Published in final edited form as:

J Pain. 2007 December ; 8(12): 906–913. doi:10.1016/j.jpain.2007.06.009.

A systematic review of measures used to assess chronic musculoskeletal pain in clinical and randomized controlled clinical trials

Leighann Litcher-Kelly¹, Sharon A. Martino², Joan E. Broderick¹, and Arthur A. Stone¹

¹*Applied Behavioral Medicine Research Institute, Department of Psychiatry and Behavioral Science, Stony Brook University*

²*School of Health Technology and Management, Department of Physical Therapy, Stony Brook University*

Abstract

There are many types of pain assessments available to researchers conducting clinical trials, ranging from simple, single-item Visual Analog Scale (VAS) questions through extensive, multidimensional inventories. The primary question addressed in this survey of top-tier medical journals was: Which pain assessments are most commonly used in trials? Articles addressing chronic musculoskeletal pain in clinical trials were identified in seven major medical journals for the year 2003. A total of 50 studies (total original research articles reviewed: 1,476) met selection criteria, and from these we identified 28 types of pain assessments. Selected studies were classified according to the dimensions of pain assessed, the type of scale and descriptors/anchors used, and the reporting period specified. The most frequently used assessments were the single-item VAS scale and the Numeric Rating Scale (NRS); multidimensional inventories were used infrequently. There was considerable variability in the instructions patients received about the period to consider when evaluating their pain, and many studies provided only cursory information about their assessments in the methods. Overall, it appears that clinical trials utilize simple measures of pain and that there is no widely accepted standard for clinical pain assessment that would facilitate comparison of outcomes across trials.

Perspective—This review highlights the heterogeneity of pain outcome measures used and the abundance of single-item measures in clinical trials. While there are many pain outcome measures available to clinical researchers, more consistency in the field should be encouraged so that results between studies can be compared.

Keywords

Pain; Self-report; Patient Reported Outcomes; Clinical Trials

Pain is a leading public health problem in the United States affecting over 50 million Americans at an annual cost exceeding \$100 billion dollars.⁴⁶ This translates into 70 million healthcare visits a year, making pain the leading cause for health care utilization. Fifteen percent of adults in the United States have had chronic low back pain at some point in their lives, resulting in 250 million lost days of work, and conferment of disability to over 10 million Americans.²⁰

Corresponding author: Leighann Litcher-Kelly, Putnam Hall, South Campus, Stony Brook, NY 11974-8790, Phone: 631-632-3094, Fax: 631-632-3165, E-mail: Leighann.Litcher@sunysb.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

This staggering figure only accounts for one of many chronic pain diagnoses. The American Pain Society and the Joint Commission on Accreditation of Healthcare Organizations have designated pain as the “fifth vital sign” in an effort to enhance the awareness and need of assessment.¹⁸

Pain assessments rely on self-report measures intended to quantify qualities of pain like intensity, sensory characteristics, affective responses, and coping.³⁰ Many multi-dimensional measures are available including the McGill Pain Questionnaire (MPQ40) and the Western Ontario and MacMaster Universities Osteoarthritis Index (WOMAC2). Alternatively, simpler single-item measures are available, such as the Visual Analog Scale (VAS) and Numeric Rating Scale (NRS), which are meant to measure particular qualities of pain such as intensity or unpleasantness. A great deal of information is available about the psychometric qualities of these questions and scales, and many are considered reliable and valid (for a review see⁶²). It is also noteworthy that researchers are actively pursuing more accurate alternatives for the measurement of pain including advanced scaling,²¹ real-time assessments,⁵⁹ and alternatives to self-report.^{31,43} In addition, initiatives in the pain research community have recently generated recommendations on the types of outcome measures that should be used in future research.¹²

Current review

With the large number of options available to clinical researchers, our interest was in determining which self-report pain measures were employed in clinical trials. One of our motivations for posing this question is based on the possible disconnect between the availability of sophisticated, well-validated, multidimensional instruments and the instruments that are actually used in trials. Our impression of this literature is that simple assessments (e.g., one-item questions addressing pain intensity) are often used as outcomes in important clinical trials. Knowing that casual impressions can be mistaken, we conducted a systemic review of the trials published in 2003 that used pain as an outcome.

Because randomized controlled clinical trials (RCTs) and clinical trials (CTs) are used to determine clinical effectiveness of treatments and are, therefore, particularly influential studies, we chose to review pain outcome measures used in RCT and CT research published in several leading medical journals. We assumed these journals would be indicative of best research practice. We limited our review to studies of patients with chronic pain and only included studies that examined chronic musculoskeletal, non-malignant pain to prevent a high degree of heterogeneity in the measures observed.

Methods

A systematic review was conducted for all original articles in three types of medical journals (pain, general medicine, and musculoskeletal disease) for the year 2003. Our goal was to choose top-tier journals for each the three groups, thus, we reviewed the AIM (Abridged Index Medicus) Core Clinical Journal list and the impact factors (from 2004) for each journal. For the pain and general medicine journals, we chose the two most influential in their respective fields, and for the disease-specific journals, we chose those with high impact factors that focused on musculoskeletal, non-malignant, non-headache pain. Thus, the seven journals reviewed and their impact factors were *Journal of Pain* (impact factor 2.0), *Pain* (4.1), *Journal of the American Medical Association (JAMA)* (24.8), *New England Journal of Medicine (NEJM)* (38.57), *Arthritis and Rheumatism* (7.4), *Rheumatology* (4.1), and *Spine* (2.7).

