



# DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

## Are 2 Questions Enough to Screen for Depression and Anxiety in Patients With Chronic Low Back Pain?

The Harvard community has made this article openly available. [Please share](#) how this access benefits you. Your story matters.

<b>Citation</b>	Reme, Silje Endresen, Stein Atle Lie, and Hege R. Eriksen. 2014. "Are 2 Questions Enough to Screen for Depression and Anxiety in Patients With Chronic Low Back Pain?" <i>Spine</i> 39 (7): E455-E462. doi:10.1097/BRS.0000000000000214. <a href="http://dx.doi.org/10.1097/BRS.0000000000000214">http://dx.doi.org/10.1097/BRS.0000000000000214</a> .
<b>Published Version</b>	<a href="https://doi.org/10.1097/BRS.0000000000000214">doi:10.1097/BRS.0000000000000214</a>
<b>Accessed</b>	February 16, 2015 10:37:05 AM EST
<b>Citable Link</b>	<a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:12406597">http://nrs.harvard.edu/urn-3:HUL.InstRepos:12406597</a>
<b>Terms of Use</b>	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a>

*(Article begins on next page)*

## EPIDEMIOLOGY

OPEN

# Are 2 Questions Enough to Screen for Depression and Anxiety in Patients With Chronic Low Back Pain?

Silje Endresen Reme, PhD,\*†‡ Stein Atle Lie, PhD,\* and Hege R. Eriksen, PhD\*

**Study Design.** Cross-sectional study.

**Objective.** To examine the sensitivity of 2 single-item questions compared with 2 longer questionnaires for screening depression and anxiety among patients with chronic low back pain (CLBP).

**Summary of Background Data.** Psychosocial factors are frequently identified as risk factors for developing CLBP and as predictors for treatment, and questionnaires are often used to screen for this. Shorter instruments may be easier to use in clinical practice settings.

**Methods.** A total of 564 patients with 2 to 10 months of at least 50% sickness absence due to nonspecific low back pain were assessed for depression and anxiety with the Mini-International Neuropsychiatric Interview (MINI). Single-item questions for depression and anxiety from the Subjective Health Complaint Inventory and 2 longer questionnaires, the Hospital Anxiety and Depression Scale and Hopkins Symptom Checklist-25, were compared with MINI results, considered the “gold standard” in this study. Sensitivity and specificity of single-item and longer questionnaires and receiver operating characteristic curves were compared.

**Results.** According to MINI, the prevalence of anxiety disorders was 12% whereas that of depressive disorders was 4%. The screening questions showed 95% sensitivity and 56% specificity for depressive disorders and 68% sensitivity and 85% specificity for anxiety disorders. The longer questionnaire, Hospital Anxiety and Depression Scale, showed 91% sensitivity and 85% specificity for depressive disorders and 58% sensitivity and 83% specificity for anxiety

disorders. Hopkins Symptom Checklist-25 showed 86% sensitivity and 74% specificity for depressive disorders and 67% sensitivity and 87% specificity for anxiety disorders. For 3 of the anxiety disorders and 2 of the depressive disorders, a perfect sensitivity was found between the screening questions and MINI.

**Conclusions.** A single-item screening question was sensitive for depression but less sensitive for anxiety. The screening questions further performed equal to 2 widely used questionnaires. Validation of these results in other populations and compared with other short-item screeners is needed.

**Key words:** chronic low back pain, screening, depression, anxiety, comorbidity sick leave, single-item screening, psychiatric symptoms, diagnostic tool, risk screening.

**Level of Evidence:** 3

**Spine 2014;39:E455-E462**

Comorbid psychiatric symptoms are frequently reported in patients with chronic low back pain (CLBP).<sup>1</sup> Current psychiatric symptoms in CLBP imply poor prognosis,<sup>2</sup> poor outcome,<sup>3</sup> and high health care utilization.<sup>4</sup> Previous research with a focus on comorbidity has primarily been on depression,<sup>5</sup> but recent findings also emphasize the importance of anxiety in this patient group.<sup>6-8</sup>

Physicians are not always successful in identifying and recognizing psychiatric symptoms. Between 30% and 50% of patients with depression fail to be recognized by their primary care physicians.<sup>9</sup> Unrecognized psychiatric symptoms bear consequences on an individual level as well as at the societal level. Undertreatment of psychological disorders and increased disability costs are 2 such examples.<sup>10,11</sup> Clinical guidelines emphasize the importance of screening for psychiatric symptoms in patients with CLBP,<sup>12</sup> and systematic reviews from primary care further indicate that screening for depression can improve outcomes by reducing the risk of persistent depression.<sup>13-15</sup>

Several short and accurate instruments for detecting depression are available,<sup>16-19</sup> and systematic reviews have found short instruments to be as good as longer questionnaires in detecting depression and, possibly, anxiety.<sup>13,18</sup> In a study of patients with low back pain (LBP), a 2-item screening test was even more accurate in screening for depressive symptoms than physical therapists' ratings were.<sup>19</sup> Previous research related to screening of depression has primarily been conducted

From the \*Uni Health, Uni Research, Bergen, Norway; †Department of Environmental Health, Harvard School of Public Health, Boston, MA; and ‡Department of Pain Management and Research, Oslo University Hospital, Oslo, Norway.

