Comparison between the Beck Depression Inventory and psychiatric evaluation of distress in patients on long-term sick leave due to chronic musculoskeletal pain

Patricia Olaya-Contreras1
Torgny Persson2
Jorma Styf1

1Department of Orthopaedics, Institute of Clinical Sciences, The Sahlgrenska Academy at The University of Gothenburg, Sweden; 2The Lundby Hospital, Gothenburg, Sweden

Background: Chronic musculoskeletal pain (CMP) is associated with psychological distress and long-term disability. Underlying diagnoses causing long-term sickness absence due to CMP have not been explored enough. In a somatic health care setting, it is important to identify mental health comorbidity to facilitate the selection of appropriate treatment. The objectives of this study were to compare the scores of depressed mood obtained on the Beck Depression Inventory (BDI) with the diagnosis of depression made by a psychiatrist, and to study the prevalence of undiagnosed mental health comorbidity in these patients.

Methods and patients: 83 consecutive patients on sick leave (mean duration 21 months) due to CMP who had been referred by the Social Insurance Office to an orthopedist and a psychiatrist for assessment of the patient’s diagnoses and capacity to work. The mean age was 45 (23–61) years, 58% were women and 52% were immigrants. The accuracy of measurements was calculated using the Diagnostic Statistical Manual of Mental Disorders IV as the Gold standard.

Results: Psychiatric illness was diagnosed in 87% of the patients. The diagnosis was depression in 56%, other psychiatric illnesses in 31%, whereas 13% were mentally healthy. Of all the patients, only 10% had a previous psychiatric diagnosis. The median value of the BDI score was 26 points in depressed patients, whereas it was 23 in patients with other psychiatric diagnoses. The sensitivity of the BDI to detect depression was 87.5%. We found good agreement between the BDI score and a diagnosis of depression.

Conclusion: Undiagnosed psychiatric disorders were commonly seen in patients with CMP. The high sensitivity of the BDI scores enables the screening of mental health comorbidity in patients with a somatic dysfunction. The test is a useful tool for detecting distress in patients who are on long-term sick leave due to CMP and who need additional treatment.

Keywords: agreement, disability, underlying diagnoses

Introduction
Chronic musculoskeletal pain (CMP) has been strongly associated with high test scores for psychological distress and fatigue.1 The main defining attributes of psychological distress are inability to cope, change in emotional state, experienced emotional discomfort, experienced harm, and symptoms of depression and/or anxiety.2 Between 40% and 100% of patients who attend primary health care with distress express more than 1 type of physical symptoms related to pain.3 In particular, depression predicts a poor outcome of treatment and disability in patients with CMP.4-6 Further, the perception of pain by patients with CMP determines a patient’s activity level and mood.7,8
Changing social and economic conditions, stress-related illness, and psychological problems have all been shown to explain the increasing number of patients with long-term sickness such as back pain in Sweden.\textsuperscript{9} In the presence of CMP and disability, underlying diagnoses causing long-term sick leave has not been explored enough.\textsuperscript{10,11} Therefore, in a somatic health care setting, it is important to identify mental health comorbidity, to select the appropriate treatment, and to facilitate return to work of patients with CMP.

The Beck Depression Inventory (BDI) provides health care professionals with a practical tool to assess depression.\textsuperscript{12–16} The BDI score gives a measure of the severity of symptoms, which can be used to improve diagnosis and to select treatment.\textsuperscript{17} The instrument has excellent construct validity and internal consistency when assessing depression in chronic pain populations.\textsuperscript{18} However, little is known about the performance of the instrument to assess depression in patients on long-term sick leave due to CMP seen in a somatic health care setting.

Given that CMP is a complex complaint of major magnitude in Sweden, causing high cost, disability, and psychosocial detriment in these patients, it is important to screen the associated factors that contribute to delay the recovery of patients with CMP.

The aim of this study was to compare the scores obtained on the BDI with the diagnosis of depression made by a psychiatrist in this group of patients. Furthermore, we studied the prevalence of undiagnosed mental health comorbidity among patients on long-term sick leave due to CMP.