Inclusion Criteria

All original research articles in these journals for the year 2003 were reviewed. The following inclusion criteria were required: (1) a randomized, controlled clinical trial (RCT; patients were randomized to treatment arms) or a clinical trial (CT; prospective treatment study where patients were non-randomly assigned to one or more groups), or a secondary analysis of an RCT or CT (2) participants were adults diagnosed with chronic (more than 12 weeks) musculoskeletal, non-malignant pain and (3) self-reported pain was an outcome (not necessarily the primary outcome).

Review of articles

The first and second authors conducted preliminary screening for inclusion criteria and coded all articles that were eligible. To ensure consistency of article coding, the first and second authors each reviewed several issues of *Pain* and found that there were very few discrepancies in their results. The articles were then all independently coded; the few discrepancies were resolved by re-reviewing the article together.

Information was coded about study characteristics and about the pain measure(s) employed. For each study, we coded: 1) the duration of the study protocol; 2) the chronic pain diagnosis of the sample; 3) the length of time patients were in pain or had a diagnosed chronic pain condition; 4) the definition of chronic pain reported in the study; and, 5) the outcome measure (s) used to assess pain. The following information was coded for each pain outcome: 1) the type of scaling used (e.g., Visual Analog Scale, Numeric Rating Scale); 2) the response options for the scale (e.g., “no pain” to “most pain” as the endpoints for a VAS); 3) the type of quality of pain assessed (e.g., pain intensity or level of usual pain); 4) the reporting period assessed for making the rating (e.g., right now, or over the last week); 5) whether or not the measure is a standard validated measure (e.g., the MPQ, the WOMAC); and, 6) whether or not the article reports/uses the measure as it was originally developed or was revised to meet the needs of the authors' protocol. If the measure used was a standard measure (i.e., MPQ) and the article did not provide information to the contrary, we assumed it was used as designed.

Results

Characteristics of Eligible Studies

From a total of 1476 original research articles reviewed, 50 met our eligibility criteria (see Table 1 for a description of article selection and distribution of articles by journal).^{1,5-8,10,11,13-17,19,22-29,33-38,41,44,45,47-58,60,61,63-68} Of the 50 studies selected for review, 66% were RCTs or secondary analyses of RCTs; the other 34% were CTs. The duration of the study protocol was less than 6 months for almost half (42% of the 50 eligible studies), and it was 6 months to 12 months for 20%, >12 months to 2 years for 26%, and > 2 years for 10% (one study did not report the duration of the protocol). The most frequent diagnoses in these studies were chronic low back pain (n=13; 26%), rheumatoid arthritis (n=9; 18%), chronic pain not specified (n=7; 14%), osteoarthritis (n=5; 10%), fibromyalgia (n=3; 6%), and thirteen other diagnoses that were reported by only one study. The mean duration of time patients were in pain (or with a chronic pain diagnosis) was greater than 3 years for 48% of the studies. With regard to the definition of chronic pain, only 23 of 50 (46%) studies provided descriptions. Of these, 9 used American College of Rheumatology (ACR) criteria for rheumatoid arthritis, osteoarthritis, or fibromyalgia, 4 used a criterion of pain for at least 3 months, and 2 used pain for at least 6 months. No other definitions of chronic pain were used more than once.

Characteristics of pain measures

Forty-four percent of the 50 eligible studies used one measure to assess pain as an outcome, 36% used two measures, 12% used three measures, and 8% used four measures. The majority of the eligible studies included pain as their primary outcome measure (64%). The pain outcome measures that were coded from the 50 studies were broken into 28 distinct classifications for measuring pain (e.g., VAS, McGill Pain Questionnaire, Pain Frequency). Table 2 lists each of the measures and their frequencies. Some articles did not provide sufficient information to determine under which classification it should be grouped. For example, when a measure was simply described as a “5-point categorical scale,” it may have been a NRS, but could have been another type of scale (i.e., VRS). We always coded scales as specifically as possible, but when in doubt, we created a separate category that was exactly as the authors report it in the methods section of the article (in the above example, the measure was coded as a “5-point categorical scale”).

Table 2 also classifies measures as single-item, multi-item or multidimensional. In order to determine which category to place each measure, we used the information provided in the methods section as well as the citation used for the measure (when provided). As the name implies, single-item measures consist of only one item and only measure one dimension of pain. Examples are VAS, Verbal Rating Scales (VRS), Numeric Rating Scales (NRS) and single-item categorical variables. Similarly, multi-item measures also only measure one dimension of pain, but use more than one item to assess this construct. An example is the Roland-Morris Disability Questionnaire which measures pain-related disability. Another example of a multi-item measure is the SF-36 pain subscale since it is only assessing one dimension of pain (namely bodily pain). One article stated in the methods section that Pain Frequency was measured using the sum of four items and we categorized this as multi-item. Finally, we also categorized measures as multidimensional if they assessed more than one dimension of pain. For example, the MPQ assesses sensory and affective qualities of pain along with pain intensity. There were measures that we couldn't categorize based on the information in the methods section and the citations, and we include them under the heading ‘unknown.’

