Published in final edited form as:
Pain Med. 2010 March ; 11(3): 337–346. doi:10.1111/j.1526-4637.2009.00774.x.

The Brief Pain Inventory and its "Pain at its Worst in the last 24 Hours" Item: Clinical Trial Endpoint Considerations*

Thomas M. Atkinson, PhD[1,2], **Tito R. Mendoza, PhD, MS**[3], **Laura Sit, BA**[2,4], **Steven Passik, PhD**[1], **Howard I. Scher, MD**[4], **Charles Cleeland, PhD**[3], and **Ethan Basch, MD, MSc**[2,4]

[1]Department of Psychiatry and Behavioral Sciences, Memorial Sloan-Kettering Cancer Center, New York, New York

[2]Health Outcomes Research Group, Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, New York

[3]Department of Symptom Research, The University of Texas M. D. Anderson Cancer Center, Houston, Texas, USA

[4]Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, New York

Abstract

Context—In 2006, the United States Food and Drug Administration (FDA) released a draft Guidance for Industry on the use of Patient-Reported Outcomes (PRO) Measures in Medical Product Development to Support Labeling Claims. This draft guidance outlines psychometric aspects that should be considered when designing a PRO measure, including conceptual framework, content validity, construct validity, reliability, and the ability to detect clinically meaningful score changes. When finalized, it may provide a blueprint for evaluations of PRO measures which can be considered by sponsors and investigators involved in PRO research and drug registration trials.

Objective—In this review we examine the short form of the Brief Pain Inventory (BPI) and particularly the "pain at its worst in the last 24 hours" item in the context of the FDA draft guidance, to assess its utility in clinical trials that include pain as a PRO endpoint.

Results and Conclusions—After a systematic evaluation of the psychometric aspects of the BPI, we conclude that the BPI and its "pain at its worst in the last 24 hours" item generically satisfy most key recommendations outlined in the draft guidance for assessing a pain-reduction treatment effect. Nonetheless, when the BPI is being considered for assessment of pain endpoints in a registration trial, sponsors and investigators should consult with the appropriate FDA division early during research design to discuss whether there is sufficient precedent to use the instrument in the population of interest or whether additional evaluations of measurement properties are advisable.

Keywords

Pain Measurement; Patient Outcome Assessment; United States Food and Drug Administration; Drug Labeling

* *Reprint requests to:* Thomas M. Atkinson, PhD, Memorial Sloan-Kettering Cancer Center, Health Outcomes Research Group, 307 E. 63rd St. 2nd Floor, New York, NY 10065, USA. Phone: 646-735-8132. Fax: 646-735-0032. atkinsot@mskcc.org.

Introduction

Pain is an important characteristic of many medical conditions and is widely used by clinical trialists and pharmaceutical sponsors as a patient reported outcome (PRO) endpoint in registration track research [1]. Although a number of pain instruments have been developed and used in clinical trials [2-3], the short form of the Brief Pain Inventory (BPI) [4-6], and particularly the BPI's single “pain at its worst in the last 24 hours” item (also referred to in prior publications as the “11-point pain intensity numerical rating scale” or “11-point NRS” item) [7-8], are frequently used. The United States Food and Drug Administration (FDA) released a draft Guidance for Industry in 2006 on the use of PRO Measures in Medical Product Development to Support Labeling Claims [9]. This draft guidance contains specific recommendations for defining the conceptual framework of a PRO instrument, as well as the instrument's reliability, validity, and ability to detect clinically relevant changes. Although this guidance should facilitate use of PRO endpoints by providing a blueprint for instrument developers and trialists, it is important to assure that existing instruments and endpoint models can fulfill this new regulatory standard. This review evaluates the psychometric properties of the BPI and its “pain at its worst in the last 24 hours” item in terms of the FDA draft guidance, first by summarizing specific methodologic recommendations of the guidance and then by considering measurement properties of the BPI in light of each.

A PRO is defined in the draft guidance as a measurement of any portion of a given patient's health status that is derived directly from the patient, in the absence of interpretation of the response by a physician or any other individual [9]. There are a number of conceptual challenges related to PROs, including the identification of outcomes that are considered important to patients themselves, as well as the issue of accurately and reliably capturing a given patient's unique perception of multidimensional health-related information [10]. These challenges, historical inconsistencies in methods used to develop and administer PRO measures in trials [11], and the importance of direct symptom measures in judging treatment effects were incentives leading to the FDA guidance [9]. We will review the FDA recommendations that pertain to the assessment of pain endpoints, then consider whether the measurement properties of the BPI fulfill the specifications in the draft guidance. These FDA recommendations are summarized in Table 1.

FDA Recommendations

FDA Recommendations for the Conceptual Framework of a PRO Instrument

The FDA recommends that the conceptual framework of a PRO instrument be confirmed using empirical evidence during instrument development [9]. The conceptual framework should include an explicit description of the purported relationship between the instrument's concepts, domains, and items. The guidance notes that a single concept can be measured with a single item, multiple items, or multiple items for multiple domains of that particular concept. Responses to these items must be clear and appropriate, with example response option formats provided in the draft guidance: visual analog scale (VAS), anchored or categorized VAS, Likert scale, rating scale, recording of events as they occur, pictorial scale, or checklist.