**Methods**

**Patients**

We carried out an observational study, namely a cross-sectional study. The Social Insurance Office in Göteborg referred 83 consecutive patients during the period 2003–2004, at the Capio Lundby Hospital, for an analysis of their biological and psychological functioning. The patients were recruited from the Västra Götaland region that includes the city of Göteborg. All patients were on sick leave due to a somatic (orthopedic) diagnosis. Most patients had been extensively investigated by their primary physician and by other specialists because of CMP. The Social Insurance Office had requested that the patients undergo a team assessment by an orthopedist and a psychiatrist to establish their diagnoses and capacity to work (Figure 1). The mean age of the patients was 45 (23–61) years, 47 (57%) were women and 43 (52%) were patients with a non-Swedish background. Of them, 25 (58%) were women. There was no difference for age by origin of the patients (t-test, $P < 0.05$). All the patients had been on sick leave for an average duration of 21 (3–96) months, due to CMP.

**Evaluation of sick-leave allowance**

In Sweden, patients on long-term sick leave (which is defined as sick leave that exceeds 3 months) may be referred to one of the several diagnostic centers to establish the cause of their sickness absence, to estimate their capacity to work, to determine the sick leave allowance, and to follow a rehabilitation plan (www.socialstyrelsen.se).\textsuperscript{19} Normal work capacity is defined as either the ability to perform the same task, or the ability to earn the same income, as prior to sickness. Our study was part of a more extensive study that was approved by the Ethics Committee at the University of Göteborg Dnr 7-94. Only patients who could read and write Swedish were invited to complete all questionnaires before undergoing the clinical evaluation (Figure 1).

**Orthopedic evaluation**

All patients underwent a thorough orthopedic evaluation made by the orthopedist (Jorma Styf). Measurements of physical function included the range of motion (ROM) of the cervical and lumbar spine, and the motion of all major joints and all painful joints of the upper and lower extremities. Furthermore, imaging methods were taken into account in the physical evaluation. All patients had a diagnosis according to the International Classification of Diseases (ICD)-10. The main locations of pain causing disability were the neck and shoulders (76% of patients), the back (18%), and other locations in 6%. Most patients suffered from nonspecific cervicalgia (M54.2). Seventy-eight patients (94%) experienced pain also in a subsidiary location. The results of the physical examination and pain location have been published elsewhere.\textsuperscript{20}

**Assessment of depressed mood (BDI score)**

The degree of depressed mood was measured by the BDI-1A.\textsuperscript{21–23} This version was a revision of the original instrument, published by Beck in 1971. The BDI provides a quantitative assessment of depressive symptoms, considering components of cognitive, affective, and behavioral distress in different populations as well as in patients with CMP.\textsuperscript{24,25} The BDI has excellent construct validity and internal consistency for psychiatric and nonpsychiatric populations.\textsuperscript{15–18,26} This score can be used to improve diagnosis and the selection of treatment.\textsuperscript{26} The BDI-1A version of the test is more popular than the BDI-II version, and its psychometric performance is satisfactory in different study populations.\textsuperscript{27} Nevertheless,
several authors have argued that the BDI total score can be inflated, overestimating depression in patients with chronic pain because of the over-representation of certain somatic symptoms reported by these patients.

One question in the full questionnaire that dealt with sexual activity was excluded for the whole study population. This has also been considered in previous studies. Thus, an item mean substitution method for the whole sample was performed.

We compared the BDI scores, with the Structured Clinical Interview for DSM-IV-TR Disorders (SCID), the Swedish version, as the gold standard. The orthopedist and the psychiatrist were both blinded for each other’s findings and for the BDI-score. We applied the following BDI cut-off scores to classify depression in agreement with previous studies below 13 for none or minimal depression, 14–20 for mild depression, 21–30 for moderate depression, and above 30 for severe depression.

Twelve patients of the 83, 5 were women and 5 were immigrants, did not answer the BDI questionnaire, leaving 71 patients in the sample (Figure 1). Five of the 12 patients were diagnosed with psychiatric conditions other than depression and 3 of the 12 patients with depression. Four of the 12 had no psychiatric illness.

**Psychiatric assessment**

Psychiatric illness was assessed by a psychiatrist (Torgny Persson) employing the SCID which investigated whether the patient’s syndrome fulfills the ICD-10 diagnostic criteria for a depressive episode. The main clinical diagnoses found in our patients were major depressive episode (n = 40). Other main psychiatric diagnoses found: post-traumatic stress disorder (n = 6), psychosis (n = 1), stress syndrome (n = 2), adjustment disorder (n = 1), alcohol dependence (n = 3), pain disorder (n = 6), conversion disorder (n = 1), borderline personality disorder (n = 1), attention deficit hyperactivity disorder (n = 1), and 9 patients had no psychiatric illness.