Acknowledgment date: January 7, 2013. First revision date: September 15, 2013. Second revision date: November 29, 2013. Acceptance date: January 2, 2014.

The manuscript submitted does not contain information about medical device(s)/drug(s).

The Research Council of Norway grant funds were received in support of this work.

No relevant financial activities outside the submitted work.

This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 3.0 License, where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially.

Address correspondence and reprint requests to Silje Endresen Reme, PhD, Uni Health, Uni Research, Krinkelkroken 1, 5020 Bergen, Norway; E-mail: sreme@hsph.harvard.edu

DOI: 10.1097/BRS.0000000000000214

in primary care settings,<sup>20</sup> with little focus on screening for psychiatric symptoms in secondary care. Yet, patients with CLBP are frequently seen in secondary care settings where psychiatric symptoms are either not acknowledged or the patients are subjected to a range of different questionnaires as part of the standard medical procedures. Recent findings have indicated that one simple question might be enough to screen for depression in different chronic pain populations when a validated questionnaire was used as the “gold standard.”<sup>21</sup> However, these results need to be tested with a diagnostic interview as the “gold standard” comparator. In the current study, the Mini-International Neuropsychiatric Interview (MINI) was used as the “gold standard.” MINI is a short-structured interview with high reliability and validity for assessing psychiatric disorders.<sup>22</sup> Furthermore, few studies have focused on anxiety as well as depression screening in patients with CLBP despite the fact that anxiety and pain-related fear may be more strongly correlated with LBP disability than with depression,<sup>8</sup> and more disabling than the pain itself.<sup>23</sup>

The aim of this study was to compare the sensitivity of 2 single-item screening questions for depression and anxiety, with 2 longer, widely used, and validated questionnaires, against a diagnostic interview to determine whether the single-item screening questions could be enough to screen for depression and anxiety in patients with CLBP.

## MATERIALS AND METHODS

### Sample and Procedure

A total of 564 patients sick-listed 2 to 10 months for nonspecific LBP were included in the study. They were all consecutive participants in an ongoing clinical trial (the CINS trial) for LBP, recruited through written invitation from the National Insurance Administration.<sup>24</sup> The inclusion criteria of the trial were age 20 to 60 years, sick-listed at least 50% for 2 to 10 months due to nonspecific LBP, and at least 50% employed. Exclusion criteria were pregnancy, osteoporosis (known osteoporotic fracture), hemophilia, cancer, recent back trauma, serious psychiatric disorder (mainly due to ongoing psychosis, high suicide risk, and/or serious depression), cardiovascular disease, ongoing insurance trial (workers' compensation claim), and not fluent in the Norwegian language. Sixty-three participants were excluded because of 1 or more of these criteria. At baseline, all patients completed validated questionnaires, had a clinical examination by a medical doctor, and were screened for psychiatric diagnoses.

### Psychiatric Interview

MINI, which is a structured diagnostic interview<sup>22</sup> for *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV)*,<sup>25</sup> and *International Classification of Diseases, Tenth Revision (ICD-10)*,<sup>26</sup> assessing psychiatric disorders, was used as the “gold standard” comparator in the current study. It is based on “yes” and “no” answers and covers 23 axis 1 disorders. In the multiaxial system of *DSM-IV*, axis I disorders include all major mental disorders.

The interview usually takes 15 to 20 minutes and has high reliability and validity.<sup>22</sup> We used the Norwegian version of MINI plus.<sup>27</sup> The MINI plus is a version of MINI that is particularly designed for research. Various time frames are used for the different disorders: current, past, or lifetime. “Current” refers to past 2 weeks (major depression), past month (panic disorder, social phobia, obsessive-compulsive disorder, post-traumatic stress disorder [PTSD]), or ongoing at the time of the interview with no further time reference (agoraphobia, generalized anxiety disorder). In this study, we analyzed only current depressive disorders (consisting of 3 diagnoses: major depressive episode, major depressive episode with melancholia, and dysthymia) and current anxiety disorders (consisting of 6 diagnoses: generalized anxiety disorder, agoraphobia, panic disorder, social phobia, obsessive compulsive disorder, and PTSD). The interviews were conducted by 8 medical doctors specialized in physical medicine and rehabilitation who had been trained in the administration of MINI. All interviews were audio taped, and 90 audiotapes were randomly selected for evaluation. The evaluations showed satisfactory quality of the majority of interviews. For interviewers with lower quality ratings, immediate feedback and supervision were provided. The interviewers administering MINI were blinded (masked) for the results of the 2 longer questionnaires and the single-item questions.

