The Journal of Pain, Vol 9, No 2 (February), 2008: pp 105-121
*Available online at www.sciencedirect.com*

ELSEVIER

# Consensus Statement

## Interpreting the Clinical Importance of Treatment Outcomes in Chronic Pain Clinical Trials: IMMPACT Recommendations

Robert H. Dworkin,* Dennis C. Turk,† Kathleen W. Wyrwich,‡ Dorcas Beaton,§ Charles S. Cleeland,‖ John T. Farrar,¶ Jennifer A. Haythornthwaite,# Mark P. Jensen,† Robert D. Kerns,** Deborah N. Ader,†† Nancy Brandenburg,‡‡ Laurie B. Burke,§§ David Cella,‖‖ Julie Chandler,¶¶ Penny Cowan,## Rozalina Dimitrova,*** Raymond Dionne,††† Sharon Hertz,§§ Alejandro R. Jadad,‡‡‡ Nathaniel P. Katz,§§§ Henrik Kehlet,‖‖‖ Lynn D. Kramer,¶¶¶ Donald C. Manning,### Cynthia McCormick,**** Michael P. McDermott,†††† Henry J. McQuay,‡‡‡‡ Sanjay Patel,§§§§ Linda Porter,‖‖‖‖ Steve Quessy,¶¶¶¶ Bob A. Rappaport,§§ Christine Rauschkolb,#### Dennis A. Revicki,***** Margaret Rothman,#### Kenneth E. Schmader,††††† Brett R. Stacey,‡‡‡‡‡ Joseph W. Stauffer,§§§§§ Thorsten von Stein,‖‖‖‖‖ Richard E. White,¶¶¶¶¶ James Witter,§§ and Stojan Zavisic#####

**Departments of Anesthesiology and Neurology, University of Rochester School of Medicine and Dentistry, Rochester, New York.*
*†University of Washington, Seattle, Washington.*
*‡Saint Louis University, St. Louis, Missouri.*
*§Institute for Work and Health, University of Toronto, Toronto, Canada.*
*‖M.D. Anderson Cancer Center, Houston, Texas.*
*¶University of Pennsylvania, Philadelphia, Pennsylvania.*
*#Johns Hopkins University, Baltimore, Maryland.*
***VA Connecticut Healthcare System and Yale University, West Haven, Connecticut.*
*††National Institute of Arthritis and Musculoskeletal and Skin Diseases, Bethesda, Maryland.*
*‡‡Pfizer, Inc., New York, New York.*
*§§United States Food and Drug Administration, Rockville, Maryland.*
*‖‖Evanston Northwestern Healthcare and Northwestern University, Evanston, Illinois.*
*¶¶Merck & Company, Blue Bell, Pennsylvania.*
*##American Chronic Pain Association, Rocklin, California.*
****Allergan, Inc., Irvine, California.*
*†††National Institute of Nursing Research, Bethesda, Maryland.*
*‡‡‡University Health Network and University of Toronto, Toronto, Canada.*
*§§§Analgesic Research, Needham, Massachusetts.*
*‖‖‖Rigshospitalet, Copenhagen University, Copenhagen, Denmark.*
*¶¶¶Purdue Pharma, Stamford, Connecticut.*
*###Celgene Corporation, Warren, New Jersey.*
*****McCormick Consultation, Bethesda, Maryland.*
*††††University of Rochester, Rochester, New York.*
*‡‡‡‡University of Oxford, Oxford, United Kingdom.*
*§§§§Elan Pharmaceuticals, Inc., San Diego, California.*
*‖‖‖‖National Institute of Neurological Disorders and Stroke, Bethesda, Maryland.*
*¶¶¶¶GlaxoSmithKline, Research Triangle Park, North Carolina.*
*####Johnson & Johnson, Raritan, New Jersey.*
******Health Care Analytics Group, United Biosource Corporation, Bethesda, Maryland.*
*†††††Duke University Medical Center and GRECC, VA Medical Center, Durham, North Carolina.*
*‡‡‡‡‡Oregon Health and Science University, Portland, Oregon.*
*§§§§§Alpharma, Elizabeth, New Jersey.*

|||||||Tercica, Brisbane, California.
¶¶¶¶¶Endo Pharmaceuticals Inc., Chadds Ford, Pennsylvania.
#####AstraZeneca, Södertälje, Sweden.

**Abstract: A consensus meeting was convened by the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) to provide recommendations for interpreting clinical importance of treatment outcomes in clinical trials of the efficacy and effectiveness of chronic pain treatments. A group of 40 participants from universities, governmental agencies, a patient self-help organization, and the pharmaceutical industry considered methodologic issues and research results relevant to determining the clinical importance of changes in the specific outcome measures previously recommended by IMMPACT for 4 core chronic pain outcome domains: (1) Pain intensity, assessed by a 0 to 10 numerical rating scale; (2) physical functioning, assessed by the Multidimensional Pain Inventory and Brief Pain Inventory interference scales; (3) emotional functioning, assessed by the Beck Depression Inventory and Profile of Mood States; and (4) participant ratings of overall improvement, assessed by the Patient Global Impression of Change scale. It is recommended that 2 or more different methods be used to evaluate the clinical importance of improvement or worsening for chronic pain clinical trial outcome measures. Provisional benchmarks for identifying clinically important changes in specific outcome measures that can be used for outcome studies of treatments for chronic pain are proposed.**

***Perspective:** Systematically collecting and reporting the recommended information needed to evaluate the clinical importance of treatment outcomes of chronic pain clinical trials will allow additional validation of proposed benchmarks and provide more meaningful comparisons of chronic pain treatments.*



***Key words:** Chronic pain, randomized clinical trials, outcome measures, clinical importance, assessment, quality of life, physical functioning, emotional functioning, global ratings.*

*". . . a difference is a difference only if it makes a difference."*

—Darrell Huff, 1954, p. 58[66]

There is widespread agreement that efforts to develop improved treatments for patients with chronic pain are a research priority. Variability in the outcome measures used in clinical trials hinders evaluations of the efficacy and effectiveness of treatments. In recognition of this, the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) recommended core outcome domains[143] and specific outcome measures[34] for chronic pain trials. Including a standard set of outcome measures in clinical trials facilitates the process of developing research protocols, permits pooling of data from different studies, and provides a basis for systematic reviews and meaningful comparisons among treatments.

A critical consideration in evaluating existing outcome measures as well as when developing improved measures[144] is the clinical importance or meaningfulness of the change in scores that occurs following treatment. In analyzing clinical trial outcome data, establishing the statistical significance and confidence intervals of treatment responses is a pivotal step. However, because statistical significance reflects both the magnitude and variability of the treatment effect as well as the sample size, a statistically significant improvement may reflect a benefit that is clinically meaningless. For this reason, it is generally acknowledged that determinations of statistical significance must be supplemented by consideration of the clinical importance of changes in outcome measures.[76,155] Such information provides a basis for evaluating and comparing the impact of treatments on symptoms, functioning, well-being, and overall health-related quality of life (HRQoL). Depending on the specific outcome, clinical importance and meaningfulness can be assessed by patients, clinicians, significant others, and representatives of society at large,[14,49,134] for example, third-party payors. For chronic pain, however, most measures of treatment response involve patient-reported outcomes (PROs), for which the patient is the most important judge of whether changes are important or meaningful.[1,144] The objective of the present article is

Received June 29, 2007; Revised August 21, 2007; Accepted September 28, 2007.

Supported by unrestricted grants to the University of Rochester Office of Professional Education from Allergan, Alpharma, AstraZeneca, Celgene, Elan, Endo, GlaxoSmithKline, Johnson & Johnson, Merck, NeurogesX, Pfizer, Purdue, and Schwarz Biosciences.

The views expressed in this article are those of the authors. No official endorsement by the US Department of Veterans Affairs, US Food and Drug Administration, US National Institutes of Health, or the pharmaceutical companies that provided unrestricted grants to the University of Rochester Office of Professional Education should be inferred.

Address reprint requests to Robert H. Dworkin, PhD, University of Rochester School of Medicine and Dentistry, 601 Elmwood Avenue, Box 604, Rochester, NY, 14642. E-mail: robert_dworkin@urmc.rochester.edu

1526-5900/$34.00



doi:10.1016/j.jpain.2007.09.005

to present IMMPACT consensus recommendations for determining clinically important changes for outcome measures for chronic pain trials.

## Methods

An IMMPACT consensus meeting was held that had 2 goals: (1) To develop general recommendations for determining clinically important changes for chronic pain outcome measures; and (2) to propose provisional benchmarks for identifying clinically important changes in the specific outcome measures for chronic pain clinical trials previously recommended by IMMPACT.[34] The recommendations of specific outcome measures were, in part, based on commissioned literature reviews that were prepared by individuals who had not been involved in the development of outcome measures for the domains that they reviewed. These literature reviews are available on the IMMPACT-II meeting page at www.immpact.org/meetings.html and should be consulted for detailed reviews and discussions of the measures that were considered, the evidence on which the recommendations were based, and the reasons for selection or rejection of specific measures.

On the basis of these background literature reviews and discussion and debate at a consensus meeting, specific measures were recommended for chronic pain trials in a previous publication.[34] Among the criteria used in evaluating these and other measures were (1) appropriateness of the measure's content and conceptual model; (2) reliability; (3) validity; (4) responsiveness; (5) interpretability; (6) precision of scores; (7) respondent and administrator acceptability; (8) respondent and administrator burden and feasibility; (9) availability and equivalence of alternate forms and methods of administration (eg, self-report, interviewer); and (10) availability and equivalence of versions for different cultures and languages. In evaluating the extent to which the various measures reviewed in the background presentations fulfilled these criteria, appropriateness of content, reliability, validity, responsiveness, and participant burden were given the greatest weight. In particular, measures for which published information on these specific criteria were lacking were not recommended, and when such information was available for 2 or more relevant measures, recommendations were primarily based on comparisons of these 5 attributes.[34]

The IMMPACT consensus meeting on clinical importance on which the present article is based included an international group of 40 participants from universities, governmental agencies, a patient self-help organization, and the pharmaceutical industry. Participants were selected on the basis of their research, clinical, or administrative expertise relevant to the design and evaluation of chronic pain treatment outcomes. An attempt was made to include broad representation of various disciplines while limiting the size of the meeting to promote frank discussion. Because not all attendees were familiar with recent advances in the determination of clinically important change, several articles were circulated prior to the meeting[7,18,101,153,161] and 3 background lectures were presented at the meeting that examined general methodologic issues: (1) Clinically meaningful change: An overview (KWW); (2) clinical importance: The rheumatology perspective (DB); and (3) potential designs for studies of minimal clinically important differences (DB and KWW).

