Author's response to reviews

Title: The assessment of developmental status using the Ages and Stages questionnaire-3 in nutritional research in North Indian young children.

Authors:

Ingrid Kvestad (Ingrid.Kvestad@cih.uib.no)
Sunita Taneja (sunita.taneja@sas.org.in)
Tivendre Kumar (drtivendre@gmail.com)
Nita Bhandari (CHRD@sas.org.in)
Tor A Strand (Tor.Strand@cih.uib.no)
Mari Hysing (Mari.Hysing@uib.no)

Version: 2 Date: 2 April 2013

Author's response to reviews: see over
Regarding the review of Kvestad et al., “The assessment of developmental status using the Ages and Stages questionnaire - 3 in nutritional research in North Indian young children”.

Thank you for the excellent and thorough review, and for the invitation to revise and resubmit our paper. We have done our best to follow all suggestions, in addition to some grammatical changes throughout the manuscript. Below you will find our response to each of your comments:

Comments reviewer 1:
It is important to have a culture-free, easy to administer, comprehensive developmental assessment tool, particularly for young children with age range zero to five years. Ages and Stages questionnaire (ASQ) is a screening tool that addresses several domains of development of young children through mother’s report. In this study the authors made a good attempt by assessing the feasibility of the translated version of ASQ in a North Indian research setting through mixed approach (direct assessment/observation and mother’s report) in a clinic setting as the mothers were illiterate and inexperienced to fill out the questionnaire. Despite presenting moderately good internal consistency of the adapted instrument, lack of some reliability and validity measures limit interpretation. The manuscript in its current form needs additional work. The authors are encouraged to address the following concerns:

1. Minor essential revisions:
   a. Please show how ASQ scores correlate with children’s nutritional (Breast feeding and Z scores) and socio-demographic variables (income/ parental education/occupation). This information will explain some validity of the measures.

   Response: Thank you, we agree that the relation between ASQ-scores and related variables could give information on the validity of the measure. However, we have assessed the relations between several child characteristics and ASQ scores in multiple linear and cox proportional hazards models. Including these analyses in the present report makes the paper too comprehensive. These results will be presented in a separate publication.

   b. Is the change of item “Does your child eat with a fork?” to “Does your child take chapatti with Dal (lenses)?” picking up the underlying construct? The original question is focusing on the ability of the child to eat with a fork and the changed one is focusing on what he eats.

   Response: Thank you for this remark. The “fork question” is a question in the Personal Social subscale. In our understanding, the underlying construct is not the fine motor ability to use a fork, but rather the ability to eat with fork. In the age range (approx. 2 years) for this
item, children in Delhi will be eating chapatti and Dal everyday. In the discussions during training with field supervisors and scientific staff in Delhi, we decided to replace to eat with fork with the ability to use chapatti to take Dal. On page 11 rule 10, we have put in “in the Personal Social subscale” to frame the fork item better.

c. Test –retest reliability score is very important information, particularly for this paper, please provide the information if available or discuss this point as a weakness.
Response: Thank you, we agree that test-retest reliability is important when evaluating an assessment tool. Within the framework of this study, this was not possible. The lack of test-retest has been discussed as a weakness in the discussion (see under).

d. Please mention about limitations of the study in the “Discussion Section”
Response: We have added a paragraph on limitations in the discussion section:

We have added the following paragraph:
“When evaluating the transference of an assessment tool to a new cultural context, test-retest reliability is of importance. Within the framework of this study, such evaluation was not possible. This is a definite weakness of the study. Furthermore, piloting of the translated questionnaire prior to the study would be preferable, and give room for further adjustments ahead of the study start based on preliminary calculations of internal consistencies, variability and constant items. These limitations of the study, together with other remarks in the discussion section should set the groundwork for further attempts to transfer the ASQ-3 to new cultural settings.”

e. In figure 1, different percentage of observation in same subscale introduced tester’s bias. Please explain what exactly this figure is representing. It is not clear how one subscale can have different percentage of observation.
Response: Thank you for your comment. The figure shows how the different subscale contains variable amount of items that the examiners could observe during sessions in contrast to items that were answered based on mothers report.

We have added text to the figure to clarify:
“For different subscales, there were variations on the number of items observed during sessions. The figure shows these differences for each subscale in percentages.”

