Depression in Chronic Pain Patients: Prevalence and Measurement

Helen Poole, PhD*;†; Susie White, DClinPsych‡; Chantal Blake, DClinPsych‡,§; Peter Murphy, PhD‡,¶; Ros Bramwell, PhD**

*Faculty of Science, School of Psychology, Liverpool John Moores University, Liverpool; †Pain Research Institute, Liverpool; ‡Lancashire Care Foundation Trust, Blackpool; §University of Liverpool, Liverpool; ¶The Walton Centre for Neurology and Neurosurgery, NHS Trust, Liverpool; **University of Chester, Chester, U.K.

Abstract: This study aimed to: (1) determine prevalence of depression in patients referred to specialist pain services using the Structured Clinical Interview (SCID) diagnostic interview, (2) compare results on the Beck Depression Inventory II (BDI-II) with the SCID to determine the utility of the BDI-II as a screening tool in this population.

Thirty-six participants were recruited, mainly women, with a mean age = 47.83 years (standard deviation = 12.85 years), who were heterogeneous with regard to their pain. All completed the BDI-II and SCID. The SCID diagnosed 26 (72%) cases of depression. BDI-II scores showed 31 (86%) that reported at least mild depression. Agreement between BDI-II scores over threshold for mild depression and SCID diagnosis were assessed by Cohen’s kappa (= 0.6). ROC analysis for BDI-II scores against SCID diagnosis gave a large area under the curve (0.97, 95% confidence interval 0.93 to 1.02), suggesting BDI-II is an excellent screen for this population, although the curve was unusual in that sensitivity was high even when the false positive rate was zero. ROC analysis suggested 22 or above as an optimum cut-off score for depression on the BDI-II—higher than for a general population sample.

It has been suggested that the BDI overestimates incidence of depression in pain patients, but this study confirmed through diagnostic interview the very high incidence of depression in this population. It is therefore questionable whether there is value in screening referrals for depression. When using BDI-II for screening, audit or evaluation purposes with a pain clinic population, we suggest a cut-off of 22 or above.

Key Words: depression, chronic pain, prevalence, assessment

BACKGROUND

Depression is common in patients with chronic pain. Prevalence rates, however, have been shown to vary between 1.5% and 87% depending upon the assessment method used.1 As Morley et al.2 note, being in pain can be very stressful and can cause or worsen symptoms of depression. Similarly, it is known that depression may also amplify the pain experience, with worsening depression linked to increased pain behavior, reduced activity levels, deteriorating social and occupational functioning, as well as greater use of medical services.1,3,4

However, not all patients with chronic pain suffer from depression. When considering the prevalence of
depression, investigators must not only consider the presence of pain in patients, but also its duration, intensity, and the number of pain conditions present. Chronic pain is not a single entity, but a heterogeneous group of conditions comprising multiple etiologies and within these, varying degrees of chronicity and severity. Evidence demonstrates that multiple pains are more strongly associated with psychiatric disorders and for a variety of pain conditions, increased duration and severity are linked with a greater risk of psychological distress.6–9

Importantly, depression can interfere with treatment for chronic pain, encouraging drop out and relapse, making those who have both conditions particularly difficult to treat.10–12 For this reason, screening for depression in this patient population has been advocated. Identification of depression can lead to the provision of additional and more appropriate adjuvant treatment, which may have implications for the success or otherwise of treatment directed specifically at the pain. Certainly, the accurate assessment of depressive symptomatology is essential to inform patient management and treatment planning.13

Structured Clinical Interviews (SCID) represent the gold standard means of assessment and they are designed to enable health professionals to make standardized, reliable, and accurate diagnoses of mental disorders. The SCID is perhaps the most commonly used. It is a comprehensive interview designed to make axis I diagnoses for the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) and its reliability and validity are well established.15

However, SCID are time consuming and require specialist training to administer, and therefore not suited to routine clinical practice in the pain clinic. For this reason, self-report scales have frequently been employed to screen patients for depression in order to identify those who require further investigation. A number of such measures are available and have been used in clinical and research settings in patients with chronic pain, eg, CES-D,16 Zung Self Rating Depression Scale,17 Hospital Anxiety and Depression Scale,18 and the Beck Depression Inventory (BDI).19

