Published in final edited form as:
J Pain Symptom Manage. 2013 January ; 45(1): 71–82. doi:10.1016/j.jpainsymman.2011.12.284.

Longitudinal Comparison of Three Depression Measures in Adult Cancer Patients

Shelley A. Johns, PsyD, ABPP, **Kurt Kroenke, MD**, **Erin E. Krebs, MD, MPH**, **Dale E. Theobald, PhD, MD**, **Jingwei Wu, MS**, and **Wanzhu Tu, PhD**
School of Nursing (S.A.J) and School of Medicine (K.K., E.E.K., J.W., W.T.), Indiana University, Indianapolis; Center for Implementing Evidence-Based Practice (K.K., E.E.K.), Richard Roudebush VA Medical Center, Indianapolis; Regenstrief Institute, Inc., (K.K., E.E.K., W.T.), Indianapolis; and Community Health Network (D.E.T.), Indianapolis, Indiana, USA

Abstract

Context—Although a number of depression measures have been used with cancer patients, longitudinal comparisons of several measures in the same patient population have been infrequently reported.

Objective—To compare the Hopkins Symptom Checklist-20 (HSCL-20), SF-36 Mental Health Inventory (MHI-5), and Patient Health Questionnaire-9 (PHQ-9) in adults with cancer.

Methods—Of 309 cancer patients enrolled in a telecare management trial for depression, 247 completed the three depression measures at both baseline and at three months, as well as a retrospective assessment of global rating of change in depression at three months. Internal consistency and construct validity of each measure was evaluated. Responsiveness was compared by calculating standardized response means (SRM) and receiver operating characteristic area under the curve (AUC), using global rating of change as the external comparator measure. Differences between intervention and control groups in depression change scores were compared by calculating standardized effect sizes (SES).

Results—Internal reliability coefficients for the three measures were ≥ 0.77 at baseline and ≥ 0.84 at three months. Construct validity was supported with strong correlations of the depression measures among themselves, moderately strong correlations with other measures of mental health, and moderate correlations with vitality and disability. In terms of responsiveness, SRMs for all measures significantly differentiated between three groups (improved, unchanged, worse) as classified by patient-reported global rating of change in depression at three months. The three measures were able to detect a modest treatment effect in the intervention group compared with the control group (SES ranging from 0.21 to 0.43) in the full sample, whereas detecting a greater treatment effect in depressed participants with comorbid pain (SES ranging from 0.30 to 0.58). Finally, the three measures performed similarly in detecting patients with improvement.

Address correspondence to: Shelley A. Johns, PsyD, ABPP, Regenstrief Institute, 5th Floor, 1050 Wishard Blvd., Indianapolis, IN 46202, USA, sheljohn@iupui.edu.

Disclosures

The authors declare no conflicts of interest.

Conclusion—The HSCL-20, MHI-5, and PHQ-9 were established as reliable, valid, and responsive depression measures in adults with cancer. Given current recommendations for measurement-based care, our study shows that clinicians treating depressed cancer patients have several measures from which to choose.

Keywords

Cancer; depression; assessment; psychometrics; HSCL-20; MHI-5; PHQ-9; responsiveness

Introduction

Depression is an underdiagnosed, undertreated, and common complication of cancer at any stage.[1–4] The prevalence of major depressive disorder and depressive symptoms comorbid with cancer appears to be 10% to 25%[5] and up to 50% among adults with advanced cancer receiving palliative care.[6] Untreated depression is associated with functional impairment, diminished quality of life, poor treatment adherence, an increased burden of physical symptoms, and higher health care utilization.[7, 8] Depression also has been found to be a significant predictor of mortality in cancer patients.[9, 10] Given this potential for grave consequences, effective treatment of the depressed patient across the cancer trajectory is essential.[11, 12] An important first step in the provision of this necessary care is the accurate and timely detection of depressive symptoms.

There are a number of self-report depression measures available for use in cancer care and oncology research. Several papers have been published comparing the psychometric performance of two depression measures in cancer, with the Hospital Anxiety and Depression Scale (HADS) being the most commonly studied measure;[13] however, few studies compare more than two measures in the same sample. In recent reviews of the psychometric performance of depression measures in cancer, the large majority of included studies have been cross-sectional.[5,13–17] The paucity of longitudinal psychometric data makes the choice among measures difficult for the researcher planning an intervention trial for depression in the cancer setting[13] or the clinician providing measurement-based care in the treatment of depression.[18]

The objective of the current study was to directly compare the longitudinal psychometric performance (i.e., internal consistency reliability, construct validity, and comparative responsiveness) of three widely-used measures—the Hopkins Symptom Checklist 20-item depression severity scale (HSCL-20),[19] the SF-36 Mental Health Inventory (MHI-5),[20] and the Patient Health Questionnaire 9-item depression module (PHQ-9).[21] Although the HSCL-20 has been used to assess depression severity and outcomes in large clinical trials in oncology patients,[22, 23] its psychometric characteristics were not formally reported in those trials. Likewise, the MHI-5 has been used in cancer research,[24–28] but its psychometric characteristics in oncology populations have been inadequately characterized. The PHQ-9 has performed well as a depression screening measure in a large sample of cancer patients in the U.K.[29] and has served as the primary outcome measure in a depression trial targeting cancer patients;[30] however, its responsiveness in this population has not been reported. Moreover, in the most comprehensive systematic review of distress screening measures in cancer to date, Vodermaier et al.[16] concluded that the PHQ-9 "deserves further empirical evaluation of its value for distress screening of cancer patients" based on its psychometric strength in large samples of primary care patients.

To compare the HSCL-20, MHI-5, and PHQ-9 in the current study, data were drawn from a longitudinal trial of adult cancer patients being systematically treated for depression. Our primary aim was to assess the responsiveness of the three measures when compared to patient-reported global rating of change and, in particular, to determine whether the measures could differentiate between patient subgroups who reported *improvement, no change*, or *worsening* in their depressive symptoms. Additionally, we assessed the construct validity of the three measures by testing their cross-sectional and longitudinal associations with multiple domains of health-related quality of life. Finally, we determined the internal consistency of each measure.