Table 2 demonstrates the propensity for using single-item measures to assess pain in clinical trials. Over half (54%) of all the pain outcome measures we coded from the 50 eligible studies were single-item measures, and the vast majority were VASs. In contrast only 16% were multidimensional scales such as the MPQ, and 27% were multi-item scales that measured one dimension of pain (i.e. the Neck Disability Index). Note that one article reported using only one of the pain subscales from the multidimensional pain scale, the MPI and it was unclear whether the whole measure was administered. For this case, we classified it as “one subscale from a multidimensional pain measure.” We were unable to classify two measures in terms of single-item, multi-item, or multidimensional because not enough information was provided within the article (one we assumed was multi-item; see footnote for Table 2). Finally, three articles reported they assessed the “Pain Frequency,” however, the three assessed this construct in different ways and, thus, it appears in different categories in Table 2. One article stated that they assessed pain frequency using a single item, another wrote that it was the sum of 4 6-point Likert questions and the last did not give enough information about the measure for us to determine which category was appropriate.

Visual Analog Scale (VAS)

The pain outcome measure most frequently used in the 50 eligible studies was the VAS, with over half of the studies (60%) including a VAS to measure pain as an outcome in their trial. Remarkably, the VAS was used as the only pain outcome measure in 20% of the 50 studies. Since it was the most frequently used measure, we have provided details about the characteristics of the VAS reported in the studies. Table 3 presents the breakdown of the

dimensions assessed, the description of the scale provided, the absence or presence of scale anchors, and the reporting periods for the VAS measures.

Discussion

The purpose of this review was to increase understanding about how pain has been assessed in recent clinical trials. To accomplish this task, we reviewed over 1476 papers published in 7 major medical journals for a single year. Fifty studies used a pain measure as an outcome in the context of a clinical trial (RCT or CT) and met our inclusion criteria. This study was primarily exploratory; we did not undertake this review with a set of hypotheses about what we would find, although we had several impressions and general questions about pain measurement in published studies. For example, would the results show that one of the sophisticated, multidimensional pain questionnaires available to researchers, such as the McGill Pain Questionnaire, dominates the field?

Before addressing these questions, we report an unanticipated finding that emerged from the coding task. The task of coding the studies was quite challenging because many articles provided sparse descriptions of the assessments of their outcome measure(s). While this clearly made it more difficult to conduct the review, it suggests a more important point. A good methods section of an article should contain enough detail to allow for the study to be replicated, but the descriptions we found would often have been inadequate for conducting a replication study. Furthermore, comparison of results with other trials would be very difficult. Unfortunately, we originally planned to examine the ways that pain outcome measures were modified for specific protocols, but because of the frequently inadequate descriptions provided in the methods sections, it was difficult to ascertain whether measures were modified and if so, how. Comments during the review process by journal editors and reviewers could remedy this problem, and we recommend such improved reporting of assessment methods in future studies.

This survey plainly showed that one or two instruments did not dominate the pain outcome scene. Of the 50 articles included in this review, 28 different classifications of pain measures were derived. We did not expect this degree of assessment heterogeneity, and we contrast this finding to another area we are familiar with – the assessment of depression. In that area of research, there are just a handful of instruments (e.g., the Beck Depression Inventory and the Center for Epidemiological Studies Depression Scale) used in clinical trials. We interpret this as signifying that there is not an accepted standard for the assessment of chronic pain in clinical trials. An unfortunate consequence of this situation is that there is an inherent lack of comparability among the trials, because so few studies assessed pain in the same way. This observation is consistent with a recent review article examining the psychometric properties of osteoarthritis (OA) questionnaires. It included 37 articles (after reviewing 1,930) and consistent with our results, found 32 measures to assess pain and physical functioning in those papers.⁶² Clearly, this shows that there is not a ‘gold standard’ for measuring pain in either the broad category of chronic pain or within the specific chronic pain disease literature of OA.

The most compelling finding was that single-item measures were used as the only pain outcome measure in 34% of the trials (17 out of 50). Another important finding is that 60% of all eligible studies (30 out of 50) used a VAS as their pain outcome measure, and 10 of these used a VAS as their only pain outcome measure. This is in contrast to our expectation at the outset of the review when we thought that multidimensional pain scales, such as the MPQ, would be typical. Instead, we found that only 16% of the trials utilized a multidimensional pain measure. Thus, despite the vast amount of conceptual and empirical development that has gone into the multidimensional scales, they do not appear to be the choice of clinical pain researchers.

Our conclusion is that clinical researchers are opting for very brief pain assessments, and we speculate that this may be driven by the need to minimize patient assessment burden. We question whether this is a good decision in all cases. In some trials, pain may not be a primary outcome and therefore a comprehensive assessment may not be warranted. However, when pain is an outcome of importance to the trial, researchers need to ask what aspects of pain their treatment is expected to effect and then consider if that outcome will be adequately measured with single items or scales.

