Author's response to reviews

Title: Validity of a self-reported diagnosis of depression among participants in a cohort study using the Structured Clinical Interview for DSM-IV (SCID-I)

Authors:

Almudena Sanchez-Villegas (asanchez@dcc.ulpgc.es)
Javier Schlatter (schlatter@unav.es)
Felipe Ortuno (fortunos@unav.es)
Francisca Lahortiga (fhortiga@unav.es)
Jorge Pla (jpla@unav.es)
Silvia Benito (sbenito@alumni.unav.es)
Miguel A Martinez-Gonzalez (mamartinez@unav.es)

Version: 2 Date: 15 February 2008

Author's response to reviews: see over
Reviewer's report
Validity of a self-reported diagnosis of depression among participants in a cohort

General
This is an interesting paper, validating the use of self reported depression as diagnosed by a physician as a measure of depression. I have a number of general points:

1. Background
This section is too general, and need to focus more specifically on the area addressed by the paper. The points that probably need to be made are:

- there is a need to be able to estimate prevalence and incidence of depression in population studies

Following the reviewer’s suggestion we have added this new paragraph in the revised version of the manuscript.

- depression scales have their limitations (list them) and one alternative is asking the participant if they have been diagnosed with depression

The main limitations of the use of scales are:
- limited sensitivity to change in depression severity
- heavy weighting toward behavioral and somatic symptoms,
- low item level reliability

In fact we have asked participants for a physician-made diagnosis of depression, this is the variable that we are considering the appropriate outcome (see, for example, Sanchez-Villegas et al. Eur J Nutr 2007;46:337-46). Given the high educational level of the participants in our cohort, we think that this self-report of a depression diagnosis can be used as a reliable outcome. Nevertheless, we also acknowledge that we will have (as in any other outcome for a cohort, such as cancer or atherosclerosis) a fair number of false-negative cases that eventually will be reported in future ascertainment by our participants (or their families) if the depression is sufficiently severe.

- how well does this measure a) estimate the prevalence of depression and b) agree with formal psychiatric assessment?

This question is precisely what we want to answer through this analysis. Other self-reported diagnoses have been used in several studies with very good results. We state this issue in a paragraph of the discussion section:

“The validity of the self-reported diagnosis of different diseases has been sufficiently studied in several populations in which the cultural level of the participants was as high as it is in the SUN Project”

2. Methods
2.1 The abbreviations for the different follow-up examinations are confusing, and the authors would be better referring the examinations as baseline, first follow-up etc.

As the reviewer suggested we have replaced the abbreviations by the words “assessment at baseline and in the follow-up questionnaires”.

2.2 The authors used the SCID-CV. They should refer to the interview was such throughout.
The SCID-I is a structured clinical interview designed to diagnose DSM-IV Axis I disorders. Though the interview, diagnostic criteria were applied for any of the possible diagnoses of depression: major depressive episode, dysthymia, mood disorder or adaptative disorder. We have explained along the whole new version of the manuscript that the interview was done according to the SCID-I every time that we mention the interview in the text.

2.3 Validation study: why were there fewer controls, and why were they from a different region (Navarra only) than the cases? (Or were the two groups from the same region?)

Following the reviewer’s comment we have added a new paragraph in the methods section explaining the selection procedure.

2.4 The study has a low participation rate, and this weakens the argument significantly. The authors should return to this issue in the discussion. We have added the following sentence to the discussion section of the new version: “We acknowledge that participation in the validation study was low (24.2% and 23.3% for depressed and non-depressed subjects, respectively) and therefore a selection bias cannot be excluded, conferring an artificially high validity to our results. To prevent this bias, we put special care to blind participants about the aim of the psychiatrist interview. Thus we tried to avoid that those participants who may be aware of incorrectly classifying themselves as depressed may be embarrassed of participating. Moreover, when we compared the socio-demographic characteristics of participants (n= 104) and non-participants (n=333) we did not find any systematic difference regarding key variables (sex, age, smoking status, BMI, or physical activity) in their baseline assessment”.

2.5 Were all the non-contacted eligible participants abroad? Some of the non-contacted participants were abroad whereas others agreed to participate although finally they did not attend to the University Clinic to make the interview.