FDA Recommendations for the Reliability of a PRO instrument

The FDA draft guidance recommends that a PRO instrument must demonstrate test-retest reliability, as well as internal consistency. Test-retest reliability refers to the degree to which obtained scores remain stable over a given time period, with no expected change in the target concept. A test-retest interval can vary based on the concept being assessed and the practical needs of the researcher; this time period must be long enough to minimize practice/learning effects. Internal consistency is a measure of the degree to which items within a

scale are measuring the same hypothetical concept, as well as the degree of relatedness among these particular items. While the guidance has noted the value of internal consistency, it also notes that this measure, in the absence of test-retest reliability, may not be sufficient for clinical trial purposes. Acceptable reliability estimates for both internal consistency and test-retest reliability have been suggested as an alpha value greater than or equal to 0.75 [9].

FDA Recommendations for the Content Validity of a PRO instrument

Items in a PRO instrument must measure relevant and important (i.e., clinically meaningful) aspects of each concept or domain contained in the instrument. These concepts and domains must encompass what patients consider the most important and comprehensive outcomes of the condition and its therapy. In addition, evidence should be provided by specific documentation of patient input in item generation as well as evaluation of patient understanding through cognitive interviewing. The guidance recommends that saturation be reached – the point at which no new relevant or important information emerges and collecting additional data will not add to the understanding of how patients perceive the concept of interest and the items in the questionnaire.

FDA Recommendations for the Construct Validity of a PRO instrument

The guidance states that documented relationships between results obtained using the instrument and results obtained using other measures must be consistent with pre-existing hypotheses concerning these relationships. In addition, the PRO instrument must be demonstrated to have the ability to differentiate between clinically distinct groups in the population of interest.

FDA Recommendations for the Clinical Relevance of Score Changes of a PRO instrument

A PRO instrument should be equally sensitive to gains and losses in health status, as well as being sensitive to change at all points within the entire range expected for the clinical trial population. When change is expected, change should be detected by the instrument. Pre-existing subgroup differences in ability to detect change must be accounted for when interpreting results. The draft guidance cautions investigators against basing claims on the demonstration of statistical significance and instead recommends examination of a cumulative distribution function (CDF) of responses that characterizes possible treatment effects. The specific score changes that define a responder should be established and agreed upon prior to performing a pivotal trial of a PRO endpoint, both in order to adequately power the study and to establish a clinically meaningful responder definition.

BPI Properties In Relation to the FDA Draft Guidance

BPI Conceptual Framework

The short form of the Brief Pain Inventory (BPI) was originally developed as the Wisconsin Brief Pain Questionnaire and was designed to assess pain and its impact in cancer. It has been used to assess other pain conditions as well [5-6]. Designed by the Pain Research Group at the University of Wisconsin – Madison, which was also the World Health Organization (WHO) Collaborating Center for Symptom Evaluation in Cancer Care, the BPI was developed as an instrument that would quantify and assess pain using patient self-reported information [12]. Based on the premise that pain is multidimensional [13], the BPI was designed to measure directly two key aspects of pain: sensory pain and reactive pain, as reported by the subjects. The sensory pain dimension is characterized by pain intensity and is measured in four items of the BPI using a numeric rating scale (NRS). Notably, the NRS is one of the response options advocated in the FDA draft guidance [9]. The NRS utilizes a linear scale from 0-10, with 0 representing “no pain” and 10 being indicative of “pain as bad

as you can imagine". Patients are asked to rate their pain along the number continuum for items that query their pain: (1) at its worst in the last 24 hours, (2) pain at its least in the last 24 hours, (3) average pain, and (4) pain right now. Use of the 11-point NRS of pain intensity is consistent with published recommendations for core outcome measures in clinical chronic pain trials [13].

The reactive pain component is measured in terms of the degree to which pain interferes with everyday patient function. The NRS for this dimension also uses a 0-10 scale, with 0 representing "does not interfere" and 10 indicating "completely interferes." Patients are asked to quantify the degree to which pain interferes with functioning along the 11-point scale on seven items (i.e., general activity, mood, walking ability, normal work, relations with other people, sleep, and enjoyment of life). The published psychometric properties of the BPI refer to the severity and interference items measuring the sensory and reactive components of pain, respectively.

The BPI also contains supplemental items that allow a patient to indicate treatments or medications they are receiving to treat their pain, the percentage of relief obtained in the past 24 hours from the treatments or medications, and the anatomical location of their pain on a body diagram.

BPI Reliability

Internal consistency of the BPI has been demonstrated in a series of studies. Cronbach alpha coefficients for the pain intensity scale have ranged from 0.78 to 0.96 [14-27]. For the pain interference scale, Cronbach alpha coefficients have been shown to range from 0.83 and 0.95 [14-28]. Throughout BPI validation, the item "pain at its worst in the last 24 hours" has consistently shown the highest degree of internal consistency, with Cronbach alpha coefficients ranging from 0.77 to 0.90 [17, 21-23, 25, 29]. Table 2 is a display of the internal consistency coefficients reported in prior validation studies.

For test-retest reliability, results obtained have shown a wide range of variability. For example, during validation of the Spanish version of the BPI, the pain intensity factor was found to have below acceptable test-retest validity (0.62), while the pain interference factor had acceptable reliability (0.77) after an interval of seven days [24]. Within a Taiwanese sample, three items from the BPI pain intensity index were found to have test-retest reliability below the acceptable range when using a one to ten day interval between testing (i.e., least pain, average pain, pain now [23]). Another investigation in a sample of German subjects found the test-retest reliability of the BPI to be at 0.97 after an interval of 30 to 60 minutes [22]. As part of validation in a sample of patients with osteoarthritis, test-retest validity estimates ranged from 0.67 to 0.93 for pain intensity, with pain interference estimates ranging from 0.68 to 0.93 [20].

