"Cognitive and anxiety symptoms in screening for clinical depression in diabetes A systematic examination of diagnostic performances of the HADS and BDI-SF"

Sultan, Serge ; Luminet, Olivier ; Hartemann, Agnès

Abstract
BACKGROUND: Little systematic research into the diagnostic performance of instruments used to screen for clinical depression is available for people with diabetes. The objective of this study was to compare performances of the HADS and BDI-SF and their components in association with a standard diagnostic interview. METHODS: In a sample of 298 French outpatients from a diabetes clinic (165 men, aged 59.4+/-10.7 years), we assessed diagnoses of clinical depression (CD, n=42) and major depression (MD, n=30) using the MINI and administered the HADS and BDI-SF. RESULTS: Cognitive symptoms from the BDI-SF (BDI cog) were more closely associated with MD than CD. BDI cog and HADS total scores performed best overall in identifying clinical depression (AUCs under ROC curve 85%). For identification of CD, the sensitivity/specificity of BDI cognitive symptoms was 88/71% (cutoff 3+) and for the HADS 83/65% (cutoff 13+). For identification of MD, BDI cog scored 83/80% (cutoff 4+) and HAD-A 80/76% (cuto...

Document type: Article de périodique (Journal article)

Référence bibliographique
DOI : 10.1016/j.jad.2009.09.022

Available at:
http://hdl.handle.net/2078.1/28862
[Downloaded 2018/11/21 at 15:06:24]
Cognitive and anxiety symptoms in screening for clinical depression in diabetes
A systematic examination of diagnostic performances of the HADS and BDI-SF

Serge Sultan, Olivier Luminet, Agnès Hartemann

Université Paris Descartes, France
Université Catholique de Louvain, Belgium
Belgian National Fund for Scientific Research (FNRS-FRS), Belgium
Université Pierre et Marie Curie, France
Assistance Publique-Hôpitaux de Paris, Hôpital Pitié-Salpêtrière, France

Article history:
Received 21 July 2009
Received in revised form 29 September 2009
Accepted 29 September 2009
Available online 25 October 2009

Keywords:
Depression
Diabetes
HADS
BDI
Sensitivity
Specificity

1. Introduction

Although studies have consistently emphasized the impact of depression in people with diabetes, little systematic research has as yet focused on the assessment procedures used to identify depression in this population (McHale et al., 2008). Depression is frequent in people with diabetes, with prevalence ranging from 8% to 20% (Anderson et al., 2001; Chou and Chi, 2005). Compared with patients with diabetes alone, patients with diabetes and depressive symptoms exhibit higher risks of morbidity and mortality (Katon et al., 2005). It is recommended to screen systematically for depression in patients with this condition (ADA, 2009).

Research into the taxonicity of depression suggests that symptoms making this diagnosis are not equally important when it comes to diagnostic performance (Beach & Amir, 2003). The task of identifying depression reliably might be made easier by focusing on specific symptoms such as depressed mood and cognitive symptoms from the BDI-SF (BDIcog). The objective of this study was to compare performances of the HADS and BDI-SF and their components in association with a standard diagnostic interview.

Methods: In a sample of 298 French outpatients from a diabetes clinic (165 men, aged 59.4 ± 10.7 years), we assessed diagnoses of clinical depression (CD, n = 42) and major depression (MD, n = 30) using the MINI and administered the HADS and BDI-SF.

Results: Cognitive symptoms from the BDI-SF (BDIcog) were more closely associated with MD than CD. BDIcog and HADS total scores performed best overall in identifying clinical depression (AUCs under ROC curve 85%). For identification of CD, the sensitivity/specificity of BDI cognitive symptoms was 88/71% (cutoff 3+) and for the HADS 83/65% (cutoff 13+). For identification of MD, BDIcog scored 83/80% (cutoff 4+) and HADS-A 80/76% (cutoff 9+). Logistic regression analyses further suggested that BDIcog and HADS-A discriminated better between depressed and non-depressed patients than the somatic and anhedonia items present in the same scales. The depression subscale of the HADS performed poorly.

Limitations: The consecutive nature of the sample may limit the generalizability of our findings.

