Author's response to reviews

Title: Factorial validity and internal consistency of the PRAFAB-questionnaire in women with stress urinary incontinence

Authors:

Erik JM Hendriks (erik.hendriks@epid.unimaas.nl)
Arnold TM Bernards (bernards@paramedisch.org)
Bart J Staal (bart.staal@epid.unimaas.nl)
Henrica CW de Vet (hcw.devet@vumc.nl)
Rob A de Bie (ra.debie@epid.unimaas.nl)

Version: 2 Date: 17 November 2007

Author's response to reviews: see over
To the Editor

BMC Urology

Dear Editor,

I am pleased to submit our revisions of our manuscript ‘Factorial validity and Internal consistency of the PRAFAB questionnaire in women with stress urinary incontinence’ (MS: 9231009991583336).

Thank you for giving me the opportunity to address the concerns of the reviewers. I have listed the changes we have made in a detailed point-by-point response and underlined the changes in our revised and uploaded manuscript including all additional (tables 1-5, one figure and one appendix).

We would like to thank the reviewers for their comments and remarks to our paper. In our reply we addressed all comments. We have justified the need for this analysis and paper and we further explained what we did and why. Furthermore we did an English-language ‘editorial’ service review. We believe that the revisions have resulted in a more consistent and readable manuscript.

All the authors have approved this revised manuscript.

Thank you for considering publication of our manuscript in BMC Urology.

Yours sincerely,

Erik Hendriks (PhD, MSc), Health Scientist, Clinical Epidemiologist
Senior Researcher at the Department of Epidemiology and
Co-director at the Centre for Evidence Based Physiotherapy and clinical guidelines (CEBP) Maastricht University, the Netherlands
**Responses to the reviewer's comments**

**Title:** Factorial validity and internal consistency of the PRAFAB questionnaire in women with stress urinary incontinence

**1. Reviewer:** Jozé Braspenninck

**RESEARCH QUESTION**
The authors’ question concerns the factor structure of the PRAFAB questionnaire and they also tried to replicate their findings on the construct validity in a broader sample. The research question is not new; some of the authors have posed it before [7]. However, the sample in the earlier study was too small to provide a definitive answer. The necessity to come up with two distinctive factors has not been addressed before. The PRAFAB is a short 5-item questionnaire. Two questions deal with the loss of urine (amount and frequency), one question is about the protection used and two questions are on the impact of the loss of urine on daily activities and self-image. All five items seem to be important clinical outcome measurements.  

**Reviewer’s comment:** Why do the authors want to group them?

**Response:** Before giving an answer we will provide a brief introduction. Urinary incontinence (UI) is a common but often distressing condition. Objective measures concentrate on the number of UI episodes, weight of pads or more costly invasive procedures in combination with standardized bladder volume assessments. Although valid, there is an increasing recognition that objective measurements of UI incontinence symptoms alone are poor indicators of the extent of limitations in activities and the psychosocial impact as a result of UI (or quality of a patient’s individual life). The impact of UI will be different from patient to patient and not necessarily related to the severity of UI (leakage) symptoms.  

Temporary or permanent stress UI symptoms are consistently associated with embarrassment, distress and anxiety, which may negatively affect social participation, intimate relationships and self-esteem and may have severe repercussions on the patient’s health-related quality of life. However, the associations are highly variable. Measuring *UI leakage severity* and its *perceived symptom impact* on health-related quality of life is therefore important both when assessing outcomes in clinical practice and in clinical decision-making.