In addition, reviews of the literature relevant to determining the clinical importance of changes for the specific outcome measures previously recommended by IMMPACT[34] for the following 4 core chronic pain outcome domains were presented at the meeting. The individuals who prepared these reviews and delivered these presentations were selected on the basis of their research expertise regarding the specific outcome measures. As opposed to the reviews that provided the basis for recommending specific outcome measures,[34] these reviews were sometimes prepared by individuals who had been involved in the development of the specific outcome measure as it was believed that they would have the greatest knowledge regarding what would constitute a clinically important change in the measure: (1) Pain intensity, assessed by a 0 to 10 numerical rating scale (MPJ); (2) physical functioning, assessed by the Multidimensional Pain Inventory Interference Scale (DCT) and by the Brief Pain Inventory Interference Scale (CSC); (3) emotional functioning, assessed by the Beck Depression Inventory (RDK) and by the Profile of Mood States (JH); and (4) participant ratings of overall improvement, assessed by the Patient Global Impression of Change scale (JTF).

The presentations are available on the IMMPACT-IV meeting page at www.immpact.org/meetings.html and should be consulted for the literature reviews that provided the background for the discussions among the participants that occurred at the consensus meeting and the preparation of this article. The recommendations included in this article are based on the consensus that emerged from consideration of these literature reviews, the extensive discussion and debate that took place during the consensus meeting, and the continued discussion that occurred during the preparation of this article, which was revised incorporating feedback from all of the authors until consensus was reached on the text and table.

## General Considerations in Determining Clinically Important Differences

In considering the determination of clinically important differences, 2 different aspects of the interpretation of clinical trial results must be distinguished. The first is establishing what change in the outcome measure represents a clinically important difference for patients. The second is establishing the difference in the magnitude of response between the treatment and control groups that will be considered large enough to establish the scientific or therapeutic importance of the results. This difference between groups is also used to calculate the sample size required for the clinical trial, and can involve

group differences in either central tendency (eg, means) or in the proportions of responders (eg, percentages of patients that obtain a defined response). Such responder analyses require knowledge of what magnitudes of individual change can be considered clinically important so that patients can be categorized as responders and nonresponders. Although this is a crucial step in understanding and interpreting the results of a clinical trial, it does not identify the magnitude of differences between treatment groups that should be considered important. In this article, we emphasize the determination of clinically important changes for individuals and not the determination of the importance of group differences, which can only be established in the broader context of the disease being treated, the currently available treatments, and the overall risk-benefit ratio of the treatment.

The outcomes of clinical trials of treatments for chronic pain include mean changes on 1 or more measures of pain, as well as mean changes on various measures of physical and emotional functioning.[143] From these results, the multiple stakeholders in these studies, such as clinicians, regulators, payors, and, ultimately, people with pain, must then determine the efficacy of the treatment.[7,8,49,87,134] Trials with negligible mean benefits may be sufficiently powered for the results to be statistically significant, as noted above. In addition, trials demonstrating a relatively large mean change may also be difficult to interpret because it cannot be assumed that all participants in the active arm of a clinical trial uniformly experienced the magnitude of benefit reflected by the mean improvement. Indeed, depending on the sample size, a few individuals with large improvements in a trial's active treatment group can dramatically increase the overall mean improvement even if others in the same group demonstrated little improvement or even a worsening of their condition. More importantly, because pain relief occurs and is appreciated by individuals differently, the magnitude of a statistically significant group mean change may bear little relation to an important improvement for the person with pain.

The development of criteria for determining what are important changes in individuals' scores on the outcome measures used in chronic pain trials would provide clinicians and researchers with essential methods for evaluating treatment responses of individuals in clinical trials and clinical practice. Such individual-level criteria make it possible to conduct responder analyses that classify each trial participant as "improved," "stable," or "worse" on the basis of validated criteria of important change.[158]

In research on the importance of changes in PROs, patient-based, clinician-based, and laboratory-based assessments have all been used. In the present article, we use the term "clinically important" not only to distinguish clinically important changes from those that are statistically significant but also to emphasize that we are referring to changes in clinical conditions that are important to patients as well as others. In addition, when using the term clinically important in this article, we have not always distinguished "minimally" important changes from those changes that are more substantial[126]; the identification of different magnitudes of important change has received less attention than the determination of criteria for minimally important changes.

## *Methods for Determining Criteria for Important Change*

In an important review of approaches used for interpreting change in HRQoL measures, Lydick and Epstein[89] classified these methods as either "anchor-based" or "distribution-based." Anchor-based methods relate changes in scores on a measure to a standard that is different from the specific measure itself, whereas distribution-based methods use statistical parameters associated with the measure (eg, effect size, standard error of measurement) to interpret the magnitude of changes in the measure's scores over time.[29,32,111,127]

### Anchor-Based Methods

Many anchor-based approaches for establishing criteria for identifying important change rely on a global item completed by the patient as the anchor for within-person changes. For example, to determine whether there were important changes in pain over the course of treatment, patients' pain could be assessed at baseline and again at the end of the trial, at which point they would also be asked if they were "better," "about the same," or "worse," compared with the beginning of the trial. These improvement ratings would then serve as the standard with which to evaluate the importance to the patient of whatever changes in pain had occurred during the course of the trial.

Using the anchor-based approach in patients with heart and lung disease, Jaeschke et al[69] anchored the amount of change that had occurred in several HRQoL domains by using 7-point scales with which patients assessed their improvement or worsening. Patients who reported their change to be 1 ("almost the same, hardly any better/worse"), 2 ("a little better/worse"), or 3 ("somewhat better/worse") were considered to have had small but important changes, and the means of the HRQoL score differences that corresponded to these differences were considered the minimal clinically important difference (MCID). Later criteria for minimal change excluded ratings of 1 ("almost the same, hardly any better/worse"),[75] and, most recently, the term minimal important difference (MID) has been defined as "the smallest difference in score in the domain of interest that patients perceive as important, either beneficial or harmful, and that would lead the clinician to consider a change in the patient's management."[58]

Anchor-based methods for interpreting changes in PRO measures have also included between-subjects studies. For example, social comparison studies anchor HRQoL score changes to patients' rating of their HRQoL compared with another patient with whom they have interacted.[107-109] In addition, anchor-based approaches have used clinician-based or laboratory-based anchors to provide criteria for PRO changes. For example, Kosinski

et al[83] used 1% to 19% decreases over time in the number of swollen and tender joints as anchors in patients with rheumatoid arthritis. In oncology, anchors such as tumor response,[20] performance status,[19-21] and hemoglobin values[22] have been used.

Despite the ease of simply including an anchoring measure in studies evaluating change over time, there are several problematic aspects of this approach.[102] In the case of global items, patients are asked to provide a retrospective comparison of their change since an earlier time point and may not accurately remember previous levels of pain or other PROs unless there was a salient event, like labor and delivery.[114] In addition, it has been demonstrated that such global items can be associated with patients' conditions when they are making these ratings[57,128,129]; for example, if patients feel good at Time 2, they may report being better than they were at Time 1, even if they were doing just as well or even better at Time 1. Hence, retrospective anchors may sometimes lack validity as a criterion of important change, although the extent to which this occurs probably varies and should be carefully evaluated.[57]

Furthermore, if important changes in PRO measures such as pain and HRQoL can be identified by global ratings or clinical standards such as swollen joints serving as the "gold standard," then why are the PRO measures needed at all?[102,127] Although the specific anchors that have been used to determine clinically important changes all have limitations, they make it possible not only to evaluate whether patients themselves believe they have improved or not but also to determine the extent to which such patient assessments are associated with clinician-based and laboratory-based measures.

### Distribution-Based Methods

Two key distribution-based methods that have emerged for determining clinically important changes are the effect size and the standard error of measurement (SEM). The group effect size is determined by subtracting the mean of the Time 1 scores from the mean of the Time 2 scores and then dividing this difference by the standard deviation at Time 1. Cohen[28] proposed criteria, as a convention, for the magnitude of effect sizes for group differences. On the basis of the literature in behavioral science, 0.20 was proposed as the lower bound for a small effect (or change), 0.50 as the threshold for a moderate effect, and 0.80 and above as reflecting a large effect. Kraemer et al[84] suggested that these thresholds might be more accurately termed "smaller than typical," "typical," and "larger than typical" to reflect the fact that they were meant to be relative to typical findings in behavioral science research.

In a review of published MCID and MID studies for HRQoL measures, Norman et al[101] calculated the effect size of anchor-based criteria for change in 38 separate studies. The mean effect size across these studies was 0.495 (with a standard deviation of 0.155), which did not vary as a function of whether studies used 7-point vs other scales, minimal improvements vs clinically important improvements, or generic vs disease-specific HRQoL measures. To explain these consistent results, the authors referred to Miller's[95] "magical number 7 plus or minus 2" as the limit on human information processing capacity, which can be shown to be consistent with a change of 1 point on 5-, 7-, or 9-point scales reflecting an effect size of approximately 0.50.

It is important to note, however, that not all of the effect size estimates for MCID and MID studies found in this meta-analysis approached 0.50.[101] Indeed, for studies that used very small time increments, like a 2-day period after chemotherapy, patients believed that the magnitude of a small but important PRO change corresponded to an effect size of 0.12,[123] whereas the effect size corresponding to the amount of change that physical therapy patients expected to see from HRQoL improvements in back or shoulder pain ranged from 0.86 to over 1.00.[63,132] These dramatic departures from the mean result confirm that 1 effect size does not fit all.[159] Nevertheless, a 0.50 effect size (ie, one-half the standard deviation) may be a reasonable criterion to use when beginning to investigate important changes in PRO measures, including pain and physical and emotional functioning.[127]

Important concerns about the effect size criterion for identifying important differences involve the fact that the standard deviation of a measure is specific to a particular sample[77] and that reliability of PRO measures can be modest. For example, if a sample is heterogeneous or if a measure has limited reliability (ie, substantial error), the standard deviation may be large, and the corresponding value for a 0.50 effect size will be much larger than with a homogeneous sample or a more reliable measure. The standard error of measurement (SEM) provides a measure of within-person change that is less dependent on a specific sample because it incorporates both the standard deviation and the reliability.[3,103] To determine how many SEMs constitute an important change, Wyrwich et al[162-164] computed the SEM for various chronic heart and respiratory disease outcome domains and compared the results with the anchor-based criteria described above by Jaeschke et al.[69] For all domains, 1 SEM approximated the MCID threshold. Likewise, Cella et al[20] found that 1 SEM corresponded to anchor-based criteria for small but important differences in the HRQoL of patients with lung cancer.