Conclusion: Results suggest that, in addition to depressed mood, both negative thoughts and anxiety are core elements for the correct identification of clinical depression in chronic illnesses such as diabetes. It may be more appropriate to use the total score when applying the HADS and distinguish non-somatic symptoms within the BDI.
mood, anhedonia or suicidality, which seem to represent the core disturbances observed in clinical depression.

But assessing depression in the medically ill is a challenge because any increase in prevalence may reflect either a genuine depression or physical symptoms associated with the medical illness. One method used to correct this type of confound has been to exclude features common to ill patients in the assessment of depression: weight loss, sleep disturbances, fatigue and so on. This strategy was adopted in the development of the Hospital Anxiety and Depression Scale (HADS, Zigmond & Snaith, 1983). Another method is to give symptoms that seem to characterize depression better (e.g. cognitive symptoms or anxiety) a higher weighting than less discriminating symptoms such as distress. This was the approach adopted when developing the Beck Depression Inventory-Short Form (BDI-SF) with the aim of providing clinicians with an instrument capable of detecting depression in medically ill patients in primary care (Beck and Beck, 1972; Beck et al., 1997). A focus on cognitive symptoms and anxiety is justified in diabetes since the illness is known to encourage a negative self-image and worry about the future (Rubin and Peyrot, 2001).

The objective of this research was to compare the diagnostic performances of the BDI-SF and HADS in the identification of clinical depression, with a particular focus on anxiety and the cognitive symptoms of depression.

2. Method

2.1. Participants

The sample comprised a group of 302 consecutive type 2 diabetes outpatients visiting the Diabetes Department at Pitié-Salpêtrière Hospital, Paris, France between September 2006 and November 2007. Analyses were based on 298 patients, since questionnaires were missing for 4. The study was proposed to 370 patients but 68 refused (18.4%) because of lack of time (Table 1). The patients who refused did not differ from the final sample on age or gender. Inclusion criteria corresponded to a longitudinal follow-up approach currently employed in the department: Type 2 diabetes identified at least one year prior to inclusion (ADA, 2006), age between 20 and 75 years and no major comorbidity.
apart from diabetes-related complications. The study protocol received full Institutional Review Board approval.

2.2. Materials and procedure

2.2.1. Demographic and clinical variables

Physicians from the department invited eligible patients to participate. Participants were then directed to a psychologist intern who obtained their informed consent and administered an initial structured clinical interview. The examiner was unaware of the somatic or psychological status of the patient. Self-reports were then completed by the patients and given back the same day. The demographic items were assessed by means of specially designed questions. Clinical variables were obtained from the patients’ medical records after other data had been collected.

2.2.2. MINI

All patients were administered the Mini-International Neuropsychiatric Interview version 5.0.0 for the DSM-IV (MINI, Modules A and B) as a measure of clinical depression (Sheehan et al., 1998). Interns had been trained in the administration of clinical depression by the Neuropsychiatric Interview version 5.0.0 for the DSM-IV. The highest level of discrimination was achieved by the HADS (Wald = 40.28; $\chi^2 = 62.93$) and BDlcog (Wald = 36.19; $\chi^2 = 49.89$) followed by BDI-SF (Wald = 37.45; $\chi^2 = 49.55$), HAD-D (Wald = 35.85; $\chi^2 = 48.49$), HAD-A (Wald = 35.82; $\chi^2 = 46.26$) and BDIsom (Wald = 17.30; $\chi^2 = 19.48$). We performed the same analyses to predict the probability of MD and found that the highest level of discrimination were obtained by BD-I-SF (Wald = 39.45; $\chi^2 = 61.17$) and BDlcog (Wald = 38.33; $\chi^2 = 61.56$). These scales discriminated MD even better than they did CD. The values for the other scales were lower.

When comparing ROC curves in association with the diagnosis of CD, BDI-SF and BDlcog (AUCs = 0.85 95%CI .78–0.91) outperformed BDlsom (AUC = 0.72 95%CI .64–0.80) ($z > 2.95$, $p < .01$), whereas the performance of BDI-SF was similar to that of BDlcog. HADS (AUC = 0.85 95%CI .79–0.92) outperformed HAD-A (AUC = 0.81 95%CI .73–.87) ($z = 2.10$, $p < .05$). We also found a tendency for HADS to outperform HAD-D ($p < .08$). This suggests that the HADS tends to be more accurate in identifying clinical depression when anxiety was considered in addition to depression items. We also found that the performances of BDI-SF and HADS were similar, as were those of BDlcog and HAD-D, on the one hand, and BDlcog and HAD-D on the other.