Of these, the BDI is the most widely used. Its psychometric properties have been extensively studied in chronic pain populations and it is one of the two measures recommended by the IMMPACT consensus group for the measurement of emotional functioning in chronic pain clinical trials.20 A common conclusion from these studies is that the BDI may overestimate depression in this population as it includes items concerning the somatic aspects of depression which are also common features of having a chronic pain condition, eg, difficulty sleeping. As such, patients may score highly on these items because of their pain rather than their mood.21–23

The most recent revision of the BDI, the BDI-II was designed to reflect changes in DSM-IV criteria for depression and this resulted in the removal of some somatic items. Despite this, some transdiagnostic items remain. Compared with other groups, a different pattern of scoring on the BDI-II for patients with chronic pain has been identified25 and the developers advise caution when using the measure with medical patients.

The setting for the research was a chronic pain management program (PMP) within a specialist U.K. center. Patients are generally referred to this type of center after other treatments have failed and typically, they present with pain of a longer duration than those in primary care settings.

The aims of this study were to use the SCID firstly to assess the prevalence of depression in patients referred to specialist pain services, and secondly, as a criterion against which to evaluate the utility of the BDI-II as a screening tool in this population.

**METHOD**

**Design**

A cross-sectional design using both questionnaires and clinical interviews addressed the study goals.

**Participants**

A total of 36 participants were recruited for the study. As the SCID is a gold standard measure and the BDI-II a validated screening tool, a large correlation between the two was expected. Cohen notes that 35 participants provide a power of 88% to detect a correlation of 0.50 at \( P < 0.05 \), or 98% power to detect a correlation of 0.60 at the same \( P \) value.

All participants were patients with chronic pain attending the assessment clinic of a specialist pain center to determine their suitability to take part in a multidisciplinary PMP based on cognitive behavioral principles. Not all were accepted onto the program. Consistent with other studies, the generally accepted operational definition of chronic pain as an unresolved episode of pain greater than 12 weeks duration was adopted.27–29
This study was nested within a larger questionnaire-based study. During the recruitment period, 225 patients attended assessment clinics and were targeted for recruitment. Of these, 158 (70%) agreed to take part and of these, 36 completed the SCID. This represented an opportunity subsample, as participation in SCID interviews was fitted around clinical assessments and as a maximum of two researchers were available at any one time. Tests confirmed there were no significant differences in age, pain duration, sex, marital status, age on leaving full-time education, or occupational status between the larger sample and this subsample.

The majority were white British (94%) and 64% were female (n = 23). Their mean age was 47.83 (standard deviation [SD] = 12.85) years and mean pain duration was 9.25 (SD = 7.29) years. They were heterogeneous with regard to their pain, with most reporting more than one pain site; back (58%), neck and shoulders (44%), head (25%), whole body (14%), and others (e.g., arms, legs) (50%). Over half the sample were married (58%), with the remainder single (19%), separated (17%), or cohabiting (6%). Around three-quarters of the sample were not working because of pain and/or disability (67%) or for other reasons (11%).

All data were anonymized and the study was approved by the relevant local research ethics committees.

INSTRUMENTATION

The SCID

The SCID\textsuperscript{14} is a semistructured interview designed to be administered to either psychiatric or medical patients for the reliable diagnosis of DSM-IV axis I diagnoses\textsuperscript{30} and takes between 45 and 90 minutes to complete. The SCID comprises an overview section that allows participants to describe their current symptoms. Following this, six modules explore current or past mood episodes, psychotic symptoms, differential diagnosis of psychotic disorders, mood disorders, alcohol and other substance use disorders, and anxiety and other disorders. In the mood episodes section, major depression disorder (MDD), recurrent major depression disorder (RMDD), Bipolar I and II Disorder, and Dysthymic Disorder symptoms are considered. Further, MDD, RMDD, and Bipolar I Disorder can be categorized for severity as mild, moderate, or severe. Inter-rater reliability of the SCID is described as between 0.70 and 1.00 (kappas) and the tool demonstrates good reliability and validity.\textsuperscript{14}

The BDI-II

The BDI-II\textsuperscript{24} is a 21-item self-report instrument that assesses the severity of depressive symptoms in adolescents and adults over the last 2 weeks. Each item is rated on a 4-point scale (0–3) with total scores ranging from 0 to 63.