Methods

Study Design and Sample

This paper uses data from the Indiana Cancer Pain and Depression (INCPAD) study, a randomized controlled trial that compared a telecare management intervention to usual care for the treatment of depression and pain in a heterogeneous sample of cancer patients. Patients were enrolled from 16 urban and rural oncology practices. Patients presenting for oncology clinic visits were eligible if they had clinically significant depression, defined as a PHQ-9 score of 10 or greater and endorsement of depressed mood and/or anhedonia;[31, 32] or cancer-related pain that was scored as 6 or greater on the “worst pain in the past week” item of the Brief Pain Inventory.[33, 34] Excluded were individuals who did not speak English; had moderately severe cognitive impairment as defined by a validated six-item cognitive screener;[35] had schizophrenia or other psychosis; had a pending pain-related disability claim; were pregnant; or were enrolled in hospice care. Further details of the INCPAD study design[36] and primary outcomes of the trial[37] have been published. The INCPAD protocol was approved by the Indiana University and Community Hospitals institutional review boards, and all patients included in the study provided written informed consent.

The INCPAD trial enrolled 131 (32%) participants with depression only, 96 (24%) with pain only, and 178 (44%) with both depression and pain. The 309 subjects enrolled for depression (with or without pain) constitute the sample for the current study. Participants were followed for 12 months, with outcome assessments at baseline and at one, three, six, and 12 months conducted by research assistants blinded to treatment group. Whereas the SCL-20 was assessed at all five time points, the PHQ-9 and MHI-5 were assessed at baseline, three and 12 months only. Baseline and three-month data are analyzed in the current study because the most intense phase of the INCPAD depression intervention as well as the greatest change in depression scores occurred during the initial three months of the trial. Moreover, this period was least affected by attrition as a result of death, disease progression, drop-out, or loss to follow-up that had occurred by 12 months.

Depression Measures

HSCL-20—The HSCL-20[19] has been used extensively to assess depression severity and outcomes in large clinical trials in noncancer medical populations[38–49] and in cancer.[16, 22, 23] It was the primary depression outcome measure in the INCPAD trial. The measure is well validated and has demonstrated sensitivity in detecting differences in depression severity between treatment groups.[44, 46] The HSCL-20 includes the full 13-item depression scale plus seven additional items from the Hopkins Symptom Checklist-90 Revised (SCL-90-R), which were added to better represent all diagnostic symptoms of major depression and improve the instrument’s responsiveness.[50] There are two versions of this questionnaire; INCPAD used version B.[51] The HSCL-20 inquires about the presence/extent of each of 20 symptom over the past four weeks, with response options being *not at all, a little bit, moderately, quite a bit,* and *extremely*, ranging from 0 to 4. The average score across all items is calculated so that the HSCL-20 scale score ranges from 0 to 4, with higher scores representing more severe depressive symptoms. Specifically, HSCL-20 mean scores less than 0.5 indicate no depression, scores from 0.5 to 1.7 or less indicate mild to moderate depression, and scores greater than 1.7 indicate severe depression.[52]

MHI-5—The MHI-5[53] constitutes the mental health subscale of the Medical Outcomes Study (MOS) Short Form Health Survey, and is included in both the 20-item version (SF-20) and the 36-item version (SF-36).[20] The MHI-5 consists of three depression and two anxiety items. The measure asks respondents how much of the time in the last month they felt nervous, downhearted and blue, calm and peaceful, happy, or so down in the dumps that nothing could cheer them up. Answers are scored on a 1 to 5 scale ranging from *all of the time* to *none of the time*. The two items that ask about positive feelings are reverse scored. An MHI-5 score is calculated by summing the scores for each item and then transforming the raw scores to a scale ranging from 0–100, with higher scores representing less severe psychological distress.[54] In medically ill populations, MHI-5 scores of 52 or less suggest the presence of significant depressive symptomatology.[26, 55] We also calculated a total score for the MHI-d (the three depression items of the MHI-5).[56–58] MHI-d scores can range from 3 to 15, with higher scores representing more severe depression.

PHQ-9—The PHQ-9.[21, 31] is well validated as a multipurpose diagnostic, severity, and outcome measure.[59] It includes nine questions that correspond to the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (DSM-IV) symptoms of depression. The frequency of each symptom over the previous two weeks is assessed, with response options including *not at all, several days, more than half the days,* and *nearly every day*. These responses are scored 0, 1, 2, or 3, respectively, resulting in a total score that can range from 0 to 27, with higher scores representing more severe depressive symptoms. PHQ-9 scores of 10 or greater indicate clinically significant depression. The PHQ-9 has consistently demonstrated strong psychometric properties in general medical samples[31] and in a range of settings, countries, and populations,[60, 61] and its responsiveness has been demonstrated in primary care patients.[59, 62] The first two items of the PHQ-9 assess the two cardinal symptoms of depressive disorders (i.e., anhedonia and depressed mood). An abbreviated measure comprising those two items, the PHQ-2, has been shown to be a reliable and valid ultra-brief tool to assess depression diagnosis, severity, and outcome.[63, 64] The PHQ-2 has been shown to be sensitive to change in a sample of medical outpatients without cancer;[64] however, its responsiveness has not been assessed in oncology patients.

Measures of Construct Validity

The SF-12 Mental Component Summary (MCS-12)[65] includes 12 items aggregating four subscales from the SF-36 assessing psychological distress and social and role disability resulting from emotional problems. Scores range from 0 to 100, with higher scores representing better mental health functioning. The seven-item Generalized Anxiety Disorder Scale (GAD-7[66]) is a validated screening severity measure for generalized anxiety disorder as well as a good screening measure for panic, social anxiety, and post-traumatic stress disorders. Scores range from 0 to 21, with cut-points of 5, 10, and 15 representing mild, moderate, and severe levels of anxiety, respectively. The four-item SF-36 Vitality Scale[54] is a validated measure of fatigue. Scores range from 0 to 100, with higher scores indicating greater vitality. The three-item Sheehan Disability Scale (SDS[67]) assesses impairment in the domains of work/school activities, social functioning, and family life/home responsibilities. Scores range from 0 to 10, with higher scores representing greater disability. Finally, we included a single *general health perceptions* item that asks respondents to rate their health on a five-point scale from excellent to poor and which has been shown to predict long-term health outcomes.[68]

Assessing Responsiveness

We used a patient-reported retrospective *global rating of change* as an independent measure for evaluating responsiveness. At three months, participants were asked, "Overall, since starting the study, would you say that your depression is *worse, about the same,* or *better*?" Depending on their response, participants were then grouped into three mutually exclusive categories of *worse, unchanged*, and *improved.* Those rating their depression as *better (improved)* were asked if their depression was *a little better, somewhat better, moderately better, a lot better,* or *completely better.* Although retrospective global ratings do not represent a criterion standard for the assessment of change in depression, they have been advocated as a patient-centered approach to defining clinically meaningful change[69] and have been applied to a range of oncology patient populations.[70–73]

Analysis

We examined each measure's internal consistency reliability, construct validity, and responsiveness. Internal consistency was measured with Cronbach's alpha (α), representing the extent to which items comprising a scale measure the same concept.[74] Construct validity was evaluated by examining Spearman correlation coefficients and was demonstrated when relationships among scales or items that we hypothesized *a priori* to be similar were stronger than relationships among scales or items thought to be less similar. Specifically, we expected that the three depression measures would correlate most strongly with each other and with other mental health measures (MCS-12 and GAD-7), moderately strongly with measures of vitality (SF Vitality Scale) and disability (SDS), and least strongly with general health perceptions.

