Author's response to reviews

Title: Inter- and intrarater reliability of Minimal Eating Observation and Nutrition Form - version II (MEONF-II) nurse assessments among hospital inpatients

Authors:

Albert Westergren (Albert.Westergren@hkr.se)
Ólina Torfadóttir (olina.torfadottir@gmail.com)
Peter Hagell (Peter.Hagell@hkr.se)

Version: 2
Date: 12 May 2014

Author's response to reviews: see over
Point-by-point response to reviewer’s comments

Title: Inter- and intrarater reliability of Minimal Eating Observation and Nutrition Form - version II (MEONF-II) nurse assessments among hospital inpatients

ALL CHANGES ARE MARKED IN YELLOW IN THE MANUSCRIPT

Reviewer: Mary Hickson
Reviewer's report:
This is a clear well written report of a well designed study. Well done. This is a useful topic as the tools available for screening different patient groups in hospital is not defined and a range of tools are used. Many of which do not have adequate reliability data. thus, this contributes to the body of knowledge.

RESPONSE: Thank you.

- Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

Please include the sensitivity and specificity data in the introduction. This is important for people to decide whether to use the tool or not.

RESPONSE: This information is now provided in the manuscript.
MEONF-II has shown acceptable sensitivity (0.61-0.73), specificity (0.79-0.88) and accuracy (0.68-0.82) compared to the Mini Nutritional Assessment (MNA, 18 item version)

Please reference your supplementary data and explain what it is and link it to your paper in some way - I did not spot a reference to it.

RESPONSE: This was done in the previous version. We have highlighted it in green in the manuscript, page 3.
The MEONF-II (Appendix)

- Major Compulsory Revisions (which the author must respond to before a decision on publication can be reached)
You need to be clearer about the sample size – you decided for good reasons to use a small sample size and justified this in your methods. Then you raise small sample size as a weakness – if it is
a weakness why did you not use your fullsample of 56 patients. This is confusing. What readers want to know is whether this data is robust – on one hand you say it is, then you say it isn’t. Take a definitive line and that will help readers with less knowledge of reliability study design decide whether to use this tool confidently.

RESPONSE: Thank you, this was a really good comment. We have deleted the part in the discussion about the sample size as a weakness. Further on, we added the following in methods, pages 3-4:

The other patients were included for the purpose of estimating the pointprevalence of undernutrition risk at the hospital (not presented here).

No other changes required

Reviewer: Kylie Lange

Reviewer’s report:
A well presented article on the inter- and intra-rater reliability of a malnutrition screening tool. The authors have previously published on validity and diagnostic performance of the tool; this is the first paper to review its inter- and intra-rater reliability.

RESPONSE: Thank you.

Major Compulsory Revisions:

Abstract. It should be made clear in the abstract that while a total of 8 nurses participated in the study, each patient was scored by k=2 of these. At the moment the description gives the impression that each patient was scored by all 8 nurses.

RESPONSE: This has now been clarified.

Inpatients (n=24; median age, 69 years; 11 women) were assessed by eight nurses (intrarater reliability, two nurses scored each patient independently) using the MEONF-II on two consecutive days (intrarater reliability, each patient was scored by the same nurse day 1 and day 2).

Methods, paragraph 2 (Participants). I would like more information on the rationale behind only including every second of the
consenting subjects in the study. Why were not all 56 patients included? Or, if the aim was always n=28, why were 56 consented?

RESPONSE: Good comment. We added the following in methods, pages 3-4:

The other patients were included for the purpose of estimating the pointprevalence of undernutrition risk at the hospital (not presented here).

Methods, paragraph 4 (Data collection) & Table 2. What distinguished the "first group" of raters from the "second group"? The methods state that the assessments were done in parallel, so it's not clear to me how the nurses in each pair were designated as number "1" or "2". Was this done randomly after the assessments were completed? If it reflects the order that the assessments were completed in, could this be specified in the methods along with some details on how the ordering and timing was determined (ie, what determined which nurse completed their assessment first). Within each pair of nurses was it the same individual who went first for each patient?

RESPONSE: We have made the following clarifications, on page 4 and also made clarifications in table 2 (foot note). We hope that this will make the process description more clear.

The nurses in each pair agreed upon who should be designated to be the “first” and “second nurse” (for the purpose of interrater reliability). During day 2 the MEONF-II assessment was repeated, by the same nurse who made the assessment of the same patient on day 1, to assess the intrarater reliability.

Table 2

The “first” (n=4) and “second” nurses (n=4) conducted parallel but independent assessments of patients on day 1 and repeated their assessments on day 2 (the same nurses for the same patients as on day 1).

Methods, paragraph 11 (Analysis). Please clarify the method of ICC used. Was it really a two-way fixed model? The subjects factor is always random if we expect that the selected sample tells us something about a larger population of interest (for example, see page 574 in the referenced Schuck paper). So the choices for a two-way model would be either two-way mixed (subjects random,
raters fixed) or two-way random (subjects random, raters random). Can you please clarify what was used here?

**RESPONSE:** Good comment, thank you. We have made the following clarifications in the manuscript.

Page 6:
Regarding the ICC we used a two-way mixed model (subjects random, ratings fixed, single measurement, absolute agreement)

**Table 3:**

| g) Two-way mixed model for single measures, absolute agreement criterion. |

**Minor Essential Revisions:**

Table 1. Second footnote is incomplete or has a typo.

**RESPONSE:** It was a typo that now has been changed.

**Discretionary Revisions:**
The authors may want to consider including in the Discussion a comment on the performance of the tool at both the group and individual level. See for example, section 3.5.1. in the GRRAS guidelines.

**RESPONSE:** This was a really good point. We have added the following in the discussion and in conclusions:

> According to Kottner et al. (12), reliability values of 0.60-0.80 may be regarded as sufficient for group-level applications, whereas values of at least 0.90 are required for individual patient assessments, when important decisions are based on the assessment [12]. Thus, from this perspective the MEONF-II performs well at both a group and (in most instances) individual level and can be used for both for research purposes as well as in clinical practice.

The MEONF-II can be used in a reliable way in research and clinical practice.

**Thank you to both reviewers for valuable input!**