Author’s response to reviews

Title: Detecting depressive and anxiety disorders in distressed patients in primary care; comparative diagnostic accuracy of the Four-Dimensional Symptom Questionnaire (4DSQ) and the Hospital Anxiety and Depression Scale (HADS)

Authors:

Berend Terluin (berendt@kpnplanet.nl)
Evelien PM Brouwers (e.p.m.brouwers@uvt.nl)
Harm WJ van Marwijk (hwj.vanmarwijk@vumc.nl)
Peter FM Verhaak (p.verhaak@nivel.nl)
Henriëtte E van der Horst (he.vanderhorst@vumc.nl)

Version: 2 Date: 14 July 2009

Author’s response to reviews: see over
Dear Editor of BMC Family Practice

I herewith submit, for your consideration, our revised manuscript on the detection of depressive and anxiety disorders in distressed patients in primary care. We would like to thank the reviewers, Adomas Bunevicius and Enric Aragonès, for their thorough and constructive comments, which we feel have helped us improve the manuscript significantly. The reviewers asked for extra information and clarification to allow readers to better interpret our findings. We have supplied this. They also commented on the composition of our study sample. We have discussed this issue in more detail. In addition, and contrary to Aragonès’ opinion, we have explained why our study sample is relevant for the diagnostic problems experienced in daily general practice. We disagree with Bunevicius about the use of an area under the curve (AUC) statistic to determine optimal cut-off points. Sensitivity and 1-specificity values of all possible cut-off points of a scale determine the ROC curve that allows an AUC estimate.

Below, we will address the reviewers’ concerns point-by-point. We will copy the comments (in Italic print) and subsequently provide our responses. In the manuscript all changed or new pieces of text are marked by red print.

**Reviewer: Adomas Bunevicius**

**Methods. Patients and procedure.**

1) Page 5, first paragraph. One of the inclusion criteria was that patients “had a ‘nervous breakdown’ according to the GP...”. Authors should explain in more details in the Methods section what is considered a “nervous breakdown” is and support their explanation with references.

Response: We have added a few sentences and references to explain what Dutch GPs mean by “nervous breakdown”.

2) Page 5, first paragraph. Among the inclusion criteria to the study were that patients were employed and were 18-60 years old. A great part of primary care patents are the ones that are retired. Indeed in the present study the mean age of study population was found to be 39.5
years (SD 9.2). It shows that older patients that comprise a great part of primary care patients were not included to the study. Therefore results of this study cannot be generalized in all GP patients. That needs to be discussed in the Discussion section.

Response: We should have mentioned this indeed. Moreover, this comment made us aware of other limitations of the study. We have addressed the limitations in more detail in the Discussion section (p. 15).

3) Page 5, first paragraph. Authors state that: “The GPs were instructed not to include patients who, in their opinion, had a depressive or anxiety disorder”. This statement should be explained in more specifically. What does it mean ‘in their (GPs) opinion’? Because depression and anxiety are clinical diagnoses.

Response: We changed the sentence into: “The GPs were instructed not to include patients with obvious depressive and anxiety disorders (i.e. patients in whom they had clinically diagnosed such disorders)” (p. 5).

**Methods. Measurements.**

1) Page 6 and 7, paragraphs 2 and 3. Not sufficient information on the 4DSQ and the HADS. Authors should provide scoring ranges of all of the subscales of the 4DSQ and the HADS, explain what are high scores and what are low scores in the HADS scale, and also provide the cutoff points of both scales (and subscales) used (or recommended).

Response: We have added the requested information and provided the recommended cut-off points in a new Table.

**Results.**

1) Pages 10 and 11. Authors are looking for an optimal cut-off value of the 4DSQ subscales and the HADS. In my opinion this part is a major minor point of the study, because to my knowledge the Area under the ROC curve (AUC) is the most valuable value to determine the optimal cutoff point because it provides information on the balance between sensitivity and specificity of the test at a certain cutoff value and an optimal cutoff point of the scale is the one with the highest AUC. Therefore, I think it necessary to perform the ROC analyses (calculate AUCs) for all cutoff values of the 4DSQ subscales and the HADS first and then determine the optimal values.