Because there is no consensus on the single best method to evaluate responsiveness,[75] we used three methods to enhance the robustness of our results. The standardized response mean (SRM) has been advocated as one of the better indices of responsiveness because it incorporates the response variance in the denominator,[76] permits comparisons across different instruments and patient subgroups, and allows confidence intervals (CI) to be calculated to facilitate testing for significant differences in responsiveness between outcome measures. For each measure, we calculated the SRM as the change score (baseline mean minus three-month mean) divided by the standard deviation (SD) of the change score. SRM values can be interpreted with Cohen's guidelines for two-group effect sizes (i.e., 0.2 = small, 0.5 = moderate, and ≥ 0.8 = large).[77] We also determined each measure's responsiveness to the intervention by evaluating the standardized effect size (SES), calculated as the difference between change scores in the intervention and control groups divided by the pooled SD of change scores.[78] As with SRM, SES values can be interpreted with Cohen's guidelines for two-group effect sizes. Finally, we conducted receiver operating curve (ROC) analysis,[79] calculating the area under the curve (AUC) for each depression measure using patient-reported global improvement as the comparator measure. We calculated the AUC for detecting *any improvement*, as well as *at least moderate improvement*.

Analyses were performed using SAS® software version 9.1 (SAS Institute, Inc., Cary, NC).

Results

Patient Characteristics

Of the 309 depressed patients enrolled in the INCPAD trial, the mean (SD) age was 58.3 (11.0) years, with 67.6% being female. Our sample was largely white (80.9%), with 16.5% black, and 2.6% from other racial groups. The type of cancer was breast in 30.4% of the participants, lung in 20.4%, gastrointestinal in 17.5%, lymphoma or hematological malignancy in 12.0%, genitourinary in 9.4%, and other cancers in 10.4%. The phase of cancer was newly-diagnosed in 119 (38.5%), disease-free or maintenance therapy in 132 (42.7%), and recurrent or progressive in 58 (18.8%). The number of subjects with complete three month data for each of the depression measures and the global rating of change item ranged from 243 (78.6%) to 247 (79.9%) of those enrolled. Twenty percent of subjects (n=62) either dropped out of the study (n=23; 7.4%), were lost to follow-up (n=27; 8.7%), or died (n=12; 3.9%) between baseline and three-month follow-up. These patients did not differ in age, sex, or baseline depression severity from the three-month completers.

Mean (SD) depression scores at baseline were 1.6 (0.6) on the HSCL-20; 50.2 (19.7) on the MHI-5; 8.8 (2.7) on the MHI-d; 15.9 (4.2) on the PHQ-9; and 4.1 (1.6) on the PHQ-2. These means all represent at least moderate levels of depression severity on each measure. At three months, 43.4% of the sample reported *improvement* in their depression compared with baseline on the global rating of change item, whereas 36.5% reported that their depression was *unchanged* and 20.1% reported their depression was *worse*.

Internal Consistency Reliability

At baseline, internal consistency was good for the HSCL-20 (α=0.85) and PHQ-9 (α=0.81); acceptable for the MHI-5 (α=0.77) and MHI-d (α=0.77); and lowest for the PHQ-2 (α=0.62), which may be a function of its brevity (two items). At three months, internal consistency on all measures ranged from 0.80 for the PHQ-2 to 0.89 for the HSCL-20.

Construct Validity

As shown in Table 1, all depression measures were strongly correlated with each other at magnitudes ranging from 0.70 to 0.95 at three months. Correlations were assessed at three months because, at this assessment, all depression measures were included in the same interview, whereas at baseline, the PHQ-9 was administered in the eligibility interview and the HSCL-20 and MHI-5 were administered in the baseline interview conducted up to one week later.

As summarized in Table 2, there was a gradient in the strength of association of the three full depression measures with other measures of construct validity, being greatest for other mental health measures (MCS-12 and GAD-7), intermediate for fatigue (SF-Vitality Scale) and disability (SDS), and least for general health perceptions. As expected, the MHI-5 had the strongest correlation with the MCS-12, most likely because the MHI-5 comprises five of the 12 items of the MCS-12.

Responsiveness

As shown in Fig. 1, the SRMs calculated on observed change scores (baseline minus three months) differentiated between the three patient groups for each measure. Specifically, the patient group reporting symptoms as *improved* on the global rating of change reported the greatest improvement in depression scores and the largest SRM for each of the three scales, whereas the *worse* group reported the least improvement. The *unchanged* group was intermediate between the *improved* and *worse* groups. Moreover, the 95% CIs of the SRMs for each depression measure show no overlap among the *improved, unchanged,* and *worse* groups.

Table 3 depicts the responsiveness of each measure in the full sample and when using a subgroup analysis based on the presence or absence of comorbid pain. In the full sample, all three full depression measures detected a modest treatment effect (SES ranging from 0.21 to 0.43) when comparing depression change over three months in the intervention vs. control groups of the INCPAD trial. The effect sizes were smaller for the abbreviated measures

(MHI-d and PHQ-2), suggesting they may be less sensitive than the full measures for detecting treatment effects. In participants without pain, the effect sizes were smaller than in either the full sample or the subgroup with comorbid pain. Notably, patients with comorbid pain had more severe depression at baseline and showed more movement with the intervention.

Table 4 shows that all the depression measures were fair in detecting patient-reported improvement in depression, with AUC for the full measures ranging from 0.69 to 0.71 for detecting *any improvement* (which includes those who rated their depression as *a little better, somewhat better, moderately better, a lot better,* or *completely better*) and from 0.74 to 0.78 for detecting *at least moderate improvement* (which includes those who rated their depression as *moderately better, a lot better,* or *completely better*). Three points are noteworthy. First, these AUCs are consistent between measures, but lower than what might be demonstrable when comparing with a criterion standard, such as a structured clinical interview; however, these values are reasonable considering the comparison was to patient-reported change. Second, the AUCs are modestly higher for detecting *at least moderate improvement* vs. *any improvement*, suggesting that accuracy of the measures may be somewhat influenced by the degree of clinical change. Third, the AUC for the full measures is somewhat higher than for their abbreviated versions (except the MHI-d for detecting *any improvement*), suggesting a slight superiority for full measures in detecting clinical improvement.