### Screening Questionnaires

All patients filled out 2 longer, validated screening questionnaires assessing psychiatric symptoms. The sensitivity of these questionnaires was further compared with the single-item questions against the psychiatric interview. The following 2 questionnaires were used:

#### *Hospital Anxiety and Depression Scale*

The Hospital Anxiety and Depression Scale (HADS) is a measure of anxiety and depression<sup>28</sup> where participants are asked to answer the questions on the basis of how they have been feeling over the past week. It has been found to be a reliable instrument for detecting states of depression and anxiety and is also a valid measure of severity of the emotional disorders.<sup>29</sup> No somatic items or items concerning sleep problems are included, and the scale thus avoids overlap with somatic symptoms of physical illness. It consists of 7 items for depression and 7 items for anxiety. Scores for each item range from 0 to 3, and the higher the score, the more severe the disorder. Cronbach  $\alpha$  coefficients of internal consistency have been documented across several studies, with a mean of 0.83 for anxiety and 0.82 for depression.<sup>29</sup> A cutoff score of 8 or more on both subscales was used because this has been found to give an optimal balance between sensitivity and specificity (~0.80 for both subscales) according to *DSM-III*, *DSM-IV*, *ICD-8*, and *ICD-9*.<sup>29</sup>

#### *Hopkins Symptom Checklist-25*

Hopkins Symptom Checklist-25 (HSCL) measures anxiety and depressive symptoms<sup>30</sup> and has particularly been found to

be a reliable and valid questionnaire in identifying depressive symptoms.<sup>31</sup> The scale consists of 25 questions about the presence and intensity of the most common psychiatric symptoms of anxiety and depression, with response alternatives ranging from 1 (not bothered) to 4 (extremely bothered). The patients received the following instructions: "To what extent have you been bothered by the following symptoms in the last 14 days including today?" Internal consistency has been found to be high, with Cronbach  $\alpha$  coefficients ranging from 0.84 to 0.87.<sup>30</sup> We used the widely used cutoff of 1.75 to define depression and anxiety on the respective subscales, which has been found to give an optimal balance between sensitivity and specificity, particularly for females (0.82).<sup>32,33</sup>

### Screening Questions

The screening questions were taken from the Subjective Health Complaints (SHC) inventory. The SHC inventory records complaints, without asking for attributions or medical diagnosis.<sup>34</sup> The selection of questions is not based on any specific theory but covers, instead, frequent health complaints and reasons for encounter with a general practitioner.<sup>35</sup> The SHC inventory consists of 29 questions concerning severity and duration of common health complaints. The participants received the following instructions: "On the following page you find some common health problems and complaints. We want you to look at each and every one of them and report to *what extent you have been affected* during the last month." The respondents report whether, and to what extent, they have been affected in the last 30 days, ranging from 0 to 3 (0 = not at all; 1 = a little; 2 = some; 3 = serious). In the current study, we used only 2 questions concerning anxiety/nervousness and depression/low mood. The exact wording of the anxiety and depression items on the SHC was "anxiety" for the anxiety item and "sad, depressed" for the depression item. They are further referred to as the screening questions.

### Statistical Analysis

The primary outcome of the study was the reporting of sensitivity and specificity compared with the "gold standard." Receiver operating characteristic (ROC) curves were used to examine the discriminating ability of the single-item screening questions compared with the "gold standard" (MINI). The area under the ROC curve (AUC) was calculated as a measure of the ability of the 2 single-item screening questions to discriminate between patients with and without a diagnosis of depression or anxiety. According to a rough guide,<sup>36</sup> an AUC between 0.70 and 0.80 represents a fair test, between 0.80 and 0.90 represents a good test, and between 0.90 and 1 represents an excellent test. Sensitivity and specificity and positive and negative predictive values were calculated. To study whether HADS and HSCL had additional information compared with the screening questions, we made ROC curves based on predicted probabilities from logistic regression models (generalized linear models), including the screening questions as continuous variables and then entering HADS and HSCL subsequently, first, as categorical variables and, second, as continuous variables. Testing of additional statistical signif-

icant contribution to the AUC was based on likelihood ratio test from the generalized linear models. The statistical analyses were performed in SPSS (version 18) and the statistical program "R" (version 2.13; <http://www.r-project.org>) using the libraries verification and gam.

### Ethical Considerations

The Regional Ethical Committee and the Norwegian Social Science Data Services approved the study. All principles of the Helsinki declaration were followed. Informed consent was signed by each participant, with emphasis on the right to withdraw from the study at any time without any explanation. The trial is registered at <http://www.clinicaltrials.gov>, with registration number NCT00463970.

### RESULTS

The study population comprised 50.5% females. Mean age of the population was 45 years, and the participants reported an average of 11 years' duration of their back pain (Table 1). The prevalence of current depressive disorders was 4% (n = 21) and the prevalence of current anxiety disorders was 12% (n = 69), as assessed by the "gold standard" (MINI). Missing data were largest on the single anxiety question (6.2%; n = 35) and depression question (5.9%; n = 33), with somewhat less on HSCL (3.0%; n = 17), HADS (2.2%; n = 12), and MINI (1.8%; n = 10).