2.6 Definition of depression. The SCID-CV established the diagnosis of current and lifetime depressive disorder. The authors need to explain more fully how they used it to validate depressive episodes which occurred in the past. Which depressive diagnoses did the authors regard as depression? DSM-IV includes adjustment disorders with depression or with depression and anxiety, dysthymia as well as well as a 'not otherwise specified' category, in addition to major depressive disorder. So we need a little detail on the definition of depression. Our aim was not to validate previous episodes of depression, but only those recently reported because the questionnaires are sent every other year. As we explained above (point 2.2), SCID-I a structured clinical interview designed to diagnose DSM-IV Axis I disorders. We classified the diagnoses in major depressive episode, dysthymia, mood disorder or adaptative disorder.

3. Results
3.1 SCID-CV gives diagnosis of depression according to DSM-IV. It would be very useful to see the breakdown of diagnoses. The main diagnosis was major depressive episode (42% of the cases) followed by adaptive disorders (30%) and dysthymia (14%).
Following the reviewer’s suggestion we have added a sentence in the new version of the manuscript.

3.2 The prevalence of depression in the population was 26.1%. This is very high for depression confirmed by diagnostic interview, and casts doubt on the validity of the sample unless the authors have independent confirmation that this figure is likely to be accurate. **The figure 26.1 is referred to lifetime depression up to the completion of the baseline questionnaire. This is the cumulative percentage of depression referred by the participants of our cohort.**

3.3 The agreement between self-reported depression and SCID-CV depression should be assessed using the kappa statistic. Using the method of Fleiss [Fleiss, J. L. 1981. Statistical Methods for Rates and Proportions. 2nd ed. New York: Wiley.] I get the following agreement: Kappa (95% CI) = 0.535 (0.374 - 0.696) **Kappa statistics is a good approach. Nevertheless, if we consider the SCID-CV as gold standard, the use of sensitivity, specificity and predictive values (percentage of confirmed diagnosis) are the best alternative.**

3.4 I am not sure what the extensive analysis of factors associated with accuracy of diagnosis adds. The findings are peculiar, and the number of factors tested must surely have resulted in some spurious findings. **Although many factors have been analysed, the reported results might complete information about validity of self-reported diagnosis of depression. In fact, the results obtained for differences in diagnosis accuracy according to smoking or physical activity status are logical and adequately explained in the discussion section. Nevertheless, following the reviewer’s comment we have added a sentence in the discussion section addressing the possibility of some spurious findings.**

4. Discussion

4.1 The first point, the high depression prevalence, needs to be discussed more fully. It throws significant doubt on the validity of the study.
4.2 I don’t understand the second sentence. **The first sentence refers to the percentage of confirmed depression derived from the validity analysis (the proportion of confirmed cases of depression was calculated as the number of those who reported a physician diagnosis of depression and had depression according SCID-CV, divided by all those who reported a physician diagnosis of depression). The sentence does not refer to the prevalence of depression in our cohort. We have re-written these sentences to improve their understanding.**

4.3 The main thrust of the discussion should be to evaluate the authors’ method of screening for depression against the alternatives. Questionnaire scales are not, in fact, time-consuming. The Beck Depression Inventory FastScreen, for example, is only 7 questions. And the authors should refer to the European guidelines on
measuring mental health in population surveys, which recommend the CIDI-SF. The discussion does not really sum up the important questions: whether the authors' approach results in an accurate measure of the prevalence of depression or not. Certainly, the sensitivity is very poor.

**Our screening method for depression was a self-reported diagnosis of a depression done by any physician. We believe that this is a better alternative than to use some kind of self-administered questionnaire because we prefer to have false negative than to have false positive reports.**

In fact, the Beck Depression Inventory FastScreen analyzes the presence of depressive symptoms (mainly cognitive symptoms) and not a clinical diagnosis of depression. A subject could present depressive symptoms but not for that reason a depressive disorder.

重大修稿建议（作者必须在决定是否接受出版前回复）

The paper should be rewritten, focusing on the issues highlighted above. We have improved the manuscript incorporating the suggestions made by the reviewer.

次要修稿建议（例如，图的标签缺失，或是术语的使用错误，作者可以纠正）

As noted above, use of SCID-I for SCID-CV, and unnecessary abbreviations. **We have substituted the terms Q_0; 1_2 AND Q_4.**

We considered correct to use the term SCID-I, because this is a clinical version to diagnose DSM-IV Axis I disorders.

随意修稿建议（作者可以选择忽略）

The English throughout is excellent. There are some instances where a better word might be chosen, such as 'cultural level' in the discussion, which might better read 'educational level' or 'socioeconomic level'. **We have substituted the term cultural level by the term educational level.**