We have also done some slight changes in the document under the paragraph in the result section:
“For different subscales, there were variations on the number of items observed during sessions, and on the number scored based on caregivers report, see figure 1 for details on the percentage of observed items during sessions. The Gross Motor, Fine Motor and the Problem Solving subscale were the subscales where most items were observed by the examiners. In the Communication and Personal Social subscales, more items are based on caregiver’s report of children’s relevant skills. The Personal Social is the subscale where the least of the items were based on the examiners observations during sessions.”

We agree that some of the differences in observations may be an observer’s bias, although other analysis (inter observer agreement) indicate that this is not the case.
1. The conclusion in the abstract is formulated stronger than in the text:

**Abstract:**
The translated and adjusted ASQ-3 “home procedure” shows promise as an easily administered assessment tool for the collection of reliable data on the developmental status in infants and young children in large epidemiological studies in low and middle-income countries, although particular care is needed to obtain satisfying alpha values in all subscales, and to ensure variability in all items when transferred to a given cultural context.

**Text:**
Our results are promising in terms of the possibility of effectively train examiners to collect reliable data in a large study, although for future utility in a North Indian research setting, particular attention must be held to adjustments of items in order to enhance internal consistency and ensure variability in all items. The report underlines the significance of close awareness to cultural adjustments when transferring an assessment tool to a new cultural context, both in terms of translation and adaptation of items and in terms of cultural appropriate administration.

Response: Thank you for this remark, see below.

2. In my opinion the conclusion should be that the procedure was found to be feasible, but also that the ASQ should be adjusted further, for instance by deleting non-varying items from the scales, or regroup items age appropriately for Indian children, in order to improve the alpha’s of the scales (not of the items, as it says in the text).

Response: Thank you for this input. We agree that the conclusions should be coherent.

The abstract conclusion now reads:
“**We found that the translated and adjusted ASQ-3 “home procedure” was a feasible procedure for the collection of reliable data on the developmental status in infants and young children. Examiners were effectively trained over a short period of time, and the total ASQ scores showed adequate variability. However, further adjustments are needed to obtain satisfying alpha values in all subscales, and to ensure variability in all items when transferred to a North Indian cultural context.**”

The conclusion in the manuscript reads:
“The present study has evaluated the feasibility of the ASQ-3 “home procedure” as an easily administered and inexpensive assessment tool for the collection of data on developmental status in infants and young children in an epidemiological study in a North Indian urban setting. Our results are promising in terms of the possibility of effectively train examiners to collect reliable data in a large study. However, for future utility in similar research setting, particular attention must be held to further adjustments of items, as well as the possibility of re-grouping items more age-appropriately, in order to enhance the internal consistency of the scales. The report underlines the significance of close awareness to cultural adjustments when transferring an assessment tool to a new cultural context, both in terms of translation and adaptation of items and in terms of cultural appropriate administration.”

In addition, we have put in your point of “regroupe age appropriately” in a paragraph in the discussion page 20:
“This last assumption gives rise to the idea that the 18 constant items are not developmentally appropriate items for this North Indian sample of infants and young children, and should be adjusted and/or regrouped age appropriately prior to further use.”

3. The information may become more interesting for the readership of this journal if some information was also presented regarding the subgroups in relation to the physical growth data – but perhaps the authors want to present this information more extensively in a different paper.

Response: Yes, we are presenting this information in a separate publication. See also response to reviewer one.

But the authors should provide some information on the % of children that showed developmental delay, as an illustration to what extent the assessment used here indeed identifies children with developmental risk or delay.

Response: Due to the lack of Indian norms on the ASQ-3 we find that the identification of delay based on the American cut-off should be done with caution. We did not have any other measures of developmental status in the study that would give information of delays among the children.

4. The authors describe that they needed to adapt the ASQ; they explain that the mirror item needed to be adjusted as mirrors are not used daily in this part of India. Nevertheless they still continue to use a mirror as one of the materials. Also, a zipper is not familiar in this area, but instead a board with magnets is used to assess comprehension of ‘up and down’; is this material familiar to the children?