When examining test-retest reliability, the "pain at its worst in the last 24 hours" item had acceptable reliability during validation of the BPI in German [22] and Taiwanese [23] subjects (0.80 and 0.96, respectively). Although there may be variability among studies, these results demonstrated that the test-retest reliability of the "pain at its worst in the last 24 hours" item is highest when administered over a short time span (i.e., hourly or daily), suggesting these may be the optimal intervals during future drug trials. Table 3 contains a summary of reported test-retest reliability coefficients from German, Spanish, and Taiwanese BPI validation samples.

BPI Content Validity

As part of instrument development, patients with breast, prostate, colon, rectum, or gynecologic cancer ($n = 667$) as well as patients with rheumatoid arthritis ($n = 32$) were

interviewed with questions related to their level of pain intensity and the degree to which pain was interfering with their everyday activities [5-6]. The cancer patients were more likely to attribute their pain to cancer than to unrelated causes, and patients with metastatic cancer, specifically of the breast or prostate, were more likely to report pain than those with non-metastatic cancer. In addition, the BPI has been validated in many languages [17, 21-27, 29-34], indicating its global applicability. During instrument development, “pain at its worst in the last month” was tested as an item rather than “pain at its worst in the last 24 hours.” The “pain at its worst in the last month” item was found to be highly related to pain interference items, consistent with the pattern of relationship between “pain at its worst in the last 24 hours” and pain interference items in subsequent studies [30, 35-36].

BPI Construct Validity

During development of the BPI, items were designed and tested to measure sensory and reactive pain: pain intensity and pain interference, respectively [12]. The test developers validated this pain model during development using exploratory factor analysis (EFA; [12]). Through a series of validation studies, Korean [17], Greek [29], Norwegian [21], German [22], Taiwanese [23], Spanish [24], Italian [25], Japanese [26], French [25], Chinese [33], Filipino [30], Russian [34], Vietnamese [31], and Hindi [27] versions of the BPI also yielded a similar two factor structure (Figure 1). The test developers have provided evidence that the pain interference factor can be further divided into affective (i.e., enjoyment of life, relations with others, and mood) and activity (i.e., walking, general activity, working, and sleep items) sub-components [37]. The two factor (i.e., pain intensity and pain interference) representation was replicated in cardiac patients [18], in patients with non-cancer pain [14-15], in patients with Multiple Sclerosis [28], in osteoarthritic patients [20], and in surgical patients [38]. The “pain at its worst in the last 24 hours” item was an indicator of pain intensity during validation studies, with factor loading coefficients ranging from 0.34 to 0.90 [17-18, 21-23, 25-26, 28-29]. Table 4 shows reported factor loading coefficients across the spectrum of validation studies. As recommended by the FDA, sponsors should assure that any instrument planned for use in a registration trial has been evaluated previously in a population or condition pertinent to the planned research, or the instrument should be evaluated as part of the research.

Approaches for Determining Clinically Relevant BPI Score Changes

To address individual patient variation while determining clinically meaningful outcomes for a PRO pain endpoint, investigators should consider responder analyses [39]. In drug trials, a responder analysis allows for the evaluation of the pain endpoint in terms of a drug's perceived effectiveness for each patient. Using this approach in a study to determine the impact of a self-care intervention on cancer pain management [40], the BPI was administered to patients in standard care and in intervention groups at baseline and then six weeks later. Response categories were prespecified and used to divide the intervention group into three subgroups, with clinically significant differences found among these subgroups on the BPI interference scale, as well as for mood and quality of life measures. Cleeland et al. (2005) also used a responder analysis to demonstrate the effectiveness of an analgesic protocol in controlling pain. Responders were defined as patients who changed group membership, defined by their pain levels post-intervention compared to baseline. For example, patients who reported moderate or severe pain (ratings of 5-10 on a 0-10 numeric rating scale) at baseline but then reported no pain or mild pain (ratings of 0-4) after the intervention were considered responders.

Farrar and colleagues proposed the use of an analytical technique known as the cumulative proportion of responder's analysis (CPRA) in an attempt to further enhance understanding of PRO data [41]. The CPRA is a descriptive technique that displays a full range of clinical

results in graph form, by treatment group, with all possible patient responder levels along the *x*-axis and the *y*-axis representing the proportion of patient responders at each level. While this technique does not include statistically significant change, it does illustrate possible trends that can be further explored by investigators and clinicians and is similar to the CDF suggested as part of FDA guidance [9]. As use of the CPRA has been demonstrated in studies of pain [41], this technique could be applied in studies of the BPI.

Guyatt and colleagues have described two approaches for determining the clinical significance of health status measures [42]. An anchor-based method is an investigation of the assessment instrument as it relates to an independent measure, while distribution-based methods focus on the actual distribution of resulting scores [42]. For distribution-based methods, a variety of statistical criteria have been used in studies of PRO measures, depending on how the lowest degree of clinically relevant change, or minimally important difference is defined(MID [43]).

A meta-analysis of health-related quality of life studies yielded evidence that, on average, a 0.5 standard deviation (SD) change is indicative of the MID [44]. In an investigation comparing the MID levels of 0.2 SD, 0.5 SD, and 1 standard error of measurement (SE), no differences were found between levels for the number of metastatic breast cancer patients improving after supportive-expressive group therapy. It was concluded that the MID should be derived from patient opinion-defined clinical anchors [45] in a representative sample [46]. Evidence supporting the use of 1 SE was provided in a study of change from baseline in a group of patients with advanced non-small cell lung cancer [47].

