Published in final edited form as:
*J Pain*. 2012 August ; 13(8): 764–772. doi:10.1016/j.jpain.2012.05.004.

# Comparative Responsiveness of Pain Measures in Cancer Patients

**Kurt Kroenke**[*,†,‡], **Dale Theobald**[§], **Jingwei Wu**[†], **Wanzhu Tu**[†,‡], and **Erin E. Krebs**[*,†,‡]

[*]Veterans Affairs Health Services Research and Development Center of Excellence for Implementing Evidence-Based Practice, Indianapolis, Indiana

[†]Department of Medicine, Indiana University, Indianapolis, Indiana

[‡]Regenstrief Institute, Inc, Indianapolis, Indiana

[§]Community Home Health Hospice and Symptom Management Group, Indianapolis, Indiana

## Abstract

Brief measures to assess and monitor pain in cancer patients are available, but few head-to-head psychometric comparisons of different measures have been reported. Baseline and 3-month data were analyzed from 274 patients enrolled in the Indiana Cancer Pain and Depression (INCPAD) trial. Participants completed the Brief Pain Inventory (BPI), the PEG (a 3-item abbreviated version of the BPI), the short form (SF)-36 pain scale, and a pain global rating of change measure. The global rating was used as the criterion for standardized response mean and receiver operating characteristic curve analyses. To assess responsiveness to the trial intervention, we evaluated standardized effect size statistics stratified by trial arm. All measures were responsive to global improvement, discriminated between participants with and without improvement, and detected a significant intervention treatment effect. Short and longer measures were similarly responsive. Also, composite measures that combined pain severity and interference into a single score (BPI total, PEG, SF-36 pain) performed comparably to separate measures of each domain (BPI severity and BPI interference).

**Perspective—**Pain measures as brief as 2 or 3 items that provide a single score are responsive in patients with cancer-related pain. Ultra-brief measures offer a valid and efficient means of assessing and monitoring pain for the clinical management as well as research of cancer-related pain.

### Keywords

Cancer; pain; measures; assessment; diagnosis; psychometrics

Pain is one of the most prevalent symptoms in patients with cancer, experienced by one-third to two-thirds of all patients at some point during their illness.[54] Pain can occur in all phases of cancer and may be caused by the tumor itself, diagnostic or therapeutic

Address reprint requests to Kurt Kroenke, MD, Regenstrief Institute, 5th Floor, 1050 Wishard Blvd, Indianapolis, IN 46202. kkroenke@regenstrief.org.

None of the authors have any financial or other conflicts of interest to disclose relevant to this paper.

procedures, or cancer treatment.[43,44] Cancer-related pain has a major adverse impact on functional status and quality of life[25,38,43,51,57] and, moreover, is undertreated in a substantial proportion of patients.[10,15,47] A national survey of a geographically representative sample of medical oncologists in the U.S. found that 2 of the greatest barriers to optimal cancer pain management were patient reluctance to report pain and inadequate assessment of pain by physicians and nurses.[2] This survey also revealed inadequate training in pain management, poor treatment decisions in simulated cases, and low referral rates to pain or palliative care specialists.

Valid and efficient ways of assessing pain are essential to optimize detection, monitoring, and management.[34] Multidimensional approaches to assessing pain are recommended for clinical trials and other types of research.[13,14] In clinical practice, however, brief and simple measures may be more desirable. Krebs et al[31,33] have found that ultra-brief pain measures that yield a single composite score may be comparable to longer, multidimensional measures for evaluating and treating chronic noncancer pain. Whether this is also true for cancer-related pain, however, is not known. Despite literature reviews and the convening of expert groups, there is not yet a consensus on the optimal tool for assessing pain in cancer patients.[5,19,21,23,26,28]

The Indiana Cancer Pain and Depression (INCPAD) study, a randomized clinical trial implemented with patients who had different types and phases of cancer, tested a 12-month telephone care management intervention to decrease pain and/or depression.[37] The INCPAD trial demonstrated moderate pain improvement in the intervention group compared to the usual-care group ($P < .001$).[39] The pain measures that were serially administered in INCPAD included the Brief Pain Inventory (BPI), the 3-item PEG scale (an abbreviated version of the BPI), and the 2-item Short-Form (SF)-36 pain scale. While the BPI is the most commonly used measure for assessing both pain severity and interference in adults with cancer and the SF-36 pain scale is a frequently used composite pain measure, there is a limited amount of research doing head-to-head psychometric comparisons of these or other pain measures in patients with cancer.[26] Such comparisons are important for researchers in selecting measures for studies where pain is either a primary or secondary outcome, and for clinicians in choosing a measure they can use in practice to monitor response to pain treatments and to adjust therapy as needed. Factor analysis of brief and longer pain measures in a small Italian study of patients with cancer-related pain revealed a high degree of association among pain dimensions as well as a single factor explaining most of the different scales variability.[9] Limitations of this study included its small sample size (n = 53) and the inclusion of some measures infrequently used in clinical research and practice.

One important aspect of a measure's validity is responsiveness, the ability to detect a meaningful change in a clinical state. Responsiveness, also referred to as sensitivity to change, can be assessed in several ways. In this paper, we therefore compare the responsiveness of these 3 measures to cancer-related pain over 3 months by assessing the performance of the measures in 1) differentiating (using standardized response means) between subgroups of patients who report global pain improvement, lack of change, or worsening over 3 months; 2) discriminating (using receiver operating curve [ROC] analysis) between patients who did and did not report improvement of their pain at 3 months; and 3)

detecting (using standardized effect sizes) an intervention treatment effect, as well as assessing whether the treatment effect differed between depressed and nondepressed patients.