For interpretation of the BDI-II, Beck et al.\textsuperscript{24} present a table of scores indicative of: severe (29+); moderate (20–28), and mild (14–19) depression. Scores of 13 and below suggest an absence of depression. However, they do not specify cut-off scores for research purposes and recommend caution when selecting them, depending upon the sensitivity and specificity required. They suggest a conservative score of 17, providing 93% true positive rate, and an 18% false positive rate. The BDI-II has been found to have excellent internal consistency and test–retest reliability with a diverse range of samples.\textsuperscript{24,31,32}

PROCEDURE

Potential participants received an information sheet about the study with their appointment letter. On arrival at the clinic, patients were approached by two members of the research team who discussed the study with them and obtained informed written consent. All participants completed the BDI-II as part of the standard assessment procedure at the clinic. Interviews took place in a private room at the clinic and were tape recorded with permission. At the beginning of the interview, participants were reminded that they could terminate the interview at any time, refuse to answer questions, and move around as much as they needed to remain comfortable.

Interviews were conducted by two of the authors, who were both trainee clinical psychologists at the time. Each had training on the use of the SCID-101 and had gained experience in its use. As a check on reliability, six of the interviews were independently rated by both trainees then compared. The level of agreement was assessed using a Cohen’s kappa, a statistic which corrects for chance when assessing the level of agreement between two raters or measures. Kappa takes values between 0 and 1, with 1 indicating higher agreement.\textsuperscript{33} For current and past diagnoses as well as severity of the diagnoses, an inter-rater reliability kappa of 1.00 (P < 0.001) was gained, demonstrating complete agreement. Therefore, subsequent interviews were rated by the person who conducted the interview only.
Interviewers were blinded with regard to the BDI-II scores and categories prior to completion of the interview and reporting its outcome.

Statistical analyses were completed using SPSS (v14; SPSS Inc., Chicago, IL, USA).34

RESULTS

SCID Outcomes

The SCID allowed for the assessment and diagnosis of current and past psychological disorders (Table 1). We were particularly interested in depressive disorders which could also be given a severity rating.

Of the 36 patients interviewed, 26 (72%) were diagnosed with depression; of these, 17 had MDD and 9 had RMDD. None of the participants were found to have dysthymia, bipolar depression, depression with psychosis, or any other form of not specified depression. The participants diagnosed with RMDD were found mainly to have this in a mild to moderate form. Participants diagnosed with MDD were more evenly distributed across the mild, moderate, and severe categories.

Of the 26 who were currently depressed, 25 had secondary depression, ie, they reported their depression was a consequence of their pain condition. The two patients who were not depressed, but were found to have other disorders, also felt these were a consequence of their pain condition. Half of the patients reported previous psychological problems and these were predominantly anxiety or depressive disorders.

BDI-II Outcomes

BDI-II scores were classified using the cut offs in the manual and the number and percentage in each category are shown in Table 2. These data suggest that 86% of patients displayed some degree of depressive symptomatology. The mean score on the BDI-II was 27.47 (SD = 12.06), which is in the moderately depressed category.

Comparison Between BDI-II and SCID Categories

Categorical data for the four categories of the SCID and BDI-II were cross tabulated (Table 3). The level of agreement was assessed using a Cohen’s kappa. Agreement between the SCID and BDI-II was very low, Cohen’s kappa = 0.2. The BDI-II misclassifies 5 patients as depressed (3 mild, 2 moderate) who were not depressed according to the SCID.

In order to determine the utility of the BDI-II for identifying potential cases of depressed/not depressed patients, the symptomatic categories of the SCID and BDI-II were collapsed into one “depressed” category and these were compared with the “not depressed” category (Table 4). Agreement between these categories was slightly better, Cohen’s kappa = 0.6.

