Reviewer's report

**Title:** Psychometric properties of the IDS-SR30 for the assessment of depressive symptoms in Spanish population

**Version:** 1  **Date:** 8 July 2011

**Reviewer:** Edwin de Beurs

Reviewer's report:

The authors investigated the psychometric properties of the Spanish self-report version IDS in a large sample of 1595 patients with MDD. Specifically, they evaluated the performance of individual items, its factor structure, reliability of the total scale and subscales, and its convergent validity. All in all the manuscript is a valuable addition to the literature on the psychometrics of the IDS.

- **Major Compulsory Revisions**

1. The abstract is too long and offers rather detailed information, which belongs to the main body of the text.

2. In the abstract (and again in the method section of the manuscript) the HDRS is presented as “the gold standard”. There is no need to do so: the HDRS can (and should) be presented as the measure used to assess convergent validity. Twice is mentioned that patients had been in treatment for at least 6 weeks. Why was this a requirement? And it makes me wonder whether the first assessment can truly be considered a baseline assessment as the baseline was at least 6 weeks earlier. Some clarification should be offered (in the text, not in the abstract).

3. On page 10 the authors present means for baseline and follow-up scores for the IDS. The baseline score could be included in Table 1 (accompanied by the pretest scores on the HDRS or the proportion of patients scoring beyond the cut-off for the HDRS, as this information is essential in describing the sample. The findings regarding item 4 suggest that a reformulation (“Sleeping more than usual”? ) is in order. Is this item also problematic in other language versions?

4. Convergent validity: In Table 6 the correlation between change scores could be added. In addition, the pre and posttest scores on the IDS and the HDRS allow for a comparison of the sensitivity to change of both measures. I suggest an analysis with SPSS GLM for a 2 by 2 design with time and instrument as factors. A significant interaction effect would reveal a difference between both measures in sensitivity to change. Finally, the data allow for an investigation of relevant cut-off points for statistical and clinical significant change. For the HDRS cut-off values are published. Which cut-off value on the IDS yields the best
correspondence with HDRS defined outcome of treatment?

- Minor Essential Revisions

1. Information is lacking on how these patients were diagnosed prior to inclusion in the study. Was a diagnostic interview part of the intake procedure? Was diagnosis based on clinical impression of the psychiatrist? The authors should offer some clarification.

2. The HDRS is a rating scale and was probably completed by the psychiatrist. Was any training with the instrument provided and/or required?

3. The authors mention some adaptations to the South American version of the IDS. To what extent? Were they far-reaching? Is this the first evaluation of the Spanish version and, if not, how do the findings compare to previous evaluations of Spanish (or South American) versions.

4. Page 12 top: items 9 and 24 were assigned to the anxiety/somatic subscale, 30 tot the Mood/Cognition subscale.

5. In the discussion, the authors mention that excluding item 4 may improve the internal consistency. It does or it doesn’t. The present data do not favour the position that it does, as the internal consistency only improves marginally according to the baseline data and not in the follow-up data.

- Discretionary Revisions

1. Self-report measures can be used to screen for the presence of a disorder or to assess the severity of a condition, but not to diagnose. As such the IDS is not a diagnostic tool, but rather an assessment instrument.

2. The manuscript could benefit from a thorough check by a native speaker as grammatical errors abound throughout the text. Furthermore, the text could be formulated more succinctly. An example: “Among the strengths of the present study it must be highlighted the large representative sample included, which avoids patient bias and makes the results…” could read: “A strength of the present study is the large and representative sample which makes the results…”. Another one: “The main limitation of this study is that we did not assess the questionnaire in primary care.” should read: “… that we did not administer the questionnaire to primary care patients.”

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

**Quality of written English:** Needs some language corrections before being published
Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:

I declare that I have no competing interests