In an investigation of chronic pain intensity in individuals with painful diabetic neuropathy, postherpetic neuralgia, lower back pain, fibromyalgia, or osteoarthritis measured with the 11-point pain intensity NRS (i.e., the “pain at its worst” item), it was found through the analysis of receiver operating characteristic (ROC) curves that a raw point decrease of 1.74 (27.9%) was indicative of clinically meaningful improvement [48]. Another study of pain intensity using an 11-point NRS in patients with osteoarthritis, rheumatoid arthritis, or ankylosing spondylitis found through the use of the ROC method that decreases in 2 points (33%) could differentiate between patients describing their pain as “slightly better” and “much better” [49]. In a study of patients with spinal cord injury or amputation, decreases from pretreatment to posttreatment of 1.8 points (36%) on an 11-point NRS were found to be evidence of clinically meaningful changes in pain intensity [7]. The Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) group summarized the results of these three studies which suggest that changes of 2 points (30-36%) along the NRS are indications of much better, much improved or meaningful decreases in chronic pain, with a 4 point (50%) change indicating “very much improved” pain [8]. IMMPACT cautioned that further research is necessary before a consensus can be reached on the clinical meaningfulness of score changes.

Cutpoint Analysis in the BPI

For data analysis and interpretation purposes, it is sometimes advantageous to reduce the 0-10 NRS into fewer categories, such as mild, moderate, and severe. Specific cutpoint scores for these delineations can be determined analytically through the use of cutpoint analyses by examining the relationship between pain severity and pain interference, and then evaluating for clinically meaningful score changes over a series of given time points [50]. Given and colleagues have established that cutpoints can vary by symptom when categorizing patient responses as mild, moderate, or severe [51]. Serlin and colleagues conducted such an evaluation in patients with cancer pain from the United States, France, the Philippines, and China to determine whether such pain could be classified into categories based upon the relationship between pain severity and pain intensity [30]. After administering the BPI to

this sample of 1897 individuals, the investigators first established that when compared with the three other pain intensity items (i.e., “pain at its least in the last 24 hours”, “pain now”, and “average pain”), the “pain at its worst in the last 24 hours” item had the strongest relationship with the pain interference scale (coefficient alpha range 0.76 to 0.85). The “pain at its worst in the last 24 hours” item was then used to stratify patients along the NRS into three pain cutpoints based on the established criteria the cutpoints were determined to be 1-4 (mild), 5-6 (moderate), and 7-10 (severe). These cutpoint determinations were consistent with results from a study that employed the use of interviews with community members to determine the relationship between an 11-point NRS and mild, moderate, and severe pain cutpoints [52]. Findings were similar in two additional studies of individuals with symptomatic bone metastases [35, 50]. Notably, a lower cutpoint of 4 was suggested by evidence from Farrar et al. finding that a 2-point reduction in pain from a baseline of 4 was considered clinically meaningful by patients [48]. Therefore, a low end cutpoint level of either 4 or 5 on the 11-point NRS appears to have a reasonable basis in patient-based studies.

Wang et al. reported that there were no significant differences in functional health and well being of patients with no pain/mild pain compared to those with moderate/severe pain [33]. Cleeland et al. determined using multidimensional scaling in a multicultural sample of cancer patients that moderate and severe levels of pain are most strongly associated with pain interference [37]. Investigators used these criteria for patient inclusion in a study of short-course versus long-course radiotherapy in individuals with painful bone metastases and determined that those in the short-course radiotherapy group had less acute toxicity than those in the long-course group [36].

Discussion

The recently released FDA draft Guidance for the use of PRO Measures in Medical Product Development to Support Labeling Claims focuses on the need for measurement instruments that are psychometrically sound in terms of conceptual framework, content validity, construct validity, reliability, and the ability to detect clinically meaningful score changes. This review demonstrates that the short form of the Brief Pain Inventory and particularly its “pain at its worst in the last 24 hours” (11-point pain intensity NRS) item, despite being developed over two decades ago, can fulfill many of the key recommendations of the FDA draft guidance, and support its use in future clinical trials of pain as a PRO endpoint.

For clinical trialists, there are a number of future directions for examining the BPI as it relates to the FDA draft guidance for PRO measures. Optimal recall periods may differ depending on the population of interest and the setting and may merit individual evaluation in varied populations or settings of interest. In addition, as with all PRO endpoints in blinded controlled treatment trials, the possibility of unintentional unblinding should be addressed. For example, a single item can be administered at time of disenrollment that queries patients on which study arm they believed they were enrolled. The BPI should be validated using alternative methods of administration, such as tablet laptop computers, interactive voice response systems, or personal digital assistant devices [53-54]. In this paper we have highlighted the “pain at its worst in the last 24 hours” item, but other measures of pain intensity, such as “average pain” and “pain right now” have been used in clinical trials and are included in the BPI. These items have also demonstrated measurement properties consistent with the FDA draft guidance [15, 17, 20-27, 29, 31, 34], and are shown in Tables 2, 3, and 4.

