**Author's response to reviews**

**Title:** Validation of the Turkish version of the Centre for Epidemiologic Studies Depression Scale (CES-D) in patients with Type 2 diabetes mellitus.

**Authors:**

Vicky Lehmann (v.Lehmann@uvt.nl)
Ceylan Makine (ceylan_makine@hotmail.com)
Çagatay Karşıdağ (ckarsidag@yahoo.com)
Pinar Kadıoğlu (kadioglup@yahoo.com)
Kubilay Karşıdag (karsidag@istanbul.edu.tr)
François Pouwer (F.Pouwer@uvt.nl)

**Version:** 2  **Date:** 31 May 2011

**Author's response to reviews:** see over
Validation of the Turkish version of the Centre for Epidemiologic Studies Depression Scale (CES-D) in patients with Type 2 diabetes mellitus.

Vicky Lehmann, Ceylan Makine, Çagatay Karıdağ, Pinar Kadioğlu, Kubilay Karıdağ, François Pouwer

Comments to the editor and reviewers in general:
We are very pleased to get the opportunity to improve our manuscript and would like to thank the reviewers for their useful comments.

The highlighted text is also highlighted in the resubmitted manuscript.

Reviewers' comments for the authors:

Reviewer #1: Mirjana Pibernik-Okanovic

1.1
-A rationale for determining a latent structure of the questionnaire is not sufficiently clear. Why did the authors decide to perform a factor analysis in a relatively small and homogeneous sample of diabetic patients? Using an oblique rotation „based on the assumption that a test measuring depression will have correlated factors“ does not sufficiently clarify the underlying theoretical prerequisites.

Author’s Response: We have clarified this by adding the following text to the introduction:
“We aimed to explore the latent structure of the CES-D (Turkish version) as earlier research has described different factor structures for this scale. As a result, it is unclear which factor structure should be used in Turkish participants (and specifically in diabetic Turkish patients).” (page 3)

The statement regarding the use of oblique rotation is based on the assumption that a test measuring depression will have correlated factors was solely given from a statistical point of view, clarifying we used oblique instead of varimax rotation which produces orthogonal factors (we added this aspect on page 6). In other words, in case correlated factors can be expected (which is the case with depression questionnaires), researchers can best use oblique rotation.

We further agree with the reviewer that our sample was relatively small for the factor analysis, but the factor analysis in only one part of our study, besides we studied a relatively homogenous sample and the KMO and Bartlett’s test suggest reliable results.

The disadvantage mentioned by the reviewer of having a homogenous sample can also be regarded as an advantage, as an important disadvantage of heterogeneous samples, is that they provide little evidence for certain subsamples, for example patients with insulin naive diabetes.

We wanted to keep the explorative nature of our paper, but since the second reviewer also suggested to include some hypotheses, we added our (careful) expectations on page5 & 6.

Furthermore, Tatar and Saltukoglu [26] conducted extensive analyses in a huge sample of 1143 Turkish students and healthy adults, supporting the initial four-factor structure, but also reporting problems concerning the dimension of somatic complaints and especially with item 7. Given these findings, we anticipated to find a four-factor structure in our dataset.

[...]

Since the WHO-5 measures a positively connoted construct as well-being, we suspect negative associations with the outcome of the CES-D.

[...]
If problems related to diabetes coincide with depressive symptoms, their association is consequently suspected to be positive.

1.2.
adding some details about the translation process
- The Statistical analysis section should specify which correlations were used to test convergent validity (the last two lines in this section). The correlational method used should also be mentioned in Table 3.

**Author’s Response:** Information was added on page 4 and 7 respectively for these two points (and Table 4)
The Turkish version of the CES-D, developed by Spijker and colleagues [18], was used in the present study.
“Correlations between the CES-D with the other self-report measures were calculated, using Pearson’s correlation, to test its convergent validity.”

1.3.
-Table 1: the percentages should be referred to the number of examinees in the subsamples. Could the authors check the standard deviations of the PAID scores? Are they as high as visible from the table? How many patients were above the score of 40 which indicates serious diabetes-related problems? The sum of percentages indicating male and female subjects across centres is not 100% but 99%. Please correct this.

