
**Available from Deakin Research Online:**

[http://hdl.handle.net/10536/DRO/DU:30029366](http://hdl.handle.net/10536/DRO/DU:30029366)

Reproduced with the kind permission of the copyright owner.

**Copyright:** 2010, British Journal of General Practice
Identification of depression in diabetes: the efficacy of PHQ-9 and HADS-D

Prasuna Reddy, Benjamin Philpot, Dale Ford, James A Dunbar

ABSTRACT

Background
Clinical guidelines advise screening for depression in patients with diabetes. The Patient Health Questionnaire (PHQ-9) and the depression subscale of the Hospital Anxiety and Depression Scale (HADS-D) are commonly used in primary care.

Aim
To compare the efficacy of HADS-D and PHQ-9 in identifying moderate to severe depression among primary care patients with type 2 diabetes.

Design of study
Self-report postal survey, clinical records assessed by GPs.

Setting
Seven metropolitan and rural general practices in Victoria, Australia.

Method
Postal questionnaires were sent to all patients with diabetes on the registers of seven practices in Victoria. A total of 561 completed postal questionnaires were returned, giving a response rate 47%. Surveys included demographic information, and history of diabetes and depression. Participants completed both the PHQ-9 and HADS-D. Clinical data from patient records included glycosylated hemoglobin (HbA1c) levels and medications.

Results
The proportion of the total sample completing HADS-D was 96.8% compared with 82.4% for PHQ-9. Level of education was unrelated to responses on the HADS-D but was related to completion of the PHQ-9. Using complete data (n = 456) from both measures, 40 responders showed HADS-D scores in the moderate to severe range, compared with 103 cases identified by PHQ-9. Only 35 cases were classified in the moderate to severe category by both the PHQ-9 and HADS-D. Items with the highest proportions of positive responses on the PHQ-9 were related to tiredness and sleeping problems and, on the HADS-D, feeling slowed down.

Conclusion
It may be that the items contributing to the higher prevalence of moderate to severe depression using the PHQ-9 are due to diabetes-related symptoms or sleep disorders.

Keywords
depression; diabetes; Hospital Anxiety and Depression Scale; Patient Health Questionnaire; sleep disturbance.

INTRODUCTION

Diabetes is a major contributor to the global burden of disease and a growing number of studies show links between depression and diabetes.1-3 The negative impact depression can have on quality of life for people with diabetes, together with the increased healthcare costs of comorbid depression have been recognised.4

In the UK, the Quality and Outcomes Framework provides incentives for GPs to use validated questionnaires to identify people with depression, including those with existing heart disease or diabetes.5 The Patient Health Questionnaire (PHQ-9)6 and the depression subscale of the Hospital Anxiety and Depression Scale (HADS-D)7 are increasingly used to improve the identification and management of people with depression among those with diabetes or heart disease,8-10 but doctors’ responses to the category scores vary, depending on which of the questionnaires is used.11,12

Several studies show that PHQ-9 and HADS-D differ in the proportion of people classified with mild, moderate, or severe depression.13,14 A previous study15 by the current authors identified the prevalence of psychological disorders among people with type 2 diabetes in Victoria, Australia, and used...
several indices of depression, distress, and anxiety. This study reports the results of administering both the PHQ-9 and HADS-D measures to all participants and compares the performance of each in identifying depression.

**METHOD**

**Participants**
Between February 2007 and March 2008, 1200 postal questionnaires were sent to 10 practices in rural and metropolitan Victoria, Australia to distribute to potential participants. Clinical data were recorded by practice staff. Adults with type 2 diabetes from seven general practices participated. A total of 561 completed questionnaires that could be matched with clinical records were received, giving a response rate of 46.8%.

**Measures**
The questionnaires asked about demographics, diabetes, and depression. Clinical data included glycated hemoglobin (HbA1c), body mass index (BMI), and medications. The depression questionnaires used were HADS-D and PHQ-9. Standard cut-off scores were used with HADS-D to classify minimal (0–7), mild (8–10), and moderate to severe (≥11) levels of depression. For PHQ-9 the cut-off scores were: minimal (0–4), mild (5–9), and moderate to severe (≥10).

