Author's response to reviews

Title: Validation of the PHQ-9 as a screening instrument for depression in diabetes patients in specialized outpatient clinics.

Authors:

Kirsten M van Steenbergen-Weijenburg (Ksteenbergen@trimbos.nl)
Lars de Vroege (Lvroege@trimbos.nl)
Robert R Ploeger (R.ploeger@zgt.nl)
Jan W Brals (J.brals@zgt.nl)
Martijn G Vloedbeld (M.vloedbeld@zgt.nl)
Thiemo F Veneman (T.veneman@zgt.nl)
Leona Hakkaart-van Roijen (Hakkaart@bmg.eur.nl)
Frans FH Rutten (F.Rutten@bmg.eur.nl)
Aartjan TF Beekman (A.beekman@ggzingeest.nl)
Christina M van der Feltz-Cornelis (Cfeltz@trimbos.nl)

Version: 2 Date: 2 July 2010

Author's response to reviews: see over
Dear Editor,

Thank you for your letter in which you indicate that you will consider a resubmitted paper. We are very grateful with the useful comments of reviewer 1. Please find in red our revisions according to the comments of reviewer 1. We do hope that you will consider this revised version for publication.

On behalf of all authors,

yours sincerely,

Kirsten van Steenbergen-Weijenburg
Reviewer's report

Title: The PHQ-9 is a valid screening instrument for depression in diabetes patients in outpatient clinics.

Version: 1 Date: 8 February 2010

Reviewer: Richard M. Schulz

Reviewer's report:
The need to validate the PHQ-9 as a screening instrument is important, and the process used is appropriate.

* Thank you very much for your compliments.

The following issues should be addressed prior to a decision in support of publication.

1. Many terms are used but not explained. Current descriptions are insufficient. The assumption that readers will understand can not be made. Examples of this issue are a) algorithm score,

* Thank you for noticing. We added the following information to the 'methods' section:

"This 'algorithm', developed by Kroenke et al (4) is positive for MDD when a total of five questions on the PHQ-9 have a score of two or more points, with exception for question nine: scoring at least 1 point is sufficient. Besides that, question one ("in the past two weeks I had less interest and fun in doing activities"), or question two ("in the past two weeks I felt dejected, depressed or desperate"), have to be answered positively."

b) why the flattening of the ROC curve is important,

* We agree fully that the concept of a ROC curve needs to be explained more extensively. Therefore, we added the following sentences to the methods section:

"When a higher criterion value is selected, the false positive fraction will decrease with increased specificity but on the other hand the true positive fraction and sensitivity will decrease, as described by Zweig et al (9):

"In a Receiver Operating Characteristic (ROC) curve the true positive rate (Sensitivity) is plotted in function of the false positive rate (100-Specificity) for different cut-off points. Each point on the ROC plot represents a sensitivity/specificity pair corresponding to a particular decision threshold. A test with perfect discrimination (no overlap in the two distributions) has a ROC plot that passes
through the upper left corner (100% sensitivity, 100% specificity). Therefore the
closer the ROC plot is to the upper left corner, the higher the overall accuracy of the
test" (9). The flattening of the curve shows when there is no additional benefit of the
screening method.

To answer the research question on the criterion validity of the PHQ-9 a Receiver
Operating Characteristic curve (ROC curve) is made with SPSS version 15.0."

To answer the research question on the criterion validity of the PHQ-9 a Receiver
Operating Characteristic curve (ROC curve) is made with SPSS version 15.0."

c) acceptable risk of a falsely screened patient. An editor should help with these
issues.

"We agree with the reviewer that the acceptable risk has to be mentioned. Therefore,
we added the following sentences to the methods section:

"There is always a possibility that a patient is falsely screened positive or negative.
Therefore, it is important to reduce this possibility by identifying the most optimal
combination of sensitivity and specificity. This way, the clinically acceptable risk of a
falsely screened patient can be determined."

2. The manuscript itself does not seem to be well constructed. Examples of
issues are: a) Table 3 is incorrectly attributed, b) Table 5 was not available or
accessible (this may be due possibly to issues with the interface, c) figure 1
reports "summed scores" in the title, but not on the vertical axis of the graph, d)
the manuscript title reports conclusions, which it should not; appropriate title
should be "Validation of the PHQ-9 as a screening ...."