### Prevalence of Psychiatric Symptoms

According to HADS, 18% of the patients (n = 100) scored above the cutoff for depression and 23% (n = 125) scored above the cutoff for anxiety. According to HSCL, 29% (n = 156) scored above the cutoff for depression and 20% (n = 111) scored above the cutoff for anxiety. On the basis of the 2 single-item screening questions, 46% (n = 243) reported to have been bothered by depression whereas 22% (n = 115) reported to have been bothered by anxiety during the last month.

### Discriminating Abilities of the Screening Questions to Identify Depression and Anxiety

#### Depression

The screening question for depression showed good discriminating abilities in identifying depressive disorders as assessed by MINI. The single-item screening question showed 95% sensitivity in detecting current depressive disorders but somewhat low specificity (56%) (Table 2). For 2 of the 3 depressive diagnoses, major depressive episode (100%; 95% confidence interval [CI] = 0.77–1.0) and dysthymia (100%; 95% CI = 0.31–1.0), a perfect sensitivity was observed, meaning that the screening question could identify every patient with an ongoing depressive episode or dysthymia. The AUC when the screening question was included as a continuous variable (0, 1, 2, 3) was 0.90 (95% CI = 0.82–0.97), whereas AUC values for the different response alternatives were somewhat lower (Table 2), implying that the single-item screening question is an excellent screening item to identify patients

**TABLE 1. Baseline and Clinical Characteristics (N = 564)**

Continuous Variables	Mean	SD	Median
Age, yr	45	9.8	45
Duration of back pain, yr	11	10.5	7.5
Back pain intensity (0–10)	6.5	1.9	7
Pain during activity (0–10)	5.9	2.1	6
Pain while resting (0–10)	4.0	2.3	4
Subjective Health Complaints, no. complaints	10	4.9	10
Categorical Variables	n	%	
Sex			
Male	273	49.5	
Female	278	50.5	
Civil status			
Married/cohabitant	396	71.9	
Single/widow/divorced	155	28.1	
Education			
Primary school 1–12 yr	385	70.1	
University/college	164	29.9	

with depressive symptoms that may or may not represent a depressive disorder.

One of the longer questionnaires (HADS) showed similar discriminating abilities as the single-item screening question, with 91% sensitivity and 85% specificity (Table 2). The AUC

was 0.88 (95% CI = 0.80–0.95). The other longer questionnaire (HSCL) showed similar but poorer discriminating abilities, with 86% sensitivity and 74% specificity (Table 2). The AUC was 0.80 (95% CI = 0.71–0.89). We found no statistical significant gain from including information from HSCL when it was added to the screening question in the logistic regression model, neither as a categorical variable nor as a continuous variable. HADS had a statistically significant impact on the relation to MINI, but the change in the AUC was low and may not have practical or clinical consequences. The discriminating abilities of HADS, HSCL, and the screening question were further visualized in Figure 1.

**Anxiety**

The screening question for anxiety did not show as good discriminating abilities in identifying anxiety disorders determined by MINI (Table 3). The screening question showed a somewhat low sensitivity (68%) but a high specificity (85%). However, for 3 of the anxiety diagnoses, a perfect sensitivity was found: panic disorder (100%; 95% CI = 0.56–1.0), obsessive-compulsive disorder (100%; 95% CI = 0.52–1.0), and PTSD (100%; 95% CI = 0.31–1.0). The AUC when included as a continuous variable (0, 1, 2, 3) was 0.79 (95% CI = 0.72–0.86), whereas AUC values for the different response alternatives were somewhat lower (Table 3), implying that the screening question is a fair/good screening question to identify patients with anxiety disorders.

One of the longer questionnaires (HADS) showed poorer discriminating abilities than the single-item screening question, with only 58% sensitivity and 83% specificity, which gives an AUC of 0.76 (95% CI = 0.63–0.78). The other longer questionnaire (HSCL) showed similar discriminating abilities as the screening question, with 67% sensitivity and 87% specificity. The AUC indicated that it is a fair test in

**TABLE 2. Discriminating Abilities of the Single-Item Depression Question, HADS, and HSCL When Compared With MINI (N = 525)**

