as the current results are similar to those of a previous review, the added value of the current results to the existing literature is limited.

**Discretionary Revisions**
1. Methods, page 7: The authors indicate that the PRISMA guidelines were used as a guideline for the reporting of this review. PRISMA was originally developed for the reporting of intervention studies. Cochrane is currently developing a handbook for the reporting of systematic reviews of Diagnostic Test Accuracy. Although not yet finalized, drafts of chapters can be obtained from their website.
2. Methods, page 8: “Studies which included the same ....were only included once –in– the meta-analysis.” Please add ‘in’.
3. Methods, page 8: “This tool was modified to ensure that –is– was applicable .....”: ‘is’ should be ‘it’.

**Minor Essential Revisions**
4. Although the writing style is generally good, throughout the paper, the tense changes between present tense and past tense, even within paragraphs. Most journals prefer past tense throughout except for when recommendations are made for future studies and clinical practice.
5. Background, page 5: It would be helpful to the reader to have an idea of the internal validity of the STRATIFY tool at this stage of the paper: please add here the sensitivity and specificity obtained in the original study of Oliver et al.
6. Methods, page 7: Have you considered using a selection criterion regarding the measurement of falls? How falls were measured has a important impact on accuracy and subsequently prevalence. Both the accuracy and prevalence influence the diagnostic values (sensitivity and specificity). At this stage of the
paper it is probably too much to ask to add this criterion for the selection of your papers, but perhaps a sensitivity analysis could be added including only those studies using the best available method to measure falls? It would also be helpful to add the method of measurement for each paper to Table 2 and elaborate on this topic in your Discussion.

7. Methods, page 8: Quality assessment, please add which items were modified or omitted and/or summarize which items were used.

8. Methods, page 9: “We evaluated heterogeneity visually ..... and statistically by using the variance of logit transformed sensitivity and specificity, with smaller values indicating less heterogeneity among studies.” What is a small value? In addition, the interpretation of these values is not discussed in the results section.

9. Methods, page 9: Please describe in the statistics paragraph which sensitivity analyses you did. The results section can then be reduced to describing the results of the analyses.

10. Results, page 10: The paragraph ‘Study identification’ is difficult to follow. How many papers were included: 16, 17 or 18? Please rewrite/structure to improve the flow and readability of this paragraph.

11. Results, page 10: The (mean?) prevalence of fallers reported in the included studies was much lower than I would expect in institutionalized older adults (6.27%) and the range was wide (1.1-41.3%). This is likely to be due to the variety in length of follow-up and many studies with a short follow-up. Perhaps it is worth considering calculating a weighted prevalence rate which takes into account the differences in follow-up duration. This would affect the probabilities presented in Table 4.

12. Results, page 11: “However, the CPR is more useful at ruling out rather than ruling in falls, with a higher pooled sensitivity than specificity.” I believe this conclusion is incorrect: a higher sensitivity than specificity means that the proportion correctly classified fallers is higher than the proportion correctly classified non-fallers, and thus, ruling in is more accurate than ruling out. However, both sensitivity and specificity are not great, so how important is it that one is slightly better than the other? (Please change in Discussion, page 15, accordingly)

Major Compulsory Revisions

13. Background, page 5: Please explain why the STRATIFY tool was chosen over all the other fall prediction tools available in the literature? As described in the Discussion, there were some limitations in the development phase of this tool and a previous review showed only limited accuracy. Other tools have been shown to have better internal validity. Why is it important to do another systematic review? Were there any limitations in the previous review that were overcome in this one?

14. Discussion, page 13-16: The larger part of the discussion elaborates on the limitations of the STRATIFY tool. Again, why was this instrument chosen as the topic of this paper? Also, the recommendations for improvement of this instrument include examining: (1) optimal cut-off point, (2) which predictors
should be included, and (3) optimal weighing of the items. To me, this sounds like developing a new tool. Why not recommend abandoning this tool and using one with more promising diagnostic properties?

15. Discussion, page 13-16: The discussion lacks an interpretation of the current results in the context of accuracy of other falls prediction instruments.

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