Discussion

Our longitudinal study in cancer patients comparing the psychometric performance of three commonly-used measures for monitoring depression outcomes has several important findings. All three measures showed good internal consistency as well as construct validity. Responsiveness of the measures was generally supported in three ways. First, all three measures discriminated between patients classified by a global self-reported response measure as *improved, unchanged* or *worse*. Second, all measures were able to detect a treatment effect when comparing intervention to control patients in the trial. Third, the measures were reasonably accurate in distinguishing improved from non-improved patients by ROC analysis. Notably, abbreviated versions of two of the measures (i.e., MHI-d and PHQ-2) generally appeared less responsive than the corresponding full measures (i.e., MHI-5 and PHQ-9).

As illustrated in Fig. 1, the depression measures showed moderate to large improvement by SRM analysis in the patients who rated their depression as *improved* at three months. However, all scales showed modest improvement by SRMs in patients who rated their depression as *unchanged* and, except for the MHI-5, even slight improvement by SRMs in patients who rated their depression as *worse*. For this reason, evaluating whether each particular measure differentiates between the three subgroups (i.e., *improved, unchanged,* and *worsened*)—which each measure did in our study—may be more helpful than overinterpreting the absolute values of the SRMs. One reason for the discrepancy between SRMs and global rating of change could be that the depression measures are more sensitive to improvement than to deterioration in depression. A second reason may be because distribution-based methods (i.e., standardizing change scores into metrics such as SRM or SES) and anchor-based methods (i.e., comparing change scores to a retrospective assessment of global change) give somewhat different “windows” into assessing responsiveness of a measure, with neither one necessarily more accurate than the other.[69, 75, 80] Anchor-based methods that rely on patient-rated global estimates of change may be subject to varying degrees of recall bias depending upon length of the recall period, current severity of symptoms, and other potential patient- and disease-specific confounding factors.[57] At the same time, such retrospective perceptions of change have face validity, may be more sensitive than serial symptom assessments, and can correlate strongly with patient satisfaction with treatment.[69] For these reasons, serial assessment with a validated measure coupled with patient-rated global assessment of change may be preferable to an “either-or” approach to assessment. This is the expert consensus when caring for or investigating pain,[57] and also is a reasonable approach in caring for depression. Discordant results may warrant further probing. For example, if a depression measure suggests improvement whereas the patient globally assesses his or her status as *unchanged* or *worse*, the clinician should explore the reasons for the discrepancy, such as a few especially bothersome residual depressive symptoms, cancer-related or social factors, or other comorbid psychological or physical conditions.

There has been ongoing debate about whether or not to use measures that include somatic symptoms when assessing depression in individuals with cancer.[81, 82] Certain symptoms used as core criteria for diagnosing DSM-IV depressive disorders, such as appetite/weight changes, sleep disturbance, and fatigue, are commonly observed in the context of cancer or its treatment and have been thought to confound the assessment of depression. Therefore, some experts have suggested deleting somatic symptoms from measures assessing depression in cancer.[81] Although some researchers have noted an increased prevalence of false-positive depression when somatic items are included in depression screening in the oncology setting,[83] others have not found this to be true.[29, 84, 85] Of note, a number of chronic medical disorders besides cancer are associated with an increased prevalence of DSM-IV somatic symptoms. Some evidence suggests that retaining these somatic symptoms in depression measures and as diagnostic criteria does not lead to overdiagnosis of depression but indeed may lead to more reliable diagnoses.[84, 86–88] This is consistent with the SES and AUC analyses in the current study, where the PHQ-9 and HSCL-20, which include somatic symptoms, performed as well or better than the MHI-5, which does not include somatic symptoms.

Our study has two primary limitations. First, an independent structured psychiatric interview at baseline and follow-up would have allowed us to estimate more precise and absolute responsiveness metrics. Second, our results suggest the depression measures studied may perform better at detecting depression improvement than in assessing stable or deteriorating symptoms. We are uncertain whether this is because global ratings are an imperfect comparator measure or, alternatively, because global ratings are more sensitive to deterioration than depression scales. Of interest, a study of pain likewise showed that serial administration of pain scales agreed more closely with global ratings that indicated improvement than with global ratings indicating worsening.[89]

Although our study provides preliminary evidence supporting the psychometric performance of three depression measures in cancer, future research is needed to further establish the responsiveness of these measures and other widely used depression and anxiety measures in cancer care (e.g., Hospital Anxiety and Depression Scale, Center for Epidemiological Studies Depression Scale, Profile of Mood States, Spielberger State-Trait Anxiety Inventory, GAD-7). Additionally, future research is needed to determine the minimally clinically important difference (MCID) for distress measures used in cancer care. Löwe et al.[59] suggest determining MCID using seven-day test-retest reliability coefficients of each measure, ideally assessed without influence from treatment, and our design did not allow for this. Löwe et al.[59] found a five-point MCID for the PHQ-9 in a primary care sample; however, we are unsure if this MCID is valid in the cancer setting.

In summary, there are a number of instruments available for detecting depression in the oncology setting and monitoring its severity over time. The HSCL-20, MHI-5, and PHQ-9 may all be useful measures for monitoring depression therapy and outcomes in patients with cancer. These three instruments are brief (5–20 items), feasible for serial symptom assessments (one to five minutes to complete), appropriate for those with a sixth to ninth grade reading ability,[90] and easily scored. Moreover, the PHQ-9 is freely available (public domain) as well as DSM-IV- based and valid for establishing provisional diagnoses, assessing severity, and monitoring therapy.[59] Given the recent emphasis on measurement-based care for depression,[18] our findings provide support for the use of three commonly-used depression measures to not only screen for depression but also to monitor treatment outcomes.

Acknowledgments

This study was supported by grants from the National Cancer Institute to Dr. Kroenke (R01 CA-115369) and Dr. Johns (R25 CA-117865) as well as a training grant to Dr. Johns from the Walther Cancer Foundation, Inc. Dr. Krebs' work on this manuscript was supported by a Career Development Award from the Department of Veterans Affairs, Health Services Research and Development. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the United States government. Dr. Kroenke has received honoraria as a consultant to Eli Lilly and Forest Pharmaceuticals.