Our review allowed additional insights into how VAS scales are used in trials, and there was considerable heterogeneity in the way they were employed (i.e., time frames of the assessments) and how they were reported in method sections. The VAS was originally designed to measure current pain with either a 10cm or 100mm vertical or horizontal line with anchors on either end representing the extremes of the pain dimension being assessed. Almost half of the studies using VAS (47%) assessed the concept of pain intensity or severity, 27% measured “general pain,” “global pain,” or “pain” in general. In terms of the reporting period provided for rating the question, 17% asked about current pain, 17% about pain over the last week, and 7% about “estimated daily pain.” Of concern, 50% did not specify the reporting period used in the VAS. This further underscores the observation that there is a lack of consensus regarding measurement of pain in clinical trials, including the dimensions of pain and a recommended time frame. It also is another example of the inadequate level of description in the methods.

A challenging question emerges from the prior observation: are empirical evaluations of new treatments being shortchanged by the use of simple scales or are the simple VAS scales adequate? For at least 40 years, the multidimensional nature of the pain experience has been recognized,³⁹ and measures that incorporate assessment of several pain factors have been developed.^{9,32,40} Pain intensity, the most commonly measured factor, is but one of these dimensions. Sensory, affective, behavioral, social and attitudinal factors are among the other dimensions that have demonstrated importance for pain measurement and that are not captured through the measurement of pain intensity. Morley and Williams (2002) discuss the commonly observed gap between the dimensions of pain that are expected to be impacted by a treatment and the measures selected to index treatment outcome.⁴² When a treatment is designed to reduce the experience of pain and consequently to improve functioning and quality of life, then it does seem that we are shortchanging the impact of a clinical trial by limiting pain assessment to a single item of pain intensity. Yet, this appears to be the current state of pain assessment in the majority of clinical trials.

Despite the potential short-comings of single-item measures of pain, a group of pain researchers recently published recommendations on how chronic pain should be assessed.¹² Interestingly, the outcome measures they recommend include a single-item measure of pain intensity (a NRS as opposed to VAS) and assessments of rescue analgesic use. The report also listed several questionnaires to assess the other dimensions of pain (i.e. physical functioning, emotional impact, global improvement) to be used depending on the requirements of a particular protocol. While these recommendations agree with the current findings in that most pain researchers are already relying on single-item assessments of pain, it is notable that multidimensional measures were not recommended as the gold-standard outcome measure for chronic pain. Perhaps this is because pain researchers acknowledge these lengthier measures are more burdensome to patients or perhaps because clinical trials are more interested in assessing pain intensity and the use of rescue pain medication.

In addition to these recommendations, two studies that examined the utility and responsiveness of several pain measures in patients with OA and RA taking a non-steroidal anti-inflammatory medication reported that single-item measures like a Likert item or VAS were more responsive than more complex multidimensional measures.^{3,4} Therefore, perhaps single-item measures

are appropriate for certain clinical trials as a way to assess pain relief and/or changes pain severity/intensity. Thus the current finding that single-item measures were used in the majority of the eligible studies included in this review may not be as alarming as we first thought. However, we still maintain that for trials where pain is the primary outcome, understanding how the treatment changes all aspects of the pain experience (i.e. affective, sensory, analgesic, etc) is important.

Finally, it is worth noting that increased rigor in the measurement of pain in clinical trials is now expected in some research settings. In 2006, for the first time, the US Food and Drug Administration issued preliminary guidance for trials submitted for their review using patient reported outcomes (FDA Docket No. 2006D-0044). The issues covered in that guidance (including a variety of psychometric and assessment standards) will have far-reaching implications for the acceptability of pain assessments utilized in trials submitted to this regulatory agency. While investigators conducting pharmaceutical trials and pain researchers are undoubtedly apprising themselves of these new expectations, this review highlights the importance of more consistent measurement of pain for all clinical trials.

Acknowledgments

This work was supported by a grant from the NIH PROMIS Network (grant number here). Special thanks to Doerte Junghaenel and Natascha Santos for their assistance with data entry and to Dr. Dagmar Antmann for her comments and encouragement. Arthur A. Stone is the Associate Chair of the scientific advisory board of invivodata, inc.