**Statistical analysis**
Cronbach’s α and corrected item-total correlations were used to examine internal consistency of items on the two depression measures. Homogeneity and structure of both scales were assessed using principal components analysis. Additional analyses included calculation of χ², t-tests, and analysis of variance to determine relationships between clinical characteristics and depression scores (Stata version 10).

**RESULTS**

**Sample characteristics**
The sample comprised 309 males (median age 67 years) and 252 females (median age 70 years). Most were married (63%), from rural areas (75%), and either retired or not in full-time employment (75%). Males and females in the sample did not differ significantly in age or level of education. There were no significant differences between urban and rural responders with regard to sex, age, level of education, employment status, or health concessions.

Mean duration of diabetes for the sample was 8.84 years (standard deviation = 7.35 years). About half the sample reported no diabetes complications and a quarter reported two or more complications.

Most managed diabetes with oral medication only (53.7%), or no medication (27.3%), but 107 (19.1%) participants were taking insulin or insulin plus oral medication. Previous history of depression was reported by 161 participants; 48 of these had had an episode within the previous 12 months, and 59 within 1–5 years. Current antidepressant usage was reported by 65 of the 239 participants who had some depression.

How this fits in
Clinical guidelines recommend screening patients with diabetes for depression because of poor clinical outcomes when there is comorbid depression. The depression subscale of the Hospital Anxiety and Depression Scale (HADS-D) and the Patient Health Questionnaire (PHQ-9) are commonly used for screening in primary care, but there is poor agreement between the measures in categorising moderate to severe depression. When used to assess depression among patients with type 2 diabetes, PHQ-9 may overestimate moderate to severe depression because of items that include symptoms of diabetes or sleep disorders. HADS-D is a better screening tool for depression in patients with diabetes.
Measures exceed 0.4 (HADS-D range 0.43 to 0.72, PHQ-9 range 0.56 to 0.76). The figures given for Cronbach’s α if items are deleted show that removal of any of the individual items on HADS-D or PHQ-9 would not substantially improve the internal reliabilities of the scales.

**Factor structure**

Factor analysis of the separate scales used principal components. For HADS-D, the analysis gave one factor with an eigenvalue of 3.49, which explained 49.8% of the total variance. For PHQ-9, a single factor was obtained with an eigenvalue of 5.1, explaining 56.8% of the total variance. Most items within each scale had a substantial loading on the primary factor (HADS-D range 0.56 to 0.83, PHQ-9 range 0.66 to 0.83).

A factor analysis including all items from both HADS-D and PHQ-9 was also performed (Appendix 2.) The analysis yielded two factors with eigenvalues of 7.8 and 1.2 respectively, which together explained 56.5% of the total variance. Items with substantial loadings (>0.65) on factor 1 were: H4 ‘I feel as if I am slowed down’; P3 ‘Trouble falling or staying asleep, or sleeping too much’; P4 ‘Feeling tired or having little energy’; and P5 ‘Poor appetite or overeating’. The PHQ-9 items P1 ‘Little interest or pleasure in doing things’ and P2 ‘Feeling down, depressed, or hopeless’, loaded on both factors.

**Comparison of severity ratings between depression measures**

The correlation between HADS-D and PHQ-9 was significant at the 0.001 level ($r = 0.78$). The threshold scores for mild and moderate to severe depression were 8 and 11 respectively for HADS-D, and 5 and 10 respectively for PHQ-9. Although 117 (21.5%) of those who completed HADS-D had some depression (score >8), 186 (40.3%) of those who completed PHQ-9 showed depression (score >5). There were no significant sex differences across severity categories on either of the measures, but there were age differences. A higher proportion of those aged 65 years and under (compared with those over 65 years) reported moderate to severe depression on both HADS-D ($P = 0.029$) and PHQ-9 ($P = 0.008$).