## *Methodologic Considerations in Determining Criteria for Important Changes*

Although anchor-based and distribution-based approaches apply a single criterion for important change across all points on the PRO measures being investigated, it is important to recognize that all points on pain scales may not be equal. For example, the results of a study of labor epidural analgesia[11] suggested that, at least in some circumstances, a change in pain intensity from 3 to 1 on a 0 to 10 numerical rating scale (NRS) may be of greater importance than a change from 6 to 4. Similarly, using item response theory analyses of 0 to 10 NRS data

in cancer patients, Lai et al[85] found large gaps in pain ratings between 0 and 1, 7 and 8, and 8 and 9, with 10 never used. These results suggest that pain ratings of 2 through 6 on such a measure are considered much closer to each other in intensity than ratings at the ends of the scale.

Another important consideration in determining important change is the impact that baseline status has on patients' assessments of differences.[32,61,133,136] For example, the magnitude of pain reduction that an individual with severe pain would consider minimally important might be greater than the magnitude of reduction considered minimally important by a patient with mild or moderate pain.[59,120] Although using percentage change has the potential to correct for this,[133] the role of baseline pain would need to be evaluated in each specific situation. Patient characteristics, such as age, sex, education, and the specific clinical condition, may also play an important role in determining what magnitude of change is important.[136] Furthermore, the magnitude of change considered an important improvement might be different from what is considered an important worsening;[61,155] for example, in some circumstances, small improvements might be more important to patients than small deteriorations,[21] whereas in other circumstances, the opposite might be true.[59]

Assessments of minimally important change depend to a great extent on the definition of "minimal importance."[32] Not surprisingly, different criteria may be obtained depending on whether the emphasis is placed on determining minimally detectable versus minimally important changes. Similarly, different criteria are likely to result when evaluating minimally important changes versus moderately important changes versus substantial or definitive changes. Moreover, for certain conditions it is possible to determine what change is needed to achieve "that state which is deemed a useful target of treatment by both physician and patient, given current treatment possibilities and limitations."[154]

As noted above, patients, clinicians, third-party payors, and others may have very different perspectives regarding what benefits constitute clinically important improvement (and what changes constitute clinically important worsening). It is generally acknowledged that evaluations by patients are critical in determining the importance of changes in PROs. However, clinician perspectives can also provide valuable information, and may lead to estimates of the magnitude of important changes that are not only different from those based on patient evaluations but that also classify patients differently with respect to whether they have improved or not.[161]

The anchors used in evaluating change by patients and clinicians and their interpretation can also vary greatly. Patients base their evaluations of change on their own experience, whereas clinicians base such evaluations on their experiences across multiple patients with the same condition. Clinicians may also place greater emphasis on anchors that reflect disease processes and prognosis (eg, swollen joint counts in arthritis, glycemic control in diabetes), whereas symptoms, quality of life, and overall treatment satisfaction may be of greater importance to patients. The use of such different anchors highlights the importance of considering patterns of change, both positive and negative, across a variety of different outcomes. In such an approach, a clinically important benefit could reflect a pattern of changes in, for example, patient reports of pain, physical and emotional functioning, and side effects, and could also include clinician-based measures of disease progression. Such patterns of benefits and harms would take into account the clinical reality that what is an important change in a single outcome can differ depending on changes that have occurred in other outcomes. Unfortunately, few studies have investigated the clinical importance of patterns of different outcomes.

## *Determining Criteria for Important Changes for Groups*

To this point, we have discussed methods for determining criteria for important change in individuals. There are many situations, however, in which it is important to evaluate changes in groups to determine what magnitudes of changes over time or differences between treatment and placebo (or 2 different treatments) should be considered clinically important.[8,19] It is crucial to recognize that criteria for clinically important change in individuals cannot be directly applied to the evaluation of clinically important group differences. For example, in evaluating a new analgesic, if a 2-point decrease on a 0 to 10 NRS of pain intensity is considered a clinically important improvement for an individual, it should not be inferred that a 2-point difference in pain reduction between the analgesic and placebo must occur before the treatment benefit can be considered clinically important.[56,136] In this example, there could be a sizable percentage of patients who have a clinically important pain reduction of 2 points with the new analgesic even if there is only a 1-point mean difference between the groups.

One approach to determining the importance of group differences in a clinical trial is to compare the percentages of patients who have clinically important changes in the treatment groups. There are many approaches for performing such "responder analyses" and for evaluating treatment effect sizes more generally, including calculating the number-needed-to-treat (NNT). The NNT is the reciprocal of the absolute risk reduction and reflects the "number of patients who must be treated to generate one more success or one less failure than would have resulted had all persons been given the comparison treatment."[84] NNTs have often been used to evaluate the efficacy of treatments for pain.[45,93] In addition, a range of $\pm 0.5$ NNT has been used to determine whether an NNT has "clinical relevance" (ie, whether the NNT is within acceptable bounds of clinical "accuracy"), and such criteria could also be used to compare differences between NNTs associated with different treatments.[97] Many of the same considerations discussed above for

interpreting individual-level change criteria also apply to the interpretation of group differences in NNTs or other measures of effect size.[160]

## Determining Clinically Important Differences in Pain Intensity

In clinical trials designed to evaluate the efficacy of chronic pain treatments, the primary efficacy analysis typically involves reduction in pain intensity.[143] There has been increasing attention to identifying the magnitude of pain reduction that would constitute an important benefit of pain treatment, and several studies have examined the importance of changes in 0 to 10 NRS pain intensity scores to patients with chronic pain. In the largest of these studies, Farrar et al[43] examined data from 10 clinical trials in which 2,724 patients with painful diabetic neuropathy, postherpetic neuralgia, low back pain, fibromyalgia, or osteoarthritis completed a 0 to 10 pain intensity NRS before and after treatment and a 7-point categorical scale of global impression of change (ranging from "very much improved" to "very much worse") after treatment. Pre- to post-treatment decreases in pain intensity of 2 points or 30% were associated with patient ratings of "much improved." These thresholds did not differ as a function of diagnostic group, trial duration, treatment condition (placebo vs pregabalin), or demographic characteristics. Decreases of ≥4 points or ≥50% were associated with patient ratings of "very much improved." In receiver operating curve (ROC) analyses, a decrease of ≥1.7 points or ≥28% best distinguished patients who rated their improvement in pain as "much improved" or greater from those who rated their change as "minimally improved" or less.

Salaffi et al[120] also used ROC analyses to identify the absolute and percentage changes in 0 to 10 pain intensity scores that differentiated global outcome ratings of pain improvement in 825 patients with osteoarthritis, rheumatoid arthritis, or ankylosing spondylitis. Decreases in individuals' pain intensity ratings of ≥1.0 point or ≥15% over the course of 3 months best differentiated patients who described their pain as being at least "slightly better" from those who reported no change or a worsening in pain, and decreases of ≥2.0 points or ≥33% best differentiated patients who described their pain as "much better" from those who described the change as only "slightly better" or worse. Hanley et al[60] examined the changes in 0 to 10 pain intensity scores associated with ratings of change in patients with a physical disability and chronic pain. Decreases in individuals' pain intensity ratings of 1.8 points or 36% corresponded to reports of a "meaningful" change in pain, and decreases of 1.0 point or 20% were associated with a "noticeable, but not meaningful" decrease in pain. There were no significant differences as a function of diagnosis, sex, or treatment condition, but older individuals required a larger pain decrease (2.4 points) to rate their change as meaningful than did younger patients (1.2 points).

Considering the results of these 3 studies together, raw score changes of approximately 1 point or percentage changes of approximately 15% to 20% represent minimally important but perhaps not very important decreases using a 0 to 10 NRS of chronic pain intensity. Changes of approximately 2.0 points or 30% to 36% represent "much better," "much improved," or "meaningful" decreases in chronic pain, and a decrease of ≥4 points or ≥50% appears to represent a substantial ("very much improved") change in pain, 1 which patients have also considered "treatment success"[112] or "satisfactory improvement."[135] In a study of patients with complex regional pain syndrome, type I, larger decreases in pain were required for ratings of both "little" and "much" improvement,[48] but research is needed to replicate these results and determine whether they are unique to the procedures used or patients examined.

Several studies have examined the clinical importance of changes in chronic pain as assessed by a 10-cm visual analogue scale (VAS), and research has also been conducted on the clinical importance of changes in acute pain using either an NRS or a VAS. These studies provide additional support for the generalizability of the finding that decreases in individuals' pain intensity of approximately 1 cm (or 1.0 point) or 15% to 20% represent "minimal" or "little" change,[17,23,50,51,59,65,71,78,79,138] whereas decreases of 2.0 to 2.7 points or 30% to 41% have more meaning to patients, for example, being associated with not requesting rescue medication[40,42] or ratings of "much" or "some" change.[23,71] This research also supports the importance of taking baseline pain into account when evaluating these change scores.[17,23,40,43,71,79,139]

On the basis of this body of research, it is possible to propose provisional benchmarks[126,127] for evaluating the magnitude of changes in pain intensity and comparing the results of different chronic pain clinical trials or different treatment groups within trials. Reductions in chronic pain intensity in individuals of at least 10% to 20% appear to reflect minimally important changes. Reductions of ≥30% appear to reflect at least moderate clinically important differences, and it is recommended that the percentages of patients responding with this degree of pain relief be reported in clinical trials of chronic pain treatments. In addition, because reductions in chronic pain intensity of ≥50% appear to reflect substantial improvements, it is also recommended that the percentages of patients responding with this degree of improvement be reported.