**Statistical analysis**

Data from the 71 patients who completed the BDI questionnaire and who underwent psychiatric and orthopedic evaluations were analyzed by statistical methods. Results are given as mean and median values and by categories (percentages). Groups were compared using the Mann–Whitney U test and the Kruskall–Wallis test, and categorical data were compared using the $x^2$ test. All $P$ values reported are two-sided and significant at the 5% level ($P < 0.05$).
Assessment of the accuracy of measurements

The following indices of the accuracy, reliability, and validity were calculated in two groups: case 1 (total = 49): patients with only clinical depression (n = 40) compared with patients diagnosed not to have a psychiatric illness (n = 9); case 2 (total = 31): patients with psychiatric illness, ie, other than depression (n = 22) compared with patients diagnosed not to have a psychiatric illness (n = 9).

We calculated accuracy values, sensitivity, specificity, and likelihood ratios for dichotomous tests (positive and negative) of the BDI questionnaire. The percentage agreements were calculated for the two cases described above, in order to test the diagnostic concordance between clinical depression or psychiatric illness diagnosed by the psychiatrist and mood determined by the BDI scores. The ratio between the number of items that agreed to the total number of observations gave the degree of agreement. Inter-rater agreement was determined by comparing the psychiatric diagnosis and the BDI scores, calculating Cohen’s kappa (κ) for categorical judgments. Inter-rater agreement was determined by carrying out the McNemar’s test which compares two diagnostic tests in the same sample of individuals (where $P > 0.05$ indicates inter-rater agreement). SPSS (version 17; SPSS Inc., Chicago, IL) was used in all statistical analyses.

Results

Orthopedic and psychiatric assessment

Normal neuromuscular function was observed in 69% (49/71) of the patients at clinical investigation. Minor impairment was found in 31% (22/71). Decreased joint ROM and muscular weakness were the impairments most commonly seen.

Psychiatric illness was diagnosed in 87% (62/71) of the patients. Different grades of major depressive episode, was the main diagnosis, being made for 56% (40/71) of these patients. Other psychiatric diagnoses were given for 31% (22/71) of the patients, whereas 13% (9/71) had no psychiatric illness (Table 1). The prevalence of psychiatric illness was not statistically different between patients with normal neuromuscular and joint function and patients with impairment (Fisher’s exact test, $P = 0.816$).

Despite signs of prior presence of psychological illness in most of our patients, only 10% (7/71) of them had been diagnosed with psychiatric illness before this investigation. Thus, psychiatric illness was from the diagnostic point of view a hidden comorbidity in 89% (55/62) of the patients on sick leave due to CMP.

| Psychiatric diagnosis (number of patients) | BDI score (mean/median values) | Percentage of the total (%)
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression (n = 40)</td>
<td>28/26</td>
<td>17</td>
</tr>
<tr>
<td>Mild (n = 12)</td>
<td>20/20</td>
<td>33</td>
</tr>
<tr>
<td>Moderate to severe (n = 24)</td>
<td>29/29</td>
<td>6</td>
</tr>
<tr>
<td>Extreme depression (n = 4)</td>
<td>48/44</td>
<td></td>
</tr>
<tr>
<td>Other psychiatric diagnoses (n = 22)</td>
<td>24/23</td>
<td>31</td>
</tr>
<tr>
<td>No psychiatric illness (n = 9)</td>
<td>18/21</td>
<td>13</td>
</tr>
<tr>
<td>Total (n = 71)</td>
<td>26/23</td>
<td>100</td>
</tr>
</tbody>
</table>

Notes: *Comparisons between degrees of clinical depression: Kruskall–Wallis and median test, $P < 0.005$; **Group comparisons: Kruskall–Wallis test, $P < 0.01$ (clinical depression and other psychiatric illness vs no psychiatric illness).

Abbreviation: BDI, Beck Depression Inventory.

Assessment of depressed mood by the BDI

Depressed mood was seen in 83% (59/71) of the patients. The median values of the BDI score were 26 points for patients with clinical depression and 23 for patients with other psychiatric diagnoses. Patients with no psychiatric illness had 21 points. Patients with severe or moderate depression had higher scores of BDI than patients with mild or no depression (Kruskall–Wallis and median test, $P < 0.005$; Table 1). According to the BDI scores, 28% (20/71) of the patients were classified with the equivalent of severe depression, 27% (19/71) with moderate depression, 28% (20/71) with mild depression, and 17% (12/71) were not depressed nor had minimal depression (Table 2). Furthermore, scores on BDI exceeding 21 points were seen in 39/40 patients diagnosed by the psychiatrist with moderate or severe depression (Table 2).