To explore the underlying factor structure of the BPI and to provide additional evidence for construct validity, future investigations using this instrument should utilize the structural

equation modeling technique of confirmatory factor analysis (CFA). Like EFA, CFA allows a researcher to reduce the number of variables used based on commonalities within the data. CFA differs from EFA in that it requires an *a priori* hypothesis of the pattern of relationships between variables [55]. Since the BPI was developed based upon a two factor structure (i.e., pain intensity and pain interference), it follows that CFA be used to test construct validity rather than EFA, which does not require a pre-specified factor structure. CFA can also be used to statistically compare the factor structure of two or more groups (e.g. languages or disease conditions), statistically investigate alternative models (e.g., 2 vs 3 factors structure), or develop other models by specifying correlated measurement errors. In addition, while data from validation studies of the BPI in various languages are displayed in Tables 2-4, investigators considering administration of the BPI in additional languages are advised to confirm that validation studies have been considered in the languages of interest; else such evaluations should be conducted.

While this paper is an overview of uses for the BPI, an essential design consideration is the effects of ongoing analgesic use on pain measurements in the setting of treatment trials. For example, if a cancer drug is anticipated to improve pain related to bone metastases, but a patient is also taking a narcotic analgesic, the relative impact on pain of the cancer drug versus the analgesic may not be clear. Therefore, separate collection of analgesic use and standardization of analgesic dosing (e.g., through the use of an analgesic protocol [56]) may be advisable.

In summary, the measurement properties of the BPI and the item “pain at its worst in the last 24 hours” can fulfill the expectations of the FDA draft guidance for PRO instruments in terms of conceptual framework, reliability, construct validity, and ability to detect clinically meaningful change. For research intending to result in FDA drug approval or labeling, it is essential to assure that content and construct validity, as well as evaluations of clinically meaningful changes, have been previously established in a patient population and/or condition pertinent to the planned study; or that confirmation of these measurement properties be included as a design component of the planned research. In addition, discussion of the planned PRO and its role in the overall endpoint model is recommended with the appropriate FDA division early in the drug development process.

Acknowledgments

The authors wish to thank Robert Kane, Laurie B. Burke, and Ann Marie Trentacosti for their helpful feedback during the preparation of this manuscript.

References

1. Turk DC, Dworkin RH, McDermott MP, et al. Analyzing multiple endpoints in clinical trials of pain treatments: IMMPACT recommendations. Pain. 2008; 139:485–493. [PubMed: 18706763]
2. Tannock IF, de Wit R, Berry WR, et al. Docetaxel plus Prednisone or Mitoxantrone plus Prednisone for Advanced Prostate Cancer. N Engl J Med. 2004; 351(15):1502–1512. [PubMed: 15470213]
3. Cella D, Nichol MB, Eton D, Nelson JB, Mulani P. Estimating clinically meaningful changes for the Functional Assessment of Cancer Therapy - Prostate: Results from a clinical trial of patients with metastatic hormone-refractory prostate cancer. Value Health. 2009; 12(1):124–129. [PubMed: 18647260]
4. Cleeland CS. Research in cancer pain. What we know and what we need to know. Cancer. 1991; 67(suppl 3):823–827. [PubMed: 1986852]
5. Daut RL, Cleeland CS. The prevalence and severity of pain in cancer. Cancer. 1982; 50:1913–1918. [PubMed: 7116316]
6. Daut RL, Cleeland CS, Flanery RC. Development of the Wisconsin Brief Pain Questionnaire to assess pain in cancer and other diseases. Pain. 1983; 17:197–210. [PubMed: 6646795]

