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. 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, b) why the flattening of the ROC curve is important, c) acceptable risk of a falsely screened patient. An editor should help with these issues.

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 ...."

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 demogrpohic data obtained. If yes, this group can be compared to those who completed the study. 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? 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.

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.

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.

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.

**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.