The developers’ questionnaire (Vierhout, 1990; Mulder and Vierhout, 1989) combines these important objective and subjective aspects of UI severity in order to measure both aspects of *UI leakage severity* and *perceived symptom impact* in the 5-item PRAFAB questionnaire. The questionnaire consists of three objective items “Protection (the use of pads)”, “Amount of urine loss” and “Frequency of UI”, summarized in our study as the *Urinary Leakage Severity* items. The two subjective items of the PRAFAB questionnaire were “Adjustment of behaviour (e.g. limitations in activities) due to the UI leakage symptoms, and “Body (or self) image (e.g. bother)” as a result of the *UI leakage symptoms*, summarized in our study as the “*perceived symptom impact*” items. Thus suggesting a two-factor structure of the PRAFAB questionnaire as also proposed by the developers of the questionnaire!
Response: The purpose of our study was to confirm that the developers succeeded in their intention to include the two different aspects. Secondly, in our previous psychometric study, too small for appropriate factor analysis, we were unable to demonstrate this proposed two-factor solution as was theoretically expected: the more objective ‘urinary leakage severity’ items and the more subjectively ‘perceived symptom impact’ items.

We believe that it is important from both the practitioner’s and patient’s point of view to evaluate the severity and impact of UI on both subscales for targeted interventions, differentiated responses and informed decision-making. Multiple item questionnaires are often reduced to a single total score (as was done by using the total PRAFAB questionnaire score) but using the subscales will give detailed and relevant information on UI baseline leakage severity and its perceived symptom impact, which we consider to be important information, specifically when we are able to determine the outcome of the intervention(s) per subscale. Patients with a substantial improvement on the leakage severity scale may not be satisfied with this outcome, while other patients with the same improvement will be very satisfied. Thus the perceived symptom impact of the amount of improvement on the leakage items will be different between individual patients. As a result, these subscales may be used as separate outcomes in future research.

We have added this in the introduction and discussion section and underlined the changes in the manuscript.

Introduction section, last paragraph:

“The data in our larger prospective cohort study allowed us to more appropriately investigate the multi dimensionality and expected two-factor structure of the PRAFAB questionnaire score in a large cohort of women with stress UI (N=279). Summary indices of the two subscales ‘leakage severity’ and its ‘perceived symptom impact’ will emphasis this construct and may be used as separate outcomes in future research, instead of the total PRAFAB questionnaire score. The aim of this study was twofold. First, to investigate the factor structure and dimensionality of the (5-item) PRAFAB questionnaire score and secondly to replicate the construct validity of the PRAFAB questionnaire based on our previous findings.”

Discussion section

“The correlation between these two scale scores was moderate (r = 0.31) indicating that the scale scores measure distinct but related dimensions of the PRAFAB questionnaire and should be used as separate outcomes in future research.”

METHODS
The method is appropriate and well described. It is described in sufficient detail to replicate the work.
Reviewer’s comment: Describing a Cronbach’s alpha for two items is very unusual, an ordinary correlation will do.
**Response:** Internal consistency is a measure of the homogeneity of a scale. It indicates the extent to which items in a scale are intercorrelated and thus measure the same construct. For a useful factor analysis there should be some strong correlation among the original variables. We agree with the reviewer that describing a Cronbach’s alpha for two items is unusual and item-total correlations might also be too optimistic. Although it can be considered unusual (because it might be a rare event) it is not an obstacle for doing this statistical analysis. However, giving more detailed information and adding the results of the Spearman inter-item correlation might be a valuable suggestion, but may also be confusing if this is only done for the subscale “perceived symptom impact”. We have added the ordinal inter-item correlations of the subscales to the text of the manuscript (methods and result section) and the related table (Table 4). But this point needs also more discussion.

The main reason why we did a factor-analysis in a 5-item questionnaire is that, if we do not, a total score will be calculated, which is, based on the hypothesis not correct because the questionnaire contains two different concepts: a potential two-factorial structure (UI leakage severity and perceived symptom impact) vs. the total PRAFAB questionnaire score! Consequently, we looked at the internal consistency of the subscales by using Cronbach’s alpha. Using the Spearman correlation as an alternative for Cronbach’s alpha (maybe suggesting that ordinal correlation in this particular case might serve as a substitute for Cronbach’s alpha (in case of only two items) is not verified in our analysis or maybe not correct). The results of the Spearman inter-item correlations on both subscales were rather different when compared with the Cronbach’s alphas in both subscales. For example, the Cronbach’s alpha for ‘perceived symptom impact’ in Sample A and B is 0.84 and 0.82 resp. whereas the Spearman inter-item correlation is 0.46 in Sample B.