## Methods

### Participant Enrollment

The design and baseline participant characteristics in the INCPAD study have been published elsewhere, including participant enrollment and follow-up.[37,39] Briefly, patients with cancer-related pain and/or depression were recruited from 16 urban and rural outpatient oncology clinics. Potential participants were patients who had moderately severe cancer-related pain (a BPI worst-pain severity score ≥6), or depression (a Patient Health Questionnaire 9-item depression scale [PHQ-9] score ≥0, with depressed mood and/or anhedonia). Cancer-related pain had to be in the region of the primary tumor or cancer metastases and/or present after the onset of the cancer treatment. Patients with cancer-related pain had to have tried at least 1 analgesic but still be experiencing pain. The study excluded those who did not speak English, had moderately severe cognitive impairment, had schizophrenia or other psychoses, had a disability claim currently being adjudicated for pain, were pregnant, were in hospice care, or had preexisting pain conditions unrelated to cancer.

A total of 405 participants with depression, cancer-related pain, or both were enrolled in the INCPAD trial. For this secondary analysis, data from 274 participants who had cancer-related pain (with or without depression) were analyzed: 137 in the intervention group and 137 in the usual-care group. These 274 participants were enrolled from a group of 444 patients who met all entry criteria for cancer-related pain, thus yielding an enrollment rate of 62% of potentially eligible patients. The intervention group received centralized telephone care management (telecare) focusing on optimizing medications to treat their cancer-related pain and/or depression, while the usual-care group received care by their oncologists without attempts by study personnel to influence their pain and/or depression treatment unless a psychiatric emergency arose. Compared to the usual-care group, patients in the telecare intervention group had significantly greater pain improvement as assessed by the primary outcome of BPI pain severity (effect sizes of .36, .67, .46, and .39 at 1, 3, 6, and 12 months, respectively, $P < .001$).[39]

### Outcome Assessment

Research assistants who conducted the phone interviews to collect data at baseline and follow-up were blinded to group assignment. Medical record reviews from the oncology practice were performed during the study. The study was approved by the institutional review boards at Indiana University and Community Hospital.

We assessed responsiveness of the following pain outcome measures:

- The BPI has been validated as a measure of chronic pain in cancer and other clinical populations.[7,8,31] The BPI includes 4 items that assess the intensity of pain currently and at its least, worst, and average during the past week (rated from 0, “no pain,” to 10, “pain as bad as you can imagine”). An additional 7 items assess

pain-related functional interference (rated from 0, “does not interfere,” to 10, “interferes completely”). We assessed responsiveness of the 4-item BPI severity and 7-item BPI interference scales, as well as that of a BPI total score that includes all 11 items. For each scale, the total score is the average of all items (range, 0–10), where a higher score indicates worse pain.

- The PEG pain scale is a 3-item abbreviated scale of the BPI that includes 1 intensity item and 2 interference items (Pain intensity during the past week, pain interference with Enjoyment of life, and pain interference with General activity).[33] Like the BPI, the total score is the average of all items (range, 0–10).
- The SF-36 pain scale is a validated subscale of the Medical Outcomes Study SF-36 questionnaire.[45,56] It includes 2 items, one that assesses intensity of pain (scale range: 1, “none,” to 6, “very severe”) and one that assesses how much pain has interfered with work (scale range: 1, “not at all,” to 5, “extremely”) over the past 4 weeks. Responses are transformed into a 0 to 100 score, where a lower score indicates worse pain.

We used a patient-reported retrospective global rating of change as the reference standard for change in pain.[16,31] At 3 months, participants were asked, “Overall, since starting the study, would you say your pain is worse, about the same, or better?” Those who reported their pain was better were asked a second question, “How much better is your pain?” with the following response options: a little, somewhat, moderately, a lot, or completely better.

Other study measures used in our analyses include depression severity as assessed by the Hopkins Symptom Checklist 20-item depression scale (HSCL-20); medical comorbidity as assessed by an 8-disease checklist; and the Socioeconomic Disadvantage (SED) index, a 3-point composite score that assigns 1 point each for low education, low income, and unemployment.[40]

## Analysis

Baseline and 3-month data are used for the analyses in the present paper because the most intense phase of the INCPAD pain intervention as well as the greatest change in pain scores occurred during the initial 3 months of the trial. Moreover, this period was least affected by attrition due to death, disease progression, drop-out, or loss to follow-up that had occurred by 6 and 12 months. For the analyses that required only baseline data, we used data from all 274 participants enrolled with pain, whereas responsiveness analyses included the 230 participants with data available at both baseline and 3-month follow-up.

We calculated standardized response means (SRM) and ROC curves to analyze responsiveness. Retrospective patient-reported global ratings of change are the most common anchors used for responsiveness analysis in pain research; in this study, we used the 3-month global rating of change. Additionally, to assess responsiveness of each measure to group differences in the randomized trial, we evaluated standardized effect size (SES) statistics stratified by trial arm (intervention and control). Statistical analyses were performed using SAS v.9.1 (SAS Institute Inc, Cary, NC).

**Standardized Response Means—**The SRM is an effect size index that includes response variance in the denominator.[29] SRM values are unitless and therefore directly comparable between measures. Effect sizes of .2 are generally considered small, .5 are moderate, and .8 are large.[30] We categorized participants according to their 3-month global rating of change (worse, same, or better). For each measure, we calculated the SRM (3-month mean change score/SD of change) stratified by category of change.[41] A bootstrapping procedure was used to derive 95% confidence intervals (CIs) for the SRM.[42]

**Area Under the ROC Curve—**Area under the ROC curve (AUC) values are interpreted as the probability that a measure correctly discriminates between patients who are improved and those who are not; the possible range of values is .5 (no ability to discriminate) to 1.0 (perfect ability to discriminate). Because ROC analysis requires a binary outcome, we classified patients who stated that their pain was "better" as improved and those who stated that their pain was "the same" or "worse" as not improved.[31] The rationale for collapsing the latter 2 categories is that clinicians would typically consider patients who stay the same as treatment failures (ie, nonresponders), because the goal of treatment is to make patients better rather than have their pain remain unchanged or worsen. We calculated the AUC for each outcome measure using improvement on the 3-month global rating of change as the anchor.[12] We secondarily calculated AUCs for a moderate improvement threshold (at least "moderately" better).[14,31] Finally, we conducted pairwise comparisons between pain measures to determine if they were significantly different in their ability to detect improvement.[11]

**Standardized Effect Sizes by Randomized Trial Arm—**To assess measures' responsiveness to intervention effects, we evaluated SES according to trial arm (telecare intervention versus usual care). We calculated change scores (3 months minus baseline) and standardized effect sizes (SES = [intervention group change minus usual-care group change]/SD of pooled change scores)[4] to assess whether measures performed similarly among participants in the trial overall as well as in the depressed and nondepressed subgroups, stratified by intervention group.