Response: Thank you for making this excellent point. There will always be a balance to find the right level of cultural adaption when transferring an assessment tool to a new cultural context. During our discussions in Delhi prior to the study, we decided pragmatically to change the zipper item, although in retrospect this might seem unnecessary. However, we were already aware of the discussions on the mirror items from the ASQ manual, and based on discussions with one of the ASQ constructors (Jane Squire), we decided to keep the mirror in the item, but change the level of interaction required with the mirror.

Do the authors recommend further adaptation in this regard?

Response: Again, there will always be a balance between the need for adjustments, and
the aim of keeping as much of the original in the assessment tool. The two items causing negative average covariance were items that we during training discussed if were appropriate for the culture. Children in this area would not be used to vehicles to steer with, and the supervisors were unsure of how the stuffed animal questions would work. However, based on experiences during practice and standardization, we decided to keep these items as in the original. When looking at the calculations after completing the study, we see that these items should have been adjusted in order to enhance the internal consistency for this age group.

This has been addressed in the manuscript in the discussion:
“The problematic items should be scrutinized further in order to get an understanding to why certain items in this cultural setting show inconsistency. With further adjustments to certain items there might be a possibility to improve the internal consistency of the scales, and then increase the level of reliability.”

Did they study the alpha’s without the adapted items?
Response: Yes, we did study this. This has now been addressed in the result and discussion section of the manuscript.

In the result section we have added in page 17 and 18:
“To assess the effect of the four adapted items on the internal consistency, further analysis were performed in the relevant subscales and age groups. Removing the adapted item increased the alpha value of Personal Social, 27 month from 0.36 to 0.45, while in the communication subscale for 33 and 36 months, the values decreased from 0.81 to 0.64 and 0.92 to 0.87 respectively. For Personal Social 16, 18, 20 and 24 months the values remained unchanged.”

And in the discussion section:
“Analysis on relevant subscales when removing adapted items does not consistently lead to improved internal consistency, and thus indicating that these are not the primary cause of the poor internal consistencies.”

We have also added under the Variability paragraph in page 17, first rule:
“None of the constant items were items that were adjusted during the translation process.”

5. It is also described that the children were given time to practice with some of the materials before the assessment was done; would it not have been better if materials were used that were already familiar to the children so that practice time would not be needed?
Response: Thank you for this remark. Although it would have been preferable that all items should have been familiar to all children in the study, this would be a major challenge in the poor districts of Tigri and Dakshinpuri. Many homes are characterized by the lack of resources, and children may have very few items to play with in their daily life. Items such as blocks and beads, where therefor decided to be used in the study even though not all children were familiar with these in advanced. Due to unfamiliarity, we put in some time to practice and play with the items.

6. In the results correlations are described of subscale results with the total score; it is unclear
to me if the subscale score was excluded from the total score before doing so.

Response: The subscale score was not excluded before doing the calculations. This is in accordance with the calculations described in the reliability section of the ASQ-3 manual. We wish to follow the manual in order to compare with other studies, although we see that to exclude the subscale score may be more appropriate.

7. Discretionary Revisions
   a. Should it not be India instead of Indian in the Title?
      Response: The title has been changed to: The assessment of developmental status using the Ages and Stages questionnaire-3 in nutritional research in North Indian young children.
   b. In general the manuscript is somewhat long with repetition of several remarks, e.g. page 13 rule 14 and page 16 rule 1 concerning noting of observation of skills or scoring based upon information; or description of difficulties with the mirror items at page 11 and 19.
      Response: Thank you for these remarks. We have done changes according to your comments, as well as small adjustments throughout the manuscript in case of other redundancies.
   c. Page 5: rule 5 and 7: offers should be offer.
      Response: Thank you.
   d. Page 6, rule 10; how high was that correlation? Rule 17; interventional, should be intervention
      Response: Thank you. The correlations (from 0.76 to 0.80) have been written in the document. And rule 17 has been changed.
   e. Page 14, rule 11; suddenly a gold standard is introduced; this later turns out to be the trainer (main author); please clarify.
      Response: In page 13 and 14 (last and first rule), the first author is introduced as the gold standard during the standardization exercises. To clarify, we have put “first author” in parenthesis at page 14, rule 11 after “gold standard”.
   f. Page 17, last rule; please describe the items that cause the negative average.
      Response: We have put in the descriptions of the items.

Thank you, we hope that you find this manuscript suitable for publication.