Reviewer’s report

Title: Delivering stepped care: an analysis of implementation in routine practice

Version: 2 Date: 27 June 2011

Reviewer: Elizabeth Yano

Reviewer’s report:

MAJOR COMPULSORY REVISIONS

1. The utility of defining evidence based medicine and health services research for a journal whose audience is already very familiar with these concepts is limited, and does not substantively contribute to the argument being made for studying implementation of stepped care in the context of NICE guidelines. This section should be significantly shortened to get to the point of depression guidelines sooner and with greater focus and intent. The material in Box 1, in contrast, is essential to the argument being made, and should probably be in the narrative rather than segmented off in a Box (though the Figure within the Box is useful).

2. The authors suggest that the traditional delivery of care for depression has "performed adequately" and that the main issue is one of "demand for care." This reviewer is not aware of any such evidence that has been population-based in nature, especially in view of the substantial barriers to access to effective depression care management, the stigma associated with seeking/obtaining care in specialty mental health settings, and the high volumes of patients with depression who are seen in primary care rather than mental health settings. The authors need to take better account of the range of depression care modalities and approaches that stem from guidelines.

3. The abstract needs to be rewritten substantively. On the surface of it, it suggests that the paper represents a modest and limited contribution to the literature, presents conclusions that are not described in the results section, and makes broad and generic recommendations. The abstract would benefit from a stronger anchoring in the details of stepped care and the extent to which the NICE guidelines are so specific that the resulting variation observed is problematic vs. reflective of appropriate local adaptations.

4. The paper would benefit from better integration of examples throughout. (e.g., page 8, "steps relating to treatments of different intensity" would benefit from provision of at least one example; page 8, "minimal interventions" should be defined and/or an example given). Clarifying the "tension" between "stepped vs. stratified" and defining how they differ would be important. No references for the "tension" or issues the authors are trying to frame are provided, so the foundation for some of these statements is unclear. Actually, starting the "tension" sentence off as a new paragraph may be helpful and/or defining stepped care carefully and
then getting into the importance of the contrast with stratified care may be useful. If the relative importance of the two mechanisms was unclear according to NICE guidelines, why would variations be a surprise in terms of their implementation? And if the evidence per the authors for a decision between the two is limited, isn't this as it should be?

5. Background, page 9, "the extent to which" -- needs to be refined and clarified as to what the "critical issue" is. Similarly, the authors have not made an adequate case for what the "HTA/HSR effectiveness/delivery dilemma" is. Currently, the questions that the paper aims to answer do not seem particularly well linked to the introductory statements leading up to them. The questions themselves are however reasonably important ones that may provide useful information for certain audiences. For example, the introduction does not make a case for why knowing more about how patients access and flow through stepped care is important (this reviewer thinks it likely is important, but the logical flow in the introduction doesn't get us there).

6. Methods, page 10, if the authors "assisted sites in designing their stepped care models...", then wouldn't some of the variation be due to their procedures? There seems to be an undercurrent in how the paper is currently framed that "blames" the guidelines for being too specific and not specific enough at the same time, which distracts from what the paper actually may have to offer.

7. Methods, page 10, the types of data collected are listed, but they are not linked to any sort of conceptual or theoretical framework that serves as a foundation for determining whether they are the right variables and/or where there may be gaps. Analysis of 'patient pathways' using "start and end points" is not clear -- maybe an example would help. How the authors determined step directions should be explained -- again an example may help. Did the authors have any hypotheses about the structure and patient flow patterns going into this work?

8. Results, page 14, it may not be uncommon for there to be data acquisition problems in operations projects, however, as written, there is no information provided to interpret "large amounts of data were missing for many variables." Are missing data chiefly among those variables needed to describe the various samples within and across sites? Or was missing data a theme for all measures, including flow patterns? If so, the utility of the available information is admittedly difficult to interpret, limiting the potential takeaway from what is presented as findings. For example, in the absence of better patient-level data or practice-level patient information, how should readers interpret the variations found? Maybe all of these variations are appropriate, maybe not, but either way, as written, they appear to be a list of different sites' patterns, and what readers are supposed to take away is not sufficiently clear.

9. Results, page 14, "multiple routes with multiple inputs and outcomes" -- again one or more examples would anchor this statement in the narrative (in addition to what is provided in the diagrams).
10. Results, page 15, ratios of intensity -- this section is interesting but requires some anchoring in an example or other information to help readers interpret what a 22:1 ratio of intensity means at a practice and/or provider and/or patient level.

11. Discussion, page 16, paragraph 2, seems like this should be in the Results section (seems like new information). Actually, much of the Discussion section as written is easier to understand, follow and interpret than much of what is in the Results section. In fact, integration and anchoring of findings in the literature does not start in the Discussion section until page 18. Better integration and framing of the findings in what else is known about stepped care models and about implementation in general would be useful in the Discussion section.

12. Discussion, page 20, the authors suggest that the "most significant implication" of their work is "the failure of guidelines to be implemented in a manner consistent with the authors' intentions." The dilemma here is that they set up the argument up front that the guidelines are non-specific with respect to details of stepped care implementation, so consistency with them would by definition create variability as local providers insert their own interpretation and preferences into how they implement the guidelines.

MINOR ESSENTIAL REVISIONS

1. Background, end of first paragraph, "consider these..." does not fit the rest of the sentence. Greater clarity in the meaning of "wider service context" is needed.

2. Background, page 7, bottom paragraph, "as" missing between "seen" and "a reasoned."

3. Methods, page 10, "four volunteer NHS providers" -- as Implem Sci is an international journal, it would be helpful to clarify early whether this means four individual clinicians or four individual practices. It is not clear in the 2nd paragraph which it is (e.g., "assisted sites"). While this becomes clear under Results, it should be made clear under Methods as well.

4. Methods, page 11, "anonymised data were submitted..." This section needs to be clarified with respect to the steps between identifying data from existing information systems vs. those collected by clinicians. To whom were data submitted? And how? And how were data anonymised? Not sure if a table that clarifies these distinctions and procedures would be helpful, but as written, it is unclear what data were obtained from which sources by whom.

5. Results, page 13, again examples would help readers better understand study findings (e.g., "low intensity treatment steps").

6. Results, page 13, please define the "voluntary sector" for your non-UK colleagues. What is a "senior mental health worker"? Please describe the "self-directed computerised treatment programme" more.

DISCRETIONARY REVISIONS
None.

**Level of interest:** An article of limited interest

**Quality of written English:** Not suitable for publication unless extensively edited

**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.