Response: We don’t quite understand this comment. After all, the ROC-curve is a graphical representation of the sensitivity and 1-specificity values of all possible cut-off values of a
continuous diagnostic variable [1]. The area under this curve is, therefore, a single value belonging to the instrument, not to individual cut-off points. We have added some explanation and a reference to the manuscript (p. 9).

2) Tables 3 and 4. The names of the tables could be more specific. Need explanation on what is written in bold type in the table and explain all the abbreviations used. As I mentioned earlier I think that the AUC values are essential in determining an optimal cutoff values, and this information should be included in tables. Also what does abbreviation “4DKL-distress” in Table 3 mean?

Response: We agree that the Tables titles could perhaps be more informative. However, in our opinion, the combination of Table titles and Table legends supply ample explanation to understand the Tables’ content. Nevertheless, we are open for suggestions for improvement. We have added an explanation of what is displayed in bold type. “4DKL” is the Dutch abbreviation for “4DSQ”. We have corrected this error.

Discussion
1) Authors should discuss more about their findings and compare them with the results from other validation studies: what were the AUCs of the HADS and 4DSQ found in other studies, what were the optimal cutoff values of the HADS and 4DSQ found in other studies etc.

Response: We have added a discussion of previous validation studies of the 4DSQ and HADS in primary care samples. This discussion made us aware of an important difference between our study and previous studies regarding the study population selected. Previous studies have all used unselected patients, whereas we have used a particular high risk group (one of the three worth mentioning). This shed a new light on the significantly higher cut-off points found in our study and we have addressed this issue in the Discussion section (p. 14).

Conclusions
1) Page 16. Authors state that: “The recommended cut-off points for the 4DSQ anxiety scale and both HADS scales probably need revision, but the optimal cut-off points should first be established in independent samples”. But cutoffs points of the HADS are being revised, including in a sample of primary care patients (for example in the paper by Bunevicius et al. paper in the Depression and Anxiety). As I have mentioned previously this needs to be discussed.
Response: Following on the discussion above, we changed the final sentences into: “Future research should focus on the diagnostic accuracy in high-risk groups, in particular in patients presenting their distress to their GP, patients with medically unexplained physical symptoms and patients with chronic physical conditions. Special attention needs to be given to optimal cut-off points in these groups as these may very well differ from those determined in unselected primary care patients.”

Minor essential revisions

Background

1) Page 4, first paragraph. References are not sufficient. The authors state that: “Depressive and anxiety disorders are prevalent in primary care patients, although these disorders are often not recognized as such by the general practitioner (GP)” and cite Verhaak, 1995. More and newer references are needed would be useful in supporting this statement.
Response: We have added more and newer references.

2) Page 4, first paragraph. References are needed to support the description of the “distressed patient”. Also, authors state that: “These patients present with psychological complaints, such as nervousness or depression…” Are patients with a diagnosis of depression also considered distressed?
Response: Regarding the “distressed patient”, we do not consider these patients as belonging to a special category. We just meant patients who are distressed and express their distress in contact with their GP. We have reformulated the sentence and changed “depression” into “feeling depressed” to avoid confusion about this symptom.

3) Page 5, second paragraph. Authors write that: “In distressed patients we do not need questionnaires to detect just any ‘depression’ or ‘anxiety’, or distress in general, since the psychological character of the problems is already apparent both to the doctor and the patient”. What authors mean by ‘depression’ or ‘anxiety’? Is it clinical diagnoses of depression and anxiety disorders or is it symptoms of depression or anxiety?
Response: “Any ‘depression’ or ‘anxiety’” referred to symptoms and disorders of all severity levels. We changed the sentence into: “In distressed patients we do not need questionnaires to
detect mild depressive or anxiety disorders (let alone depressive or anxiety symptoms) …” (p. 4-5).