Table 1. Structured Clinical Interview (SCID) Results

<table>
<thead>
<tr>
<th>SCID Diagnostic Category</th>
<th>n</th>
<th>Mild</th>
<th>Moderate</th>
<th>Severe</th>
<th>Other Current Diagnoses</th>
<th>Past Diagnoses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not depressed</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td>Generalized anxiety disorder (n1)</td>
<td>Major depressive disorder (n3)</td>
</tr>
<tr>
<td>Major depressive disorder</td>
<td>17</td>
<td>2</td>
<td>9</td>
<td>6</td>
<td>Social phobia (n1)</td>
<td>Anxiety disorder (n1)</td>
</tr>
<tr>
<td>Recurrent depressive disorder</td>
<td>9</td>
<td>3</td>
<td>6</td>
<td></td>
<td>Panic disorder (n4)</td>
<td>Generalized anxiety disorder (n3)</td>
</tr>
</tbody>
</table>

Table 2. Number and Percentage in Each Category of the Beck Depression Inventory-II (BDI-II)

<table>
<thead>
<tr>
<th>BDI-II Category (Score Range)</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not depressed (0–13)</td>
<td>5</td>
<td>14</td>
</tr>
<tr>
<td>Mild (14–19)</td>
<td>5</td>
<td>14</td>
</tr>
<tr>
<td>Moderate (20–28)</td>
<td>9</td>
<td>25</td>
</tr>
<tr>
<td>Severe (29 and over)</td>
<td>17</td>
<td>47</td>
</tr>
</tbody>
</table>

Table 3. Cross-Tabulation of Structured Clinical Interview (SCID) and Beck Depression Inventory-II (BDI-II) Categories

<table>
<thead>
<tr>
<th>SCID Categories</th>
<th>Not depressed</th>
<th>Mild</th>
<th>Moderate</th>
<th>Severe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not depressed</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Mild</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Moderate</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>Severe</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
</tr>
</tbody>
</table>
**ROC Curve Analysis**

ROC curves may be used to calibrate a screening measure (such as the BDI-II) against a diagnostic (such as the SCID). The ROC graphs the true positive rate (sensitivity) against the false positive rate (1 specificity). ROC curves may be used to determine an optimum cut-off score on the screening tool. The shape of the curve can also be used as an indication of the utility of a screening tool. A useful screening measure is characterized by a large area under the curve—if the screening tool were no better than chance, then the ROC curve would look like a straight line from bottom left to top right and the area under the curve would equal 0.5.

In this case, an ROC curve was prepared for the total score on the BDI-II against a SCID diagnosis of depression (all categories). As can be seen from Figure 1, the curve produced was unusual in that sensitivity was high even when the false positive rate was zero. The area under the curve was large (0.97, 95% confidence interval 0.93 to 1.02), indicating an excellent screening test. Table 5 shows the sensitivity and specificity values for the key range. The ROC curve graph and associated table suggest a cut off in the range 18–24. Based on this sample, the optimum cut off suggested is a score of 22 or above on the BDI-II. At this score, 89% of cases would be correctly classified and only 10% of cases missed.

**DISCUSSION**

This study used a gold standard assessment for mental disorders, the SCID, firstly, to assess the prevalence of depression in patients referred to specialist pain services, and secondly, as a criterion against which to evaluate the utility of the BDI-II as a screening tool in this population. The SCID revealed that the majority of patients (26 out of 36, 72%) were depressed.

Scores on the BDI-II suggested 31 out of 36 (86%) patients had symptoms of depression. The BDI-II is not a diagnostic instrument and agreement between it and the SCID on severity of depressive symptoms was poor. Nevertheless, concordance between the two types of assessment regarding “caseness,” i.e., whether or not an individual is depressed, was reasonably good; kappa = 0.6. It should also be noted that as kappa adjusts for chance, it provides a very conservative measure of agreement. Given the correspondence between the SCID and BDI-II on “caseness,” it is reasonable to suggest that in situations where it is not possible to use the SCID, e.g., because of time constraints or a lack of trained personnel, then the BDI-II represents a good alternative.