We would like to thank Drs. Patrick O. Monahan and Linda F. Brown for their insightful reviews of this manuscript and Donna Burgett for her administrative support.

References

1. Fallowfield L, Ratcliffe D, Jenkins V, Saul J. Psychiatric morbidity and its recognition by doctors in patients with cancer. Br J Cancer. 2001; 84:1011–1015. [PubMed: 11308246]
2. Greenberg DB. Barriers to the treatment of depression in cancer patients. J Natl Cancer Inst Monogr. 2004; 32:127–135. [PubMed: 15263054]
3. Passik SD, Dugan W, McDonald MV, et al. Oncologists' recognition of depression in their patients with cancer. J Clin Oncol. 1998; 16:1594–1600. [PubMed: 9552071]
4. Sharpe M, Strong V, Allen K, et al. Major depression in outpatients attending a regional cancer centre: screening and unmet treatment needs. Br J Cancer. 2004; 90:314–320. [PubMed: 14735170]
5. Pirl WF. Evidence report on the occurrence, assessment, and treatment of depression in cancer patients. J Natl Cancer Inst Monogr. 2004; 32:32–39. [PubMed: 15263039]
6. Hotopf M, Chidgey J, Addington-Hall J, Ly KL. Depression in advanced disease: a systematic review Part 1. Prevalence and case finding. Palliat Med. 2002; 16:81–97. [PubMed: 11969152]
7. Fann JR, Thomas-Rich AM, Katon WJ, et al. Major depression after breast cancer: a review of epidemiology and treatment. Gen Hosp Psychiatry. 2008; 30:112–126. [PubMed: 18291293]
8. Rodin G, Lloyd N, Katz M, et al. The treatment of depression in cancer patients: a systematic review. Support Care Cancer. 2007; 15:123–136. [PubMed: 17058100]
9. Pinquart M, Duberstein PR. Depression and cancer mortality: a meta-analysis. Psychol Med. 2010; 40:1797–1810. [PubMed: 20085667]
10. Satin JR, Linden W, Phillips MJ. Depression as a predictor of disease progression and mortality in cancer patients: a meta-analysis. Cancer. 2009; 115:5349–5361. [PubMed: 19753617]
11. Adler, NE.; Page, AEK., editors. Institute of Medicine (IOM). Cancer care for the whole patient: meeting psychosocial health needs. Washington, DC: National Academies Press; 2008.
12. Dy SM, Lorenz KA, Naeim A, et al. Evidence-based recommendations for cancer fatigue, anorexia, depression, and dyspnea. J Clin Oncol. 2008; 26:3886–3895. [PubMed: 18688057]
13. Luckett T, Butow PN, King MT, et al. A review and recommendations for optimal outcome measures of anxiety, depression and general distress in studies evaluating psychosocial interventions for English-speaking adults with heterogeneous cancer diagnoses. Support Care Cancer. 2010; 18:1241–1262. [PubMed: 20596731]
14. Mitchell AJ. Short screening tools for cancer-related distress: a review and diagnostic validity meta-analysis. J Natl Compr Canc Netw. 2010; 8:487–494. [PubMed: 20410338]

15. Trask PC. Assessment of depression in cancer patients. J Natl Cancer Inst Monogr. 2004; 32:80–92. [PubMed: 15263046]
16. Vodermaier A, Linden W, Siu C. Screening for emotional distress in cancer patients: a systematic review of assessment instruments. J Natl Cancer Inst. 2009; 101:1464–1488. [PubMed: 19826136]
17. Ziegler L, Hill K, Neilly L, et al. Identifying psychological distress at key stages of the cancer illness trajectory: a systematic review of validated self-report measures. J Pain Symptom Manage. 2011; 41:619–636. [PubMed: 21310585]
18. Gaynes BN, Rush AJ, Trivedi MH, et al. Primary versus specialty care outcomes for depressed outpatients managed with measurement-based care: results from STAR*D. J Gen Intern Med. 2008; 23:551–560. [PubMed: 18247097]
19. Derogatis, LR.; Rickels, K.; Uhlenhuth, EH.; Covi, L. The Hopkins Symptom Checklist: a measure of primary symptom dimensions. In: Pihot, P., editor. Psychological measurement in psychopharmacology: problems in pharmacopsychiatry. Basel, Switzerland: Kargerman; 1974. p. 79-110.
20. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care. 1992; 30:473–483. [PubMed: 1593914]
21. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. JAMA. 1999; 282:1737–1744. [PubMed: 10568646]
22. Fann JR, Fan MY, Unutzer J. Improving primary care for older adults with cancer and depression. J Gen Intern Med. 2009; 24 (Suppl 2):S417–424. [PubMed: 19838842]
23. Strong V, Waters R, Hibberd C, et al. Management of depression for people with cancer (SMaRT oncology 1): a randomised trial. Lancet. 2008; 372:40–48. [PubMed: 18603157]
24. Clough-Gorr KM, Ganz PA, Silliman RA. Older breast cancer survivors: factors associated with change in emotional well-being. J Clin Oncol. 2007; 25:1334–1340. [PubMed: 17312327]
25. Gore JL, Krupski T, Kwan L, Fink A, Litwin MS. Mental health of low income uninsured men with prostate cancer. J Urol. 2005; 173:1323–1326. [PubMed: 15758790]
26. Kroenke CH, Bennett GG, Fuchs C, et al. Depressive symptoms and prospective incidence of colorectal cancer in women. Am J Epidemiol. 2005; 162:839–848. [PubMed: 16207809]
27. Cull A, Gould A, House A, et al. Validating automated screening for psychological distress by means of computer touchscreens for use in routine oncology practice. Br J Cancer. 2001; 85:1842–1849. [PubMed: 11747324]
28. Silliman RA, Dukes KA, Sullivan LM, Kaplan SH. Breast cancer care in older women: sources of information, social support, and emotional health outcomes. Cancer. 1998; 83:706–711. [PubMed: 9708934]
29. Thekkumpurath P, Walker J, Butcher I, et al. Screening for major depression in cancer outpatients: the diagnostic accuracy of the 9-item patient health questionnaire. Cancer. 2011; 117:218–227. [PubMed: 20737537]
30. Ell K, Xie B, Quon B, et al. Randomized controlled trial of collaborative care management of depression among low-income patients with cancer. J Clin Oncol. 2008; 26:4488–4496. [PubMed: 18802161]
31. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med. 2001; 16:606–613. [PubMed: 11556941]
32. Kroenke K, Spitzer RL. The PHQ-9: a new depression and diagnostic severity measure. Psychiatric Annals. 2001; 32:509–521.
33. Cleeland, CS. Pain assessment in cancer. In: Osaba, D., editor. Effect of cancer on quality of life. Boca Raton, FL: CRC Press; 1991. p. 293-305.
34. Cleeland CS, Gonin R, Hatfield AK, et al. Pain and its treatment in outpatients with metastatic cancer. N Engl J Med. 1994; 330:592–596. [PubMed: 7508092]
35. Callahan CM, Unverzagt FW, Hui SL, Perkins AJ, Hendrie HC. Six-item screener to identify cognitive impairment among potential subjects for clinical research. Med Care. 2002; 40:771–781. [PubMed: 12218768]
36. Kroenke K, Theobald D, Norton K, et al. The Indiana Cancer Pain and Depression (INCPAD) trial: design of a telecare management intervention for cancer-related symptoms and baseline