All of these proposed benchmarks must be confirmed in future studies that directly assess patient evaluations of what is noticeable, important, and major improvement. Moreover, whether or not a particular change in pain represents an important change can depend on the clinical and situational context. For example, the level of change in pain that is considered important is influenced by baseline pain, and may also vary by age, the patient's clinical condition, and prior treatment response. In addition, much of the research reviewed above was based on studies that evaluated the efficacy of a single treatment, and it would be important to determine whether the magnitudes of clinically important changes vary depend-

ing on whether monotherapy or add-on therapy is being considered. Furthermore, the role played by the costs and side effects of treatment and the anticipated duration of the change (eg, a 10% decrease that lasts for several years might be more important than 1 that lasts for a few months) should be carefully evaluated.

It is also recommended that all chronic pain clinical trials report a cumulative proportion of responder analysis. In this approach, the entire distribution of treatment response is depicted in a graph of the proportion of responders for all percentages of pain reduction from 0% through 100%.[41] Using such a graph, it is possible to compare treatment groups with respect to the percentages of patients achieving any percentage of pain reduction, not only the benchmarks discussed above but also any others that might be more informative depending on the specific circumstances. Such an analysis can also be extended to include the percentages of patients whose pain has increased over the course of the clinical trial, which makes it possible to compare the extent to which worsening has occurred in the different treatment groups.[105]

## Determining Clinically Important Differences in Physical Functioning

Physical functioning is 1 of 2 outcome domains that are recommended as core components of HRQoL that should be assessed in all clinical trials of treatments for chronic pain.[143] The Interference Scale of the Multidimensional Pain Inventory[81] (MPI) and the Interference Scale of the Brief Pain Inventory[26] (BPI) have been recommended by IMMPACT for the assessment of physical functioning.[34] This recommendation applies to all chronic pain conditions, unless well-validated disease-specific measures are available, for example, the Roland-Morris Disability Questionnaire[113] and Oswestry Disability Index[38] for low back pain, for which clinically important differences have been presented,[18] and the Western Ontario and McMaster Universities Osteoarthritis Index[12] and other measures in patients with arthritis.[8,153]

### *Multidimensional Pain Inventory Interference Scale*

The MPI is a 60-item self-report inventory, designed to assess pain patients' cognitive, behavioral, and affective responses to their condition.[81] The MPI consists of 12 empirically derived scales that are grouped into 3 sections (pain and its impact; responses by significant others; activities). The Interference Scale is included in the section on pain and its impact and consists of 9 questions (eg, "In general, how much does your pain interfere with your day-to-day activities?"; "How much has your pain changed your ability to take part in recreational and other social activities?"), which are rated on 7-point scales ranging from 0 ("no interference/change") to 6 ("extreme interference/change").

The MPI has been translated into several languages,[44,47,88,96,157] and its psychometric adequacy has been demonstrated in diverse types of chronic pain, including chronic low back pain,[146] headache,[122] fibromyalgia,[145] systemic lupus erythematosus,[54] and cancer.[147] It has been used as an outcome measure in clinical trials of diverse treatments, including rehabilitation,[2,148] physical exercise,[82] percutaneous electrical nerve stimulation,[152] pharmacological treatments,[67,116] radiofrequency lesioning,[53,150] and psychological treatments.[137]

To date, no studies have used anchor-based methods to examine criteria from clinically important changes on the MPI Interference Scale. In the absence of such data, the distributional characteristics of the scale can be used to provide estimates for important differences. Normative data from a representative sample of published and unpublished studies that used the MPI to assess pain and functioning in patients with diverse chronic pain syndromes suggest that based on the scale's standard deviation, a change of approximately 0.6 points would be a reasonable benchmark for future studies designed to identify to minimal clinically important changes on this measure.[15,16,36,73,74,110,117,131] This criterion is consistent with the SEMs that have been calculated for this measure across diverse chronic pain conditions and treatments, which range from 0.4 to 0.8.[142] The variability among the standard deviations and SEMs found in these studies, however, suggests that criteria used for the clinical importance of changes on the MPI Interference Scale may differ depending on the specific pain condition being examined in a given trial, a possibility that must be considered in the further development of such criteria.

### *Brief Pain Inventory Interference Scale*

The BPI Interference Scale is a 7-item self-report measure, designed to assess the extent to which pain interferes with various components of functioning, including physical and emotional functioning and sleep.[25,26] The items in this scale can be grouped into those that assess physical functioning (general activity; walking ability; normal work, including both work outside the home and housework), those that assess emotional functioning (mood; relations with people; enjoyment of life), and a single item that assess the extent to which pain interferes with sleep. The BPI has been translated into many languages, and its psychometric adequacy was first established in patients with cancer pain but has now been demonstrated in multiple types of chronic non-cancer pain.[5,94,104,106] It has been used as an outcome measure in clinical trials of diverse treatments, including both pharmacological and psychological treatments.[24]

Several studies have examined the magnitude of treatment-associated change in BPI Interference Scale scores, and the results have generally demonstrated that improvements in open-label and randomized clinical trials range from 1 to 3 points, depending on the specific pain conditions and treatments studied.[24] In other studies, the relationships between these scores and patient reports of global satisfaction and improvement with treatment were examined.[24] Patients who are more satisfied with their treatment or current situation report lower levels of interference, and the results of several studies

indicate that the differences in Interference Scale mean scores between patients who report being satisfied or improved with treatment and those who are less satisfied or not improved range from 1 to 2 points, depending on the specific measures of global satisfaction or improvement, pain conditions, and treatments studied.[24,39]

The distributional characteristics of the BPI Interference Scale can also be used to provide ranges for identifying important differences on this measure.[24] Available data suggest that a change of 1 point on the Interference Scale, which is approximately one-half its standard deviation, would be a reasonable benchmark for future studies designed to identify to minimally clinically important changes.

## Determining Clinically Important Differences in Emotional Functioning

Emotional functioning is the second component of HRQoL recommended as a core outcome domain that should be assessed in all clinical trials of the efficacy and effectiveness of treatments for chronic pain.[143] Two measures have been recommended by IMMPACT for the assessment of emotional functioning in such trials.[34]

### *Beck Depression Inventory*

The Beck Depression Inventory[9,10] (BDI) was recommended because of its excellent psychometric properties and its extensive use in pain clinical research and responsiveness to change in pain clinical trials. The BDI consists of 21 groups of 4 statements designed to assess severity of current symptoms of depressive disorders, with total scores on the measure ranging from 0 to 63. An extensive empirical literature reveals generally acceptable internal consistency (Cronbach $\alpha$s = .73–.95), test-retest reliability (Pearson rs = .80–.90), convergent validity (mean Pearson r = .60), and responsiveness to change, which has been demonstrated in numerous pharmacotherapy and psychotherapy clinical trials in patients with depression, as well as a relatively large number of pain clinical trials.[80] The availability of multiple translations of the BDI and its brevity (ie, 5–10 minutes required for completion) and low reading level requirements (ie, fifth or sixth grade) are additional strengths of this measure.

Several types of data were considered for identifying criteria for clinically important change in BDI scores during a pain clinical trial, including consideration of normative data from psychiatric and substance-abusing populations.[80] Mean BDI scores from multiple studies ranged from a low of 27.8 for heroin users to a high of 38.5 for a sample of persons with major depressive disorder. Studies reporting BDI scores for samples of persons with pain found mean BDI scores ranging from 7.5 (lumbar surgery patients) to 25.5 (depressed, nontreatment-seeking individuals). Based on such data, Beck and Steer[9] recommended that scores below 10 should be considered to reflect "minimal or no" depression, with score ranges of 10 to 18, 19 to 29, and 30 to 63 reflecting "mild to moderate," "moderate to severe," and "severe" depression, respectively.

Geisser et al[52] recommended that a score of 21 on the BDI distinguishes chronic pain patients with and without major depressive disorder. Morley et al[99] found a mean BDI score of 17.6 (SD = 8.7) in a sample of nearly 2000 persons entering chronic pain treatment. Approximately 18% of these patients had "minimal," 46% "mild," 27% "moderate," and 10% "severe" depression using the criteria recommended by Beck and Steer,[9] and 28.5% of the sample was "depressed" using the cutoff recommended by Geisser et al.[52]

In 2 randomized trials of pharmacological treatments for chronic low back pain, mean changes on the BDI were 3.5[70] and 3.8,[6] and for 8 psychological intervention trials, the pre- to post-treatment BDI changes for the active treatment conditions ranged from 1.4[91] to 12.3.[100] Morley et al[98] reported a mean effect size of .52 for emotional functioning measures in a meta-analysis of the benefits of cognitive-behavioral or behavioral therapies for chronic pain versus waiting list controls, but a recent meta-analysis of a broader range of psychological interventions failed to find a significant effect on emotional functioning.[64]

Few studies have specifically considered what changes on the BDI would constitute important improvement in pain clinical trials. Vlaeyen et al[151] used a criterion of a greater than 4-point decrease on the BDI and a post-treatment score of less than 12 as evidence of "clinically significant" improvement. In 2 other studies, important improvement was considered to have occurred when patients with a baseline BDI score $>$10 reported a score of $\leq$10 after treatment.[72,125]

Considered together, available data suggest that 3 different strategies could be used to determine clinically important changes in BDI scores. One is to consider a patient to have shown important improvement when the BDI score falls into the "normal" range, that is, a score below 10. Given what may be limited effects of existing pain interventions on emotional functioning, requiring normal levels of depression as an outcome appears to be too conservative and is likely to be insensitive to important, but smaller degrees of improvement. A second approach would be to consider that important change has occurred when a patient shifts to a less severe category of depression following treatment (eg, from moderate to mild). Support for the use of this criterion comes from the validity of the severity categories proposed by Beck and Steer[9] and evidence of their relatively normal distribution in samples of patients with chronic pain. However, shifts between these categories in chronic pain patients may be a relatively arbitrary criterion of important change, especially when pain clinical trials do not specifically target emotional functioning.

Applying one-half standard deviation to depression severity category standard deviations ranging from 8.1 to 10.4, a change of 5 points on the BDI could be considered a reasonable estimate of a clinically important change. Advantages of this approach as an initial benchmark for chronic pain trials using the BDI as an outcome measure are the extensive data available on the psychometric properties of the BDI and that this magnitude of change

would seem to reflect a clinically moderate benefit. Smaller magnitudes of improvement on the BDI than this have been found in the few pharmacological trials that have reported such data. Smaller magnitudes of change should therefore be investigated to determine minimally important differences to patients on the BDI.