**Effect of Depression on Responsiveness of Pain Measures—**We first examined whether depression status affected responsiveness in a logistic regression model where the dependent variable was patient global rating of pain change (improved versus same/worse), the predictor variable was depression as a categorical variable (yes/no) defined by our PHQ-9 thresholds, and covariates were baseline pain severity, age, sex, race, SED index, medical comorbidity, and treatment arm (intervention versus control). Second, we ran the same model using as the predictor variable baseline depression severity as a continuous variable (HSCL-20 score). Third, we also examined whether responsiveness of each of the specific pain scales varied with depression status by running linear regression models with the 3-month change score on the pain scale as the dependent variable, baseline depression status as a categorical (yes/no) predictor variable, and adjusting for the same covariates.

## Results

### Patient Characteristics

The baseline characteristics of the participants are shown in Table 1. Two-thirds of the participants were female, 77% were white, and 65% had both cancer-related pain and depression upon enrollment in the study. Their mean age was 58 years, with a range from 23 to 85 years. At baseline, they reported moderate pain intensity on the BPI scale. Two-thirds of the patients had comorbid depression. As previously reported, randomization resulted in intervention and usual-care groups that were similar in all baseline characteristics.[37] According to the pain global rating of change, 51 (22.2%) of the 230 participants were better at 3 months, 72 (31.3%) were about the same, and 107 (46.5%) were worse.

### Internal Reliability and Interscale Correlations

The pain scales had good internal reliability with a Cronbach's alpha of .79 for the BPI severity scale, .89 for the BPI interference scale, .89 for the BPI total scale, .69 for the PEG scale, and .73 for the SF-36 pain scale.

Table 2 shows the correlations of the pain scales with one another at baseline and 3 months. All scales were strongly correlated at both timepoints, with, as expected, the highest correlations for BPI-related scales that had some items in common.

### Standardized Response Means

Table 3 shows SRM values for patients classified as worse, the same, and better according to their 3-month global rating of change. Overall, effect sizes for measures were similar within each category of change. Participants with improved pain according to the global rating had moderate-to-large improvement in scores on each scale. Also, SRMs significantly differed ($P < .001$) between the better, same, and worse groups for each of the scales, except for the same and worse groups on the SF-36 pain scale.

### Area Under the ROC Curve

Table 4 shows the AUC for each measure, first using any improvement on the global rating as the reference standard and second using moderate improvement on the global rating. Overall, AUC values for the measures were similar; most AUC values had fair discriminatory ability (range, .73–.81). There were no statistically significant differences between the AUCs of any of the measures in detecting either any improvement or moderate improvement. The responsiveness of each measure to the presence of any improvement was about as good as its responsiveness to moderate improvement, and shorter measures (PEG and SF-36 pain) performed similarly to longer measures.

### Standardized Effect Sizes

As shown in Table 5, all measures detected a moderate intervention effect, with relatively similar SES for the scales, ranging from .43 to .58 in the overall sample. SES were generally larger in the nondepressed than in the depressed subgroup, consistent with the greater difficulty of treating pain in depressed patients.[1]

### Effect of Depression on Responsiveness of Pain Measures

In logistic regression models, baseline depression status measured as either a categorical or continuous variable did not influence patient global rating of change (improved versus same/worse) in unadjusted or adjusted models. The only covariates that resulted in a greater likelihood of improvement were being assigned to the intervention arm and having fewer comorbid medical diseases. Likewise, linear regression models demonstrated that the magnitude of improvement as assessed by 3-month change scores for each of the specific pain scales was greater in individuals assigned to the intervention arm as well as those with more severe pain at baseline but was not influenced by depression status.

## Discussion

### Summary of Measures

Our study has several important findings. First, all 3 pain measures (BPI, PEG, and SF-36 pain) performed similarly in detecting improvement and in differentiating outcome categories (better, same, worse) by 2 types of responsiveness analyses: SRM and ROC. Second, the measures were comparable in detecting a moderate between-group (intervention versus control) treatment effect in the randomized trial. Third, composite measures (BPI total, PEG, and SF-36 pain) performed as well as the BPI subscales that measure pain severity and pain interference separately. Fourth, shorter scales (PEG and SF-36 pain) consisting of 2 or 3 items were nearly as responsive as longer measures.

There are several important reasons that global rating of change was selected as the standard for comparing the responsiveness of pain measures. First, comparing the responsiveness of multiple different pain measures required using a single method independent of the pain measures themselves to classify patients into the 3 groups of improved, unchanged, and worse. Had longitudinal changes on the pain measures themselves been used to classify patients, there would have been differing numbers of patients in the 3 groups for each of the measures, and it would have been impossible to do comparisons across measures. Second, global rating of change is an anchor-based method of assessing change, whereas longitudinal change on a continuous pain measure is a distribution-based method of assessing change. Both methods contribute important but complementary information in assessing responsiveness, and using an anchor-based standard is a well-accepted approach to compare responsiveness of distribution-based measures.[16,49,58] Third, global rating of change is one of the most commonly used anchor-based methods for assessing responsiveness in pain research.[14,31] Of note, the results in Table 3 provide empiric support for the use of global rating of change as the independent standard because the SRMs for each pain measure discriminated between the 3 subgroups of patients who globally rated their pain as better, the same, or worse.