characteristics of study participants. Gen Hosp Psychiatry. 2009; 31:240–253. [PubMed: 19410103]

37. Kroenke K, Theobald D, Wu J, et al. Effect of telecare management on pain and depression in patients with cancer: a randomized trial. JAMA. 2010; 304:163–171. [PubMed: 20628129]
38. Dobscha SK, Corson K, Hickam DH, et al. Depression decision support in primary care: a cluster randomized trial. Ann Intern Med. 2006; 145:477–487. [PubMed: 17015865]
39. Bair MJ, Robinson RL, Eckert GJ, et al. Impact of pain on depression treatment response in primary care. Psychosom Med. 2004; 66:17–22. [PubMed: 14747633]
40. Corey-Lisle PK, Nash R, Stang P, Swindle R. Response, partial response, and nonresponse in primary care treatment of depression. Arch Intern Med. 2004; 164:1197–1204. [PubMed: 15197045]
41. Fraser SA, Kroenke K, Callahan CM, et al. Low yield of thyroid-stimulating hormone testing in elderly patients with depression. Gen Hosp Psychiatry. 2004; 26:302–309. [PubMed: 15234826]
42. Hedrick SC, Chaney EF, Felker B, et al. Effectiveness of collaborative care depression treatment in Veterans' Affairs primary care. J Gen Intern Med. 2003; 18:9–16. [PubMed: 12534758]
43. Boudreau DM, Capoccia KL, Sullivan SD, et al. Collaborative care model to improve outcomes in major depression. Ann Pharmacother. 2002; 36:585–591. [PubMed: 11918503]
44. Unutzer J, Katon W, Callahan CM, et al. Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. JAMA. 2002; 288:2836–2845. [PubMed: 12472325]
45. Felker B, Katon W, Hedrick SC, et al. The association between depressive symptoms and health status in patients with chronic pulmonary disease. Gen Hosp Psychiatry. 2001; 23:56–61. [PubMed: 11313071]
46. Kroenke K, West SL, Swindle R, et al. Similar effectiveness of paroxetine, fluoxetine, and sertraline in primary care: a randomized trial. JAMA. 2001; 286:2947–2955. [PubMed: 11743835]
47. Williams JW Jr, Barrett J, Oxman T, et al. Treatment of dysthymia and minor depression in primary care: a randomized controlled trial in older adults. JAMA. 2000; 284:1519–1526. [PubMed: 11000645]
48. Katon W, Robinson P, Von Korff M, et al. A multifaceted intervention to improve treatment of depression in primary care. Arch Gen Psychiatry. 1996; 53:924–932. [PubMed: 8857869]
49. Katon W, Von Korff M, Lin E, et al. Collaborative management to achieve treatment guidelines. Impact on depression in primary care. JAMA. 1995; 273:1026–1031. [PubMed: 7897786]
50. Williams JW Jr, Stellato CP, Cornell J, Barrett JE. The 13- and 20-item Hopkins Symptom Checklist Depression Scale: psychometric properties in primary care patients with minor depression or dysthymia. Int J Psychiatry Med. 2004; 34:37–50. [PubMed: 15242140]
51. Walker J, Sharpe M, Kroenke K, Murray G. The HSCL-20: one questionnaire, two versions. J Psychosom Res. 2010; 68:313–314. [PubMed: 20159221]
52. Steffens DC, Snowden M, Fan MY, et al. Cognitive impairment and depression outcomes in the IMPACT study. Am J Geriatr Psychiatry. 2006; 14:401–409. [PubMed: 16670244]
53. Stewart AL, Hays RD, Ware JE Jr. The MOS short-form general health survey. Reliability and validity in a patient population. Med Care. 1988; 26:724–735. [PubMed: 3393032]
54. Ware, JE.; Snow, KK.; Kosinski, M. SF-36 health survey manual and interpretation guide. Lincoln, RI: Quality Metric Inc; 1993.
55. Pan A, Lucas M, Sun Q, et al. Bidirectional association between depression and type 2 diabetes mellitus in women. Arch Intern Med. 2010; 170:1884–1891. [PubMed: 21098346]
56. Cuijpers P, Smits N, Donker T, ten Have M, de Graaf R. Screening for mood and anxiety disorders with the five-item, the three-item, and the two-item Mental Health Inventory. Psychiatry Res. 2009; 168:250–255. [PubMed: 19185354]
57. Dworkin RH, Turk DC, Wyrwich KW, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. J Pain. 2008; 9:105–121. [PubMed: 18055266]

58. Yamazaki S, Fukuhara S, Green J. Usefulness of five-item and three-item Mental Health Inventories to screen for depressive symptoms in the general population of Japan. Health Qual Life Outcomes. 2005; 3:48. [PubMed: 16083512]

59. Lowe B, Unutzer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the patient health questionnaire-9. Med Care. 2004; 42:1194–1201. [PubMed: 15550799]

60. Gilbody S, Richards D, Brealey S, Hewitt C. Screening for depression in medical settings with the Patient Health Questionnaire (PHQ): a diagnostic meta-analysis. J Gen Intern Med. 2007; 22:1596–1602. [PubMed: 17874169]

61. Kroenke K, Spitzer RL, Williams JB, Lowe B. The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: a systematic review. Gen Hosp Psychiatry. 2010; 32:345–359. [PubMed: 20633738]

62. Lowe B, Kroenke K, Herzog W, Grafe K. Measuring depression outcome with a brief self-report instrument: sensitivity to change of the Patient Health Questionnaire (PHQ-9). J Affect Disord. 2004; 81:61–66. [PubMed: 15183601]