**Author’s Response:** All remarks were incorporated in table 1; the standard deviations of the PAID were indeed that high
-Data on the instrument’s reliability should, in my opinion, be presented within a table. A description given on page 8 is somewhat confusing, evaluating the item-total correlation to be high, but mentioning the range from 0.25-0.73. Could the authors give a more comprehensive information about the instrument’s internal consistency?

**Author’s Response:** We agree with the reviewer and information was added in a new Table 3 and slight changes were made in the text on page 9
Moreover, the item-total correlations were overall quite high (ranging from 0.42-0.73) but also showed some low values of 0.12, 0.20, and 0.25 for the items 19 (others dislike me), 13 (talk less), and 4 (as good as others) respectively (see also Table 3).

1.4.
-According to the criteria, the WHO-5 scores between 50 and 28 indicate mood below average, and scores below 28 as being indicative for depression. Please check the criteria which were used in classifying data in such a way.

**Author’s Response:** We thank the reviewer for this comment, agree, and have changed the cut-off scores

1.5.
-Do the authors think that the obtained 2-factor solution, and particularly the combination of depressed mood and somatic complaints, may be due to the specificities of the examined patients? Literature data indicate that somatic items inconsistently indicate depressive symptoms while hypothetically reflecting complaints related to hyperglycemia. Would you please comment a hypothetical relation between the sample’s characteristics (well regulated patients, homogenous with respect to diabetes therapy) and the obtained results?

**Author’s Response:** Aspects tapping this comment were added in the discussion section on page 11. Overall, we agree with the reviewer that these questions are relevant and interesting. However, to answer them we need another group of patients, these questions cannot be answered with the present sample, unfortunately.
1.6.
The authors do not seem to be clear enough when elaborating hypothetical limitations of the study (page 11, lines 9-12. Stating that „Determining the percentage of diabetes patients with clinically significant levels of depression....“ is not quite correct and should be changed into“.....elevated level of depressive symptoms”. According to the literature, only about one third of patients with elevated symptoms is classified as clinically relevant depression by using a diagnostic clinical interview.

**Author’s Response:** We agree with the reviewer and changed things according to this comment and added more aspects concerning limitations of our study on page 12

A major limitation of our study was the relatively small and homogeneous sample. However, preliminary tests as KMO and Bartlett’s test assured the appropriateness of our sample for trustworthy factor analyses. However, future research should also include more heterogeneous samples for example concerning the severity and treatment of diabetes.

1.8.
- I would suggest to replace the part of the abstract stating „The original five-factor structure proposed by Radloff was rejected“ into „.....was not confirmed“.

**Author’s Response:** We completely agree and have changed this sentence accordingly in the abstract.

Reviewer #2: S S Kazarian

2.1.
In the Abstract, they erroneously state that Radloff had proposed a five-factor solution to the English version of the

**Author’s Response:** We have made an error and are glad the reviewer recognized this. We have changed this accordingly in the abstract aligning our statement to the method section where we had indicted the right (four-)factor structure.

- While the article is empirically rather than theoretically driven, the authors nevertheless should have advanced specific hypotheses or predictions concerning the factor structure they expected for the Turkish version of the CES-D (1, 2, 3, or 4) and supported their prediction perhaps on the unique characteristics of Turkish culture.
- A similar limitation concerns specific hypotheses in relation to the correlations of CES-D scores with scores of the other measures, and the potential correlates listed in Table 1.

**Author’s Response:** We thank the reviewer for this comment and added our careful expectation on page 5 & 6

Furthermore, Tatar and Saltukoglu [26] conducted extensive analyses in a huge sample of 1143 Turkish students and healthy adults, supporting the initial four-factor structure, but also reporting problems concerning the dimension of somatic complaints and especially with item 7. Given these findings, we anticipated to find a four-factor structure in our dataset.

[...]
Since the WHO-5 measures a positively connoted construct as well-being, we suspect negative associations with the outcome of the CES-D.

[...]
If problems related to diabetes coincide with depressive symptoms, their association is consequently suspected to be positive.