In this study we demonstrated the two-factor solution as hypothesized by the developers, meaning that the total PRAFAB score is not appropriate. We recommend using the subscales as separate outcome measures or looking at the individual items. In the discussion-section we discussed and underscored the importance and clinical relevance of the evaluation on the individual items.

“In Sample B the Cronbach’s alpha coefficient for the PRAFAB questionnaire subscale ‘Leakage severity’ was 0.78 with item-total correlations ranging from 0.67 to 0.85 and Spearman inter-item correlations ranging from 0.42 to 0.68, indicating that the items were rather homogenous. The Cronbach’s alpha for the subscale ‘Perceived symptom impact’ was 0.82 with item-total correlations ranging from 0.74 to 0.84 and inter-item correlation of 0.46 (Table 4). These results were quite similar to those of Sample A. The correlation between both subscales (factors) was 0.31 in both study samples.”

**DATA**
The data are sound and well verified.

**REPORTING**
The manuscript adheres to the relevant standards for reporting and data deposition.

**Reviewer’s comment:** Several references have the status ‘submitted’.
Response: One of these submitted papers “The psychometric properties of the PRAFAB questionnaire: a brief assessment questionnaire to evaluate severity of urinary incontinence in women” was recently published in Neurourology and Urodynamics (Neurourol Urodynam 2007;26:998-1007). We will delete the references with the status (‘submitted’).

DISCUSSION AND CONCLUSION
Reviewer’s comment: The discussion is very long with repeated information. More attention should be given to why we would be interested in dividing a five-item instrument into two components. The components are rather strongly related (r=0.31). The authors themselves suggest in the 5th paragraph (last sentence) and 6th paragraph of the discussion that all five elements are meaningful in themselves.

Response: With respect to this reviewer’s comment, we decided to discuss the five individual items of the PRAFAB questionnaire from different perspectives. From the perspective of a biopsychosocial approach, the 5-item PRAFAB questionnaire also covers the key elements of the bio-psychosocial model. This discussion was set up to underscore the importance of using the subscales as separate outcomes or looking at the individual items by themselves instead of only the total PRAFAB score. If this part of the discussion is found to be redundant we can remove this part from the discussion section, which will make the discussion paragraph shorter. But we believe that this discussion is worthwhile and important. There were no comments from the other reviewers on this point. Nevertheless, we are open for a good suggestion on how to handle this or we will leave it to the discretion of the editor.

Reviewer’s comment: … the components are rather strongly related (r=0.31).
Response: We discussed the relationship of the two components in the discussion paragraph: “… The consequence of the two-factor dimensionality is that this measure is able to make a distinction between different kinds of baseline severity and consequently different kinds of outcomes: ‘leakage severity’ and ‘perceived symptom impact’ of stress UI on the patient’s life. The correlation between these two scale scores was moderate (r = 0.31) but significant (p<0.01) indicating that the scale scores are measuring distinct but related dimensions of the PRAFAB questionnaire. This is in line with what we know from the literature demonstrating that the perceived symptom impact or consequence of stress UI will be different from patient to patient and not necessarily related to the severity of the stress UI leakage symptoms. One may argue that this correlation is rather strong but in fact this result is what we expected based on the literature and summarized in the introduction part as “… objective measures of UI incontinence symptoms alone are poor indicators of the extent of activity limitations and psychosocial impact as a result of UI (or quality of the patient’s life). The impact of UI will be different from patient to patient and not necessarily related to the severity of UI (leakage) symptoms …”. The correlation of 0.31 demonstrates the rather variable and moderate relationship between ‘UI leakage severity’ and ‘perceived symptom impact’ and underscored the importance of using the subscales in the PRAFAB questionnaire or as we said earlier, looking at all individual items separately.