63. Kroenke K, Spitzer RL, Williams JB. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care. 2003; 41:1284–1292. [PubMed: 14583691]

64. Lowe B, Kroenke K, Grafe K. Detecting and monitoring depression with a two-item questionnaire (PHQ-2). J Psychosom Res. 2005; 58:163–171. [PubMed: 15820844]

65. Ware JE, Kosinski M, Keller SD. A 12-item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care. 1996; 34:220–233. [PubMed: 8628042]

66. Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med. 2006; 166:1092–1097. [PubMed: 16717171]

67. Sheehan DV, Harnett-Sheehan K, Raj BA. The measurement of disability. Int Clin Psychopharmacol. 1996; 11 (Suppl 3):89–95. [PubMed: 8923116]

68. Idler EL, Benyamini Y. Self-rated health and mortality: a review of twenty-seven community studies. J Health Soc Behav. 1997; 38:21–37. [PubMed: 9097506]

69. Fischer D, Stewart AL, Bloch DA, et al. Capturing the patient's view of change as a clinical outcome measure. JAMA. 1999; 282:1157–1162. [PubMed: 10501119]

70. Cella D, Hahn EA, Dineen K. Meaningful change in cancer-specific quality of life scores: differences between improvement and worsening. Qual Life Res. 2002; 11:207–221. [PubMed: 12074259]

71. Cella D, Yount S, Du H, et al. Development and validation of the Functional Assessment of Cancer Therapy-Kidney Symptom Index (FKSI). J Support Oncol. 2006; 4:191–199. [PubMed: 16669463]

72. Doyle C, Crump M, Pintilie M, Oza AM. Does palliative chemotherapy palliate? Evaluation of expectations, outcomes, and costs in women receiving chemotherapy for advanced ovarian cancer. J Clin Oncol. 2001; 19:1266–1274. [PubMed: 11230467]

73. Yount S, List M, Du H, et al. A randomized validation study comparing embedded versus extracted FACT Head and Neck Symptom Index scores. Qual Life Res. 2007; 16:1615–1626. [PubMed: 17922221]

74. FDA Center for Drug Evaluation and Research. Guidance for industry: Patient reported outcome measures—use in medical product development to support labeling claims. Rockville, MD: US Department of Health and Human Services; 2009.

75. Terwee CB, Dekker FW, Wiersinga WM, Prummel MF, Bossuyt PM. On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. Qual Life Res. 2003; 12:349–362. [PubMed: 12797708]

76. Wright JG, Young NL. A comparison of different indices of responsiveness. J Clin Epidemiol. 1997; 50:239–246. [PubMed: 9120522]

77. Cohen, J. Statistical power analysis for the behavioural sciences. New York, NY: Academic Press; 1977.

78. Buchbinder R, Bombardier C, Yeung M, Tugwell P. Which outcome measures should be used in rheumatoid arthritis clinical trials? Clinical and quality-of-life measures' responsiveness to

treatment in a randomized controlled trial. Arthritis Rheum. 1995; 38:1568–1580. [PubMed: 7488277]
79. Deyo RA, Centor RM. Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. J Chronic Dis. 1986; 39:897–906. [PubMed: 2947907]
80. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol. 2008; 61:102–109. [PubMed: 18177782]
81. Bukberg J, Penman D, Holland JC. Depression in hospitalized cancer patients. Psychosom Med. 1984; 46:199–212. [PubMed: 6739680]
82. Raison CL, Miller AH. Depression in cancer: new developments regarding diagnosis and treatment. Biol Psychiatry. 2003; 54:283–294. [PubMed: 12893104]
83. Dugan W, McDonald MV, Passik SD, et al. Use of the Zung Self-Rating Depression Scale in cancer patients: feasibility as a screening tool. Psychooncology. 1998; 7:483–493. [PubMed: 9885089]
84. Kathol RG, Mutgi A, Williams J, Clamon G, Noyes R Jr. Diagnosis of major depression in cancer patients according to four sets of criteria. Am J Psychiatry. 1990; 147:1021–1024. [PubMed: 2375435]
85. van Wilgen CP, Dijkstra PU, Stewart RE, Ranchor AV, Roodenburg JL. Measuring somatic symptoms with the CES-D to assess depression in cancer patients after treatment: comparison among patients with oral/oropharyngeal, gynecological, colorectal, and breast cancer. Psychosomatics. 2006; 47:465–470. [PubMed: 17116946]
86. Williams JW Jr, Noel PH, Cordes JA, Ramirez G, Pignone M. Is this patient clinically depressed? JAMA. 2002; 287:1160–1170. [PubMed: 11879114]
87. Simon GE, Von Korff M. Medical co-morbidity and validity of DSM-IV depression criteria. Psychol Med. 2006; 36:27–36. [PubMed: 16202189]
88. Koenig HG, George LK, Peterson BL, Pieper CF. Depression in medically ill hospitalized older adults: prevalence, characteristics, and course of symptoms according to six diagnostic schemes. Am J Psychiatry. 1997; 154:1376–1383. [PubMed: 9326819]
89. Krebs EE, Bair MJ, Damush TM, et al. Comparative responsiveness of pain outcome measures among primary care patients with musculoskeletal pain. Med Care. 2010; 48:1007–1014. [PubMed: 20856144]
90. Williams JW Jr, Pignone M, Ramirez G, Perez Stellato C. Identifying depression in primary care: a literature synthesis of case-finding instruments. Gen Hosp Psychiatry. 2002; 24:225–237. [PubMed: 12100833]

