Reviewer's report

**Title:** Usability Evaluation of a Clinical Decision Support Tool for Osteoporosis Disease Management

**Version:** 1 **Date:** 29 August 2010

**Reviewer:** Jim Warren

Reviewer's report:

I heartily agree with the authors about the importance, and value, of this sort of usability testing. The paper is well researched, and I find it very hard to fault any aspect of the techniques employed. I suppose the greatest limitation is that it's a little hard to generalise from the findings, but at least it puts out there a 'data point' - e.g., these folks did pragmatic usability testing and readily found ways to improve user interaction with their system.

I have the following discretionary revision suggestions:

1. In Study 1 the authors say "the moderator simulated a patient at risk for osteoporosis and completed the RAQ on a Tablet PC" - this raises questions for me. Was it always the same simulated patient (i.e. exactly same profile)? How was this profile decided? How might the session have gone differently with other risk levels and risk factors presenting? I can see why this approach was pragmatic, but I think it is a big limitation as compared to what could be learned from a range of patient profiles (simulated or actual).

2. The authors cite Nielsen (ref 24), but I think could go a bit further in explaining the rationale for this sort of usability testing with Nielsen's argument of diminishing returns in problems found as number of testers/subjects increases.

3. Also, Ash, Berg and Coiera's work on negative consequences of health IT is cited (ref 36) but I think should be brought out more explicitly, probably in the Intro, to explain the motivation for this sort of experiment. Further to this, I find the studies in BMJ by Eccles, Rousseau and others (2002, 2003) to be a great cautionary tale of running to an RCT without looking at basic usability. All in all, I just feel the authors should hammer the point more.

4. I was disappointed to realise that the physician interaction is entirely paper based. I suppose this fits the situation in Canada (as compared to the UK, Australia, NZ, Denmark, Netherlands and other locales where family doctors work in a computer-based fashion). This leads to a couple thoughts: (a) to at least recommend the progression to computerisation ('meaningful use' in the current US trend) to better support CDSS (including the recommended functionality to defer a prompted action); and (b) Kawamoto's review of guideline success factors is cited (ref 34) but the primary recommendation - that decision support be computer-based - is essentially ignored for the physicians, which seems a basis for reflection.
5. I think the Conclusion (in the body, maybe not the abstract) should recap some of the key lessons learned for this particular software (e.g., some aspects of the refinement lessons: touchscreen less error prone than stylus). As it stands it only states the meta-level lesson that problems were found. Moreover, I think the conclusion should give some reflection of the outstanding issues of physician time, control and potential diversion from more critical tasks - these were uncovered but not convincingly resolved. The reconfirmation of these critical challenges (and with perhaps the novel twist re distraction from 'real' reason for visit) is a key message as much as the more easily dealt-with usability refinement in a narrower sense.

4.

Also, just presumably a typographic error - the 11 physicians are reported "mean age ?"

**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