### *Profile of Mood States*

The Profile of Mood States[92] (POMS) is a 65-item adjective checklist that provides a total mood disturbance score and 6 subscale scores: Tension, depression, anger, vigor, fatigue, and confusion. The POMS has been used in numerous studies of patients with a variety of painful conditions. Although the scales of both long[86] and short forms[31] show satisfactory internal consistency (Cronbach $\alpha$s = .63–.96), few studies have examined the test-retest reliability of the POMS other than for coefficients reported for psychiatric patients over periods of a few weeks.[92] The validity of the POMS has been established in multiple factor-analytic studies that confirm 6 dimensions, and the convergent and discriminant validity of the subscales has also been established.[62]

There are a number of challenges to the identification of criteria for clinically important changes for the POMS total and subscale scores for use in chronic pain clinical trials. Review of the existing literature indicates that multiple variations of the POMS are in use and that these vary in the number of items included, the time frame referenced, and the reporting of scores.[35] Few studies report descriptive statistics for the total and subscale scores used and the internal consistency or stability of the subscales is rarely examined. Additionally, the POMS has been used in many short-term (across hours or days) studies of different pain treatments and various other studies lack appropriate control groups. Finally, no studies have used anchor-based approaches—for example, patients' or clinicians' ratings of improvement—to identify clinically important changes on the POMS scales.

Several studies have examined the magnitude of treatment-associated change in POMS scores, and the results have generally demonstrated that mean improvement is approximately 18 points for the total score and ranges from 1 to 4 points for the POMS subscales.[62] Other studies have reported the differences between patients with and without psychiatric disorders or chronic pain, which is approximately 26 points for the total score and range from 3 to 5 points for the POM subscale scores. On the basis of only a few studies in patients with pain, the distributional characteristics of the POMS total and subscale scores can also be used to provide ranges for identifying important differences on this measure.[115,121,130]

Available data suggest that a change of 10 to 15 points on the POMS total score, which equals approximately one-half its standard deviation and 1 SEM, would be a reasonable benchmark for future studies designed to identify to minimally important change. Available data suggest that changes of 2 to 12 points for the specific POMS subscales equal approximately one-half the standard deviation and 1 SEM of these scales and are a reasonable benchmark for future studies designed to identify to minimally important changes.[62]

## Determining Clinically Important Differences in Global Ratings of Improvement

Global ratings of improvement or treatment satisfaction provide an opportunity for clinical trial participants to integrate into 1 overall evaluation the different aspects of their response to treatment, including pain relief, improvement in functioning, and side effects. Such measures can be used to investigate participants' judgments of the importance of changes in other outcome measures[46] and, as discussed above, have served as anchors in determining clinically important differences.

The Patient Global Impression of Change scale[55] (PGIC) was recommended by IMMPACT for use in chronic pain clinical trials as a core outcome measure of global improvement with treatment.[34] This single-item rating by participants of their response during a clinical trial uses a 7-point rating scale with the options "very much improved," "much improved," "minimally improved," "no change," "minimally worse," "much worse," and "very much worse." There has been widespread use of the PGIC in recent chronic pain clinical trials,[33,156] and the measure provides a responsive and readily interpretable assessment of participants' evaluations of the importance of their improvement or worsening.

The PGIC has been used as an anchor in determining the clinical importance of improvement in pain ratings[43] and other measures,[21] which assumes that the importance of the different patient ratings on this measure is self-evident. Ratings of "much" and "very much" improved (or worse) clearly reflect what patients consider to be important changes, and it appears likely that ratings of "minimally" improved (or worse) reflect changes that patients consider less substantial but minimally important. How important a minimal improvement or worsening is to patients must depend, at least in part, on factors such as treatment convenience and cost, as well as any aspects of the side effect burden that are not considered by patients in rating their overall change. When using the PGIC in a clinical trial, it is therefore recommended that the percentages of patients endorsing each of the 7 response options in each treatment group be analyzed and reported separately, and that ratings of "minimally improved" (or "minimally worse") not be combined with the other ratings of improvement (or worsening) or the ratings of "unchanged."

## Conclusions and Recommendations

Additional research is needed on clinically important changes in chronic pain outcomes. Relatively few studies have systematically asked patients with chronic pain to specifically identify the changes—both improvement and worsening—in pain intensity, HRQoL, and overall improvement that they consider important, and it is also unknown what changes patients consider noticeable but

not important. Such studies are a priority for research on chronic pain treatment and are consistent with recent recommendations that the first step in developing new outcome measures is to determine what patients themselves consider important.[144,149] Although of secondary significance, it is also unknown what changes in pain intensity, HRQoL, and overall improvement are considered important by clinicians. Such clinician evaluations of the magnitude of important improvement and worsening should also be determined and examined with respect to their potential to provide information that complements patient assessments.

In future research evaluating the clinical importance of chronic pain outcomes, the role of baseline status and patient characteristics, such as age, sex, and education, must be carefully considered. It is also important to evaluate whether the changes in chronic pain outcomes that patients consider clinically important vary depending on the specific clinical condition, for example, chronic musculoskeletal low back pain versus spinal cord injury pain. In addition, whether the magnitude of clinically important change depends on the direction of change—that is, improvement or worsening—must be examined. Finally, and perhaps most importantly, definitions of "clinical importance" must be provided that clearly specify whether minimally noticeable, minimally important, moderately important, substantial, or definitive changes are being examined.

Recent research in arthritis and other fields has begun to investigate definitions of "low disease activity state,"[154] "patient acceptable symptom state,"[140,141] and other approaches[68] to identifying what patients (and clinicians) would consider a substantial response to treatment given current treatment possibilities.[154] Considerable research has demonstrated that pain intensity ratings of 1 to 3 or 4 on a 0 to 10 NRS (ie, "mild pain") are associated with less interference with physical and emotional functioning than higher ratings (ie, "moderate" and "severe" pain).[4,124] Although reducing pain to a mild intensity would likely be considered a substantial response to treatment by both patients and clinicians, it is unknown whether current treatments for chronic pain can achieve this end point because few clinical trials have reported the percentages of patients whose pain decreased to this level with treatment. Analyses designed to identify such "treatment success" end points should be encouraged in future research on the clinical importance of chronic pain outcomes.

In addition, few chronic pain studies have used an individualized approach to identifying the outcomes that individual patients consider most important.[27,118,119] Existing approaches to the assessment of chronic pain outcomes evaluate the same domains across all patients, and research on methods and measures that allow patients to describe what is specifically important to them and to rank the importance of treatment outcomes should also be encouraged.

Several recommendations can be made about interpreting the clinical importance of changes in the specific measures[34] of the core chronic pain outcome domains[143] recommended by IMMPACT. These recommendations also apply to evaluating clinically important changes in existing measures of other outcome domains (eg, the IMMPACT supplemental outcome domains[143]) as well as in research conducted to develop new outcome measures.[143,149] There is an emerging consensus that combi-

**Table 1. Provisional Benchmarks for Interpreting Changes in Chronic Pain Clinical Trial Outcome Measures**

| *Outcome Domain and Measure* | *Type of Improvement** | *Method†* | *Change* |
|---|---|---|---|
| Pain intensity | | | |
| 0–10 numerical rating scale | Minimally important | Anchor | 10–20% decrease |
| | Moderately important | Anchor | ≥30% decrease |
| | Substantial | Anchor | ≥50% decrease |
| Physical functioning | | | |
| Multidimensional Pain Inventory | | | |
| Interference Scale | Clinically important | Distribution | ≥0.6 point decrease |
| Brief Pain Inventory | | | |
| Interference Scale | Minimally important | Distribution | 1 point decrease |
| Emotional functioning | | | |
| Beck Depression Inventory | Clinically important | Distribution | ≥5 point decrease |
| Profile of Mood States | | | |
| Total Mood Disturbance | Clinically important | Distribution | ≥10–15 point decrease |
| Specific subscales | Clinically important | Distribution | ≥2–12 point change‡ |
| Global rating of improvement | | | |
| Patient Global Impression of Change | Minimally important | Anchor | Minimally improved |
| | Moderately important | Anchor | Much improved |
| | Substantial | Anchor | Very much improved |

*Because few studies have examined the importance of worsening on these measures, benchmarks are only provided for improvement in scores.

†Specific method used in determining benchmark provided in final column; distribution-based methods were based on use of 0.5 standard deviation or 1.0 standard error of measurement or both.

‡The magnitude of a clinically important change depends on the specific subscale, as does the direction of change that reflects an improvement.

nations of anchor-based and distribution-based approaches should be used to identify clinically important changes,[20,22,29,30,37,126,127,165] with distribution-based methods providing supportive information to supplement the results obtained from generally more informative anchor-based methods.[111] It is therefore recommended that 2 or more different approaches be used to evaluate the clinical importance of improvement or worsening for chronic pain clinical trial outcome measures, ideally including at least 1 anchor-based method supplemented by distribution-based information. Such integrated approaches to developing criteria for clinically important changes can include the use of the Delphi process to determine clinician-based anchors.[13,161]

It is likely that there will be discrepancies when using multiple methods for determining clinical importance, and it is also recommended that approaches for reconciling these differences be specified in advance.[29,30] These may include identifying a range of clinically important differences,[32,61] in addition to or instead of fixed values. Whether the degree of change meets criteria for "definitely," "probably," "possibly," or "definitely not" clinically important can also be specified.[90] Finally, because of limitations in existing knowledge regarding the clinical importance of chronic pain outcomes, the benchmarks presented in Table 1 and these recommendations should not be considered a requirement for publication, submission of grant applications, or approval of product applications by regulatory agencies. Nevertheless, greater attention to interpreting the clinical importance of chronic pain clinical trial outcomes and reporting the types of information discussed in this article will provide more meaningful comparisons of chronic pain treatments.

## Acknowledgments

The authors thank Paul J. Lambiase and Mary Gleichauf for their invaluable assistance with IMMPACT.