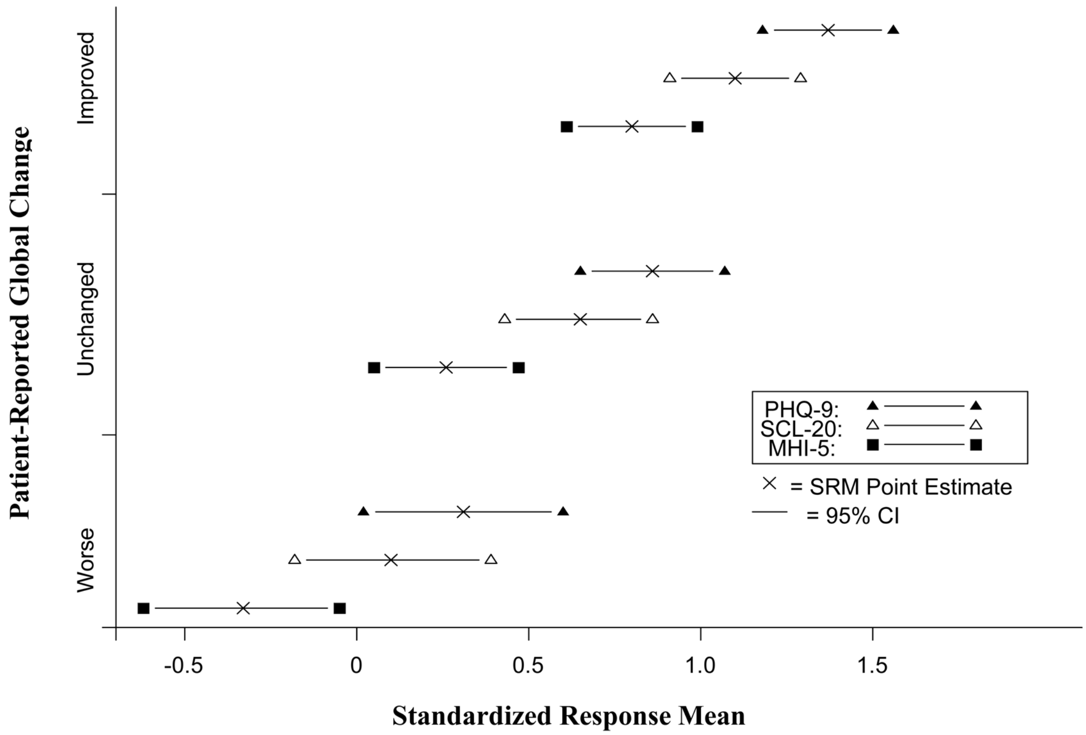

Figure 1.
Standardized response means for the full depression measures by global rating group.

Table 1

Correlations of Depression Measures[a]

Depression Measure	PHQ-9	PHQ-2	SCL-20	MHI-5[2]	MHI-d
PHQ-9		.85	.84	−.70	.70
PHQ-2			.74	−.74	.76
SCL-20				−.73	.73
MHI-5					−.95[b]
MHI-d					

[a] Data are from the three-month assessment, which included all five depression measures in the same interview. The number of subjects with complete three-month data for each of the scales ranged from 245–247. All correlations were significant at $P<0.0001$.

[b] Correlations with MHI-5 are negative since higher scores on this scale represent *less severe* depression, in contrast to other scales where higher scores represent *more severe* depression.

Table 2

Construct Validity of Three Depression Measures[a]

	MCS-12	GAD-7	SF-Vitality	Sheehan Disability	SF-General
Baseline Correlations (*N*=309)					
PHQ-9	−0.47[a]	0.45[a]	−0.28[a]	0.31[a]	−0.20[b]
HSCL-20	−0.65[a]	0.63[a]	−0.45[a]	0.41[a]	−0.18[b]
MHI-5	0.81[a]	−0.61[a]	0.46[a]	−0.38[a]	0.15
Baseline to Three-Month Change Correlations (*N*=247)					
PHQ-9	−0.56[a]	0.44[a]	−0.48[a]	0.44[a]	−0.30[a]
HSCL-20	−0.63[a]	0.54[a]	−0.53[a]	0.45[a]	−0.31[a]
MHI-5	0.79[a]	−0.53[a]	0.51[a]	−0.49[a]	0.25[a]

PHQ-9 = Patient Health Questionnaire nine-item depression scale; HSCL-20 = Hopkins Symptom Checklist 20-item depression scale; MHI-5 = Mental Health Inventory five-item distress scale; MCS-12 = Mental Component Summary 12-item scale; GAD-7 = Generalized Anxiety Disorder seven-item anxiety scale; SF-Vitality = four-item Vitality Scale from the SF-36; SF-General = one item measuring perception of general health.

[a] $P<0.0001$.

[b] $P<0.01$.

Table 3

Change Scores and Standardized Effect Size of Intervention Among Trial Participants Overall and According to Presence of Depression With and Without Pain

	Full Sample (*N* = 247)			Depression Only Subgroup (*N* = 105)			Depression + Pain Subgroup (*N* = 142)		
	Intervention group change [a] (SD)	**Control group change [a] (SD)**	**SES**	**Intervention group change [a] (SD)**	**Control group change [a] (SD)**	**SES**	**Intervention group change [a] (SD)**	**Control group change [a] (SD)**	**SES**
Full Scales									
PHQ-9	−6.43 (6.55)	−4.87 (5.85)	−0.25	−5.08 (6.88)	−5.53 (4.99)	0.08	−7.44 (6.15)	−4.36 (6.43)	−0.48
SCL-20	−0.60 (0.66)	−0.31 (0.62)	−0.43	−0.50 (0.75)	−0.37 (0.49)	−0.21	−0.66 (0.59)	−0.27 (0.71)	−0.58
MHI-5[b]	−10.1 (22.6)	−5.7 (18.3)	−0.21	−11.3 (25.4)	−9.1 (15.4)	−0.10	−9.24 (20.4)	−3.10 (20.0)	−0.30
Abbreviated Scales									
PHQ-2	−1.70 (2.30)	−1.48 (2.05)	−0.10	−1.54 (2.32)	−1.85 (1.84)	0.15	−1.82 (2.29)	−1.19 (2.17)	−0.28
MHI-d	−1.18 (3.21)	−0.98 (2.52	−0.07	−1.52 (3.46)	−1.55 (1.87)	0.01	−0.93 (3.0	−0.55 (2.87)	−0.13

SES = standardized effect size.

[a] For both the Intervention and Control groups, change = Baseline score – three-month score.

[b] MHI-5 is the only scale for which higher score indicates less depression; therefore, signs are reversed for MHI-5 to simplify comparison with the other scales.

Table 4

Accuracy of Depression Measures for Detecting Improvement at Three Months

	Accuracy for Detecting Any Improvement [a] **(*N* = 106 *yes* vs. 137 *no*)**		**Accuracy for Detecting ≥Moderate Improvement**[b] **(*N* = 54 *yes* vs. 189 *no*)**	
	AUC	**SE**	**AUC**	**SE**
Full Scales				
PHQ-9	.710	.033	.778	.035
SCL-20	.690	.034	.743	.036
MHI-5	.712	.033	.737	.039
Abbreviated Scales				
PHQ-2	.635	.035	.670	.041
MHI-d	.720	.033	.690	.041

AUC = Area under the curve, or the probability of correctly discriminating between patients who reported improvement and those who did not; SE = standard error.

[a]Any improvement = a little better, somewhat better, moderately better, a lot better, or completely better.

[b]Moderate improvement = moderately better, a lot better, or completely better.