2.2.
- It is not clear which Turkish version of the CES-D was used
The Turkish version of the CES-D, developed by Spijker and colleagues [18], was used in the presented study.

Details concerning the factor analysis are missing: what were the KMO and the Bartlett’s Test of Sphericity; what did the scree plot suggest in terms of retention of number of factors;

**Author’s Response:** This comment was very helpful and eventually made our results section stronger. Additional information as suggested by the reviewer was added on page 7.

The Kaiser-Meyer-Olkin test (KMO) was found to be .891 (thus higher than .50) and Bartlett’s test of sphericity showed a highly significant result ($\chi^2 = 1409.1$), thus both tests indicated the appropriateness of performing factor analyses in the presented sample [33].

The scree plot actually suggested a two-factor solution, but we continued our analyses forcing four factors into the final model since previous studies (e.g. [16, 26]) suggested a four-factor structure. Results of this factor analyses showed that the initial fifth factor (subsuming three items) was split […]

- the correlation between the two empirically derived factors.

**Author’s Response:** This is a very valuable suggestion and information has been added on page 9.

The two factors were moderately correlated with another $r = .329$ (p<.0001), but strongly associated to the total score with $r = .918$ (depression) and $r = .676$ (positive affect).

2.3.
- It is not clear if all the measures used in the study were in the Turkish language and if they were translated using back translation methodology. Similarly, it is not clear if the three measures were administered in a counterbalanced order.

**Author’s Response:** All measure were in the Turkish language and references were added in the method section.

- It is also not clear why the demographic, clinical and psychological data in Table 1 were included when theoretically grounded hypotheses were not advanced and when some of the data was already reported in an earlier publication (Makine et al., 2009).

**Author’s Response:** We feel that it would be inconvenient for the readers to refer them to another paper to get a description of the characteristics of our sample. Only a reference to Makine et al (2009) might be annoying, a not all readers will have access to the previous article. However, we felt that shortening Table 1 might be convenient and omitted some of the information (HbA, comorbidity)

2.5
The conclusion that ‘they both contribute and form parts of the overall construct of depression’ needs to be supported empirically, that is the correlation between the two factors and their correlations with the total score; and their differential correlations with the two other measures and the demographic correlates. The two-factor solution they obtained should also be reconciled with the four-factor solution reported by Tatar & Saltukoglu for their Turkish version of the CES-D.

**Author’s Response:** We appreciate this comment and included the correlations in our discussion (page 11).

Discussing our two-factor solution in light of the results provided by Tatar & Saltukoglu is a very good point and has also been added on page 12

Yet, they both contribute and form parts of the overall construct of depression as shown by their correlation with each other and with the total score. Their correlations with the other self-report measures suggest that the first factor (depression) is more strongly associated with the other measures yielding similar results as the total score. However, correlations of the second factor
showed to be of significant importance as well. The usefulness of the combination of depressed and somatic complaints in our first factor may be questioned though. There is the well-known problem that people with chronic diseases show elevated levels of depression measured with self-report instrument due to somatic complaints that are actually caused by their disease and not depression. This is for example also the case in the Beck Depression Inventory and a general problem in chronically ill people and not specific to the CES-D.

Thus, our expected four-factor structure was not supported by the data which was surprising, since especially the findings of Tatar and Saltukoglu [26] promised similar results due to the same cultural background of the sample. However, testing a chronically ill sample may mean a higher association of depressed and somatic symptoms leading to their combined loading on one factor instead of in a healthy sample where depressed and somatic complaints may be more fine-grained and loose from another.

2.6. Sample size and representation as limitations should be highlighted

   **Author’s Response: We agree and added this aspect in our discussion (page12)**

A major limitation of our study was the relatively small sample size. However, preliminary tests as KMO and Bartlett’s test assured the appropriateness of our sample for trustworthy factor analyses. However, future research should also include more heterogeneous samples for example concerning the severity and treatment of diabetes.

2.7 The authors do miss inclusion of relevant articles in the field (e.g., Tatar & Saltukoglu, Mansour & Jabir).

   **Author’s Response: We completely agree with the reviewer that the paper of Tatar and Saltukoglu was of major importance and added this paper to our manuscript.**