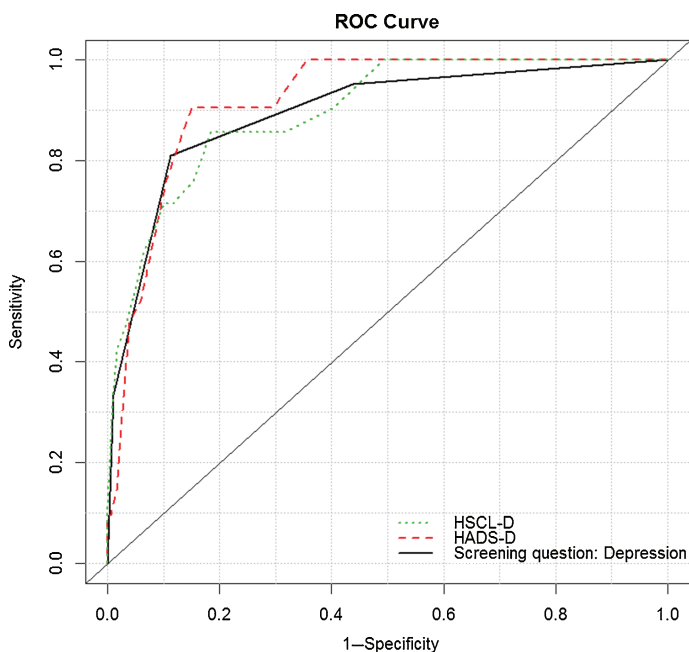
Depression Screening	Sensitivity (95% CI)	Specificity (95% CI)	Overall Accuracy	PPV (95% CI)	NPV (95% CI)	AUC 95% CI
Screening question*: 0 vs. 1 and 2 and 3	95% (77.3–99.2)	56% (52.0–60.6)	58%	8% (5.5–12.5)	100% (98.0–99.9)	0.76 (0.68–0.83)
Screening question†: 0 and 1 vs. 2 and 3	81% (60.0–92.3)	89% (85.9–91.3)	89%	23% (15.1–34.2)	99% (97.8–99.7)	0.85 (0.75–0.95)
Screening question‡: 0 and 1 and 2 vs. 3	33% (17.2–54.6)	99% (97.7–99.6)	96%	58% (32.0–80.7)	97% (95.5–98.4)	0.66 (0.52–0.81)
HADS-Depression (score ≥8)	91% (71.1–97.4)	84% (81.2–87.4)	85%	19% (12.5–27.8)	100% (98.4–99.9)	0.88 (0.80–0.95)
HSCL-Depression (score ≥1.75)	86% (65.4–95.0)	74% (69.9–77.4)	74%	12% (7.5–17.7)	99% (98.4–99.9)	0.80 (0.71–0.89)

\*Not affected by depression at all versus a little, some, seriously affected.

†Not or a little affected by depression versus some or seriously affected.

‡Not, a little or some affected versus seriously affected.

HADS indicates Hospital Anxiety and Depression Scale; HSCL, Hopkins Symptom Checklist; MINI, Mini-International Neuropsychiatric Interview; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; AUC, area under the receiver operating characteristic curve.



**Figure 1.** Receiver operating characteristic curves for HADS, HSCL, and the single-item screening question for depression. HADS indicates Hospital Anxiety and Depression Scale; HSCL, Hopkins Symptom Checklist; D, depression.

identifying anxiety disorders (0.77; 95% CI = 0.70–0.84). There was a statistically significant increase in the AUC for HADS and for HSCL when they were added to the screening question in the logistic regression model, but the changes in AUC were relatively small (Table 4). The discriminating abilities of HADS, HSCL, and the screening question are further visualized in Figure 2.

## DISCUSSION

The aim of this study was to determine whether 2 single-item screening questions would be enough to screen for depression and anxiety in patients with CLBP. The results showed that the single-item screening questions were sensitive for depression but less sensitive for anxiety. Compared with 2 widely used and validated questionnaires, the 2 screening questions performed similar in identifying patients with depression and anxiety. The AUC values indicated that the screening questions are good/excellent in detecting depression and fair/good in detecting anxiety disorders in this group of patients.

In seeking to identify patients with certain diagnoses in routine clinical practice, a 2-step process is often undertaken where a short screening test with high sensitivity is used to identify potential individuals with the condition. For those who screen positive, a second, often more detailed test with greater specificity is used to confirm. It should, however, be noted that setting the threshold of a test at high sensitivity leads to a large number of false-positive tests, which can be burdensome in clinical practice.<sup>16</sup> For that reason, we examined different cutoff points of the various response alternatives of the screening questions and found the first option listed in Tables 2 and 3 (not affected by depression at all *vs.* a little, some, seriously affected) to be the preferred cutoff. However, the optimal accuracy (maximizing the combination of sensitivity and specificity) for the depression question was the second option (not or a little affected by depression *vs.* some or seriously affected).

When each diagnosis of either depression or anxiety was analyzed separately, the results revealed that the screening questions had a perfect sensitivity for 3 anxiety disorders (panic disorder, obsessive-compulsive disorder, and PTSD) and 2 depressive disorders (major depressive episode and dysthymia).

**TABLE 3.** Discriminating Abilities of the Single-Item Anxiety Question, HADS, and HSCL When Compared With MINI (N = 545)

Anxiety Screening	Sensitivity (95% CI)	Specificity (95% CI)	Overall Accuracy	PPV (95% CI)	NPV (95% CI)	AUC (95% CI)
Screening question*: 0 vs. 1 and 2 and 3	68% (55.6–77.8)	85% (81.4–87.9)	83%	39% (30.5–48.2)	95% (92.3–96.6)	0.76 (0.69–0.83)
Screening question†: 0 and 1 vs. 2 and 3	46% (34.6–58.1)	96% (94.1–97.7)	90%	64% (49.5–76.0)	93% (90.0–94.7)	0.71 (0.63–0.79)
Screening question‡: 0 and 1 and 2 vs. 3	11% (5.3–20.6)	100% (99.2–100)	89%	100% (64.6–100)	89% (85.7–91.2)	0.55 (0.47–0.63)
HADS-Anxiety (score ≥8)	58% (46.2–68.9)	83% (79.1–85.9)	80%	33% (25.1–41.5)	93% (90.3–95.2)	0.70 (0.63–0.78)
HSCL-Anxiety (score ≥1.75)	67% (54.9–76.7)	87% (83.3–89.4)	84%	42% (33.4–51.6)	95% (92.1–96.4)	0.77 (0.70–0.84)

\*Not affected by anxiety at all versus a little, some, seriously affected.