7. Hanley MA, Jensen MP, Ehde DM, et al. Clinically significant changes in pain intensity ratings in persons with spinal cord injury or amputation. Clin J Pain. 2006; 22:25–31. [PubMed: 16340590]
8. 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(2):105–121. [PubMed: 18055266]
9. US Food and Drug Administration. [Accessed December 9, 2009] Guidance for Industry. Patient-reported outcome measures. 2009. Use in medical development to support labeling claimsAvailable from: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf
10. Rothman ML, Beltran P, Cappeileri JC, Lipscomb J, Teschendorf B. Patient-reported outcomes: Conceptual issues. Value Health. 2007; 10(Supplement 2):S66–S75. [PubMed: 17995476]
11. Revicki DA. Consistent patient-reported outcomes. Value Health. 2002; 5(4):295–296. [PubMed: 12102691]
12. Cleeland CS, Ryan KM. Pain assessment: Global use of the Brief Pain Inventory. Ann Acad Med Singapore. 1994; 23(2):129–138. [PubMed: 8080219]
13. Dworkin RH, Turk DC, Farrar JT, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain. 2005; 113:9–19. [PubMed: 15621359]
14. Tan G, Jensen MP, Thornby JI, Shanti BF. Validation of the Brief Pain Inventory for Chronic Nonmalignant Pain. J Pain. 2004; 5(2):133–137. [PubMed: 15042521]
15. Keller S, Bann CM, Dodd SL, Schein J, Mendoza TR, Cleeland CS. Validity of the Brief Pain Inventory for use in documenting the outcomes of patients with noncancer pain. Clin J Pain. 2004; 20(5):309–318. [PubMed: 15322437]
16. Tittle MB, McMillian SC, Hagan S. Validating the Brief Pain Inventory for use with surgical patients with cancer. Oncol Nurs Forum. 2003; 30(2):325–330. [PubMed: 12692666]
17. Yun YH, Mendoza TR, Heo DS, et al. Development of a cancer pain assessment tool in Korea: A validation of a Korean version of the Brief Pain Inventory. Oncology. 2004; 66:439–444. [PubMed: 15452372]
18. Gjeilo KH, Stenseth R, Wahba A, Lydersen S, Klepstad P. Validation of the Brief Pain Inventory in patients six months after cardiac surgery. J Pain Symptom Manage. 2007; 34(6):648–656. [PubMed: 17629665]
19. Williams VSL, Smith MY, Fehnel SE. The validity and utility of the BPI interference measures for evaluating the impact of osteoarthritic pain. J Pain Symptom Manage. 2006; 31(1):48–57. [PubMed: 16442482]
20. Mendoza T, Mayne T, Rublee D, Cleeland C. Reliability and validity of a modified Brief Pain Inventory short form in patients with osteoarthritis. Eur J Pain. 2006; 10:353–361. [PubMed: 16051509]
21. Klepstad P, Loge JH, Borchgrevink PC, Mendoza TR, Cleeland CS, Kaasa S. The Norwegian Brief Pain Inventory questionnaire: Translation and validation in cancer pain patients. J Pain Symptom Manage. 2002; 24(5):517–525. [PubMed: 12547051]
22. Radbruch L, Loick G, Kiencke P, et al. Validation of the German version of the Brief Pain Inventory. J Pain Symptom Manage. 1999; 18(3):180–187. [PubMed: 10517039]
23. Ger LP, Ho ST, Sun WZ, Wang MS, Cleeland CS. Validation of the Brief Pain Inventory in a Taiwanese population. J Pain. 1999; 9(2):105–121.
24. Badia X, Muriel C, Gracia A, et al. Validación española del cuestionario Brief Pain Inventory en pacientes con dolor de causa neoplásica. Med Clin (Barc). 2003; 120(2):52–59. [PubMed: 12570914]
25. Caraceni A, Mendoza TR, Mencaglia E, et al. A validation study of an Italian version of the Brief Pain Inventory (Breve Questionario Per La Valutazione Del Dolore). Pain. 1996; 65:87–92. [PubMed: 8826494]
26. Uki J, Mendoza T, Cleeland CS, Nakamura Y, Takeda F. A brief cancer pain assessment tool in Japanese: The utility of the Japanese Brief Pain Inventory – BPI-J. J Pain Symptom Manage. 1998; 16(6):364–373. [PubMed: 9879161]

27. Saxena A, Mendoza T, Cleeland CS. The assessment of cancer pain in North India: The validation of the Hindi Brief Pain Inventory – BPI-H. J Pain Symptom Manage. 1999; 18(3):27–41. [PubMed: 10439570]
28. Osborne TL, Raichle KA, Jensen MP, Ehde DM, Kraft G. The reliability and validity of pain interference measures in persons with Multiple Sclerosis. J Pain Symptom Manage. 2006; 32(3): 217–229. [PubMed: 16939846]
29. Mystakidou K, Mendoza T, Tsilika E, et al. Greek Brief Pain Inventory: Validation and utility in cancer pain. Oncology. 2001; 60:35–42. [PubMed: 11150906]
30. Serlin RC, Mendoza TR, Nakamura Y, Edwards KR, Cleeland CS. When is cancer pain mild, moderate or severe? Grading pain severity by its interference with function. Pain. 1995; 61:277–284. [PubMed: 7659438]
31. Cleeland CS, Ladinsky JL, Serlin RC, Nugyen CT. Multidimensional measurement of cancer pain: Comparison of U.S. and Vietnamese patients. J Pain Symptom Manage. 1988; 3:23–27. [PubMed: 3351345]
32. Larue F, Colleau SM, Brasseur L, Cleeland CS. Multicentre study of cancer pain and its treatment in France. BMJ. 1995; 310:1034–1037. [PubMed: 7728056]
33. Wang XS, Mendoza TR, Gao SZ, Cleeland CS. The Chinese version of the Brief Pain Inventory (BPI-C): Its development and use in a study of cancer pain. Pain. 1996; 67:407–416. [PubMed: 8951936]
34. Kalyadina SA, Ionova TI, Ivanova MO, et al. Russian Brief Pain Inventory: Validation and application in cancer pain. J Pain Symptom Manage. 2008; 35(1):95–102. [PubMed: 17980999]
35. Harris K, Li K, Chow E. Worst, average or current pain in the Brief Pain Inventory: Which should be used to calculate the response to palliative radiotherapy in patients with bone metastases? J Clin Oncol. 2007; 19(7):523–527.
36. Hartsell WF, Scott CB, Bruner DW, et al. Randomized trial of short- versus long-course radiotherapy for palliation of painful bone metastases. J Natl Cancer Inst. 2005; 97(11):798–804. [PubMed: 15928300]
37. 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]
38. Mendoza TR, Chen C, Brugger A, et al. The utility and validity of the modified Brief Pain Inventory in a multiple-dose postoperative analgesic trial. Clin J Pain. 2004; 20(5):357–362. [PubMed: 15322443]
39. Dionne RA, Bartoshuk L, Mogil J, Witter J. Individual responder analyses for pain: Does one pain scale fit all? Trends Pharmacol Sci. 2005; 26(3):125–130. [PubMed: 15749157]
40. Miaskowski C, Dodd M, West C, et al. The use of a responder analysis to identify differences in patient outcomes following a self-care intervention to improve cancer pain management. Pain. 2007; 129:55–63. [PubMed: 17257753]
41. Farrar JT, Dworkin RH, Mitchell MB. Use of the cumulative proportion of responders analysis graph to present pain data over a range of cut-off points: Making clinical trial data more understandable. J Pain Symptom Manage. 2006; 31(4):369–377. [PubMed: 16632085]
42. Guyatt GH, Osaba D, Wu AW, Wyrwich KW, Norman GR. Methods to explain the clinical significance of health status measures. Mayo Clin Proc. 2002; 77:371–383. [PubMed: 11936935]
43. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials. 1989; 10:407–415. [PubMed: 2691207]
44. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: The remarkable universality of half a standard deviation. Med Care. 2003; 41(5):582–592. [PubMed: 12719681]
45. Lemieux J, Beaton DE, Hogg-Johnson S, Bordeleau LJ, Goodwin PJ. Three methods for minimally important difference: no relationship was found with the net proportion of patients improving. J Clin Epidemiol. 2007; 60:448–455. [PubMed: 17419955]
46. Schünemann HJ, Guyatt GH. Commentary - Goodbye M(C)ID! Hello MID, where do you come from? Health Serv Res. 2005; 40(2):593–597. [PubMed: 15762909]