The cross tabulation of HADS-D and PHQ-9 scores in Table 1 shows a lack of concurrence of distribution within severity cut-offs. Of the 103 cases identified by PHQ-9 as moderate to severe, 35 were in the same category on HADS-D, but 31 had minimal depression on HADS-D. Of the 40 identified by HADS-D in the moderate to severe range, PHQ-9 classified 35 as moderate to severe and two cases as minimal depression.
**Distribution of expected and observed item responses by severity cut-offs**

On both measures, the maximum possible rating for each item was 3, indicating highest frequency of occurrence (this was ‘nearly all the time’ on HADS-D, and ‘nearly every day’ on PHQ-9). The expected proportion of maximum ratings for each item within each total score, \(y = a/b\), was calculated by determining, for each possible total score, (a) the number of ways in which the total score could be arrived at while keeping a particular item fixed at the maximum and (b) the overall number of ways in which the total score could be arrived at. Means of expected proportions were calculated for each depression severity to be compared with the observed proportion by using 95% confidence intervals.

Figures 2 and 3 show the distribution of item responses by severity cut-offs on HADS-D and PHQ-9 respectively. Only one HADS-D item and two PHQ-9 items showed over 50% of moderate to severe category responses on the maximum rating: H4 ‘I feel as if I am slowed down’; P3 ‘Trouble falling or staying asleep, or sleeping too much’; and P4 ‘Feeling tired or having little energy’. The proportions of observed responses to these items were significantly higher than their corresponding expected responses.

**Clinical characteristics and depression scores**

Table 2 shows the mean scores on the two depression measures by diabetes and depression clinical characteristics. Duration of diabetes was related to both HADS-D \((P<0.001)\) and PHQ-9 \((P = 0.026)\) scores, with greater duration (≥5 years) linked to higher depression. A greater number of diabetes complications was related to higher scores on HADS-D \((P<0.001)\) and PHQ-9 \((P<0.001)\). Use of medication, either oral or oral plus insulin, was related to higher HADS-D scores \((P<0.001)\) and higher PHQ-9 scores \((P = 0.004)\). Depression scores were unrelated to recent HbA1c levels in patients’ clinical records. Obesity (BMI >30 kg/m²) was related to higher depression on both HADS-D \((P = 0.002)\) and PHQ-9 \((P = 0.001)\). Participants who had a previous episode of depression showed higher scores on HADS-D \((P<0.001)\) and PHQ-9 \((P<0.001)\), as did those being prescribed antidepressants at the time of the study (HADS-D, \(P<0.001\); PHQ-9, \(P<0.001\)).

**DISCUSSION**

**Summary of main findings**

Both the HADS-D and PHQ-9 measures demonstrated acceptable reliability and robustness of factor structure. Both questionnaires can be self-
PHQ-9 items found the HADS-D item ‘feel slowed down’, together with the PHQ-9 items concerning sleeping, tiredness, and poor appetite or overeating, showed high loadings on the first factor. These results suggest somatic symptoms and behaviours related to diabetes may be contributing to the depression scores, particularly when measured using PHQ-9.

**Strengths and limitations of the study**

This is a singular comparison of responses to PHQ-9 and HADS-D measures among participants with diagnosed type 2 diabetes in primary care, providing a snapshot of what GPs are likely to encounter in everyday practice. The study demonstrates that both measures can be used for screening through self-completion of a postal questionnaire, but that HADS-D appears to provide a more accurate view. The prevalence of depression with HADS-D is similar to other epidemiological studies in the region.\(^{15,16}\)

Gold standards such as the Structured Clinical Interview for DSM-IV\(^{17}\) were beyond the scope of the study but could probably have further clarified which measure performed better for screening. The sample was selected because of type 2 diabetes status. This study did not have a comparison sample that had depression but not diabetes, but the factor analysis of pooled HADS-D and PHQ-9 items for this group of patients with diabetes showed a different pattern of item loadings when compared with a similar analysis of primary care patients without diabetes in Sweden.\(^{14}\)

**Comparison with existing literature**

The psychometric properties of the HADS-D and PHQ-9 depression measures in this Australian population with type 2 diabetes are similar to those reported in samples of patients in primary care in the UK\(^{11,13}\) and Sweden.\(^{14}\) Comparing the two measures, PHQ-9 identified more than twice as many in the moderate to severe category as HADS-D, a pattern that has been noted in other studies.\(^{11,13}\) Other authors reported that these large differences in categorisation are not reflected in prescribing because GPs take other things, such as history of depression, into account along with the questionnaire score.\(^{11,12}\)

One study showed that although PHQ-9 categorised 83.5% of patients as having moderate to severe depression compared with 55% by HADS-D, prescription rates were almost identical at 79%, and referral rates at 23.7% and 20.3% respectively.\(^{11}\) The study also reported lower treatment rates among patients with heart disease or diabetes.