The pain scales showed a moderate between-group treatment effect size in patients with and without comorbid depression (Table 5). This is a salient finding given the frequency of comorbid depression in cancer patients with pain and the additive adverse effects of depression and pain on quality of life and functional status.[38] The treatment effect size, although still moderate in magnitude, was somewhat less in depressed patients compared to

nondepressed patients (Table 5), which is consistent with previous research showing that pain response to therapy may be negatively influenced by depression.[1,35,59] The fact that a reasonable treatment effect was nonetheless demonstrated in the depressed group may in part have been due to the design of the INCPAD intervention, which provided treatment for depression as well as pain in patients suffering from both symptoms.[37] Importantly, our multivariable models did not find baseline depression status measured as either a categorical or a continuous variable to be a confounder of patient-reported global improvement in pain or in change scores on any of the specific pain measures. Thus, the smaller treatment effect sizes seen for depressed compared to nondepressed patients more likely represents the greater difficulty in treating pain in depressed patients rather than a biasing influence of depression on pain assessment. Future studies incorporating differential item functioning analyses to further explore the psychometric effects of depression may be worthwhile.

## Translating Results into Practice

Pain screening, assessment, and monitoring of treatment response are core quality indicators for management of cancer-related pain.[19] However, as recently noted by Hjermastad et al,[22] "Pain is still not routinely measured in cancer clinical practice. This may be because most tools are too long and cumbersome for patients and clinicians to use." For example, 2 of the most commonly used measures in research of cancer-related pain have been the McGill Pain Questionnaire and the BPI.[5,25,48] However, the length and complexity of the McGill Pain Questionnaire in particular makes it less feasible for clinical practice, in contrast to the shorter and simpler BPI. Indeed, numerical (0–10) rating scales (of which the BPI and PEG are exemplars) have been shown to have higher discriminatory capability than verbal rating scales in distinguishing differing levels of pain[3] and are increasingly being recommended by experts in palliative care.[23,28]

Our previous study of primary care patients with noncancer musculoskeletal pain also showed a similar responsiveness of the PEG, BPI severity, BPI interference, and BPI total scales.[31] In that study, the SF-36 pain scale appeared somewhat less responsive. It has been recommended that pain severity and pain interference are distinct dimensions, both of which should be assessed in pain research.[13,14] Indeed, a systematic literature review of pain assessment scales for patients with advanced cancer receiving palliative care revealed that pain severity and interference accounted for the majority of items on the scales.[24]

The BPI total, PEG, and SF-36 pain scales do include both severity and interference items but combine them into a single composite score. This may be attractive in clinical practice so that providers need not interpret and monitor 2 separate numbers but rather can follow a single score. In particular, separate scores may complicate decision-making if severity and interference ratings are discordant. Of the 2 dimensions, the interference/functional impairment associated with pain may be of somewhat greater importance than pain severity alone.[53,55] In this regard, the BPI total and PEG assign approximately a 2:1 weight to interference items, which constitute 7 of the 11 items in the BPI total and 2 of the 3 items in the PEG. Further evidence in support of a composite score is provided by a comparison of 6 pain measures in 53 cancer patients that found that a single factor explained most of the different scales' variability.[9] Although a single score may be especially attractive for clinical

practice, the fact that composite scores have responsiveness similar to separate severity and interference scales may also be useful in choosing a primary outcome in clinical trials.[39,52]

Brevity is another desirable feature of measures intended for clinical practice. For example, the validity of ultra-brief measures (2–4 items) for depression and anxiety has been established.[36,46] The fact that responsiveness of the 3-item PEG is similar to its longer parent measure for both cancer-related pain in our INCPAD trial as well as noncancer pain in a previous trial[31] suggests a promising role for ultra-brief pain measures as well. Currently, a single-item assessment of current pain on a 0 to 10 scale has been promulgated as part of the "pain as the fifth vital sign" initiative advocated by The Joint Commission and other health organizations. However, this single item was initially developed to assess acute pain in hospitalized patients, whereas studies of its performance in outpatients with chronic pain have demonstrated only modest accuracy in detecting clinically important pain.[32] In this regard, the ultra-brief PEG may achieve an optimal balance between single-item and longer pain scales. The SF-36 pain scale is also brief but is more complicated to score (requiring transformation) and was less responsive than the PEG in a previous trial of noncancer pain.[31] It is possible that the longer recall window for pain of the SF-36 pain scale compared to the PEG and BPI scales (past 4 weeks versus past 1 week) may partly account for its lower responsiveness in treatment trials.

## Example Case

A 62-year-old man with newly diagnosed lung cancer and right chest wall pain reports a PEG score of 7, indicating a relatively high level of pain (1–3, 4–6, and 7–10 represent mild, moderate, and severe levels of pain, respectively). Initiation of opioid therapy reduced the PEG score to 3; however, a subsequent increase to 8 following thoracotomy required an escalation in opioid dose. After 4 weeks, nonopioid analgesic therapy was sufficient until 6 months later when the cancer progressed and bony metastases developed. Long-acting opioids were initiated with the supplemental use of short-acting opioids for rescue pain. As the patient progressed to hospice care, localized radiotherapy for bone pain as well as higher doses of opioids were used in efforts to maintain the PEG score in the range of 3 to 5 as much as possible. To minimize unnecessary clinic visits, the patient's PEG score was monitored at home using automated technology (either interactive voice-recorded phone calls or internet-based reporting) coupled with nurse calls for PEG scores that failed to improve or worsen or for treatment-related side effects or nonadherence.[27]