47. Cella D, Eton DT, Fairclough DL, et al. What is a clinically meaningful change on the Functional Assessment of Cancer Therapy-Lung (FACT-L) Questionnaire? Results from Eastern Cooperative Oncology Group (ECOG) Study 5592. J Clin Epidemiol. 2002; 55:285–295. [PubMed: 11864800]

48. Farrar JT, Young J, J T, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. Pain. 2001; 94:149–158. [PubMed: 11690728]

49. Salaffi F, Stancati A, Silvestri CA, Ciapetti A, Grassi W. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. Eur J Pain. 2004; 8(4):283–291. [PubMed: 15207508]

50. Li KK, Harris K, Hadi S, Chow E. What should be the optimal cut points for mild, moderate, and severe pain? J Palliat Med. 2007; 10(6):517–525.

51. Given B, Given CW, Sikorskii A, et al. Establishing mild, moderate, and severe scores for cancer-related symptoms: How consistent and clinically meaningful are interference-based severity cut-points? J Pain Symptom Manage. 2008; 35(2):126–135. [PubMed: 18158231]

52. Palos GR, Mendoza TR, Mobley GM, Cantor SB, Cleeland CS. Asking the community about cutpoints used to describe mild, moderate, and severe pain. J Pain. 2006; 7(1):49–56. [PubMed: 16414555]

53. Coons SJ, Gwaltney CJ, Hays RD, et al. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO good research practices task force report. Value Health. Nov 11.2008 2008 Published Online ahead of Print.

54. Gwaltney CJ, Shields AL, Shiffman S. Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: A meta-analytic review. Value Health. 2008; 11:322–333. [PubMed: 18380645]

55. Kieffer KM. An introductory primer on the appropriate use of exploratory and confirmatory factor analysis. Research in the Schools. 1999; 6(2):75–92.

56. Portenoy RK, Thomas J, Moehl Boatwright ML, et al. Subcutaneous methylnaltrexone for the treatment of opiod-induced constipation in patients with advanced illness: A double blind, randomized, parallel group, dose-ranging study. J Pain Symptom Manage. 2008; 35(5):458–468. [PubMed: 18440447]

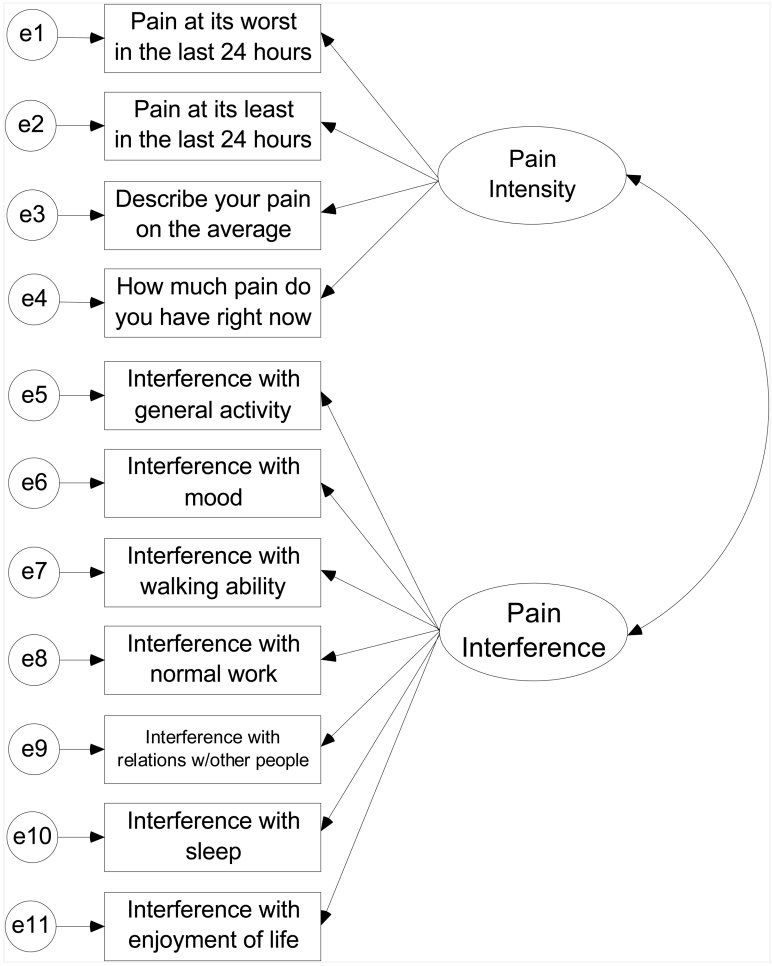

Figure 1. Two Factor Representation of the Brief Pain Inventory

Table 1

Summary of FDA Recommendations for the Psychometric Properties of a PRO Instrument