A comparison of AUCs for the diagnosis of MD showed that BDI-SF (AUC = .90, 95%CI .81–.94) and BDlcog (AUC = .90, 95%CI .82–.94) outperformed HAD-D (AUC = .80 95%CI .71–.88) and HAD-A (AUC = .81 95%CI .71–.89) with z-values over 2.13 ($p < .05$), but did not perform better than HADS (AUC = .85 95%CI .77–.91, $z = 1.31$, $p = .19$). Other comparisons produced results that were comparable to those observed when predicting CD.

Table 2 reports comparisons of scales considered at cutpoints. Among measures with high sensitivity, BDlcog 3+ had the highest screening ability. Between 78% and 98% of depressed patients had a score of 3+ at a confidence level of 95%. With regard to the published cutoffs, BDI-SF 4+ exhibited the greatest sensitivity. However, the use of this measure would lead to a high rate of false positives (specificity 47%). The next highest sensitivity rates were found for HADS 13+ and HAD-A 8+. The usual cutoff of 8+ on the depression subscale HAD-D yielded a low sensitivity of 50%. These results suggest that cognitive symptoms may be used to screen for depression and that anxiety should be included in the HADS in order to optimize the rate of depressed patients detected. When considering MD alone, we found three measures with higher levels of sensitivity: BDI-SF 4+, BDlcog 4+ and HADS 13+. BDlcog 4+ outperformed the other two measures.

To summarize, the best trade-offs between sensitivity and specificity while maximizing sensitivity for the purposes of screening for CD and MD were provided by BDlcog 3+ (CD) and...
mixed somatic samples (Golden et al., 2007). There results are also versions of the BDI over the depression scale of the HADS in studies which have suggested the relative superiority of various particularly in the case of MD. Our results are consistent with is to accurately identify depression in people with diabetes, and that anxiety symptoms are necessary in the HADS if this scale examining the ability of scales to predict the probability of

4. Discussion

Two sets of results emerge from these analyses. First, by examining the ability of scales to predict the probability of depression and systematically comparing ROC curves, we were able to confirm that cognitive symptoms are central in the BDI-SF and that anxiety symptoms are necessary in the HADS if this scale is to accurately identify depression in people with diabetes, particularly in the case of MD. Our results are consistent with studies which have suggested the relative superiority of various versions of the BDI over the depression scale of the HADS in mixed somatic samples (Golden et al., 2007). The results are also at odds with the idea that all instruments should perform equally: rate of test negatives among non-depressed; Positive Predictive Value: rate of depressed among test positives; Negative Predictive Value: rate of non-depressed among test negatives.

4+ (MD). The next best trade-offs were provided by HADS 13+ (CD and MD).

