Reviewer's report

Title: Reliability, validity and psychometric properties of the Greek translation of the Major Depression Inventory

Authors:
- Konstantinos N Fountoulakis (kfount@med.auth.gr)
- Apostolos Iacovides (kfount@med.auth.gr)
- Soula Kleanthous (kfount@med.auth.gr)
- Stavros Samolis (samolis@med.auth.gr)
- Kyrikos Gougoulias (gougoulias@med.auth.gr)
- George St Kaprinis (kaprinis@med.auth.gr)
- Per Bech (gabean@fa.dk)

Version: 1 Date: 29 Sep 2002

Reviewer: Prof Toshiaki A Furukawa

Level of interest: A paper whose findings are important to those with closely related research interests

Advice on publication: Unable to decide on acceptance or rejection until the authors have responded to the compulsory revisions

The present authors established the Greek version of the Major Depression Inventory, administered it to 30 patients with major depression and 68 normal controls, and examined its test-retest reliability, internal consistency reliability, factor structure, and ROC analysis.

The major weakness of the present study lies in the selection of the subjects. It is inappropriate to use the case-control design to examine the validity of a screening instrument. The subjects must represent the wide spectrum of patients to whom the instrument would be applied; that is, they must contain subjects at borderline severity and if not, the resulting sensitivity and specificity would be spuriously inflated. (See, for example, Jaeschke, R., Guyatt, G. & Sackett, D. L. (1994). Users’ guides to the medical literature: III. how to use an article about a diagnostic test. A. are the results of the study valid? JAMA 271, 389-391. & Ransohoff, D. F. & Feinstein, A. R. (1978). Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. New England Journal of Medicine 299, 926-929.) The study should therefore restrict itself to the examination of the reliability and factor structure, and report on the preliminary ROC analyses with appropriate caution.

a) Discretionary revisions
1) There is much redundancy between the 1st and the 2nd paragraph in the Background section. It is perhaps better to explain in the Background section that the MDI can be scored either algorithmically or as a total sum score and you mean to examine validity of these two scoring methods.
2) Do the references #6 and #7 really report the MDI to be the most valid? I do not have access to them but the titles of these references do not seem to. Moreover, in the Discussion (4th para), you state "the reliability and validity of the MDI has been tested in a limited number of studies". These two statements appear contradictory.
3) MDI is not the first one to base its items on the diagnostic criteria for major depression a la DSM-III.

4) The notation plus-minus is ambiguous when it refers to either SD or SEM. In the Abstract and Materials section, please specify which of SD or SEM the numbers are referring to.

5) Results, 7th para: Please explain "24 pts vs. 6" or "28 pts vs. 2"

6) Discussion, 6th para: You should choose to express sensitivity and specificity as 0.82 or 82 and should not mix them.

Typos
1) Method, Translation: "one of whom did not knew" is strange.
2) The number of the two figures appear to be confused either in the text or in the attached figures.
3) Results, 4th para: "a two factors" is strange.
4) Results, 5th para: "except from" is strange.
5) Results, 7th para: "not recommended at least not for Greek" is strange.

b) Compulsory revisions
1) Please give more details as to the translation and backtranslation procedure of the MDI. Did one author translate the original into Greek and the second, who did not know the original, backtranslate it into English? And who checked the adequacy of the backtranslation against the original?
2) It is wrong to assess the test-retest reliability by Pearson correlation coefficient. As the instrument is on continuous scale, you should calculate ANOVA ICC (Bartko, J. J. (1976). On various intraclass correlation reliability coefficients. Psychological Bulletin 83, 762-765.)
3) There is no way of redoing the whole study to address the major weakness mentioned above. Therefore,
3-1) You need to discuss the problem of spectrum bias in the Discussion section.
3-2) The ROC curve in Figure 1 appears strange. You should not smooth the curve, as it gives a wrong impression that the ROC curve is better (that is closer to the left upper corner) than it really is. The figures on the horizontal axis are also strange; they should near 1.0 towards right.
3-3) You need to report the Area Under the Curve in order to complete the ROC analysis.

Competing interests:

The present reviewer has received fees for speaking from several pharmaceutical companies.