References

1. Almeida TF, Roizenblatt S, Benedito-Silva AA, Tufik S. The effect of combined therapy (ultrasound and interferential current) on pain and sleep in fibromyalgia. *Pain* 2003;104:665–672. [PubMed: 12927639]
2. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of the 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* 1988;15:1833–1840. [PubMed: 3068365]
3. Bellamy N, Campbell J, Syrota J. Comparative study of self-rating pain scales in Osteoarthritis patients. *Curr Med Res Opin* 1999;15:113–119. [PubMed: 10494494]
4. Bellamy N, Campbell J, Syrota J. Comparative study of self-rating pain scales in Rheumatoid Arthritis patients. *Curr Med Res Opin* 1999;15:121–127. [PubMed: 10494495]
5. Birbara CA, Puopolo AD, Munoz DR, Sheldon EA, Mangione A, Bohidar NR, Geba GP. Treatment of chronic low back pain with etoricoxib, a new cyclo-oxygenase-2 selective inhibitor: Improvement in pain and disability-A randomized, placebo-controlled, 3-month trial. *J Pain* 2003;4:307–315. [PubMed: 14622687]
6. Christensen FB, Laurberg I, Bunger CE. Importance of the back-café concept to rehabilitation after lumbar spinal fusion: A randomized clinical study with a 2-year follow-up. *Spine* 2003;28:2561–2569. [PubMed: 14652472]
7. Chrusasik S, Model A, Black A, Pollak S. A randomized double-blind pilot study comparing Doloteffin and Vioxx in the treatment of low back pain. *Rheumatology* 2003;42:141–148. [PubMed: 12509627]
8. Ciechanowski P, Sullivan M, Jensen M, Romano J, Summers H. The relationship of attachment style to depression, catastrophizing and health care utilization in patients with chronic pain. *Pain* 2003;104:627–637. [PubMed: 12927635]
9. Daut RL, Cleeland CS, Flanery RC. Development of the Wisconsin Brief Pain Questionnaire to assess pain in cancer and other diseases. *Pain* 1983;17:197–210. [PubMed: 6646795]
10. Dernis-Labous E, Messow M, Dougados M. Assessment of fatigue in the management of patients with ankylosing spondylitis. *Rheumatology* 2003;42:1523–1528. [PubMed: 13130145]
11. Drewe E, McDermott EM, Powell PT, Isaacs JD, Powell RJ. Prospective study of anti-tumor necrosis factor receptor superfamily 1B fusion protein, and case study of anti-tumour necrosis factor receptor

superfamily 1A fusion protein, in tumour necrosis factor receptor associated periodic syndrome (TRAPS): clinical and laboratory findings in a series of seven patients. *Rheumatology* 2003;42:235–239. [PubMed: 12595616]

12. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, Kerns RD, Stucki G, Allen RR, Bellamy N, Carr DB, Chandler J, Cowan P, Dionne R, Galer BS, Hertz S, Jadad AR, Kramer LD, Manning DC, Martin S, McCormick CG, McDermott MP, McGrath P, Quessy S, Rappaport BA, Robbins W, Robinson JP, Rothman M, Royal MA, Simon L, Stauffer JW, Stein W, Tollett J, Wernicke J, Witter J. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2005;113:9–19. [PubMed: 15621359]
13. Edwards RR, Doleys DM, Lowery D, Fillingim RB. Pain tolerance as a predictor of outcome following multidisciplinary treatment for chronic pain: differential effects as a function of sex. *Pain* 2003;106:419–426. [PubMed: 14659525]
14. Ercelen O, Bulutcu E, Oktenoglu T, Sasani M, Bozkus H, Saryoglu AC, Ozer F. Radiofrequency lesioning using two different time modalities for the treatment of lumbar discogenic pain: A randomized trial. *Spine* 2003;28:1922–1927. [PubMed: 12973135]
15. Ernberg M, Lundeberg T, Kopp S. Effects on muscle pain by intramuscular injection of granisetron in patients with fibromyalgia. *Pain* 2003;101:275–282. [PubMed: 12583870]
16. Fransen M, Crosbie J, Edmonds J. Isometric muscle force measurement for clinicians treating patients with osteoarthritis of the knee. *Arthritis Rheum* 2003;49:29–35. [PubMed: 12579591]
17. Frerick H, Keitel W, Kuhn U, Schmidt S, Bredehorst A, Kuhlmann M. Topical treatment of chronic low back pain with a capsicum plaster. *Pain* 2003;106:59–64. [PubMed: 14581111]
18. Gatchel RJ. Comorbidity of chronic pain and mental health disorders: the biopsychological perspective. *Am Psychol* 2004;59:795–805. [PubMed: 15554853]
19. Giles LGF, Muller R. Chronic spinal pain. *Spine* 2003;28:1490–1503. [PubMed: 12865832]
20. Grabois M. Management of chronic low back pain. *Am J Phys Med Rehabil* 2005;84:S20–S41.
21. Gracely RH, Dubner R. Reliability and validity of verbal descriptor scales of painfulness. *Pain* 1987;29:175–85. [PubMed: 3614956]
22. Hagen EM, Grasdal A, Eriksen HR. Does early intervention with a light mobilization program reduce long-term sick leave for low back pain: A 3-year follow-up study. *Spine* 2003;28:2309–2316. [PubMed: 14560075]
23. Hakkinen A, Hannonen P, Nyman K, Lyyski T, Hakkinen K. Effects of concurrent strength and endurance training in women with early or longstanding rheumatoid arthritis: Comparison with healthy subjects. *Arthritis Rheum* 2003;49:789–797. [PubMed: 14673965]
24. Hinman RS, Bennell KL, Crossley KM, McConnell J. Immediate effects of adhesive tape on pain and disability in individuals with knee osteoarthritis. *Rheumatology* 2003;42:865–869. [PubMed: 12730546]
25. Ivar Brox J, Sorensen R, Friis A, Nygaard O, Indahl A, Keller A, Ingebrigtsen T, Eriksen HR, Holm I, Koller AK, Riise R, Reikeras O. Randomized clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration. *Spine* 2003;28:1913–1921. [PubMed: 12973134]
26. Jarvik JG, Hollingworth W, Martin B, Emerson SS, Gray DT, Overman S, Robinson D, Staiger T, Wessbecher F, Sullivan SD, Kreuter W, Deyo RA. Rapid magnetic resonance imaging vs radiographs for patients with low back pain. *JAMA* 2003;289:2810–2818. [PubMed: 12783911]
27. Jensen MP, Keefe FJ, Lefebvre JC, Romano JM, Turner JA. One- and two-item measures of pain beliefs and coping strategies. *Pain* 2003;104:453–469. [PubMed: 12927618]
28. Jensen MP, Nielson WR, Turner JA, Romano JM, Hill ML. Readiness to self-manage pain is associated with coping and with psychological and physical functioning among patients with chronic pain. *Pain* 2003;104:529–537. [PubMed: 12927625]
29. Katz N, Ju WD, Krupa DA, Sperling RS, Bozalis Rodgers D, Gertz BJ, Gimbel J, Coleman S, Fisher C, Nabizadeh S, Borenstein D. Efficacy and safety of Rofecoxib in patients with chronic low back pain: Results from two 4-week, randomized, placebo-controlled, parallel-group, double-blind trials. *Spine* 2003;28:851–858. [PubMed: 12941996]
30. Katz J, Melzack R. Measurement of Pain. *Surg Clin North Am* 1999;79:231–252. [PubMed: 10352653]

