Reviewer's report

Title: Developing integrated care: a survey study to validate a four phases Development Model for Integrated care.

Version: 2 Date: 12 December 2012

Reviewer: Hubertus Vrijhoef

Reviewer's report:

In general, this manuscript presents an interesting topic in the field of integration of care. I suggest some major compulsory revisions. First, the manuscripts is about the psychometric properties of a survey and needs to be presented as such according to the steps to be taken in doing so. Currently, this is only partially the case. Second, it seems that a major finding is neglected in the interpretation of the results. The lack of agreement between the self-assessed scores with the calculation methods actually means that the experimental data do not support the calculated phases based on the previous study. Third, the discussion of the manuscript is written without making use of what is already known about this topic. This leaves the discussion rather superficial and makes it a stand-alone paper with no attempt to contribute to the ongoing discussions about the integration of care.

In more detail:

1. Is the question posed by the authors well defined?

Partially:

authors answer more questions than presented upfront. The comparison of self-assessed scores with the calculation methods (table 3) needs to be introduced as a separate research (sub)question.

2. Are the methods appropriate and well described?

Partially:

- please describe the type of study applied; I assume this study aims to describe the psychometric properties of a survey. Describe which properties are assessed in this manuscript and how.

- please clarify the rationale for the general information and resources for these variables. This also enables authors to compare the descriptive with those in other studies and discuss the relevance of study results from an international perspective;

- please explain how many items are part of the survey and how much time it takes to fill out the survey;

- please mention that study data were collected in 2008. Discuss the validity of results 4 yrs later (2012/2013);
- please describe in detail what qualitative data were collected and how these were analyzed. In methods only quantitative data collection an analyses are described, while in results also qualitative data are presented.

- please describe how the 'crucial factors' were derived and how these were measured. Derived from which literature? Collected by means of open questions?

- please describe how selection of respondents took place and discuss the power of the study. This study is underpowered for validation of survey; thumb rule is a number of 200 independent observations.

3. Are the data sound?
Partially:
- the units of analyses are presented as services, however they seem to be individuals;
- kappa scores of group coordinators with more/less coordination time are lacking;
- qualitative data get presented at various points without being related to a research questions (e.g. last paragraph 'response and characteristics');
- the presented kappa scores of .2, referring to the level of agreement between self-assessment scores and the calculation methods seem to be neglected in the conclusion and the summary. This is a serious flaw of this manuscript, as one could argue that given this result the survey lacks sufficient psychometric value against its theoretical background. Moreover, after this result the remaining of the manuscript is no longer relevant from a psychometrics perspective.
- authors need to discuss the relative large SDs in the light of the usefulness of items and/or the scoring system;
- in presenting the 2nd research question authors do not present results from ANOVA and Kruskall Walliss test nor do the present Pearson R for Figure 3 data;
- the qualitative data seem to suffer from missing data: 16 individuals have not provided data. Who are these? Why did they not provide data? This needs to be discussed in more detail.
- table 2: please explain what ‘- -’ means; 0 or Nd?; no need to present % in cells;
- table 4: In 2 out of 3 settings phase 2 takes longer than phase 1; does this imply that in daily practice phase 1 and phase 2 are one phase?
- table 5: mention N=68; explain how to read an empty cell. Was it realistic to answer respondents about the transition from 3 to 4 given their years of experience? This could be discussed as another study limitation.

4. Does the manuscript adhere to the relevant standards for reporting and data deposition? 
Partially; please see 4.

5. Are the discussion and conclusions well balanced and adequately supported
by the data?

No. This study aims to present psychometric properties of the survey. However, the study does not sufficiently describe the steps in doing so nor the requirements (e.g. power). Moreover, given the lack of agreement between experimental data and the theoretical base (calculation methods), the survey is lacking an important property. This finding is neglected in the conclusion. Further, authors frame the survey as being a ‘generic quality management tool’ without explaining how this fits in quality management.

6. Are limitations of the work clearly stated?
Partially: please see 4.

7. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished?
Partially; not for the general characteristics, power size, and crucial factors. Further, the discussion could be strengthened by means of referring to meaningful publications regarding: psychometric testing, alternative scales, development of integrated care in an international context, relevance of the survey to support the development of integrated care in a more structured manner given such a manner is desirable; one could argue for less management and more creativity as well as attention for dimensions of integration different than organization.

8. Do the title and abstract accurately convey what has been found?
No. The first part of the title is not what this manuscript is about; this is about a survey and its psychometric properties. Results should capture the main findings of this study (poor kappa values) and the conclusion needs to be rewritten accordingly.

9. Is the writing acceptable?
Yes.

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: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:
I declare that I have no competing interests