## References

1. Acquadro C, Berzon R, Dubois D, Leidy NK, Marquis P, Revicki D, Rothman M: Incorporating the patient's perspective into drug development and communication: An ad hoc task force report of the Patient-Reported Outcomes (PRO) Harmonization group Meeting at the Food and Drug Administration, February 16, 2001. Value Health 6:522-531, 2003

2. Altmaier EM, Lehmann TR, Russell DW, Weinstein JN, Kao CF: The effectiveness of psychological interventions for the rehabilitation of low back pain: A randomized controlled trial evaluation. Pain 49:329-335, 1992

3. Anastasi A, Urbina S: Psychological Testing, (7th ed). Upper Saddle River, NJ, Prentice-Hall, 1997

4. Anderson KO: Role of cutpoints: Why grade pain intensity? Pain 113:5-6, 2005

5. Anderson KO, Syrjala KL, Cleeland CS: How to assess cancer pain, in Turk DC, Melzack R (eds): Handbook of Pain Assessment, (2nd ed). New York, NY, Guilford Press, 2001, pp 579-600

6. Atkinson JH, Slater MA, Williams RA, Zisook S, Patterson TL, Grant I, Wahlgren DR, Abramson I, Garfin SR: A placebo-controlled randomized clinical trial of nortriptyline for chronic low back pain. Pain 76:287-296, 1998

7. Beaton DE: Understanding the relevance of measured change through studies of responsiveness. Spine 25:3192-3199, 2000

8. Beaton DE, Bombardier C, Katz JN, Wright JG, Wells G, Boers M, Strand V, Shea B: Looking for important change/differences in studies of responsiveness. J Rheumatol 28: 400-405, 2001

9. Beck AT, Steer RA: Beck Depression Inventory. San Antonio, TX, Psychological Corporation, 1993

10. Beck AT, Ward CH, Mendelsohn M, Mock J, Erbaugh J: An inventory for measuring depression. Arch Gen Psychiatry 4:561-571, 1961

11. Beilin Y, Hossain S, Bodian CA: The numeric rating scale and labor epidural analgesia. Anesth Analg 96:1794-1798, 2003

12. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW: Validation study of WOMAC: A health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 15:1833-1840, 1988

13. Bellamy N, Carette S, Ford PM, Kean WF, le Riche NG, Lussier A, Wells GA, Campbell J: Osteoarthritis antirheumatic drug trials, III: Setting the delta for clinical trials—results of a consensus development (Delphi) exercise. J Rheumatol 19:451-457, 1992

14. Bellamy N, Carr A, Dougados M, Shea B, Wells G: Towards a definition of "difference" in osteoarthritis. J Rheumatol 28:427-430, 2001

15. Bergstrom KG, Jensen IB, Linton SJ, Nygren AL: A psychometric evaluation of the Swedish version of the Multidimensional Pain Inventory (MPI-S): A gender differentiation evaluation. Eur J Pain 3:261-273, 1999

16. Bernstein IH, Jaremko ME, Hinkley BS: On the utility of the West Haven-Yale Multidimensional Pain Inventory. Spine 8:956-963, 1995

17. Bird SB, Dickson EW: Clinically significant changes in pain along the visual analog scale. Ann Emerg Med 38:639-643, 2001

18. Bombardier C, Hayden J, Beaton DE: Minimal clinically important difference: Low back pain: Outcome measures. J Rheumatol 28:431-438, 2001

19. Cella D, Bullinger M, Scott C, Barofsky I: Group vs individual approaches to understanding the clinical significance of differences or changes in quality of life. Mayo Clin Proc 77:384-392, 2002a

20. Cella D, Eton DT, Fairclough DL, Bonomi P, Heyes AE, Silberman C, Wolf MK, Johnson DH: What is 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 55:285-295, 2002b

21. Cella D, Hahn EA, Dineen K: Meaningful change in cancer-specific quality of life scores: Differences between improvement and worsening. Qual Life Res 11:207-221, 2002

22. Cella D, Eton DT, Lai JS, Peterman AH, Merkel DE: Combining anchor and distribution-based methods to derive minimal clinically important differences on the Functional Assessment of Cancer Therapy (FACT) Anemia and Fatigue Scales. J Pain Symptom Manage 24:547-561, 2002d

23. Cepeda MS, Africano JM, Polo R, Alcala R, Carr DB: What decline in pain intensity is meaningful to patients with acute pain? Pain 105:151-157, 2003

24. Cleeland C, Mendoza T: The Brief Pain Inventory: meaningful changes in pain interference. Presented at the fourth meeting of the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT-IV); June 2004 [www.immpact.org/immpact4/Cleeland.pdf] Accessed June 1, 2007

25. 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 67:267-273, 1996

26. Cleeland CS, Ryan KM: Pain assessment: Global use of the Brief Pain Inventory. Ann Acad Med 23:129-138, 1994

27. Clinch J, Tugwell P, Wells G, Shea B: Individualized functional priority approach to the assessment of health related quality of life in rheumatology. J Rheumatol 28:445-451, 2001

28. Cohen J: Statistical Power Analysis for the Behavioral Sciences, (2nd ed). Hillsdale, NJ, Lawrence Erlbaum Associates, 1988

29. Crosby RD, Kolotkin RL, Williams GR: Defining clinically meaningful change in health-related quality of life. J Clin Epidemiol 56:395-407, 2003

30. Crosby RD, Kolotkin RL, Williams GR: An integrated method to determine meaningful changes in health-related quality of life. J Clin Epidemiol 57:1153-1160, 2004

31. Curran SL, Andrykowski MA, Studts JL: Short Form of the Profile of Mood States (POMS-SF): Psychometric information. Psychol Assess 7:80-83, 1995

32. de Vet HC, Terwee CB, Ostelo RW, Berkerman H, Knol DL, Bouter LM: Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. Health Qual Life Outcomes 4:54, 2006

33. Dworkin RH, Corbin AE, Young JP, Sharma U, LaMoreaux L, Bockbrader H, Garofalo EA, Poole RM: Pregabalin for the treatment of postherpetic neuralgia: A randomized, placebo-controlled trial. Neurology 60:1274-1283, 2003

34. 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 113:9-19, 2005

35. Edwards RR, Haythornthwaite J: Mood swings: Variability in the use of the Profile of Mood States. J Pain Symptom Manage 28:534, 2004

36. Endler NS, Corace KM, Summerfeldt LJ, Johnson JM, Rothbart P: Coping with chronic pain. Pers Indiv Differences 34:323-346, 2003

37. Eton DT, Cella D, Yost KJ, Yount SE, Peterman AH, Neuberg DS, Sledge GW, Wood WC: A combination of distribution- and anchor-based approaches determined minimally important differences (MIDs) for four end points in a breast cancer scale. J Clin Epidemiol 57:898-910, 2004

38. Fairbank JCT, Couper J, Davies JB, O'Brien JP: The Oswestry Low Back Pain Disability Questionnaire. Physiotherapy 66:271-273, 1980

39. Farrar JT: The global assessment of pain and related symptoms. Presented at the fourth meeting of the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT-IV); June 2004 [www.immpact.org/immpact4/Farrar.pdf] Accessed June 1, 2007

40. Farrar JT, Berlin JA, Strom BL: Clinically important changes in acute pain outcome measures: A validation study. J Pain Symptom Manage 25:406-411, 2003

41. Farrar JT, Dworkin RH, Max 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 31:369-377, 2006

42. Farrar JT, Portenoy RK, Berlin JA, Kinman JL, Strom BL: Defining the clinically important difference in pain outcome measures. Pain 88:287-294, 2000

43. Farrar JT, Young JP, LaMoreaux L, Werth JL, Poole RM: Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. Pain 94:149-158, 2001

44. Ferrari R, Novara C, Sanavio E, Zerbini F: Internal structure and validity of the Multidimensional Pain Inventory, Italian language version. Pain Med 1:123-130, 2000

45. Finnerup NB, Otto M, Jensen TS, Sindrup SH: Algorithm for neuropathic pain treatment: An evidence based proposal. Pain 118:289-305, 2005

46. 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 282:1157-1162, 1999

47. Flor H, Rudy TE, Birbaumer N, Schugens MM: Zur anwendbarkeit des West Haven-Yale Multidimensional Pain Inventory im Deutschen sprachraum. Der Schmerz 4:82-87, 1990

48. Forouzanfar T, Weber WEJ, Kemler M, van Kleef M: What is a meaningful pain reduction in patients with complex regional pain syndrome type 1? Clin J Pain 19:281-285, 2003

49. Frost MH, Bonomi AE, Ferrans CE, Wong GY, Hays RD: Patient, clinician, and population perspectives on determining the clinical significance of quality of life scores. Mayo Clin Proc 77:488-494, 2002

50. Gallagher EJ, Bijur PE, Latimer C, Silver W: Reliability and validity of a Visual Analog Scale for acute abdominal pain in the ED. Am J Emerg Med 20:287-290, 2002

51. Gallagher EJ, Liebman M, Bijur PE: Prospective validation of clinically important changes in pain severity measured on a visual analog scale. Ann Emerg Med 38:633-638, 2001

52. Geisser ME, Roth RS, Robinson ME: Assessing depression among persons with chronic pain using the Center for Epidemiological Studies—Depression Scale and the Beck Depression Inventory: A comparative analysis. Clin J Pain 13: 163-170, 1997

53. Geurts JWM, van Wijk RMAW, Wynne H, Hammink E, Buskens E, Lousberg R, Knape JTA, Groen GJ: Radiofre-

quency lesioning of dorsal root ganglia for chronic lumbosacral radicular pain: A randomized, double-blind, controlled trial. Lancet 361:21-26, 2003

54. Greco CM, Rudy TE, Manzi S: Adaptation to chronic pain in systemic lupus erythematosus: Applicability of the Multidimensional Pain Inventory. Pain Med 4:39-50, 2003

55. Guy W: ECDEU assessment manual for psychopharmacology (DHEW Publication No. ADM 76-338). Washington, DC, U.S. Government Printing Office, 1976

56. Guyatt GH: Making sense of quality-of-life data. Med Care 38(suppl II):II-175-179, 2000

57. Guyatt GH, Norman GR, Juniper EF, Griffith LE: A critical look at transition ratings. J Clin Epidemiol 55:900-908, 2002

58. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR: Methods to explain the clinical significance of health status measures. Mayo Clin Proc 77:371-383, 2002

59. Hägg O, Fritzell P, Nordwall A: The clinical importance of changes in outcome scores after treatment for chronic low back pain. Eur Spine J 12:12-20, 2003

60. Hanley MA, Jensen MP, Ehde DM, Robinson LR, Cardenas DD, Turner JA, Smith DG: Clinically significant changes in pain intensity ratings in persons with spinal cord injury or amputation. Clin J Pain 22:25-31, 2006