Table 2

<table>
<thead>
<tr>
<th>Detection of clinical depression: 256 Non-depressed vs 42 with clinical depression</th>
<th>Cut-point</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
<th>Agreement (%)</th>
<th>( \kappa )</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDI-SF</td>
<td>4+</td>
<td>95 (89–100)</td>
<td>47 (41–53)</td>
<td>22 (16–29)</td>
<td>98 (96–100)</td>
<td>54</td>
<td>.18</td>
</tr>
<tr>
<td></td>
<td>8+ (optimal)</td>
<td>71 (57–85)</td>
<td>81 (76–86)</td>
<td>38 (27–48)</td>
<td>95 (92–98)</td>
<td>80</td>
<td>.38</td>
</tr>
<tr>
<td></td>
<td>16+</td>
<td>29 (15–43)</td>
<td>98 (97–100)</td>
<td>75 (54–96)</td>
<td>90 (86–93)</td>
<td>74</td>
<td>.37</td>
</tr>
<tr>
<td>BDIcog</td>
<td>3+ (optimal)</td>
<td>88 (78–98)</td>
<td>71 (66–77)</td>
<td>33 (24–42)</td>
<td>97 (95–100)</td>
<td>74</td>
<td>.35</td>
</tr>
<tr>
<td></td>
<td>4+</td>
<td>73 (60–87)</td>
<td>80 (75–85)</td>
<td>37 (27–48)</td>
<td>95 (92–98)</td>
<td>79</td>
<td>.38</td>
</tr>
<tr>
<td>BDIsom</td>
<td>4+ (optimal)</td>
<td>69 (55–83)</td>
<td>68 (62–74)</td>
<td>26 (18–34)</td>
<td>93 (89–97)</td>
<td>68</td>
<td>.22</td>
</tr>
<tr>
<td>HADS</td>
<td>13+</td>
<td>83 (72–95)</td>
<td>65 (59–71)</td>
<td>28 (20–36)</td>
<td>96 (93–99)</td>
<td>67</td>
<td>.26</td>
</tr>
<tr>
<td></td>
<td>15+ (optimal)</td>
<td>79 (66–91)</td>
<td>79 (73–84)</td>
<td>37 (28–47)</td>
<td>96 (93–98)</td>
<td>79</td>
<td>.39</td>
</tr>
<tr>
<td></td>
<td>19+</td>
<td>52 (37–67)</td>
<td>93 (90–96)</td>
<td>55 (40–70)</td>
<td>92 (89–96)</td>
<td>76</td>
<td>.36</td>
</tr>
<tr>
<td>HAD-D</td>
<td>7+ (optimal)</td>
<td>71 (58–85)</td>
<td>79 (74–84)</td>
<td>36 (25–46)</td>
<td>94 (91–97)</td>
<td>78</td>
<td>.36</td>
</tr>
<tr>
<td></td>
<td>8+</td>
<td>50 (35–65)</td>
<td>86 (81–90)</td>
<td>36 (24–49)</td>
<td>91 (88–95)</td>
<td>81</td>
<td>.31</td>
</tr>
<tr>
<td></td>
<td>11+</td>
<td>26 (13–39)</td>
<td>97 (95–99)</td>
<td>61 (39–84)</td>
<td>89 (83–93)</td>
<td>87</td>
<td>.31</td>
</tr>
<tr>
<td>HAD-A</td>
<td>8+</td>
<td>81 (69–93)</td>
<td>64 (58–70)</td>
<td>27 (19–35)</td>
<td>95 (92–98)</td>
<td>66</td>
<td>.25</td>
</tr>
<tr>
<td></td>
<td>10+ (optimal)</td>
<td>71 (58–85)</td>
<td>84 (79–88)</td>
<td>42 (30–53)</td>
<td>95 (92–98)</td>
<td>82</td>
<td>.42</td>
</tr>
<tr>
<td></td>
<td>11+</td>
<td>57 (42–72)</td>
<td>88 (84–92)</td>
<td>44 (31–58)</td>
<td>93 (89–96)</td>
<td>84</td>
<td>.41</td>
</tr>
</tbody>
</table>