31. Keefe FJ, Bradley LA, Crisson JE. Behavioral assessment of low back pain: Identification of pain behavior subgroups. *Pain* 1990;40:153–160. [PubMed: 2137905]
32. Kerns R, Turk D, Rudy T. The West Haven-Yale Multidimensional Pain Inventory (WHYMPI). *Pain* 1985;23:345–356. [PubMed: 4088697]
33. Kirwan JR, Mitchell K, Hewlett S, Hehir M, Pollock J, Memel D, Bennett B. Clinical and psychological outcome from a randomized controlled trial of patient-initiated direct-access hospital follow-up for rheumatoid arthritis extended to 4 years. *Rheumatology* 2003;42:422–426. [PubMed: 12626791]
34. Klassbo M, Larsson G, Harms-Ringdahl K. Promising outcome of a hip school for patients with hip dysfunction. *Arthritis Rheum* 2003;49:321–327. [PubMed: 12794786]
35. Kremer JM, Westhovens R, Leon M, Di Giorgio E, Alten R, Steinfeld S, Russell A, Dougados M, Emery P, Nuamah IF, Rhys Williams G, Becker JC, Hagerty DT, Moreland LW. Treatment of rheumatoid arthritis by selective inhibition of T-cell activation with fusion protein CTLA4Ig. *N Engl J Med* 2003;349:1907–1915. [PubMed: 14614165]
36. Licciardone JC, Stoll ST, Fulda KG, Russo DP, Siu J, Winn W, Swift J. Osteopathic manipulative treatment for chronic low back pain. *Spine* 2003;28:1355–1362. [PubMed: 12838090]
37. Lin EHB, Katon W, Von Korff M, Tang L, Williams JW, Kroenke K, Hunkeler E, Harpole L, Hegel M, Arean P, Hofferding M, Della Penna R, Langston C, Unutzer J. Effect of improving depression care on pain and functional outcomes among older adults with arthritis. *JAMA* 2003;290:2428–2434. [PubMed: 14612479]
38. Maier C, Dertwinkel R, Mansourian N, Hosbach I, Schwenkreis P, Senne I, Skipa G, Zenz M, Tegenthoff M. Efficacy of the NMDA-receptor antagonist memantine in patients with chronic phantom limb pain – results of a randomized double-blinded, placebo-controlled trial. *Pain* 2003;103:277–283. [PubMed: 12791434]
39. Melzack, R.; Casey, K. Sensory, motivational, and central control determinants of pain: A new conceptual model. In: Kenshalo, D., editor. *The skin senses*. Springfield, IL: Chas C Thomas; 1968. p. 423–439.
40. Melzack R. The McGill Pain Questionnaire: Major properties and scoring methods. *Pain* 1975;1:277–299. [PubMed: 1235985]
41. Meng CF, Wang D, Ngeow J, Lao L, Peterson M, Paget S. Acupuncture for chronic low back pain in older patients: a randomized, controlled trial. *Rheumatology* 2003;42:1508–1517.
42. Morley, S.; Williams, A. Conducting and evaluating treatment outcome studies. In: Turk, DC.; Gatchel, R., editors. *Psychological approaches to pain management*. New York: The Guilford Press; 2002. p. 52–68.
43. Morrell EM, Keefe FJ. The actometer: An evaluation of instrument applicability for chronic pain patients. *Pain* 1988;32:265–270. [PubMed: 3362564]
44. Munneke M, De Jong Z, Zwilander AH, Jansen A, Karel Rouday H, Peter WFH, Boonman DCG, Van Den Ende CHM, Vliet Vlieland TPM, Hazes JMW. Adherence and satisfaction of rheumatoid arthritis patients with a long-term intensive dynamic exercise program (RAPIT program). *Arthritis Rheum* 2003;49:665–672. [PubMed: 14558052]
45. Niemisto L, Lahtinen-Suopanki T, Rissanen P, Londgren KA, Sarna S, Hurri H. A randomized trial of combined manipulation, stabilizing exercises, and physician consultation compared to physician consultation alone for chronic low back pain. *Spine* 2003;28:2185–2191. [PubMed: 14520029]
46. Washington, DC: General Printing Office; Sept. 41998 [Feb. 13, 2006]. NIH Guide. New directions in pain research I. PA-98-102. Available at: <http://grants2.nih.gov/grants/guide/pa-files/PA-98-102.html>
47. Parent E, Moffet H. Preoperative predictors of locomotor ability two months after total knee arthroplasty for severe osteoarthritis. *Arthritis Rheum* 2003;49:36–50. [PubMed: 12579592]
48. Parker JC, Smarr KL, Slaughter JR, Johnston SK, Priesmeyer ML, Donovan Honson K, Johnson GE, Hewett JE, Hewett JE, Irvin WS, Komatireddy GR, Walker SE. Management of depression in rheumatoid arthritis: A combined pharmacologic and cognitive-behavioral approach. *Arthritis Rheum* 2003;49:766–777. [PubMed: 14673962]