†Not or a little affected by anxiety versus some or seriously affected.

‡Not, a little or some affected versus seriously affected.

HADS indicates Hospital Anxiety and Depression Scale; HSCL, Hopkins Symptom Checklist; MINI, Mini-International Neuropsychiatric Interview; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; AUC, area under the receiver operating characteristic curve.

**TABLE 4. AUC for the Single-Item Screening Questions Versus MINI, and Additional Contributions From HADS and HSCL**

	Anxiety			Depression		
	AUC	(95% CI)	P*	AUC	(95% CI)	P*
Screening question	0.79	(0.72–0.86)	Ref	0.90	(0.82–0.97)	Ref
Screening question + HADS†	0.81	(0.75–0.88)	0.031	0.92	(0.85–0.99)	<0.001
Screening question + HSCL†	0.83	(0.77–0.90)	0.013	0.91	(0.85–0.97)	0.79
Screening question + HADS‡	0.85	(0.79–0.90)	<0.001	0.94	(0.91–0.98)	<0.001
Screening question + HSCL‡	0.86	(0.81–0.91)	0.010	0.92	(0.87–0.98)	0.093

For depression, there were 519 observations (21 cases), and for anxiety, there were 517 observations (65 cases).

\*P values are based on likelihood ratio tests from logistic regression analyses (generalized linear model).

†HADS and HSCL entered as categorical variables.

‡HADS and HSCL entered as continuous variables.

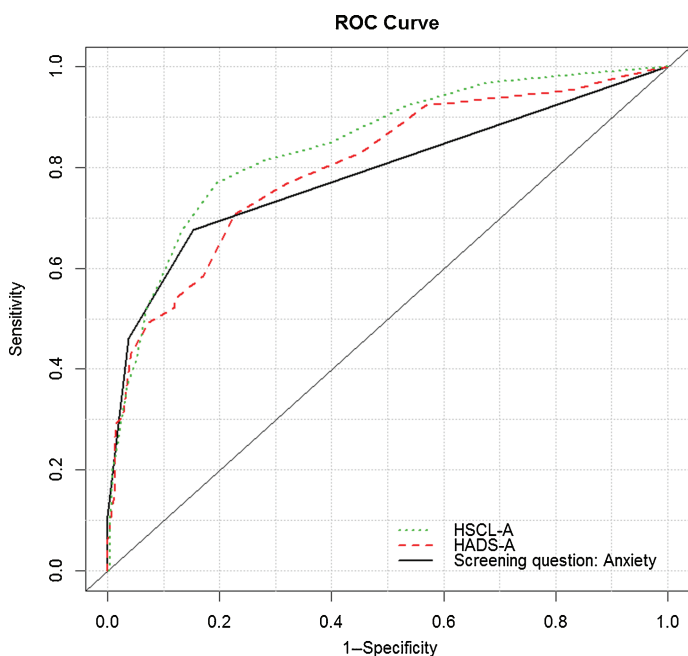
AUC indicates area under the receiver operating characteristic curve; MINI, Mini-International Neuropsychiatric Interview; CI, confidence interval; HADS, Hospital Anxiety and Depression Scale; HSCL, Hopkins Symptom Checklist.

Depression is known to be a common comorbid disorder in patients with CLBP,<sup>5</sup> and the need to screen for depression has previously been highlighted in clinical guidelines.<sup>12</sup> Anxiety, however, has received far less attention as a comorbid condition in patients with CLBP. Yet, in the current study, anxiety disorders were more prevalent than depressive disorders. In a recent study looking at a range of different factors associated with disability due to CLBP, anxiety showed the strongest association with disability.<sup>8</sup> Furthermore, pain-related fear has been found to be more disabling than the pain itself in chronic

back pain,<sup>23</sup> and the disabling consequences of anxiety are as severe as depression when combined with chronic musculoskeletal pain.<sup>37</sup> This suggests that screening for anxiety disorders may be as important as for depression in patients with CLBP. Given our results, additional work is needed to identify more sensitive instruments to screen for anxiety.