## Study Limitations

Our study has several limitations. While its generalizability was enhanced by enrolling patients with a wide range of types and phases of cancer, the target population was oncology outpatients reporting pain rather than inpatients or patients receiving hospice care. The latter groups might require even more intensive monitoring and analgesic adjustments including more frequent assessment and special attention to breakthrough pain.[6,21,50] A second limitation is the use of multiple comparisons between multiple measures. In this regard, the similar responsiveness we found for these same measures in a primary care population with noncancer pain is reassuring.[31] A third limitation is a likely ceiling effect for worsening of pain in our study population. Participants with pain in INCPAD had pain of at least

moderate severity at baseline and, therefore, a more restricted range for pain to worsen. This may partly explain why change scores among those who reported improvement were larger in magnitude than change scores among those who reported worsening. However, it is also possible that there is greater concordance between anchorand distribution-based outcome measures (ie, measures based on recall of global change from a previous time point versus those that assess the present status of pain) in terms of assessing improvement compared to assessing worsening. To this end, the recommendation that both types of measures be used to assess the outcomes of pain treatments is warranted.[13,14]

### Pain Measurement as One Component of Improving Cancer Pain Management

Although the availability of scales that are at once valid and ultra-brief is an important advance in the management of cancer-related pain, measurement alone is insufficient. By analogy, it has been established that simple screening for depression does not improve outcomes[18] unless it is coupled with systems-based interventions that include regular monitoring and adjustments of treatment.[17] Likewise, quality indicators for cancer-related pain include not only assessment but also patient education, adjustment of treatment to achieve adequate control, the use of multimodal treatments, and attention to function and side effects as well as pain intensity.[19] Nonetheless, pragmatic scales are an essential component of the measurement-based care that is necessary for the optimal treatment of pain[34] as well its concomitant psychological comorbidity.[20]

## Acknowledgments

Supported by grant R01 CA-115369 from the National Cancer Institute to Dr. Kroenke and grant MH-071268 from the National Institute of Mental Health to Dr. Kroenke.

## References

1. Bair MJ, Robinson RL, Katon W, Kroenke K. Depression and pain comorbidity: A literature review. Arch Intern Med. 2003; 163:2433–2445. [PubMed: 14609780]
2. Breuer B, Fleishman SB, Cruciani RA, Portenoy RK. Medical oncologists' attitudes and practice in cancer pain management: A national survey. J Clin Oncol. 2011; 29:4769–4775. [PubMed: 22084372]
3. Brunelli C, Zecca E, Martini C, Campa T, Fagnoni E, Bagnasco M, Lanata L, Caraceni A. Comparison of numerical and verbal rating scales to measure pain exacerbations in patients with chronic cancer pain. Health Qual Life Outcomes. 2010; 8:42. [PubMed: 20412579]
4. 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]
5. Caraceni A, Cherny N, Fainsinger R, Kaasa S, Poulain P, Radbruch L, De CF. Pain measurement tools and methods in clinical research in palliative care: Recommendations of an Expert Working Group of the European Association of Palliative Care. J Pain Symptom Manage. 2002; 23:239–255. [PubMed: 11888722]
6. Caraceni A, Zecca E, Martini C, Brunelli C, Pigni A, Gorni G, Galbiati A, Ibazeta M, Hjermstad M, Kaasa S. The validity of average 8-h pain intensity assessment in cancer patients. Eur J Pain. 2010; 14:441–445. [PubMed: 19692275]
7. Cleeland, CS. Pain assessment in cancer. In: Osoba, D., editor. Effect of Cancer on Quality of Life. Boca Raton, FL: CRC Press; 1991. p. 293-305.

8. Cleeland CS, Nakamura Y, Mendoza TR, Edwards KR, Douglas J, Serlin RC. Dimensions of the impact of cancer pain in a four country sample: New information from multidimensional scaling. Pain. 1996; 67:267–273. [PubMed: 8951920]
9. De Conno F, Caraceni A, Gamba A, Mariani L, Abbattista A, Brunelli C, La MA, Ventafridda V. Pain measurement in cancer patients: A comparison of six methods. Pain. 1994; 57:161–166. [PubMed: 8090512]
10. Deandrea S, Montanari M, Moja L, Apolone G. Prevalence of undertreatment in cancer pain. A review of published literature. Ann Oncol. 2008; 19:1985–1991. [PubMed: 18632721]
11. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. Biometrics. 1988; 44:837–845. [PubMed: 3203132]
12. 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]
13. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, Kerns RD, Stucki G, Allen RR, Bellamy N, Carr DB, Chandler J, Cowan P, Dionne R, Galer BS, Hertz S, Jadad AR, Kramer LD, Manning DC, Martin S, McCormick CG, McDermott MP, McGrath P, Quessy S, Rappaport BA, Robbins W, Robinson JP, Rothman M, Royal MA, Simon L, Stauffer JW, Stein W, Tollett J, Wernicke J, Witter J. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain. 2005; 113:9–19. [PubMed: 15621359]
14. Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, Haythornthwaite JA, Jensen MP, Kerns RD, Ader DN, Brandenburg N, Burke LB, Cella D, Chandler J, Cowan P, Dimitrova R, Dionne R, Hertz S, Jadad AR, Katz NP, Kehlet H, Kramer LD, Manning DC, McCormick C, McDermott MP, McQuay HJ, Patel S, Porter L, Quessy S, Rappaport BA, Rauschkolb C, Revicki DA, Rothman M, Schmader KE, Stacey BR, Stauffer JW, von ST, White RE, Witter J, Zavisic S. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. J Pain. 2008; 9:105–121. [PubMed: 18055266]
15. Fairchild A. Under-treatment of cancer pain. Curr Opin Support Palliat Care. 2010; 4:11–15. [PubMed: 20040878]
16. Fischer D, Stewart AL, Bloch DA, Lorig K, Laurent D, Holman H. Capturing the patient's view of change as a clinical outcome measure. JAMA. 1999; 282:1157–1162. [PubMed: 10501119]
17. Gilbody S, Bower P, Fletcher J, Richards D, Sutton AJ. Collaborative care for depression: A cumulative meta-analysis and review of longer-term outcomes. Arch Intern Med. 2006; 166:2314–2321. [PubMed: 17130383]
18. Gilbody S, Sheldon T, House A. Screening and case-finding instruments for depression: A meta-analysis. CMAJ. 2008; 178:997–1003. [PubMed: 18390942]
19. Gordon DB, Dahl JL, Miaskowski C, McCarberg B, Todd KH, Paice JA, Lipman AG, Bookbinder M, Sanders SH, Turk DC, Carr DB. American pain society recommendations for improving the quality of acute and cancer pain management: American Pain Society Quality of Care Task Force. Arch Intern Med. 2005; 165:1574–1580. [PubMed: 16043674]
20. Harding KJ, Rush AJ, Arbuckle M, Trivedi MH, Pincus HA. Measurement-based care in psychiatric practice: A policy framework for implementation. J Clin Psychiatry. 2011; 72:1136–1143. [PubMed: 21295000]
21. Haugen DF, Hjermstad MJ, Hagen N, Caraceni A, Kaasa S. Assessment and classification of cancer breakthrough pain: A systematic literature review. Pain. 2010; 149:476–482. [PubMed: 20236762]
22. Hjermstad MJ, Fainsinger R, Kaasa S. Assessment and classification of cancer pain. Curr Opin Support Palliat Care. 2009; 3:24–30. [PubMed: 19365158]
23. Hjermstad MJ, Gibbins J, Haugen DF, Caraceni A, Loge JH, Kaasa S. Pain assessment tools in palliative care: An urgent need for consensus. Palliat Med. 2008; 22:895–903. [PubMed: 18799513]
24. Holen JC, Hjermstad MJ, Loge JH, Fayers PM, Caraceni A, De CF, Forbes K, Furst CJ, Radbruch L, Kaasa S. Pain assessment tools: Is the content appropriate for use in palliative care? J Pain Symptom Manage. 2006; 32:567–580. [PubMed: 17157759]