**Implications for clinical practice and future research**

Patients appear to value these self-report measures, which they regard as objective and offering them a
means to express how they feel. GPs, however, may see depression questionnaires simply as guides to decision making.\textsuperscript{12} GPs appear more inclined to accept a HADS-D score at face value than a PHQ-9 one. This is reflected in the similar rates of prescribing and referral, even though PHQ-9 seems to be used three times more often than HADS-D.\textsuperscript{11} Perhaps GPs are taking into account the fact that PHQ-9 overestimates severity because it includes somatic symptoms that GPs are willing to attribute to the underlying condition rather than depression.

The two items forming PHQ-2 seem particularly unsuitable for screening patients with diabetes. They both had a high proportion of missing responses and so could not be used to classify patients with diabetes who also have moderate to severe depression.

There has been considerable progress in screening people with diabetes for depression.\textsuperscript{1,9} The current results suggest that PHQ-9 overestimates depression among patients with diabetes because it contains questions about tiredness, sleeping problems, and eating patterns that are common in diabetes. There is a complicated relationship between obesity, diabetes, depression, and obstructive sleep apnoea.\textsuperscript{10,11} Sleep disturbances in diabetes are frequently due to nocturia, neurogenic pain, and other causes.\textsuperscript{10,11} The PHQ-9 questions about over eating or under eating, and the somatic symptom of tiredness may be accounted for by diabetes itself, or by sleep disturbances including obstructive sleep apnoea, and could account for the high classification rate for moderate to severe depression.

HADS-D is probably a better screening instrument for patients with type 2 diabetes. If using PHQ-9, GPs should assess each of the patient’s nine answers and consider causes of sleep disorder or tiredness besides depression. Focusing on individual responses, rather than total scores, may prove more effective in reducing suffering and accurately identifying the problems encountered by the patient.\textsuperscript{20}

\textbf{Funding body}
Department of Human Services Victoria (Ref No ADF/05/8456, Project ChD/2).

\textbf{Ethics committee}
Ethical approval was obtained from DHS Victoria Human Research Ethics Committee (Application No 04/07).

\textbf{Competing interests}
The authors have stated that there are none.

\textbf{Acknowledgements}
We would like to thank the general practices of Colac, Hamilton Medical Group, Moe, Pascoe Vale, Preston, Reservoir, and Warracknabeal for their help with this study.

\textbf{Discuss this article}
Contribute and read comments about this article on the Discussion Forum: \url{http://www.rcgp.org.uk/bjgp-discuss}