The BDI-II manual proposes a score of 17 as a cut-off score for “caseness” but encourages clinicians and researchers to consider their own cut-offs dependent on

**Table 4. Cross tabulation of Structured Clinical Interview (SCID) and Beck Depression Inventory-II (BDI-II) (Not Depressed and Depressed)**

<table>
<thead>
<tr>
<th>SCID not depressed</th>
<th>BDI-II Not Depressed</th>
<th>BDI-II Depressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>

**Table 5. Sensitivity and Specificity for a Beck Depression Inventory-II (BDI-II) Total Score Against a Structured Clinical Interview (SCID) Diagnosis of Depression (For Cut-Offs in the Range of 14–25)**

<table>
<thead>
<tr>
<th>Test Positive if BDI-II Score greater than or equal to—*</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>14.5</td>
<td>1.00</td>
<td>0.60</td>
</tr>
<tr>
<td>16.0</td>
<td>1.00</td>
<td>0.70</td>
</tr>
<tr>
<td>17.5</td>
<td>0.96</td>
<td>0.70</td>
</tr>
<tr>
<td>18.5</td>
<td>0.96</td>
<td>0.80</td>
</tr>
<tr>
<td>20.0</td>
<td>0.92</td>
<td>0.80</td>
</tr>
<tr>
<td>22.0</td>
<td>0.89</td>
<td>0.90</td>
</tr>
<tr>
<td>23.5</td>
<td>0.85</td>
<td>1.00</td>
</tr>
<tr>
<td>24.5</td>
<td>0.73</td>
<td>1.00</td>
</tr>
</tbody>
</table>

* All cut-off values are the averages of two consecutive-ordered observed test values.
the population being assessed. In this study, ROC curve analysis suggests a cut-off of 22 to be the optimal for chronic pain patients, providing 89% sensitivity. Previous studies suggested that the BDI may overestimate the prevalence of depression in patients with chronic pain and a differential pattern of scoring for patients with chronic pain compared with psychiatric outpatients has been identified on the BDI-II. The use of a higher cut-off may impact this issue and lead to a reduction in false positive scores.

The BDI-II provides a current measure of depressive symptoms, with items referring to symptoms experienced in the last 14 days. Thus, it does not provide any distinction between duration and/or recurrence of symptoms. In contrast, the SCID also considers previous Axis I disorders. As stated, many of the current sample reported a history of such disorders, with the predominant being MDD (n = 11). Our data is cross-sectional and it is difficult to comment on the causal relationship between pain and depression. Nonetheless, the question of whether depression precedes or follows chronic pain has provoked much debate. A review of 37 studies by Fishbain et al. found data consistent with the view that depression follows the development of chronic pain. Our results lend support to this, as out of the 26 patients who were currently depressed, 25 felt this was a consequence of their pain condition. However, there is some limited evidence to suggest a bidirectional relationship between pain and depression. Gureje et al. found depression at baseline predicted onset of pain 12 months later and vice versa. Given the extent of premorbid episodes of MDD in our sample, it is possible that for some, depression preceded pain, but that this was not viewed as such by the patients themselves.

The high rate of depression found here may have implications for practice. The high prevalence of depression suggests that, rather than screening for depression, practitioners may assume that most pain clinic clients are depressed. It should be noted that these were patients attending a specialist pain center in the U.K., undergoing assessment for a PMP based on cognitive behavior therapy principles, whose mean duration of pain was 9.25 years. As such, they represent a minority of the chronic pain population for whom previous treatments have failed and/or who are coping less well with their pain than others who manage their pain either themselves or within the primary care sector. Good practice in this setting includes taking a thorough clinical history which would highlight the need for further input to deal with current or prepain psychiatric issues, and in this context, the use of a screening tool may be superfluous.

Nonetheless, it recognized that many PMPs do not only use such tools to identify symptomatic patients, but also as a measure to calibrate any changes that may occur pre- and post-treatment. The validity and reliability of the BDI-II for this purpose has been demonstrated, and its utility cannot be underestimated for PMPs that are designed to impact upon mood in addition to pain per se.

A number of methodological limitations are evident in the current study. We accept that convenience sampling may create possible sample bias and that the ROC analysis may be limited because of the relatively small sample and lack of variability in the BDI-II data. Participants were from a U.K. chronic pain population, and while the ethnicity is typical of the U.K. generally, where ~92% describe themselves as white British and ~7% non-White ethnic groups, we recognize that this may not be typical elsewhere. Thus, findings may not be transferable to other populations, particularly those in primary care settings. Nonetheless, the level of depression found in this sample is comparable with that found in other chronic pain populations. Patients attending chronic pain clinics are, however, only a subset of those actually experiencing chronic pain, and further research in primary care settings is recommended.

REFERENCES