25. Holen JC, Lydersen S, Klepstad P, Loge JH, Kaasa S. The Brief Pain Inventory: Pain's interference with functions is different in cancer pain compared with noncancer chronic pain. Clin J Pain. 2008; 24:219–225. [PubMed: 18287827]
26. Jensen MP. The validity and reliability of pain measures in adults with cancer. J Pain. 2003; 4:2–21. [PubMed: 14622723]
27. Johns SA, Kroenke K, Theobald D, Wu J, Tu W. Telecare management of pain and depression in patients with cancer: Patient satisfaction and predictors of use. J Ambulatory Care Management. 2011; 34:126–139.
28. Kaasa S, Apolone G, Klepstad P, Loge JH, Hjermstad MJ, Corli O, Strasser F, Heiskanen T, Costantini M, Zagonel V, Groenvold M, Fainsinger R, Jensen MP, Farrar JT, McQuay H, Rothrock NE, Cleary J, Deguines C, Caraceni A. Expert conference on cancer pain assessment and classification— the need for international consensus: Working proposals on international standards. BMJ Support Palliat Care. 2011; 1:281–287.
29. Katz JN, Larson MG, Phillips CB, Fossel AH, Liang MH. Comparative measurement sensitivity of short and longer health status instruments. Med Care. 1992; 30:917–925. [PubMed: 1405797]
30. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. Med Care. 1989; 27:S178–S189. [PubMed: 2646488]
31. Krebs EE, Bair MJ, Wu J, Damush TM, Tu W, Kroenke K. Comparative responsiveness of pain outcome measures among primary care patients with musculoskeletal pain. Med Care. 2010; 48:1007–1014. [PubMed: 20856144]
32. Krebs EE, Carey TS, Weinberger M. Accuracy of the pain numeric rating scale as a screening test in primary care. J Gen Intern Med. 2007; 22:1453–1458. [PubMed: 17668269]
33. Krebs EE, Lorenz KA, Bair MJ, Damush TM, Wu J, Sutherland J, Asch SM, Kroenke K. Development and initial validation of the PEG, a three-item scale assessing pain intensity and interference. J Gen Intern Med. 2009; 24:733–738. [PubMed: 19418100]
34. Kroenke K. Measuring and distilling the impact of pain. Pain. 2009; 142:7–8. [PubMed: 19201095]
35. Kroenke K, Wu J, Bair MJ, Krebs EE, Damush TM, Tu W. Reciprocal relationship between pain and depression: A 12-month longitudinal analysis in primary care. J Pain. 2011; 12:964–973. [PubMed: 21680251]
36. 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]
37. Kroenke K, Theobald D, Norton K, Sanders R, Schlundt S, McCalley S, Harvey P, Iseminger K, Morrison G, Carpenter JS, Stubbs D, Jacks R, Carney-Doebbeling C, Wu J, Tu W. Indiana Cancer Pain and Depression (INCPAD) Trial: Design of a telecare management intervention for cancer-related symptoms and baseline characteristics of enrolled participants. Gen Hosp Psychiatry. 2009; 31:240–253. [PubMed: 19410103]
38. Kroenke K, Theobald D, Wu J, Loza JK, Carpenter JS, Tu W. The association of depression and pain with health-related quality of life, disability, and health care use in cancer patients. J Pain Symptom Manage. 2010; 40:327–341. [PubMed: 20580201]
39. Kroenke K, Theobald D, Wu J, Norton K, Morrison G, Carpenter J, Tu W. Effect of telecare management on pain and depression in patients with cancer: A randomized trial. JAMA. 2010; 304:163–171. [PubMed: 20628129]
40. Kroenke K, Zhong X, Theobald D, Wu J, Tu W, Carpenter JS. Somatic symptoms in patients with cancer experiencing pain or depression: Prevalence, disability, and health care use. Arch Intern Med. 2010; 170:1686–1694. [PubMed: 20937930]
41. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. Med Care. 1990; 28:632–642. [PubMed: 2366602]
42. 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]
43. Mantyh PW. Cancer pain and its impact on diagnosis, survival and quality of life. Nat Rev Neurosci. 2006; 7:797–809. [PubMed: 16988655]