In research, the depression screening question could replace longer screening instruments to identify potentially depressed patients, particularly in cases where length is a concern. Although the single anxiety question was less sensitive than the depression question, and therefore may be insufficient as a screening tool in clinical practice, it still showed similar or better discriminating abilities than the 2 longer questionnaires. We thus suggest that the anxiety question could be suitable for epidemiological studies, to replace longer questionnaires, whereas the depression question could be suitable as a screening tool in clinical practice to identify patients with a possible depressive disorder. Clinical interviews take both time and resources to conduct, so a single-item screening question that could identify those in need of such clinical assessments would be highly beneficial and cost-effective. The screening question is further simple and easy to administer and therefore has good prospects for implementation in clinical practice. Previous studies from primary care have concluded that 1-question tests are unacceptable as screening tools for depression.<sup>15</sup> Other short screening tools, such as the Patient Health Questionnaire–2, have shown good promise in identifying depressive<sup>17</sup> and anxiety disorders<sup>18</sup> in various medical populations. An avenue for future research could thus be to compare the single-item screening questions with the Patient Health Questionnaire–2 to determine which may be most appropriate for specialists who care for patients with LBP conditions.

A possible limitation of the study is the surprisingly low prevalence of psychiatric symptoms in this group of patients with CLBP on sick leave. This is particularly so for depression



**Figure 2.** Receiver operating characteristic curves for HADS, HSCL, and the single-item screening question for anxiety. HADS indicates Hospital Anxiety and Depression Scale; HSCL, Hopkins Symptom Checklist; A, anxiety.

when compared with previous literature.<sup>38</sup> It raises questions whether those who administered MINI were unduly conservative in interpreting responses, which could affect the generalizability of the findings to some degree. However, in the study, thorough procedures were undertaken to ensure the quality and reliability of the interviews. This involved training and supervision of interviewers, audio taping of interviews, and evaluation of a random selection of interviews to assess adherence and quality of administration. No major protocol deviations or poor quality results were revealed in the evaluations. Possible explanations for the somewhat lower prevalence of psychiatric symptoms compared with findings from other studies are discussed elsewhere<sup>6</sup> but could be related to different population and settings of the various studies. This could be an avenue for future studies to pursue. Another potential limitation is the different time frames of the various screening instruments, which could influence their screening sensitivity: the single-item screening questions measured symptoms of the past 30 days, HSCL in the past 14 days, and HADS during the past week, whereas MINI examined current symptoms with various criteria for time frames depending on the diagnosis (14 d for current depression, 30 d for anxiety disorders). In diagnostic screening tests, we want to be as inclusive as possible to ensure a high sensitivity, as previously mentioned. A wider time frame, such as the 30-day frame applied in the single-item screening questions, might thus be preferred. Furthermore, this could possibly be part of the explanation for the high sensitivity we found for at least one of the screening questions compared with the longer questionnaires with shorter time frames.

## CONCLUSION

The results showed that the single-item screening questions were sensitive for depression but less sensitive for anxiety. They both performed similar to 2 widely used and validated questionnaires; the depression question had very high sensitivity but less specificity, whereas the anxiety question had similar or better sensitivity and specificity than the 2 longer questionnaires. In both cases, however, follow-up assessments would be needed to confirm a diagnosis of either depression or anxiety. Furthermore, the results need to be validated in other populations before any widespread use in clinical and epidemiological settings can be recommended.

### ➤ Key Points

- ❑ Simple screening tools to identify patients with CLBP and psychiatric diagnoses are needed.
- ❑ Two single-item screening questions are able to identify most patients with CLBP with depression and some with anxiety.
- ❑ The screening questions performed equal to or better than 2 longer screening questionnaires.