\textbf{REFERENCES}


### Appendix 1. Cronbach’s α and item-total correlations of the HADS-D and PHQ-9.

<table>
<thead>
<tr>
<th>Scale and items</th>
<th>Item-total correlations</th>
<th>Coefficient α (95% CI) if item deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>HADS-D (n = 543)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H1 Still enjoy things</td>
<td>0.63</td>
<td>0.79</td>
</tr>
<tr>
<td>H2 Can laugh</td>
<td>0.59</td>
<td>0.80</td>
</tr>
<tr>
<td>H3 Feel cheerful</td>
<td>0.59</td>
<td>0.80</td>
</tr>
<tr>
<td>H4 Feel slowed down</td>
<td>0.50</td>
<td>0.82</td>
</tr>
<tr>
<td>H5 Lost interest in appearance</td>
<td>0.57</td>
<td>0.80</td>
</tr>
<tr>
<td>H6 Look forward with enjoyment</td>
<td>0.72</td>
<td>0.78</td>
</tr>
<tr>
<td>H7 Can enjoy book/radio/TV</td>
<td>0.43</td>
<td>0.82</td>
</tr>
<tr>
<td>PHQ-9 (n = 462)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P1 Little interest or pleasure in doing things</td>
<td>0.76</td>
<td>0.88</td>
</tr>
<tr>
<td>P2 Feeling down, depressed, hopeless</td>
<td>0.75</td>
<td>0.88</td>
</tr>
<tr>
<td>P3 Trouble sleeping or sleeping too much</td>
<td>0.66</td>
<td>0.89</td>
</tr>
<tr>
<td>P4 Feeling tired or having little energy</td>
<td>0.75</td>
<td>0.88</td>
</tr>
<tr>
<td>P5 Poor appetite or overeating</td>
<td>0.63</td>
<td>0.89</td>
</tr>
<tr>
<td>P6 Feeling bad about self, or a failure, or have let self or family down</td>
<td>0.71</td>
<td>0.88</td>
</tr>
<tr>
<td>P7 Trouble concentrating, such as reading newspaper or watching TV</td>
<td>0.66</td>
<td>0.89</td>
</tr>
<tr>
<td>P8 Moving or speaking more slowly, or being restless, moving more than usual</td>
<td>0.61</td>
<td>0.89</td>
</tr>
<tr>
<td>P9 Thoughts of self-harm</td>
<td>0.56</td>
<td>0.90</td>
</tr>
</tbody>
</table>

HADS-D = depression subscale of the Hospital Anxiety and Depression Scale. PHQ-9 = Patient Health Questionnaire.

### Appendix 2. Factor analysis loadings with varimax rotation of pooled items from HADS-D and PHQ-9 measures of depression.

<table>
<thead>
<tr>
<th>Scale</th>
<th>Item</th>
<th>Factor 1</th>
<th>Factor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>HADS-D</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H1</td>
<td>Still enjoy things</td>
<td>0.594</td>
<td>0.338</td>
</tr>
<tr>
<td>H2</td>
<td>Can laugh</td>
<td>0.371</td>
<td>0.631</td>
</tr>
<tr>
<td>H3</td>
<td>Feel cheerful</td>
<td>0.343</td>
<td>0.670</td>
</tr>
<tr>
<td>H4</td>
<td>Feel slowed down</td>
<td>0.752</td>
<td>0.046</td>
</tr>
<tr>
<td>H5</td>
<td>Lost interest in appearance</td>
<td>0.560</td>
<td>0.299</td>
</tr>
<tr>
<td>H6</td>
<td>Look forward with enjoyment</td>
<td>0.562</td>
<td>0.568</td>
</tr>
<tr>
<td>H7</td>
<td>Can enjoy book/radio/TV</td>
<td>0.073</td>
<td>0.664</td>
</tr>
<tr>
<td>PHQ-9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>Little interest or pleasure in doing things</td>
<td>0.610</td>
<td>0.568</td>
</tr>
<tr>
<td>P2</td>
<td>Feeling down, depressed, hopeless</td>
<td>0.581</td>
<td>0.532</td>
</tr>
<tr>
<td>P3</td>
<td>Trouble sleeping or sleeping too much</td>
<td>0.657</td>
<td>0.306</td>
</tr>
<tr>
<td>P4</td>
<td>Feeling tired or having little energy</td>
<td>0.795</td>
<td>0.278</td>
</tr>
<tr>
<td>P5</td>
<td>Poor appetite or overeating</td>
<td>0.794</td>
<td>0.237</td>
</tr>
<tr>
<td>P6</td>
<td>Feeling bad about self, or a failure, or have let self or family down</td>
<td>0.531</td>
<td>0.571</td>
</tr>
<tr>
<td>P7</td>
<td>Trouble concentrating, such as reading newspaper or watching TV</td>
<td>0.267</td>
<td>0.768</td>
</tr>
<tr>
<td>P8</td>
<td>Moving or speaking more slowly, or being restless, moving more than usual</td>
<td>0.296</td>
<td>0.648</td>
</tr>
<tr>
<td>P9</td>
<td>Thoughts of self-harm</td>
<td>0.203</td>
<td>0.706</td>
</tr>
</tbody>
</table>

HADS-D = depression subscale of the Hospital Anxiety and Depression Scale. PHQ-9 = Patient Health Questionnaire. Figures in bold are items with high (>0.500) loadings on each factor.