"We are very sorry that you feel that the manuscript is not well constructed. All the
issues that you mentioned were corrected in the revised manuscript and we are very
thankful for your comments.

a) Table 3 is now correctly attributed
b) Table 5 was indeed not available, so we added table 5 to the 'figures and tables'
section.
c) The title and vertical axis now report the same
d) The conclusions are removed from the manuscript title, which has changed into:
"The PHQ-9 as a screening instrument for depression in diabetes patients in
specialized outpatient clinics."

3. Nonresponse and representative of sample is a concern. Sample went from
1,278 possible contributers to 197 in the evaluable sample. The authors claim
that there are no comparative data for nonresponders. There might be a way to
report basic data on nonresponders, it available. Of the original 1,278 completers
of the PHQ-9, were any demogrpahic data obtained. If yes, this group can be
compared to those who completed the study.

*Thank you very much for your comments. Unfortunately, no demographic data were available, as this was confidential information. This issue will be addressed in the discussion section.

The same issue can be applied throughout the process during which possible subjects did not advance. For example, is there any basic data on those who did not return the PHQ-9 within two weeks?

* Unfortunately, no data on these non responders was available, as this was confidential information. This issue is now addressed in the discussion section as follows:

"Secondly, the response rate was not high, because only patients who returned the PHQ-9 within two weeks after receiving it were included. Unfortunately, no epidemiological data on the non responders could be obtained, as this was confidential information which the hospitals were not allowed to provide."

If yes, they can be compared to completers. This is very important to determine representativeness of sample, especially important for a study assessing a screening device to be used in the broader population.

4) This is a manuscript reporting an international study. The authors make the point that their study is unique; previous studies were conducted in diabetic care clinics, unlike their specialty clinics. For this reasoning to be meaningful, the reader must know how the "specialty clinics" differ in substance from other clinics.

* Thank you very much for noticing the uniqueness of our study. The 'specialty clinics' are more now described in the introduction:

"These specialized clinics differ from general diabetic care clinics in that in these specialized clinics foremost patients with severe diabetes with complications are present and specialized clinical diabetes care is provided by a team of a diabetologist, a specialised diabetes nurse and a dietician.

5) This study relied on telephone response from a potentially depressed group to obtain data. Is this the typical use of data collection for this instrument reported in previous studies? The low response rate within this group may be due to depressed patients unable to respond. This is borne out in the low percentage of patients (18.8%) with MDD. This should be addressed.

* We agree that the telephone response might not be an optimal way to define depressed patients. However, there was almost 50% telephone response, which is average for epidemiologic research in the Netherlands. This has now been mentioned in the discussion:
Also, the telephone response of almost 50% might seem low, but this is an average response rate on epidemiologic research in the Netherlands.

6) I have great concern that the results might create considerable confusion among practitioners. A tool with different cut-off points depending on its intended use, the population for which it is used, and the setting of use, can create more uncertainty that clarity. The authors should address this issue.

* We agree that the results might create confusion among practitioners. Therefore, we have described in more detail how the PHQ-9 can be used in specialty outpatient clinics and left out the cutpoint for the use of the PHQ-9 for epidemiological research:

"Findings of the study"
In this study, for the first time the PHQ9 is validated as a screening instrument for MDD in diabetes patients visiting a specialized outpatient clinic. As such, it gives us important information about the validity and appropriate cut off scores for identifying Diabetes patients with a high possibility for having MDD.

The main finding of this study is that the PHQ-9 appears to have satisfactory criterion validity as a screening instrument for MDD in diabetes patients in specialized outpatient clinics. We recommend using a cut-off score of 12 to recognize depression in diabetes patients from specialized outpatient clinics. This is a higher cutoff score than is generally used for identification of MDD patients in the primary care setting in patients without advanced medical co-morbidity.

In summary, I believe this topic is important. The manuscript should be reworked to improve clarity and improve presentation of data. Issues surrounding nonresponse and impact on clinical practitioners should be addressed.

* We thank the reviewer for the useful comments and hope that the clarity and presentation of data is improved and that all issues are sufficiently addressed.

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

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**
I declare that I have no competing interests.