Table 1 Mean and median values of the BDI scores for groups having different psychiatric diagnoses (DMS-IV-RT) in patients with long-term sick leave due to chronic musculoskeletal pain

<table>
<thead>
<tr>
<th>Levels of depression according to the BDI cut-scores (% of all patients with and without depression, n = 71)</th>
<th>Psychiatric diagnoses (DMS-IV-RT; % of all patients with and without depression, n = 71)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severe depression: BDI &gt; 30 points n = 20 (28%)</td>
<td>Depression n = 40 (56%)</td>
</tr>
<tr>
<td>Moderate depression: BDI 21–30 points n = 19 (27%)</td>
<td>Other psychiatric diagnoses n = 22 (31%)</td>
</tr>
<tr>
<td>Mild depression: BDI 13–20 points n = 20 (28%)</td>
<td>No psychiatric diagnosis n = 9 (13%)</td>
</tr>
<tr>
<td>Minimal depression: BDI &lt; 13 points n = 12 (17%)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: BDI, Beck Depression Inventory.
Assessment of the accuracy of measurements

Table 3 presents the distribution of the BDI scores for both cases of comparison among the psychiatric diagnoses. The accuracy and agreement were higher when screening depression (case 1) than they were when screening other psychiatric illnesses (case 2) as expected. The sensitivity of the BDI questionnaire to detect depression was 87.5% in case 1 and 86% in case 2 (Table 4). The specificity remained the same in both cases (~50%), whereas the accuracy was higher in case 1 compared with case 2 (82% vs 77%). According to the Positive likelihood ratio, higher scores on BDI (≥13) were 1.6 times more likely to occur in patients with depression, and in patients with other psychiatric diagnoses than in those with no psychiatric illness (Table 4). The agreement between the BDI score and the diagnosis of major depressive episode (SCID) was 80% (39/49) in case 1 and 74% (23/31) in case 2 (Table 4).

Discussion

This study has shown that the BDI scores agree well with a psychiatric assessment of depression in orthopedic patients. Further, psychiatric disorders seem to be an important non-diagnosed comorbidity in patients on long-term sick leave due to a somatic diagnosis of CMP.

Assessment of depression by the BDI-scores

In the present study, the agreement between the psychiatric diagnosis of depression and the scores of BDI (BDI ≥ 13) was much better when depression was classified into three classes – mild, moderate, and severe – than it was when depression was defined simply as “minimal” in line with previous findings. Previous studies have shown that a psychiatric diagnosis of depression agrees well with BDI scores in patients with CMP. However, there is no previous study making this validation among patients on long-term sick leave due to CMP in a somatic health care setting.

Methodological considerations

The BDI test does not perform uniformly in nonclinical populations. Thus, higher BDI cut-off scores have been recommended to obtain the highest levels of sensitivity and specificity. We applied a cut-off value of 13 for the classification of depression rather than the original value of 10. This ensured that all groupings contained sufficient numbers of patients. Otherwise, the traditional sensitivity/specificity test methodology might have been problematic. Similar results (ie, high sensitivity and low specificity values) have been found using a BDI cut-off value of 13 points in patients with chronic pain. In this context, a good sensitivity appears to be indispensable according to several researchers. Furthermore, the high sensitivity of the BDI scores reduces the probability of false negatives, which may have serious consequences for orthopedic patients who need appropriate treatment for an undetected depression. In conclusion, our findings validate the BDI scores obtained by our patients. The BDI consists of two subscales, negative affect and physical function. However, the results of the present study are based on the total BDI score. Consequently, we expect that the percentage agreement will improve only when the negative affect subscale is correlated with the psychiatric diagnosis of depression. Using the total BDI score as well as the high prevalence of psychiatric diagnoses may reduce the inter-rater agreement. It is documented, for instance, that the value of κ depends upon the proportion of subjects (prevalence) in each category. Nevertheless, the agreement in diagnosis in this study was high considering that the BDI score is a result of self-reported depression. Certainly, the