**Methods. Measurements.**

1) *Page 6, first paragraph.* References are needed to support the statement: “The section of simple phobia was omitted because of its questionable clinical relevance in patients with a ‘nervous breakdown’, while the section of obsessive compulsive disorder was omitted because of the low prevalence of this disorder in general practice”. Why is simple phobia is of questionable clinical relevance in patients with ‘nervous breakdown’? Also, later in Table 2 last row authors name “panic/phobic disorders’. So have really patients with simple phobias were excluded from the study?

Response: We have changed the sentence into: “The specific phobia section was omitted because these problems are (if not accompanied by other mental disorders) associated with relatively little disability and impairment […] and, therefore, isolated (non-comorbid) specific phobia appears to be of relatively little importance in patients with a ‘nervous breakdown’.

The obsessive-compulsive disorder section was omitted because of the low prevalence of this disorder […]” and added some supportive references (p. 6).

2) Authors use abbreviation HADS (Hospital Anxiety and Depression Scale), but sometimes in the text use the term “HADS scale”. For example page 7, second paragraph: “Like the 4DSQ, the HADS scales employ two…”. Authors should correct that.

Response: We do not agree with this suggestion. After all, the name of the instrument is “Hospital Anxiety and Depression Scale”, abbreviated as “HADS”. This instrument comprises two (sub-)scales, depression and anxiety. It is common use to denominate these scales as “HADS depression scale” and “HADS anxiety scale”. To denominate both scales of the HADS as the “HADS scales” is therefore only logical.

**Methods. Analysis.**

1) *Authors need to explain at which cutoff values of the HADS and the 4DSQ ROC analysis was performed.*

Response: See the first point under Results above. The ROC-curve is a graphical representation of the sensitivity and 1-specificity values of all possible cut-off values of a continuous diagnostic variable.
Results

1) Table 1. Need to include information on the total number of patients. Also, a column providing information on how many percent of patients had each diagnosis would be useful. Abbreviations used in the table should be explained in the table.

Response: We added the requested information to the Table.

2) Table 2. Abbreviations used in the table 2 (4DSQ and HADS) should be explained in the table. Also it would be useful to include information at which cutoff values of the 4DSQ and the HADS the ROC analysis was performed and write significant p values in bold type. Also authors speak about panic/phobic patients in the Table 2 as well as in the text although in the Methods section they state that section of simple phobia of the CIDI was omitted. This should be stated more clearly.

Response: We provided explanation of the abbreviations 4DSQ and HADS in the Tables’ footnotes.

The ROC analyses were performed on all possible cut-off points of the scales, as is now explained in the Methods section (see also the first point under Results above).

The significant p is given in bold type.

In order to avoid any confusion about what is meant by “panic/phobic disorders” we have replaced this expression by “panic disorder, agoraphobia and social phobia”.

3) Pages 9-10 (ROC analyses subsection). Not sufficient information. The cronbach alpha values and p values should be provided in the sentence: “Only with respect to the ability to detect panic and phobic disorders, the 4DSQ anxiety scale performed significantly better than the HADS anxiety scale”.

Response: The alpha values are provided in Table 2, whereas the p-values are provided in Table 3 (which is under discussion in the cited sentence).

4) Some of the information provided in the Results belongs in the Discussion section. For example the sentence in page 10: “All these HADS total score cut-off points are considerably higher than those recommended in other studies using the HADS total score that found 11+ and 12+ as optimal cut-off points [18,19].”

Response: We have moved these comments on the results to the Discussion section.

Discussion
1) Page 12, first paragraph. Avoid repeated information. Abbreviations of the 4DSQ and the HADS are already explained in the text previously.
Response: We have changed this.