Property	FDA Draft Guidance Recommendations
Conceptual Framework	- Should be confirmed using empirical evidence during instrument development - Explicit statement of relationship between instruments concepts, domains, and items - Response options should be clear and appropriate
Reliability	- Instrument should demonstrate test-retest reliability - Instrument should demonstrate internal consistency
Content Validity	- Must encompass most important and comprehensive outcomes for patients - Patient input should be sought for item generation - Patient input should be sought until point of saturation
Construct Validity	- Obtained results should be consistent with pre-existing hypotheses - Instrument should have ability to differentiate between clinically distinct groups
Clinical Relevance of Score Changes	- Instrument should be equally sensitive to gains and losses in health status - Instrument should be sensitive to change at all points for the clinical population

Table 2

Internal Consistency Coefficients for the Brief Pain Inventory by Sample

Measure	Chinese [25]	Italian [25]	Greek [29]	Norwegian [21]	Russian [34]	Taiwanese [23]
Scale reliability if item is deleted						
Pain Intensity						
Worst Pain Last 24 Hours	0.87	0.77	0.88	0.84	0.90	0.81
Least Pain Last 24 Hours	0.89	0.73	0.86	0.85	0.94	0.78
Average Pain	0.81	0.68	0.82	0.80	0.87	0.72
Pain Right Now	0.88	0.72	0.86	0.85	0.90	0.75
Pain Interference						
General Activity	0.90	0.72	0.81	0.90	0.93	0.87
Mood	0.90	0.75	0.83	0.81	0.94	0.88
Walking Ability	0.92	0.76	0.82	0.92	0.95	0.87
Normal Work	0.89	0.72	0.82	0.91	0.94	0.88
Relations with Other People	0.89	0.76	0.84	0.91	0.95	0.89
Sleep	0.91	0.79	0.84	0.92	0.94	0.89
Enjoyment of Life	0.90	0.72	0.82	0.91	0.94	0.87

Table 3
Test-Retest Reliability Coefficients for the Brief Pain Inventory by Sample

Measure	German [22]	Osteoarthritis [20]	Spanish [24]	Taiwanese [23]
Time frame	30-60 minutes	7 days	7 days	1-10 days
Pain Intensity	----	0.87	0.62	0.79
Worst Pain Last 24 Hours	0.96	----	----	0.80
Least Pain Last 24 Hours	0.78	----	----	0.68
Average Pain	0.86	----	----	0.65
Pain Right Now	0.93	----	----	0.55
Pain Interference	----	0.92	0.77	0.81
General Activity	0.85	----	----	0.72
Mood	0.83	----	----	0.80
Walking Ability	0.91	----	----	0.69
Normal Work	0.89	----	----	0.71
Relations with Other People	0.93	----	----	0.81
Sleep	0.93	----	----	0.81
Enjoyment of Life	0.97	----	----	0.74

Table 4

Factor Loading Coefficients for the Brief Pain Inventory by Sample

Measure	Chinese [25]	Filipino [25]	French [25]	German [22]	Greek [29]	Hindi [27]	Italian [25]	Japanese [26]	Korean[17]	Non-Cancer [15]	Spanish [24]	Taiwanese [23]	USA [25]	Vietnamese [31]
Pain Intensity														
Worst Pain 24 Hours	0.74	0.74	0.64	0.76	0.82	0.44	0.57	0.90	0.59	0.65	0.73	0.34	0.68	0.86
Least Pain 24 Hours	0.81	0.83	0.82	0.90	0.81	0.82	0.63	0.83	0.94	0.76	0.86	0.78	0.87	0.81
Average Pain	0.91	0.75	0.80	0.80	1.00	0.86	0.80	0.58	0.77	0.71	0.93	0.63	0.87	0.95
Pain Right Now	0.76	0.77	0.82	0.93	0.77	0.97	0.77	0.55	0.79	0.80	0.82	0.87	0.78	0.41
Pain Interference														
General Activity	0.72	0.72	0.79	0.55	0.83	0.86	0.81	0.77	0.83	0.59	0.84	0.78	0.80	0.79
Mood	0.71	0.71	0.85	0.77	0.68	0.83	0.48	0.62	0.75	0.91	0.83	0.80	0.79	0.75
Walking Ability	0.82	0.72	0.63	0.54	0.69	0.59	0.70	0.62	0.86	0.63	0.64	0.86	0.71	0.87
Normal Work	0.86	0.79	0.73	0.76	0.79	0.85	0.88	0.61	0.91	0.73	0.64	0.73	0.80	0.91
Relations w/People	0.81	0.66	0.81	0.81	0.56	0.58	0.33	0.61	0.87	0.88	0.75	0.58	0.76	0.82
Sleep	0.62	0.60	0.56	0.83	0.49	0.56	-----	0.53	0.51	0.50	0.47	0.63	0.68	0.64
Enjoyment of Life	0.75	0.73	0.73	0.73	0.63	0.94	0.55	0.47	0.84	0.85	0.82	0.81	0.83	-----