TITLE AND ABSTRACT
No suggestions.
SUPPLEMENT 8:
Reviewer's comment: It is printed in a strange layout, but more important what is the meaning of this document?

Response: I really do not know. We have uploaded the following files: a submission letter, abstract, manuscript and tables, a figure and the PRAFAB questionnaire. It might be the endnote file. I will inform the editor about this omission.

Reviewer 2: Gerard Amarenco
No comments
No response needed

Reviewer 3: Vesna Bjelic Radisic

GENERAL COMMENT
This is an important and interesting area in clinical practice and QoL research. The results would be useful to clinicians and help patients in decision-making. I think that the application of a simple questionnaire is good, but there are at present a lot of different questionnaires.

Reviewer's comment: In my opinion there are already enough questionnaires, now is the time for using the developed questionnaires in the clinical trials.

Response: We agree with the reviewer’s statement that there are a number of valid and responsive quality of life questionnaires (QoL) for patients with stress urinary incontinence be used in research (Avery et al., J Urol 2007; 177: 39-49) and are helpful in clinical decision-making. We also agree with the reviewer’s next statement that researchers should be encouraged to only use those valid and responsive questionnaires in clinical trials, which were, unfortunately, often lacking (Avery et al., J Urol 2007; 177: 39-49). But, having said this, urinary incontinence outcome measures, including QoL questionnaires, correlate only moderately with each other at best. Furthermore, they appear to measure different aspects of the incontinence condition and are therefore not interchangeable when characterizing a patient population or assessing outcomes (Albo et al., J Urol 2007;177:1810-1814). In conclusion, it is necessary to evaluate the relationship between these severity measures and the efficacy and morbidity associated with surgical or conservative interventions, including the brief and simple outcome measures!

Despite the growing availability of valid and responsive QoL questionnaires, rates of utilization in clinical practice are still very low! Presumably these quality of life questionnaires are limited in their clinical applications for a number of reasons. For example, QoL questionnaires are time-consuming, difficult to interpret, not validated in specific patient groups or in a variety of clinical settings.
The availability of a patient’s own report with brief, simple and easy to interpret outcome measures that cover important domains of quality of life will lessen the scoring burden for clinicians and will enhance the implementation of these outcome measures in everyday clinical practice. A simple questionnaire that covers important domains of quality of life in patients with stress urinary incontinence such as the PRAFAB questionnaire or, for example, the ICIQ is rare.
Secondly, these questionnaires are often not validated in different languages. The PRAFAB questionnaire is under study in different countries as well (e.g. Brazil, Denmark).

Evaluation of the psychometric properties of these simple questionnaires is therefore crucial before being able to recommend or to implement these questionnaires for use in clinical practice (in this case, the PRAFAB questionnaire).

**REVIEWER’S SPECIFIC COMMENTS:**
As the authors said themselves, the psychometric properties of this questionnaire have gone to press. I have not read this psychometric study, but if I understand correctly, the authors have determined psychometric properties with test-retest, internal consistency etc. I suggest rejection of this study, at least until the first study is published to find out if there is an overlap.

**Response:** Our paper “The psychometric properties of the PRAFAB questionnaire: a brief assessment questionnaire to evaluate severity of urinary incontinence in women” was recently published in Neurourology and Urodynamics (Neurourol Urodynam 2007;26:998-1007). We have attached a PDF version of this paper. The PRAFAB questionnaire was developed by Vierhout (1990), has been widely implemented in the Netherlands and has high face-validity when compared to the ICIQ (3-item Q) and ISI (2-item Q) questionnaires containing the same items (in a slightly different wording) as described in the discussion section. The objective of this recently published study was to determine the psychometric properties of the PRAFAB-Q in a group of patients with stress (n=56) or urgency (n=33) UI to justify its use, which had never been evaluated before, to justify its use in both clinical practice and research. The study was designed to test the reliability, construct validity and responsiveness of the PRAFAB-Q, further defined by the minimally important change (MIC), but only in a small sample of patients. A paper in which we conducted a detailed analysis of the MIC is submitted.