2) Page 13, Limitations and strengths subsection. Sentence: “Distressed patients who are still working or who are not involved in paid employment, probably would have relatively mild distress on average, a small chance of caseness....” Should be supported by references.
Doesn’t distressed and employed comprised study population of the present study?
Response: This part of the Discussion section has been rewritten completely.

Reviewer: Enric Aragonès

Major Compulsory Revisions
In this article, the authors try to study the diagnostic performance of two different questionnaires: the HADS, a classic questionnaire with a well established usefulness; and the 4DSQ, a new instrument designed by the authors of the article. They use as gold standard the DSM-IV diagnoses obtained with the CIDI. The main limitation of this paper is the sample of patients in which the functioning of the questionnaires is studied. The sample was recruited for the purposes of a clinical trial on a psychosocial intervention in distressed patients, with inclusion and exclusion criteria probably adequate for the aims of the clinical trial but inadequate for the aims of this manuscript, because they limit very much the external validity and the utility of the results reported. For example: in a study on a diagnostic test, the exclusion of the analysis of those patients who, in opinion of the doctor, suffer the studied disorder/s is difficult to understand. The patients are highly selected and the extrapolation of the results obtained to the clinical practice is very difficult. This shortcoming is widely explained and argued by the authors in the Discussion though, in opinion of this reviewer, the recognition of this fault does not redeem it.
Response: We are glad that Dr. Aragonès brought up this point. This comment made us realize that our sample actually represents one of the major ‘patient domains’ where application of a questionnaire to help detect depressive and anxiety disorders makes sense.
According to Kelder et al. diagnostic research should be performed “where there is a diagnostic problem” [2]. When the disorder is obvious and the diagnosis easy to establish, there is no diagnostic problem [2]. With respect to depressive and anxiety disorders,
diagnostic problems are present, among other groups, in patients presenting with psychological complaints (distress). So, although our sample indeed was not primarily recruited for the purpose of a diagnostic study, the sample actually represented a clinically highly relevant “patient domain”. Without any problems, our results can be generalized to distressed patients in primary care, who have reported sick for their paid job. However, our results should be generalized with caution to distressed patients who are not employed, or who are employed and not on sick leave. Furthermore, our results should not be generalized to two other important patient domains where diagnostic problems exist regarding depressive and anxiety disorders: patients with unexplained physical symptoms and patients with chronic physical conditions. We have discussed this in the Discussion section (p. 15-16).

**Minor Essential Revisions**

*There is an extensive explanation of the principal features of the studied questionnaires but the authors must indicate if they are self-reported or administered by an interviewer.*

Response: We have mentioned in the text at page 7 and 8 that the 4DSQ and HADS are “self-rating” questionnaires.

*We can see the specific diagnoses considered in this paper in the table 1, but we would want to see these diagnoses also in the text of the Methods section.*

Response: We have now listed the disorders in the Measurements section.

*Authors must describe the number, training and expertise of the persons executing the “gold standard” (CIDI).*

Response: We have added this information.

*Authors must describe more accurately the procedure of the study: the HADS and the 4DSQ are administered before or after the CIDI? Which is the time interval between the questionnaires and the reference test? Are the CIDI interviewer blinded to the results of the questionnaires? (and, can be guaranteed that the reading of the questionnaires is independently from the criterion of reference (CIDI)?)*  

Response: We have added the missing information.

*What means 7+: > 7, or # 7?*
Response: 7+ was meant to mean $\geq 7$. To avoid confusion we have changed the notation throughout the manuscript.

*In the row of titles of the table 3 we can see the acronym 4DKL (is it the Dutch name of the questionnaire?)*

Response: It is indeed the Dutch name. We have corrected this.

*In the table 3 we can see different values of sensibility, specificity ... for the same cut-point (see 19+, and 22+)*

Response: Thank you for being so attentive. We have corrected the error.

Thank you for the opportunity to improve our work. We hope our response and revision are to your satisfaction and you can publish the paper. Any further questions or comments are more than welcome.

Yours sincerely,

Berend Terluin, MD, PhD

Reference List