44. McGuire DB. Occurrence of cancer pain. J Natl Cancer Inst Monogr. 2004; 32:51–56. [PubMed: 15263041]
45. McHorney CA, Ware JE, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care. 1993; 31:247–263. [PubMed: 8450681]
46. 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; 57:144–151. [PubMed: 17263931]
47. Mitera G, Zeiadin N, Kirou-Mauro A, DeAngelis C, Wong J, Sanjeevan T, Sinclair E, Danjoux C, Barnes E, Tsao M, Sahgal A, Chow E. Retrospective assessment of cancer pain management in an outpatient palliative radiotherapy clinic using the Pain Management Index. J Pain Symptom Manage. 2010; 39:259–267. [PubMed: 20152589]
48. Ngamkham S, Vincent C, Finnegan L, Holden JE, Wang ZJ, Wilkie DJ. The McGill Pain Questionnaire as a multidimensional measure in people with cancer: An integrative review. Pain Manag Nurs (online). 2011
49. 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]
50. Russell PB, Aveyard SC, Oxenham DR. An assessment of methods used to evaluate the adequacy of cancer pain management. J Pain Symptom Manage. 2006; 32:581–588. [PubMed: 17157760]
51. Tavoli A, Montazeri A, Roshan R, Tavoli Z, Melyani M. Depression and quality of life in cancer patients with and without pain: The role of pain beliefs. BMC Cancer. 2008; 8:177. [PubMed: 18570676]
52. Turk DC, Dworkin RH, McDermott MP, Bellamy N, Burke LB, Chandler JM, Cleeland CS, Cowan P, Dimitrova R, Farrar JT, Hertz S, Heyse JF, Iyengar S, Jadad AR, Jay GW, Jermano JA, Katz NP, Manning DC, Martin S, Max MB, McGrath P, McQuay HJ, Quessy S, Rappaport BA, Revicki DA, Rothman M, Stauffer JW, Svensson O, White RE, Witter J. Analyzing multiple endpoints in clinical trials of pain treatments: IMMPACT recommendations. Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials. Pain. 2008; 139:485–493. [PubMed: 18706763]
53. Turk DC, Dworkin RH, Revicki D, Harding G, Burke LB, Cella D, Cleeland CS, Cowan P, Farrar JT, Hertz S, Max MB, Rappaport BA. Identifying important outcome domains for chronic pain clinical trials: An IMMPACT survey of people with pain. Pain. 2008; 137:276–285. [PubMed: 17937976]
54. van den Beuken-van Everdingen MH, de Rijke JM, Kessels AG, Schouten HC, van KM, Patijn J. Prevalence of pain in patients with cancer: A systematic review of the past 40 years. Ann Oncol. 2007; 18:1437–1449. [PubMed: 17355955]
55. Von Korff M, Ormel J, Keefe FJ, Dworkin SF. Grading the severity of chronic pain. Pain. 1992; 50:133–149. [PubMed: 1408309]
56. Ware JE, Gandek B. The SF-36 Health Survey: Development and use in mental health research and the IQOLA Project. Int J Ment Health. 1994; 23:49–73.
57. Wells N, Murphy B, Wujcik D, Johnson R. Pain-related distress and interference with daily life of ambulatory patients with cancer with pain. Oncol Nurs Forum. 2003; 30:977–986. [PubMed: 14603355]
58. Wright JG, Young NL. A comparison of different indices of responsiveness. J Clin Epidemiol. 1997; 50:239–246. [PubMed: 9120522]
59. Zaza C, Baine N. Cancer pain and psychosocial factors: A critical review of the literature. J Pain Symptom Manage. 2002; 24:526–542. [PubMed: 12547052]

## Table 1

Baseline Characteristics of 274 Study Participants With Cancer-Related Pain

| Baseline Characteristics | Mean (SD) |
|---|---|
| Age | 58.08 (10.54) |
| Socioeconomic Disadvantage (SED) index* | 1.30 (.98) |
| Medical comorbidity (no. of diseases) | 2.09 (1.68) |
| BPI pain severity score (0–10 scale) | 5.22 (1.82) |
| BPI pain interference score (0–10 scale) | 5.65 (2.56) |
| BPI total score (0–10 scale) | 5.48 (2.08) |
| PEG total score (0–10 scale) | 5.87 (2.24) |
| SF-36 pain score (0–100 scale) | 31.7 (17.8) |
| | N (%) |
| Group | |
| Intervention | 137 (50.0) |
| Control | 137 (50.0) |
| Symptom | |
| Pain | 96 (35.0) |
| Pain and depression | 178 (65.0) |
| Gender | |
| Male | 93 (33.9) |
| Female | 181 (66.1) |
| Race | |
| White | 212 (77.4) |
| Black and others† | 62 (22.6) |
| Marital status | |
| Married | 130 (47.4) |
| Unmarried | 144 (52.6) |
| Type of cancer | |
| Breast | 70 (25.6) |
| Lung | 53 (19.3) |
| Gastrointestinal | 51 (18.6) |
| Lymphoma/Hematological | 40 (14.6) |
| Genitourinary | 27 (9.9) |
| Other | 33 (12.0) |
| Phase of cancer | |
| Newly diagnosed | 104 (38.0) |
| Maintenance or disease free | 110 (40.2) |
| Recurrent or progressive | 60 (21.9) |

*The SED index included variables of education (“less than high school” = 1 point), employment (“unemployed” or “unable to work due to health or disability” = 1 point), and income (“not enough” = 1 point).

†Other races constituted only 1.8% (n = 5) of the sample.

**Table 2**

Correlations at Baseline and 3 Months Among the Pain Scales in INCPAD*

| Pain Scale | BPI Total | BPI Severity | BPI Interference | PEG | SF-36 Pain |
|---|---|---|---|---|---|
| BPI total | — | .76 | .96 | .93 | −.64 |
| | | .87 | .97 | .97 | −.81 |
| BPI severity | | — | .55 | .69 | −.42 |
| | | | .71 | .81 | −.71 |
| BPI interference | | | — | .90 | −.63 |
| | | | | .95 | −.76 |
| PEG | | | | — | −.63 |
| | | | | | −.75 |
| SF-36 pain | | | | | — |

NOTE. The first correlation for each measure is the baseline correlation, and the second correlation is the 3-month correlation.