### Table 3

<table>
<thead>
<tr>
<th>Psychiatric assessment</th>
<th>Case 1</th>
<th>Case 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression</td>
<td>No psychiatric illness</td>
<td>Other psychiatric illness</td>
</tr>
<tr>
<td>BDI cut-off</td>
<td>n = 40</td>
<td>n = 9</td>
</tr>
<tr>
<td>BDI ≥ 13</td>
<td>35 (72%)</td>
<td>5 (10%)</td>
</tr>
<tr>
<td>BDI &lt; 13</td>
<td>5 (10%)</td>
<td>4 (8%)</td>
</tr>
</tbody>
</table>

Notes: Case 1: depression compared with no psychiatric illness; Case 2: other psychiatric illness irrespective of depression compared with no psychiatric illness. The figures given are percentages of the total within the group.

Abbreviation: BDI, Beck Depression Inventory.

### Table 4

<table>
<thead>
<tr>
<th>BDI cut-off</th>
<th>Sensitivity</th>
<th>LR (+)/LR (-)*</th>
<th>Percent agreement</th>
<th>Kappa (κ)**</th>
<th>McNemar test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>BDI ≥ 13</td>
<td>87.5</td>
<td>1.58/0.28</td>
<td>80.0</td>
<td>0.32*</td>
</tr>
<tr>
<td>Case 2</td>
<td>BDI ≥ 13</td>
<td>86.0</td>
<td>1.55/0.31</td>
<td>74.0</td>
<td>0.33*</td>
</tr>
</tbody>
</table>

Notes: Case 1: depression compared with no psychiatric illness (n = 49); case 2: other psychiatric illness irrespective of depression compared with no psychiatric illness (n = 31). LR (+) = sensitivity/specificity; LR (-) = 1-sensitivity/specificity; κ: Cohen’s Kappa for dichotomous variables; McNemar’s test. *Statistically significant P < 0.05; **Statistically significant P > 0.05.

Abbreviation: BDI, Beck Depression Inventory.
BDI score can be exaggerated or minimized by the person completing it, as is also the case for other self-report inventories. Four of 9 patients had BDI scores that were greater than 21 points which is equivalent of moderate depression. However, psychiatric evaluation revealed no psychiatric diagnosis in these patients. This can possibly be explained first, because of the BDI score overestimation of depression previously referred to second, when sickness-related pain and disability are at issue or in the context of symptoms medically unexplained.

Psychiatric assessment
It is well known that the experience of pain is accompanied by emotional reactions. Emotions such as anger, frustration, fear, and sadness often occur simultaneously with depression in patients with CMP. The psychiatric evaluation in this study revealed that approximately 90% of patients who were sick-listed due to somatic diagnoses had mental health comorbidity that was not diagnosed before this consultation. A high prevalence of mental health comorbidity may be expected in other somatic health care settings which treat patients with CMP, and in other clinical settings of Social Insurance Medicine. Our study raises the question whether a substantial number of patients with unspecified CMP have been labeled and sick-listed under incomplete diagnoses. In this context, Sweden has the highest rate of morbidity and duration of sick leave caused by CMP in Europe. It is possible that sick leave reported in Sweden due to a diagnosis of CMP alone may be overestimated.

Limitations
It is possible that there are more immigrant patients among our patients than there are in the usual flow of patients in general clinical settings. Furthermore, our patients constitute a special cohort that had been on long-term sick leave and referred by the Social Insurance Office for clinical evaluation and assessment of capacity to work. Therefore, our results can be extrapolated only to similar patient populations in similar circumstances.

Clinical implications
Self-reported depression predicts disability. It is, therefore, important to identify mental health comorbidity in patients on long-term sick leave due CMP, in order to be able to select the appropriate treatment and facilitate a return to work. A biopsychosocial approach is helpful in recognizing comorbidity that is affecting the recovery process of these patients. This approach is necessary in a somatic health care setting for referring properly a patient with these symptoms. The majority of our patients were suffering from nonspecific musculoskeletal pain with an additional nondiagnosed psychiatric disorder.

Conclusion
The BDI score agrees well with a diagnosis of depression in patients on long-term sick leave due to CMP. The sensitivity of the BDI is good enough, to enable the orthopedist to detect symptoms of depression in the case that psychiatric assessment is not accessible. Nondiagnosed psychiatric disorder seems to be a prevalent problem among patients sick-listed for an extended period of time due to CMP.

Disclosure
The authors report no conflicts of interest in this work.

References