## References

1. Von Korff M, Crane P, Lane M, et al. Chronic spinal pain and physical-mental comorbidity in the United States: results from the national comorbidity survey replication. *Pain* 2005;113:331–9.
2. Linton SJ. A review of psychological risk factors in back and neck pain. *Spine (Phila Pa 1976)* 2000;25:1148–56.
3. Dersh J, Mayer T, Gatchel RJ, et al. Psychiatric comorbidity in chronic disabling occupational spinal disorders has minimal impact on functional restoration socioeconomic outcomes. *Spine (Phila Pa 1976)* 2007;32:1917–25.
4. Engel CC, von Korff M, Katon WJ. Back pain in primary care: predictors of high health-care costs. *Pain* 1996;65:197–204.
5. Fishbain DA, Cutler R, Rosomoff HL, et al. Chronic pain-associated depression: antecedent or consequence of chronic pain? A review. *Clin J Pain* 1997;13:116–37.
6. Reme SE, Tangen T, Moe T, et al. Prevalence of psychiatric disorders in sick listed chronic low back pain patients. *Eur J Pain* 2011;15:1075–80.
7. Newcomer KL, Shelerud RA, Vickers Douglas KS, et al. Anxiety levels, fear-avoidance beliefs, and disability levels at baseline and at 1 year among subjects with acute and chronic low back pain. *PM R* 2010;2:514–20.
8. Moix J, Kovacs FM, Martin A, et al. Catastrophizing, state anxiety, anger, and depressive symptoms do not correlate with disability when variations of trait anxiety are taken into account. a study of chronic low back pain patients treated in Spanish pain units. *Pain Med* 2011;12:1008–17.
9. Simon GE, von Korff M. Recognition, management, and outcomes of depression in primary care. *Arch Fam Med* 1995;4:99–105.
10. Overland S, Glozier N, Krokstad S, et al. Undertreatment before the award of a disability pension for mental illness: the HUNT study. *Psychiatr Serv* 2007;58:1479–82.
11. Mykletun A, Overland S, Dahl AA, et al. A population-based cohort study of the effect of common mental disorders on disability pension awards. *Am J Psychiatry* 2006;163:1412–8.
12. Airaksinen O, Brox JI, Cedraschi C, et al. Chapter 4. European guidelines for the management of chronic nonspecific low back pain. *Eur Spine J* 2006;15(suppl 2):S192–300.
13. Pignone MP, Gaynes BN, Rushton JL, et al. Screening for depression in adults: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med* 2002;136:765–76.
14. Kroenke K, Spitzer RL, Williams JB, et al. The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: a systematic review. *Gen Hosp Psychiatry* 2010;32:345–59.
15. Mitchell AJ, Coyne JC. Do ultra-short screening instruments accurately detect depression in primary care? A pooled analysis and meta-analysis of 22 studies. *Br J Gen Pract* 2007;57:144–51.
16. Mulrow CD, Williams JW, Gerety MB, et al. Case-finding instruments for depression in primary care settings. *Ann Intern Med* 1995;122:913–21.
17. Lowe B, Kroenke K, Grafe K. Detecting and monitoring depression with a two-item questionnaire (PHQ-2). *J Psychosom Res* 2005;58:163–71.
18. Lowe B, Grafe K, Zipfel S, et al. Detecting panic disorder in medical and psychosomatic outpatients: comparative validation of the Hospital Anxiety and Depression Scale, the Patient Health Questionnaire, a screening question, and physicians' diagnosis. *J Psychosom Res* 2003;55:515–9.
19. Haggman S, Maher CG, Refshauge KM. Screening for symptoms of depression by physical therapists managing low back pain. *Phys Ther* 2004;84:1157–66.
20. Spitzer RL, Williams JB, Kroenke K, et al. Utility of a new procedure for diagnosing mental disorders in primary care. The PRIME-MD 1000 study. *JAMA* 1994;272:1749–56.
21. Reme SE, Eriksen HR. Is one question enough to screen for depression? *Scand J Public Health* 2010;38:618–24.
22. Sheehan DV, Lecrubier Y, Sheehan KH, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* 1998;59(suppl 20):22–33.



23. Crombez G, Vlaeyen JW, Heuts PH, et al. Pain-related fear is more disabling than pain itself: evidence on the role of pain-related fear in chronic back pain disability. *Pain* 1999;80:329–39.
24. Reme SE, Tveito TH, Chalder T, et al. Protocol for the Cognitive Interventions and Nutritional Supplements (CINS) trial: a randomized controlled multicenter trial of a brief intervention (BI) versus a BI plus cognitive behavioral treatment (CBT) versus nutritional supplements for patients with long-lasting muscle and back pain. *BMC Musculoskelet Disord* 2011;12:152.
25. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV*. 4th ed. Washington, DC: American Psychiatric Association; 1994.
26. World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research*. Geneva: World Health Organization; 1993.
27. Leiknes KA, Malt U, Malt EA, et al. *MINI Internasjonalt Neuropsykiatrisk Intervju, 5.0.0*. Oslo, Norway: Institutt for medisinske basalfag, Avdeling for atferdsfag, Universitetet i Oslo; 2005.
28. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983;67:361–70.
29. Bjelland I, Dahl AA, Haug TT, et al. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *J Psychosom Res* 2002;52:69–77.
30. Derogatis LR, Lipman RS, Rickels K, et al. The Hopkins Symptom Checklist (HSCL): a self-report symptom inventory. *Behav Sci* 1974;19:1–15.
31. Winokur A, Winokur DE, Rickels K, et al. Symptoms of emotional distress in a family planning service: stability over a four-week period. *Br J Psychiatry* 1984;144:395–9.
32. Nettelbladt P, Hansson L, Stefnansson CG, et al. Test characteristics of the Hopkins Symptom Check List-25 (HSCL-25) in Sweden, using the Present State Examination (PSE-9) as a caseness criterion. *Soc Psychiatry Psychiatr Epidemiol* 1993;28:130–3.
33. Sandanger I, Nygard JF, Ingebrigtsen G, et al. Prevalence, incidence and age at onset of psychiatric disorders in Norway. *Soc Psychiatry Psychiatr Epidemiol* 1999;34:570–9.
34. Eriksen HR, Ihlebaek C, Ursin H. A scoring system for Subjective Health Complaints (SHC). *Scand J Public Health* 1999;27:63–72.
35. Ursin H. Sensitization, somatization, and Subjective Health Complaints. *Int J Behav Med* 1997;4:105–16.
36. Altman DG. *Practical Statistics for Medical Research*. London: Chapman & Hall; 1991.
37. Bair MJ, Wu J, Damush TM, et al. Association of depression and anxiety alone and in combination with chronic musculoskeletal pain in primary care patients. *Psychosom Med* 2008;70:890–7.
38. Bair MJ, Robinson RL, Katon W, et al. Depression and pain comorbidity: a literature review. *Arch Intern Med* 2003;163:2433–45.