Detection of major depression: 256 Non-depressed vs 30 with major depression

<table>
<thead>
<tr>
<th>Cut-point</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
<th>Agreement (%)</th>
<th>( \kappa )</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDI-SF</td>
<td>4+</td>
<td>100 (100–100)</td>
<td>47 (41–53)</td>
<td>18 (12–24)</td>
<td>100 (100–100)</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>8+ (optimal)</td>
<td>77 (62–92)</td>
<td>81 (76–86)</td>
<td>32 (22–43)</td>
<td>97 (94–99)</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>16+</td>
<td>37 (19–54)</td>
<td>98 (97–100)</td>
<td>73 (51–96)</td>
<td>93 (90–96)</td>
<td>82</td>
</tr>
<tr>
<td>BDIcog</td>
<td>4+ (optimal)</td>
<td>83 (70–97)</td>
<td>80 (75–85)</td>
<td>33 (22–43)</td>
<td>98 (96–100)</td>
<td>80</td>
</tr>
<tr>
<td>BDIsom</td>
<td>4+ (optimal)</td>
<td>73 (58–89)</td>
<td>68 (62–74)</td>
<td>21 (13–29)</td>
<td>96 (93–99)</td>
<td>69</td>
</tr>
<tr>
<td>HADS</td>
<td>13+</td>
<td>83 (70–97)</td>
<td>65 (59–71)</td>
<td>22 (14–29)</td>
<td>97 (95–100)</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>17+ (optimal)</td>
<td>77 (62–92)</td>
<td>84 (80–89)</td>
<td>37 (25–48)</td>
<td>97 (95–99)</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>19+</td>
<td>57 (39–74)</td>
<td>93 (90–96)</td>
<td>49 (32–65)</td>
<td>95 (92–98)</td>
<td>89</td>
</tr>
<tr>
<td>HAD-D</td>
<td>7+ (optimal)</td>
<td>73 (58–89)</td>
<td>79 (74–84)</td>
<td>29 (19–39)</td>
<td>96 (94–99)</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>8+</td>
<td>53 (35–71)</td>
<td>86 (81–90)</td>
<td>30 (18–43)</td>
<td>94 (91–97)</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>11+</td>
<td>27 (11–42)</td>
<td>97 (95–99)</td>
<td>53 (28–79)</td>
<td>92 (89–95)</td>
<td>90</td>
</tr>
<tr>
<td>HAD-A</td>
<td>8+</td>
<td>80 (66–94)</td>
<td>64 (58–70)</td>
<td>21 (13–28)</td>
<td>96 (94–99)</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>9+ (optimal)</td>
<td>80 (66–94)</td>
<td>76 (69–94)</td>
<td>28 (18–37)</td>
<td>97 (95–99)</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>11+</td>
<td>67 (50–84)</td>
<td>88 (84–92)</td>
<td>40 (26–54)</td>
<td>96 (93–98)</td>
<td>86</td>
</tr>
</tbody>
</table>

Note. Numbers in parentheses are the 95% confidence interval; PPV = positive predictive value; NPV = negative predictive value. Optimal cut-points determined by ROC curve analyses minimizing the sum of squared complement of sensitivity and specificity to the unity. Sensitivity: rate of test positives among depressed patients; Specificity: rate of test negatives among non-depressed; Positive Predictive Value: rate of depressed among test positives; Negative Predictive Value: rate of non-depressed among test negatives.

1 The equation for computing chance-corrected PPV is: QPPV=(PPV−base rate)/(1−base rate).

Overall, our results suggest that depression in diabetes should be approached in terms of three core aspects: depressed mood and anhedonia (e.g. sadness, lack of satisfaction), cognitive symptoms (e.g. sense of failure, self-hate) and anxiety (e.g. worrying, feelings of panic) (see Joiner et al., 2005). They suggest that we should use not the HAD-D but the HADS which includes both depression and anxiety (Razavi et al., 1990) and that the use of the BDI-SF subscales is advisable in the medically ill since its components have different levels of validity when external diagnostic criteria are used.

However, the cutoffs that were examined here involved low PPVs with the result that a small percentage of the people screened as positives will actually be depressed. This raises the question of whether this percentage is higher than the depression base rate in people with diabetes. We computed a chance-corrected PPV (Kraemer, 1992) and found that the increases in
diagnostic value for BDICog 3+ were: 27% and 16% (8% and 20% prevalence). This means that it is more effective to use this screen simply than to rely on chance.

Some limitations must be acknowledged. First, we did not use a randomly selected sample, a fact which might limit the generalizability of our findings. Second, the instruments used were not developed to comply with the DSM-IV diagnostic criteria. Consequently, discrepancies with the structured interview might be due to differences in the time history of reported symptoms.

To conclude, we found that cognitive symptoms, e.g. a sense of failure or an experience of self-hate, appear to be an important aspect of depression which is not confounded with medical condition, and confirmed the important idea that symptoms of anxiety may constitute a central aspect of clinical depression in serious chronic illness. In clinical practice with somatic patients, it may be more appropriate to use the total score when applying the HADS and distinguish non-somatic symptoms within the BDI.

Role of funding source
Funding for this study was provided by Sanofi-Aventis Corporation; the corporation had no further role in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

Conflict of interest
No conflict declared.

Acknowledgments
We thank Alice Aouad, Lena Obrovac, Raluca Barna for their help in data collection and Delphine Gryenberg for help during data analysis. The article was written while the first author was on sabbatical, and the support of the Department of Psychology, Université Catholique de Louvain, is gratefully acknowledged.

References