Reviewer's report

Title: Quality of Prenatal Care Questionnaire: Instrument development and testing

Version: 2 Date: 19 March 2014

Reviewer: Elisabeth Svensson

Reviewer's report:

Quality of prenatal care questionnaire

Comments to the authors’ responses to my comprehensive review of the submitted manuscript

Referring to my first comments:

As a biostatistician and expert on the measurement process, development of rating scales, evaluation of the quality of assessments and questionnaires I am very much concerned about the methodological and statistical approaches used in this study.

According to the purpose explanatory factor analysis and psychometric testing were chosen. These approaches have been very popular in some applied research fields, probably because they are computer intensive and take “care of everything”. These approaches do not take care of the type of data, the knowledge of the researcher, the qualitative responses of interest, but transform data to quantified, normally distributed figures – far away from the real information collected by hundreds of pregnant women and others.

The authors’ have made their choice of using parametric statistics and the psychometric approach to their comprehensive study. This choice is clearly stated both in the response and in the manuscript, and I do understand that they keep to their choice even with the knowledge of its consequences.

However there is no reflection of their approach in the discussion regarding limitations.

Referring to my previous comments I will just add a few more.

The aim of the questionnaire is to measure quality of prenatal care. This construct is operationalized by designing items of statements to which the women should agree or disagree. The scale categories are: strongly disagree, disagree, neither agree nor disagree, agree, strongly agree and the coding of the assessments are made so that higher score is regarded to reflecting higher quality.

The final questionnaire consists of 46 statements (items) in 6 different dimensions.

In table 4 the mean scores of the dimensions are given.
Let’s take the dimension Anticipatory Guidance with 11 statements (Table 3). These statements are important to consider for the pregnant women and hopefully also for her prenatal care. Her assessments on these items will end up with a mean score of for example 2.4 – close to “disagree”. Is this mean score really helpful for the pregnant women and the staff? Does she disagree to all items?

Table 4 shows the range of mean scores (I think) and the mean of each responder’s mean score and SD. The authors stated that these mean scores ranged from 3.84 to 4.35 (3.85 in the table 4) “indicating that women rated the quality of their prenatal care towards higher end of the continuum”.

It must be noted that the women did not scored from a continuum, the used a five-point bipolar scale of (dis)agreement, even though the authors used a psychometric approach that assumes continuous normally distributed data with known magnitude and equidistance.

(Lines 530-) Irrespective of the choice of statistical approach, the results shown in Tables 6, 7, 8 regarding stability in test-retest assessments (Phase Five) must be based on paired data. The table describes the mean (SD) scored at each time point. The statistical method used is not found in the methods section.

Line 526: Only 43% of the women completed the retest version. That is an important indicator to consider.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

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

**Declaration of competing interests:**

I declare that I have no competing interests