Reviewer’s report

**Title:** Validity and reliability of a Malay version of the Brief Illness Perception Questionnaire for patients with type 2 diabetes mellitus

**Version:** 0  **Date:** 07 Mar 2017

**Reviewer:** Maritza Dowling

**Reviewer’s report:**

In this paper the authors examined the psychometric properties of a Malay translated version of the Brief Illness Perception Questionnaire (BIPQ), a widely-used instrument to measure a range of illness representations. Authors conducted sample-specific analyses to provide evidence of the translated scale reliability and validity for the use in clinical settings in Malaysia with T2DM patients.

However, the procedures employed to establish the psychometric soundness of the translated instrument and the quality of the statistical analysis and its description can be substantially improved to increase the value of the scale as a patient reported measure.

**Introduction**

It would be helpful if in the Introduction the authors present the actual instrument they are going to be translating and validating. What does it measure? What are the psychometric properties of the BIPQ as reported by the developers and other translated versions? What type (Cronbach alpha, test-retest, etc) and range of reliability indexes have been reported? In what other languages it has also been translated and validated and what had been the reported reliabilities of both the English and non-English versions? Is there evidence in the literature of the factor structure of the BIPQ? Has the BIPQ (or any of the previous versions) been examined using item response theory or Rasch approaches? Is BPIQ a unidimensional scale? That is, are the 8 items unidimensional in the original English version?

On page 6, the authors state:

"The BIPQ was developed for use in clinical practice and exists of nine items; it has been shown good psychometric properties, test-retest reliability and concurrent validity with relevant measures in 36 countries and many illness populations [2]."

Besides the obvious grammatical correction, this needs to be expanded beyond "good psychometric properties." What do you mean by "good psychometric properties?"

**Methods**
A better rationale should be provided for the selection of a "target sample size." The authors simply state:

"A target sample size of 100-150 patients was estimated to give a good precision for the reliability validity study [18]." How was this estimated? What do the authors mean by "good precision" and based on what criteria? Please provide a rationale and criteria for the selection of sample size in the validation study and how sample size is taking into account the length of the instrument.

Instrument translation

The distribution of the MBIPQ to a sample of Malay patients should be described in more detail. This should be conducted as a pilot study to evaluate the comprehensibility of the instrument items and to verify the internal consistency of the questionnaire. Ideally the respondents should have also provided feedback about the comprehensibility of the items on the questionnaire. The patients did not provide any negative feedback and filled out the forms accurately. Obviously, their responses to the questionnaire form would have been excluded from the results of the study.

Test-retest reliability

It would be helpful to cite other time intervals reported in the literature and the range of test-retest reliabilities previously reported including those reported in the development of the BIPQ scale (Broadbent et al., 2006).

How equivalent are these two samples (Negeri Sembilan vs Selengor) regarding health status and other demographic characteristics? To be able to compare results at the two time intervals, the samples should be equivalent. Test-retest reliability is more meaningful in clinically stable conditions and obviously critical sample characteristics.

Assessing construct validity

Have the instruments selected for construct validation previously validated in and used with Malaysian populations? Please provide more details.

What was the rationale for the selection of these specific instruments?

Was the content validity of the translation evaluated by both bilingual and native language speakers? Please explain in this section how content validity was established.

Statistical analyses
In addition to the Spearman correlation, Pearson product moment correlation for the test-retest reliability should also be reported and compared to the coefficients reported by the BPIQ developer's (Broadbent et al., 2006). Both coefficients provide different information on the nature of the association between two variables. The same applies for the coefficients reported for establishing construct validity. The fact that Cronbach's alpha (internal consistency) will also be evaluated/reported later should be introduced in this section.

A rationale for the selection of instruments to establish construct validity should be provided.

Information on Discriminant Validity should be examined using a Receiving Operating Curve (ROC) Analysis.

An examination of the factor structure of the translated scale (dimensionality) should be provided and compared to the English scale and other translations Results should be compared, for example, to Rasch analyses (see, e.g., Ashley, Smith, Keding, Jones Velikova, & Wright, 2013) and confirmatory factor analyses based on structural equation modeling (see, e.g., Karatas, Orzen, Kutluturkan, 2017).

Results

Information on the missing data should be provided, if possible by sample ((Negeri Sembilan vs Selengor). The equivalence of the two samples is difficult to establish. The ethnic distributions appears to be different in both sets of test-retest reports.

More importantly, given the ethnic diversity, is Malay the first language for all the participants? If not this translated instrument should be thoroughly evaluated for differential item functioning (DIF). The potential presence of DIF would be a serious threat to the validity and interpretability of the scales scores. Even if group differences due to native language is not a source of DIF, the scale should be thoroughly assessed for measurement invariance. It is well-known that lack of invariance may lead to potentially misleading group differences and inaccurate bivariate associations involving the DIF items.

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

No

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

Unable to assess

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.
Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I recommend additional statistical review

Quality of written English
Please indicate the quality of language in the manuscript:

Needs some language corrections before being published

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I have no competing interests.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.
I agree to the open peer review policy of the journal