49. Pavelka K, Recker DP, Verburg KM. Valdecoxib is as effective as diclofenac in the management of rheumatoid arthritis with a lower incidence of gastroduodenal ulcers: Results of a 26-week trial. *Rheumatology* 2003;42:1207–1215. [PubMed: 12810937]
50. Rashid S, Koller M, Haykowsky M, Jamieson K. The effect of opioid analgesia on exercise test performance in chronic low back pain. *Pain* 2003;106:119–125. [PubMed: 14581118]
51. Riemsma RP, Raal E, Rasker JJ. Group education for patients with rheumatoid arthritis and their partners. *Arthritis Rheum* 2003;49:556–566. [PubMed: 12910564]
52. Ringe JD, Dorst A, Faber H, Ibach K, Preuss J. Three-monthly ibandronate bolus injection offers favourable tolerability and sustained efficacy advantage over two years in established corticosteroid-induced osteoporosis. *Rheumatology* 2003;42:743–749. [PubMed: 12730532]
53. Romano JM, Jensen MP, Turner JA. The chronic pain coping inventory-42: Reliability and validity. *Pain* 2003;104:65–73. [PubMed: 12855315]
54. Roy EA, Hollins M, Maixner W. Reduction of TMD pain by high-frequency vibration: a spatial and temporal analysis. *Pain* 2003;101:267–274. [PubMed: 12583869]
55. Salvarani C, Cantini F, Olivieri I, Macchioni P, Padula A, Niccoli L, Grazia Catanoso M, Lo Scocco G, Boiardi L. Efficacy of Infliximab in resistant psoriatic arthritis. *Arthritis Rheum* 2003;49:541–545. [PubMed: 12910561]
56. Schuerwegh AJ, Van Offel JF, Stevens WJ, Bridts CH, De Clerck LS. Influence of therapy with chimeric monoclonal tumour necrosis factor-alpha antibodies on intracellular cytokine profiles of T lymphocytes and monocytes in rheumatoid arthritis patients. *Rheumatology* 2003;42:541–548. [PubMed: 12649401]
57. Sharpe L, Sensky T, Timberlake N, Ryan B, Allard S. Long-term efficacy of a cognitive behavioural treatment from a randomized controlled trial for patients recently diagnosed with rheumatoid arthritis. *Rheumatology* 2003;42:435–441. [PubMed: 12626793]
58. Smith MD, Wetherall M, Darby T, Esterman A, Slavotinek J, Roberts-Thomson P, Coleman M, Ahern MJ. A randomized placebo-controlled trial of arthroscopic lavage versus lavage plus intra-articular corticosteroids in the management of symptomatic osteoarthritis of the knee. *Rheumatology* 2003;42:1477–1485. [PubMed: 12867587]
59. Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient non-compliance with paper diaries. *BMJ* 2002;324:1193–1194. [PubMed: 12016186]
60. Theime K, Gromnica-Ihle E, Flor H. Operant behavioral treatment of fibromyalgia: A controlled study. *Arthritis Rheum* 2003;49:314–320. [PubMed: 12794785]
61. Tullberg M, Alstergren PJ, Ernberg MM. Effects of low-power laser exposure on masseter muscle pain and microcirculation. *Pain* 2003;105:89–96. [PubMed: 14499424]
62. Veenhof C, Bijlsma JWJ, van den Ende CHM, van Dijk GM, Pisters MF, Dekker J. Psychometric evaluation of osteoarthritis questionnaires: A systematic review of the literature. *Arthritis Rheum* 2006;55:480–492. [PubMed: 16739188]
63. Vowles KE, Gross RT. Work-related beliefs about injury and physical capability for work in individuals with chronic pain. *Pain* 2003;101:291–298. [PubMed: 12583872]
64. Watson CPN, Moulin D, Watt-Watson J, Gordon A, Eisenhoffer J. Controlled-release oxycodone relieves neuropathic pain: A randomized controlled trial in painful diabetic neuropathy. *Pain* 2003;105:71–78. [PubMed: 14499422]
65. White P, Lewith G, Hopwood V, Prescott P. The placebo needle, is it a valid and convincing placebo for use in acupuncture trials? A randomized, single-blind, crossover pilot trial. *Pain* 2003;106:401–409. [PubMed: 14659523]
66. Willoch F, Gamringer U, Medele R, Steude U, Tolle TR. Analgesia by electrostimulation of the trigeminal ganglion in patients with trigeminopathic pain: A PET activation study. *Pain* 2003;103:119–130. [PubMed: 12749966]
67. Winemiller MH, Billow RG, Laskowski ER, Harmsen WS. Effect of magnetic vs sham-magnetic insoles on plantar heel pain. *JAMA* 2003;290:1474–1478. [PubMed: 13129987]
68. Ylinen J, Takala EP, Nykanen M, Hakkinen A, Malkia E, Pohjolainen T, Karppi SL, Kautiainen H, Airaksinen O. Active neck muscle training in the treatment of chronic neck pain in women. *JAMA* 2003;289:2509–2516. [PubMed: 12759322]