*All correlations are significant at $P < .0001$.

**Table 3**

Change Scores and Standardized Response Means for Pain Measures According to Global Outcome Category at 3 Months of Better (n = 51), Same (n = 72), or Worse (n = 107)

| Pain Measure Global Category | Baseline Score (SD) | 3-Month Score (SD) | Change Score (SD)* | *P* value† | SRM‡ (95% CI) |
|---|---|---|---|---|---|
| BPI severity | | | | | |
| Better | 4.9 (1.8) | 2.5 (2.0) | 2.5 (2.2) | <.0001 | 1.13 (.94, 1.32) |
| Same | 5.1 (1.9) | 4.4 (2.1) | .7 (1.8) | — | .41 (.17, .64) |
| Worse | 5.6 (1.7) | 6.2 (1.8) | −.7 (1.8) | <.0001 | −.38 (−.66, −.10) |
| BPI interference | | | | | |
| Better | 5.2 (2.6) | 2.8 (2.5) | 2.5 (2.8) | .0004 | .91 (.70, 1.11) |
| Same | 5.5 (2.6) | 4.4 (2.4) | 1.0 (2.6) | — | .37 (.13, .61) |
| Worse | 6.2 (2.4) | 6.8 (2.4) | −.6 (2.4) | .001 | −.24 (−.52, .04) |
| BPI total | | | | | |
| Better | 5.1 (2.0) | 2.5 (2.1) | 2.6 (2.4) | <.0001 | 1.10 (.92, 1.30) |
| Same | 5.3 (2.1) | 4.4 (2.1) | .9 (2.1) | — | .42 (.18, .65) |
| Worse | 6.0 (2.4) | 6.6 (1.9) | −.6 (1.8) | <.0001 | −.33 (−.61, −.05) |
| PEG | | | | | |
| Better | 5.6 (2.3) | 2.9 (2.3) | 2.7 (2.5) | <.0001 | 1.08 (.87, 1.28) |
| Same | 5.7 (2.2) | 4.5 (2.3) | 1.1 (2.5) | — | .43 (.19, .67) |
| Worse | 6.4 (2.3) | 6.8 (2.2) | −.4 (2.5) | .001 | −.18 (−.46, .10) |
| SF-36 pain | | | | | |
| Better | 35.8 (17.6) | 54.9 (22.7) | −19.1 (23.1) | <.0001 | −.83 (−1.02, −.64) |
| Same | 34.1 (17.7) | 39.6 (17.7) | −5.6 (19.2) | — | −.29 (−.53, −.05) |
| Worse | 23.8 (15.6) | 23.2 (16.3) | .6 (15.9) | .063 | .04 (−.24, .32) |

*Change score = 3 month score – baseline score. For all measures except SF-36 pain, higher scores indicate worse pain and positive change scores indicate improvement.

†P values are from t-tests of change score in “worse” or “better” group compared with change score in “same” group for each measure.

‡SRM = change score/SD of change score.

## Table 4

Area Under the ROC Curve (AUC) for Pain Measures in Detecting Improvement*

| | Accuracy For Detecting Any Improvement† | | Accuracy For Detecting Moderate Improvement† | |
|---|---|---|---|---|
| | AUC‡ | SE | AUC‡ | SE |
| BPI severity | .783 | .030 | .809 | .032 |
| BPI interference | .726 | .035 | .732 | .040 |
| BPI total | .786 | .030 | .801 | .033 |
| PEG | .742 | .034 | .754 | .038 |
| SF-36 pain | .729 | .033 | .735 | .037 |

*Calculated in the 230 study participants who completed both baseline and 3-month assessments.

†Any improvement ≥"a little better"; moderate improvement ≥"moderately better".

‡AUC is probability of correctly discriminating between patients who have improved and those who have not. There were no statistically significant differences in pairwise comparisons of the AUCs between any of the pain measures in detecting either any improvement or moderate improvement. Likewise, within each measure, its AUC for detecting any improvement was not statistically different from its AUC for detecting moderate improvement.

**Table 5**

Change Scores* and Standardized Effect Size (SES)† of Intervention Among Trial Participants Overall and According to Depression Status

| | Overall (N = 230) | | | Nondepressed (N = 88) | | | Depressed (N = 142) | | |
|---|---|---|---|---|---|---|---|---|---|
| **Pain Scale** | **Intervention Change (SD)** | **Control Change (SD)** | **SES** | **Intervention Change (SD)** | **Control Change (SD)** | **SES** | **Intervention Change (SD)** | **Control Change (SD)** | **SES** |
| BPI severity | 1.9 (2.5) | .5 (2.0) | .58 | 1.9 (2.5) | .1 (1.9) | .76 | 1.9 (2.5) | .8 (2.0) | .47 |
| BPI interference | 2.0 (2.9) | .6 (2.8) | .46 | 1.2 (2.8) | −.6 (2.7) | .64 | 2.5 (2.8) | 1.3 (2.6) | .41 |
| BPI total | 2.1 (2.5) | .6 (2.3) | .58 | 1.5 (2.8) | −.3 (2.7) | .75 | 2.5 (2.6) | 1.2 (2.2) | .51 |
| PEG | 2.2 (2.8) | .7 (2.6) | .52 | 1.7 (3.0) | −.2 (2.8) | .62 | 2.5 (2.6) | 1.3 (2.4) | .47 |
| SF-36 pain | −15.2 (24.6) | −5.8 (17.9) | −.43 | −11.1 (26.4) | −3.5 (14.1) | −.38 | −17.4 (23.4) | −7.3 (19.8) | −.45 |

*Change = 3 month score – baseline score.

†SES = (Intervention group change – control group change)/SD of pooled change score.