**Reviewer’s comment:** This study describes only one part of psychometric testing.

**Response:** In this case we do not agree with the reviewer’s comment. In our previous paper we also studied the two-factor structure of the PRAFAB questionnaire as supposed or hypothesized by the original developers. We have now tried to replicate our findings on the construct validity in a new and broader study sample, using different patients and therapists.

Note that confirmatory factor analysis is also the method for assessing the construct validity of measures too, not a means for data reduction. Construct validity is supported if the factor structure of the scale is consistent with the constructs the instrument purports to measure. If the analysis fails to detect the underlying constructs or is inconsistent with the expectations, the construct validity is compromised.

We have added this information in the discussion section. All additions or changes are underlined throughout the manuscript.

“As a confirmatory procedure, the factor analysis is a method for assessing the construct validity of measures. The results of the confirmatory factor analysis in this study are consistent with the subscales the PRAFAB questionnaire purports to measure and supports the construct validity.”
**Reviewer’s comment:** If the first study includes all parts of psychometric testing apart from factorial validity, the present study is redundant. To my mind, a description of the factorial validity is not enough for one manuscript. I suggest that the authors put this part into the first publication.

**Response:** We have a different viewpoint with respect to this reviewer’s comment. We believe that replication of results in a new set of data, collected under different circumstances, is the proof of the pudding of ‘scientific truth’ and an important step for being confident about the measurement properties in question. Our larger study, which allows us to investigate the multi-dimensionality of the PRAFAB questionnaire more appropriately, demonstrates conflicting results compared to our previous results as might be expected because of the rather small study sample in the first (as stated in the introduction section). More explanation including references is given in our manuscript. We have discussed that already.

Based on this larger study sample we can confirm the two-factor solution as suggested by the developers of the questionnaire.

The results of our previous study demonstrated the sound psychometric properties of the PRAFAB questionnaire for both types of UI. However, we were unable to demonstrate a two-factor solution as was suggested by the developer(s).

As we wrote in the introduction to our submitted paper “The most likely explanation was that our study was underpowered for appropriate factor analyses but, on the other hand, the PRAFAB questionnaire might have only one strong underlying factor. Because of this, we used the data from our larger prospective cohort study that allowed us to more appropriately investigate the multi-dimensionality and proposed two-factor structure of the PRAFAB questionnaire score in women with stress UI (N=279). We believe that more studies are needed to replicate these findings again in new cohorts and different patient groups, etc.

**Additional response:** Factor analysis is often poorly understood by a lot of researchers and consists of a number of statistical steps, which are only possible on a large dataset. We had to start first with the exploratory analysis to explain the data and data reduction (based on our previous results the PRAFAB questionnaire might have a single underlying factor), followed by a confirmatory factor analysis (CFA) and confirmation of the a priori hypotheses and proposed two-factor structure. A CFA is primarily a method for assessing the construct validity of measures. We covered all of these steps from exploratory analysis (e.g. principal component analysis to explore meaningful factors or items to be retained), to replication analysis and finally confirmatory analysis. A cross-validation design is therefore desirable for both the exploratory and confirmatory solutions as well, which we did (having a dataset large enough for two different samples to do this analysis)!

Furthermore, as stated earlier, we also tried to replicate our findings on the construct validity. The results were quite similar to our previous study, confirming construct validity of the PRAFAB questionnaire in a different cohort of patients and therapists!
Reviewer’s comment: I suggest rejection of this study; at least until the first study is published to find out if there is an overlap.

Response: With respect to this reviewer’s comment we hope that our response(s) will change the reviewer’s initial view and recommendation to the Editor.

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

Response: We believe that all the omissions have been corrected. Furthermore, we wish to express our gratitude to Ms. Jenny de Fouw and Ms. Kay Dixon for their editorial review and valuable comments on the text of this manuscript.