Table 1

Exclusion and eligibility of articles and distribution by journal

All Articles (N=1476)	Reason for exclusion	N Excluded
	Not RCT or CT	1149 (78%)
	Non-human sample	292 (20%)
	Non-adult sample	594 (40%)
	Non-chronic musculoskeletal pain sample	1184 (80%)
	Non-pain outcomes	1230 (83%)
	Retained for analysis	50 (3.0%)
All Articles (N=1476)	N Reviewed	N Eligible
<i>Pain</i>	272	16 (6%)
<i>Journal of Pain</i>	54	1 (2%)
<i>JAMA</i>	223	4 (2%)
<i>NEJM</i>	208	1 (1%)
<i>Arthritis & Rheumatism</i>	90	9 (10%)
<i>Rheumatology</i>	173	11 (6%)
<i>Spine</i>	456	8 (2%)

Table 2
Pain measures used in RCTs and CTs¹

Measure	Number of studies that used the measure
<i>Single-item measures</i>	
Visual Analogue Scale (VAS)	30
Numeric Rating Scale (NRS)	6
Verbal Rating Scale (VRS)	5
Change in back pain intensity	1
Increased pain during or after exercise	1
Low back pain rating scale (0-10)	1
Pain Frequency	1
Pressure Pain Threshold	1
5-point categorical	1
6-point categorical	1
11-point intensity	1
<i>Multi-item measures or subscales that assess one dimension of pain</i>	
Roland Morris Disability Questionnaire	7
Oswestry Disability Questionnaire	4
SF-36 pain subscale	3
AIMS pain subscale	2
ACR-20 patients' assessment of pain	1
Low back pain bothersomeness scale ²	1
Modified Wisconsin Body Map	1
Neck Disability Index (NDI)	1
Pain Bothersomeness	1
Pain Frequency	1
Profile Pain Drawings	1
VAS 'Pain Index' - mean of score of maximum, minimum and current pain	1
<i>Multidimensional pain measures</i>	
WOMAC	6
Arhus low back pain rating scale	2
McGill Pain Questionnaire (MPQ, or SF-MPQ)	2
Multidimensional Pain Inventory (MPI; Severity or Interference subscales)	2
Diary card: 'validated patient diary card' ³	1
Neuropathic Pain Scale (NPS)	1
<i>One subscale from a multidimensional pain measure (where only one pain dimension is assessed)</i>	

Measure	Number of studies that used the measure
MPI Pain Severity subscale	1
<i>Unknown- unable to classify with information provided</i>	
No pain measure stated in the methods, but 'Pain' reported as an outcome in the results	1
Pain Frequency	1

¹ N=50 however, right column does not add to 50 because some articles used more than one pain outcome measure

² This measure is assumed to be multi-item however, in the methods section of the article it is not described fully nor is it cited.

³ The authors describe this measure in the methods section as 'multidimensional.'

Table 3Qualities of the 30 VAS measures¹

Characteristics of Assessment	Coding	N
<i>Quality of pain studied</i>	Intensity	12
	General pain	8
	Severity	2
	“Axial, peripheral, and global pain”	1
	“Intensity, steady, brief (‘shock-like’) and skin pain (‘pain elicited by non-painful stimulation of the skin’)”	1
	“Maximum, minimum, and current pain. The mean of the three measurements provided the pain index.”	1
	“Subjects’ overall perception of back pain.”	1
	Not provided	4
<i>Scale type</i>	100mm line	12
	10cm line	9
	0-100	4
	0-90mm	1
	Not provided	4
<i>Scale anchors</i>	Anchors provided	13
	Anchors not provided	17
<i>Reporting period</i>	Current	5
	Previous week / Last 7 days	5
	Daily	2
	“At rest and during activity”	1
	Hourly	1
	Morning, evening and mean	1
	Not provided	15

¹ Information in quotes are direct quotes from the text of the articles.