61. Hays RD, Woolley JM: The concept of clinically meaningful difference in health-related quality-of-life research: How meaningful is it? Pharmacoeconomics 18:419-423, 2000

62. Haythornthwaite JA, Edwards RR: Profile of Mood States (POMS). Presented at the fourth meeting of the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT-IV); June 2004 [www.immpact.org/immpact4/Haythornthwaite.pdf] Accessed June 1, 2007

63. Heald SL, Riddle DL, Lamb RL: The shoulder pain and disability index: The construct validity and responsiveness of a region-specific disability measure. Phys Ther 77:1079-1089, 1997

64. Hoffman BM, Papas RK, Chatkoff DK, Kerns RD: Meta-analysis of psychological interventions for chronic back pain. Health Psychol 25:1-9, 2007

65. Holdgate A, Asha S, Graig J, Thompson J: Comparison of a verbal numeric rating scale with the visual analogue scale for the measurement of acute pain. Emerg Med 15:441-446, 2003

66. Huff D: How to Lie With Statistics. New York, NY, W.W. Norton & Co, 1954

67. Huse E, Larbig W, Flor H, Birbaumer N: The effect of opioids on phantom limb pain and cortical reorganization. Pain 90:47-55, 2001

68. Jacobson NS, Roberts LJ, Berns SB, McGlinchey JB: Methods for defining and determining the clinical significance of treatment effects: Description, application, and alternatives. J Consult Clin Psychol 67:300-307, 1999

69. Jaeschke R, Singer J, Guyatt GH: Measurement of health status: Ascertaining the minimal clinically important difference. Control Clin Trials 10:407-415, 1989

70. Jenkins DG, Ebbutt AF, Evans CD: Tofranil in the treatment of low back pain. J Int Med Res 4:28-40, 1976

71. Jensen MP, Chen C, Brugger AM: Interpretation of visual analog scale ratings and change scores: A reanalysis of two clinical trials of postoperative pain. J Pain 4:407-414, 2003

72. Jensen MP, Turner JA, Romano JM: Correlates of improvement in multidisciplinary treatment of chronic pain. J Consult Clin Psychol 62:172-179, 1994

73. Johansson E, Lindberg P: Low back pain patients in primary care: Subgroups based on the Multidimensional Pain Inventory. Int J Behav Med 7:340-352, 2000

74. Johansson C, Dahl JA, Jannert M, Melin L, Andersson G: Effects of a cognitive-behavioral pain-management program. Behav Res Ther 36:915-930, 1998

75. Juniper EF, Guyatt GH, Willan A, Griffith LE: Determining a minimal important change in a disease-specific quality of life questionnaire. J Clin Epidemiol 47:81-87, 1994

76. Kazdin AE: The meanings and measurement of clinical significance. J Consult Clin Psychol 67:332-339, 1999

77. Kazis LE, Anderson JJ, Meenan RF: Effect sizes for interpreting changes in health status. Med Care 27:S178-S189, 1989

78. Kelly AM: Does the clinically significant difference in visual analog scale pain scores vary with gender, age, or cause of pain? Acad Emerg Med 5:1086-1090, 1998

79. Kelly AM: The minimum clinically significant difference in visual analogue scale pain score does not differ with severity of pain. Emerg Med J 18:205-207, 2001

80. Kerns RD: Beck Depression Inventory: Minimum important difference. Presented at the fourth meeting of the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT-IV); June 2004 [www.immpact.org/immpact4/Kerns.pdf] Accessed June 1, 2007

81. Kerns RD, Turk DC, Rudy TE: The West Haven-Yale Multidimensional Pain Inventory (WHYMPI). Pain 23:345-356, 1985

82. King SJ, Wessel J, Bhambhani Y, Sholter D, Maksymowych W: Predictors of success of intervention programs for persons with fibromyalgia. J Rheumatol 29:1034-1040, 2002

83. Kosinski M, Zhao SZ, Dedhiya S, Osterhaus JT, Ware JE Jr: Determining minimally important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. Arthritis Rheum 43: 1478-1487, 2000

84. Kraemer HC, Morgan GA, Leech NL, Gliner JA, Vaske JJ, Harmon RJ: Measures of clinical significance. J Am Acad Child Adolesc Psychiatry 42:1524-1529, 2003

85. Lai JS, Dineen K, Cella D, Roenn J: Can an item response theory–based pain item bank enhance measurement precision? Clin Ther 25:D34-D35, 2003

86. Lin CC, Lai YL, Ward SE: Effect of cancer pain on performance status, mood states, and level of hope among Taiwanese cancer patients. J Pain Symptom Manage 25:29-37, 2003

87. Lipscomb J, Snyder CF: The outcomes of cancer outcomes research: Focusing on the National Cancer Institute's quality-of-care initiative. Med Care 40:III3-III10, 2002

88. Lousberg R, Van Breukelen GJP, Schmidt AJM, Arnts A, Winter FAM: Psychometric properties of the Multidimensional Pain Inventory, Dutch Language version (MPI-DLV). Behav Res Ther 37:167-182, 1999

89. Lydick E, Epstein RS: Interpretation of quality of life changes. Qual Life Res 2:221-226, 1993

90. Man-Son-Hing M, Laupacis A, O'Rourke K, Molnar FJ, Mahon J, Chan KBY, Wells G: Determination of the clinical

importance of study results: A review. J Gen Intern Med 17:469-476, 2002

91. Marhold C, Linton SJ, Melin L: A cognitive-behavioral return-to-work program: Effects on pain patients with a history of long-term versus short-term sick leave. Pain 91:155-163, 2001

92. McNair DM, Lorr M, Droppleman LF: Profile of Mood States (POMS) Manual. San Diego, CA, Educational and Industrial Testing Service, 1992

93. McQuay H, Moore A: An Evidence-based Resource for Pain Relief. New York, NY, Oxford University Press, 1998

94. 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 10:353-361, 2006

95. Miller GA: The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychol Rev 63:81-97, 1956

96. Montoya P, Sitges C, Garcia-Herrera M, Izquierdo R, Truyols M, Blay N, Collado D: Abnormal affective modulation of somatosensory brain processing among patients with fibromyalgia. Psychosom Med 67:957-963, 2005

97. Moore RA, Gavaghan D, Tramèr MR, Collins SL, McQuay HJ: Size is everything: Large amounts of information are needed to overcome random effects in estimating direction and magnitude of treatment effects. Pain 78:209-216, 1998

98. Morley S, Eccleston C, Williams A: Systematic review and meta-analysis of randomized controlled trails of cognitive behavior therapy and behavior therapy for chronic pain in adults, excluding headache. Pain 80:1-13, 1999

99. Morley S, Williams AC, Black S: A confirmatory factor analysis of the Beck Depression Inventory in chronic pain. Pain 99:289-298, 2002

100. Nicholas MK, Wilson PH, Goyen J: Operant-behavioral and cognitive-behavioral treatment for chronic low back pain. Pain 48:339-347, 1991

101. 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 41:582-592, 2003

102. Norman GR, Stratford P, Regehr G: Methodological problems in the retrospective computation of responsiveness to change: The lessons of Cronbach. J Clin Epidemiol 50:869-879, 1997

103. Nunnally J, Bernstein I: Psychometric Theory, (3rd ed). New York, NY, McGraw Hill, 1994

104. 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 32: 217-229, 2006

105. Pincus T, Amara I, Koch GG: Continuous indices of core data set measures in rheumatoid arthritis clinical trials. Arthritis Rheum 52:1031-1036, 2005

106. Raichle KA, Osborne TL, Jensen MP, Cardenas D: The reliability and validity of pain interference measures in persons with spinal cord injury. J Pain 7:179-186, 2006

107. Redelmeier DA, Bayoumi AM, Goldstein RS, Guyatt GH: Interpreting small differences in functional status: The Six Minute Walk test in chronic lung disease patients. Am J Respir Crit Care Med 155:1278-1282, 1997

108. Redelmeier DA, Goldstein RS, Min ST, Hyland RH: Spirometry and dyspnea in patients with COPD. Chest 109: 1163-1168, 1996a

109. Redelmeier DA, Guyatt GH, Goldstein RS: Assessing the minimal important difference in symptoms: A comparison of two techniques. J Clin Epidemiol 49:1215-1219, 1996b

110. Reitsma B, Meijler WJ: Pain and patienthood. Clin J Pain 13:9-21, 1997

111. Revicki DA, Cella D, Hays RD, Sloan JA, Lenderking WR, Aaronson NK: Responsiveness and minimal important differences for patient reported outcomes. Health Qual Life Outcomes 4:70, 2006

112. Robinson ME, Brown JL, George SZ, Edwards PS, Atchison JW, Hirsh AT, Waxenberg LB, Wittmer V, Fillingim RB: Multidimensional success criteria and expectations for treatment of chronic pain: The patient perspective. Pain Med 6:336-345, 2005

113. Roland M, Morris R: A study of the natural history of back pain, part I: Development of a reliable and sensitive measure of disability in low-back pain. Spine 8:141-144, 1983

114. Ross M: Relation of implicit theories to the construction of personal histories. Psychol Rev 96:341-357, 1989

115. Rowbotham M, Harden N, Stacey B, Bernstein P, Magnus-Miller L: Gabapentin for the treatment of postherpetic neuralgia: A randomized controlled trial. JAMA 280:1837-1842, 1998

116. Rowbotham MC, Twilling L, Davies PS, Reisner L, Taylor K, Mohr D: Oral opioid therapy for chronic peripheral and central neuropathic pain. N Engl J Med 348:1223-1232, 2003

117. Rudy TE: Multidimensional Pain Inventory Multiaxial Assessment of Pain: Computer Program and User's Manual, Version 1.0. Pittsburgh, PA, University of Pittsburgh, 1987

118. Ruta DA, Garratt AM, Leng M, Russell IT, MacDonald LM: A new approach to the measurement of quality of life: The Patient Generated Index. Med Care 32:1109-1126, 1994

119. Ruta DA, Garratt AM, Russell IT: Patient centered assessment of quality of life for patients with four common conditions. Qual Health Care 8:22-29, 1999

120. 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 283-291, 2004

121. Sator-Katzenschlager SM, Schiesser AW, Kozek-Langenecker SA, Benetka G, Langer G, Kress HG: Does pain relief improve pain behavior and mood in chronic pain patients? Anesth Analg 97:791-797, 2003

122. Scharff L, Turk DC, Marcus DA: Psychosocial and behavioral characteristics in chronic headache patients: Support for a continuum and dual-diagnostic approach. Cephalalgia 15:216-223, 1995

123. Schwartz AL, Meek PM, Nail LM, Fargo J, Lundquist M, Donofrio M, Grainger M, Throckmorton T, Mateo M: Measurement of fatigue: Determining minimally important clinical differences. J Clin Epidemiol 55:239-244, 2002

124. Serlin RC, Mendoza TR, Nakamua Y, Edwards KR, Cleeland CS: When is cancer pain mild, moderate or severe? Grading pain severity by its interference with function. Pain 61:277-284, 1995

125. Slater MA, Doctor JN, Pruitt SD, Atkinson JH: The clinical significance of behavioral treatment for chronic low

back pain: An evaluation of effectiveness. Pain 71:257-263, 1997

126. Sloan JA, Cella D, Hays RD: Clinical significance of patient-reported questionnaire data: Another step toward consensus. J Clin Epidemiol 58:1217-1219, 2005

127. Sloan J, Symonds T, Vargas-Chanes D, Fridley B: Practical guidelines for assessing the clinical significance of health-related quality of life changes within clinical trials. Drug Inf J 37:23-31, 2003

128. Smith WB, Safer MA: Effects of present pain level on recall of chronic pain and medication use. Pain 55:355-361, 1993

129. Smith WB, Gracely RH, Safer MA: The meaning of pain: Cancer patients' rating and recall of pain intensity and affect. Pain 78:123-129, 1998

130. Speca M, Carlson LE, Goodey E, Angen M: A randomized, wait-list controlled clinical trial: The effect of a mindfulness meditation-based stress reduction program on mood and symptoms of stress in cancer outpatients. Psychosom Med 62:613-622, 2000

131. Sterner Y, Lofgren M, Nyberg V, Karlsson AK, Bergstrom M, Gerdle B: Early interdisciplinary rehabilitation programme for whiplash associated disorders. Disabil Rehabil 23:422-429, 2001

132. Stratford PW, Binkley J, Solomon P, Finch E, Gill C, Moreland J: Defining the minimal level of detectable change for the Roland Morris questionnaire. Phys Ther 76:359-365, 1996

133. Stucki G, Daltroy L, Katz JN, Johannesson M, Liang MH: Interpretation of change scores in ordinal clinical scales and health status measures: The whole may not equal the sum of the parts. J Clin Epidemiol 49:711-717, 1996

134. Symonds T, Berzon R, Marquis P, Rummans TA: The clinical significance of quality-of-life results: Practical considerations for specific audiences. Mayo Clin Proc 77:572-583, 2002

135. ten Klooster PM, Drossaers-Bakker KW, Taal E, van de Laar MAFJ: Patient-perceived satisfactory improvement (PPSI): Interpreting meaningful change in pain from the patient's perspective. Pain 121:151-157, 2006

136. Testa MA: Interpretation of quality-of-life outcomes: issues that affect magnitude and meaning. Med Care 38(suppl II):II-166-II-74, 2000

137. Thieme K, Gromnica-Ihle E, Flor H: Operant behavioral treatment of fibromyalgia: A controlled study. Arthritis Care Res 49:314-320, 2003

138. Todd KH, Funk KG, Funk JP, Bonacci R: Clinical significance of reported changes in pain severity. Ann Emerg Med 27:485-489, 1996

139. Todd KH: Patient-oriented outcome measures: The promise of definition. Ann Emerg Med 38:672-674, 2001

140. Tubach F, Dougados M, Falissard B, Baron G, Logeart I, Ravaud P: Feeling good rather than feeling better matters more to patients. Arthritis Rheum 55:526-530, 2006

141. Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, Bombardier C, Felson D, Hochberg M, van der Heijde D, Dougados M: Evaluation of clinically relevant states in patient reported outcomes in knee and hip osteoarthritis: The patient acceptable symptom state. Ann Rheum Dis 64:34-37, 2005

142. Turk DC: Multidimensional Pain Inventory Interference Scale. Presented at the fourth meeting of the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT-IV); June 2004 [www.immpact.org/immpact4/Turk.pdf] Accessed June 1, 2007

143. Turk DC, Dworkin RH, Allen RR, Bellamy N, Brandenburg N, Carr DB, Cleeland C, Dionne R, Farrar JT, Galer BS, Hewitt DJ, Jadad AR, Katz NP, Kramer LD, Manning DC, McCormick CG, McDermott MP, McGrath P, Quessy S, Rappaport BA, Robinson JP, Royal MA, Simon L, Stauffer JW, Stein W, Tollett J, Witter J: Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. Pain 106:337-345, 2003

144. Turk DC, Dworkin RH, Burke LB, Gershon R, Rothman M, Scott J, Allen RR, Atkinson JH, Chandler J, Cleeland C, Cowan P, Dimitrova R, Dionne R, Farrar JT, Haythornthwaite JA, Hertz S, Jadad AR, Jensen MP, Kellstein D, Kerns RD, Manning DC, Martin S, Max MB, McDermott MP, McGrath P, Moulin DE, Nurmikko T, Quessy S, Raja S, Rappaport BA, Rauschkolb C, Robinson JP, Royal MA, Simon L, Stauffer JW, Stucki G, Tollett J, von Stein T, Wallace MS, Wernicke J, White RE, Williams AC, Witter J, Wyrwich KW: Developing outcome measures for pain clinical trials: IMMPACT recommendations. Pain 125:208-215, 2006

145. Turk DC, Okifuji A, Sinclair JD, Starz TW: Pain, disability, and physical functioning in subgroups of fibromyalgia patients. J Rheumatology 23:1255-1262, 1996

146. Turk DC, Rudy TE: Robustness of an empirically derived taxonomy of chronic pain patients. Pain 43:27-36, 1990

147. Turk DC, Sist TC, Okifuji A, Miner MF, Florio G, Harrison P, Massey J, Lema ML, Zevon MA: Adaptation to metastatic cancer pain, regional/local cancer pain, and non-cancer pain: Role of psychological and behavioral factors. Pain 73: 247-256, 1998

148. Turner-Stokes L, Erkeller-Yuksel F, Milees A, Pincus T, Shipley M, Pearce S: Outpatient cognitive-behavioral pain management programs: A randomized comparison of a group-based multidisciplinary versus an individual therapy model. Arch Phys Med Rehabil 84:781-788, 2003

149. U.S. Department of Health and Human Services, Food and Drug Administration. Guidance for industry: Patient-reported outcome measures: Use in medical product development to support labeling claims, February 2006 [http://www.fda.gov/CDER/GUIDANCE/5460dft.pdf] Accessed June 1, 2007

150. Van Kleef M, Liem L, Lousberg R, Barendse G, Kessels F: Radiofrequency lesion adjacent to the dorsal root ganglion for cervicobrachial pain: A prospective double blind randomized study. Neurosurgery 38:1127-1131, 1996

151. Vlaeyen JW, Haazen IW, Schuerman JA, Kole-Snijders AM, van Eek H: Behavioural rehabilitation of chronic low back pain: Comparison of an operant treatment, an operant-cognitive treatment, and an operant-respondent treatment. Br J Clin Psychol 34:95-118, 1995

152. Weiner DK, Rudy TE, Glick RM, Boston JR, Lieber SJ, Morrow LA, Taylor S: Efficacy of percutaneous electrical nerve stimulation for the treatment of chronic low back pain in older adults. J Am Geriatr Soc 51:599-608, 2003

153. Wells G, Beaton D, Shea B, Boers M, Simon L, Strand V, Brooks P, Tugwell P: Minimal clinically important differences: Review of methods. J Rheumatol 28:406-412, 2001

154. Wells G, Boers M, Tugwell P: Low disease activity state in rheumatoid arthritis: Concepts and derivation of minimal disease activity. Clin Exp Rheumatol 24(suppl 43):S52-S59, 2006

155. Wells GA, Tugwell P, Kraag GR, Baker PRA, Groh J, Redelmeier DA: Minimum important difference between patients with rheumatoid arthritis: The patient's perspective. J Rheumatol 20:557-560, 1993

156. Wernicke JF, Pritchett YL, D'Souza DN, Waninger A, Tran P, Iyengar S, Raskin J: A randomized controlled trial of duloxetine in diabetic peripheral neuropathic pain. Neurology 67:1411-1420, 2006

157. Widar M, Ahlstrom G: Pain in persons with post-polio: The Swedish version of the Multidimensional Pain Inventory (MPI). Scand J Caring Sci 13:33-41, 1999

158. Wolinsky FD, Wan GJ, Tierney WM: Changes in the SF-36 in 12 months in a clinical sample of disadvantaged older adults. Med Care 36:1589-1598, 1998

159. Wright JG: Interpreting health-related quality of life scores: The simple rule of seven may not be so simple. Med Care 41:597-598, 2003

160. Wyrwich KW, Bullinger M, Aaronson N, Hays RD, Patrick DL, Symonds T: Estimating clinically significant differences in quality of life outcomes. Qual Life Res 14:285-295, 2005

161. Wyrwich KW, Finn SD, Tierney WM, Kroenke K, Babu AN, Wolinsky FM: Clinically important changes in health-related quality of life for patients with chronic obstructive pulmonary disease: An expert consensus panel report. J Gen Intern Med 18:196-202, 2003

162. Wyrwich KW, Nienaber NA, Tierney WM, Wolinsky FD: Linking clinical relevance and statistical significance in evaluating intra-individual changes in health related quality of life. Med Care 37:469-478, 1999

163. Wyrwich KW, Tierney WM, Wolinsky FD: Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. J Clin Epidemiol 52:861-873, 1999

164. Wyrwich KW, Wolinsky FD: Identifying meaningful intra-individual change standards for health-related quality of life measures. J Eval Clin Pract 6:39-49, 2000

165. Yost KJ, Cella D, Chawla A, Holmgren E, Eton DT, Ayanian JZ, West DW: Minimally important differences were estimated for the Functional Assessment of Cancer Therapy—Colorectal (FACT-C) instrument using a combination of distribution- and anchor-based approaches. J Clin Epidemiol 58:1241-1251, 2005