Pain 121 (2006) 151–157

PAIN

www.elsevier.com/locate/pain

Patient-perceived satisfactory improvement (PPSI): Interpreting meaningful change in pain from the patient's perspective

Peter M. ten Klooster [a,*], K. Wiepke Drossaers–Bakker [b], Erik Taal [a], Martin A.F.J. van de Laar [a,b]

[a] *Institute for Behavioural Research, Faculty of Behavioural Sciences, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands*
[b] *Department of Rheumatology, Medisch Spectrum Twente, P.O. Box 50.000, 7500 KA Enschede, The Netherlands*

Received 8 July 2005; received in revised form 4 October 2005; accepted 21 December 2005

Abstract

The assessment of clinically meaningful changes in patient-reported pain has become increasingly important when interpreting results of clinical studies. However, proposed response criteria, such as the minimal clinically important difference, do not correspond with the growing need for information on truly meaningful, individual improvements. The aim of the present study was to investigate satisfactory improvements in pain from the patient's perspective. Data were collected in a 2-week prospective study of 181 arthritis patients treated with a local corticosteroid injection. Baseline and follow-up pain were assessed on 100 mm visual analogue scales for pain intensity (VAS-PI). At baseline, patients also marked a hypothetical level on a VAS-PI representing a satisfactory improvement in pain. Patient-perceived satisfactory improvement (PPSI) was constructed using a 5-point categorical rating of change scale at follow-up as the anchor. PPSI was associated with a minimal reduction of 30 mm or 55% on the VAS-PI. Since absolute change in pain associated with satisfactory improvement proved highly dependent on baseline pain, percent change scores performed better in classifying improved patients. The 55% threshold for satisfactory improvement was consistent over the course of treatment and reasonably consistent across groups of patients. Our data suggest that PPSI is a clinically relevant and stable concept for interpreting truly meaningful improvements in pain from the individual perspective.

Keywords: Musculoskeletal pain; Pain measurement; Visual analogue scales; Minimal clinically important difference; Patient-perceived satisfactory improvement

1. Introduction

In recent years, both clinicians and investigators have become increasingly interested in the patient's perspective on the meaning of changes on core outcome measures (Wells et al., 2001). A commonly used method to determine thresholds for patient-perceived meaningful change is to compare changes in pain scores with patients' global ratings of the magnitude of change (Deyo and Patrick, 1995; Crosby et al., 2003). Variations on this approach have been used to define the minimal clinically important difference (MCID) in pain in various clinical settings (e.g., Stratford et al., 1998; Farrar et al., 2001; Dhanani et al., 2002; Cepeda et al., 2003; Jensen et al., 2003; Salaffi et al., 2004).

A frequently overlooked concern with this approach is that it actually mixes perspectives (Beaton et al., 2001). Whereas the patient rates the magnitude of change, the investigator determines which rating serves as the cut-off for important or satisfactory improvement. Another concern is that patients are often unable to accurately recall their initial pain and that retrospective self-reports of pain relief do not always reflect true changes in pain (Feine et al., 1998; Fischer et al., 1999; Haas et al., 2002). Instead of comparing pre-treatment

* Corresponding author. Tel.: +31 53 4893947; fax: +31 53 4894259.
E-mail address: P.M.tenKlooster@utwente.nl (P.M. ten Klooster).

0304-3959/$20.00
doi:10.1016/j.pain.2005.12.021

and current pain, these patients seem to focus mainly on the acceptability of their current status when judging the magnitude of change (Aseltine et al., 1995; Norman et al., 1997; Guyatt et al., 2002a,b).

To address these problems Tubach et al. (2005b) recently suggested to complement the MCID with the patient acceptable symptom state (PASS), an absolute value on the follow-up measure beyond which patients consider themselves well. The PASS does not deal with changes, but concentrates on the concept of achieving a satisfactory state. In this sense, the concept of the patient-derived PASS is very similar to arbitrarily defined or data driven concepts as adequate analgesia (e.g., Benedetti et al., 1998) and low disease activity state (e.g., van der Heijde et al., 1990). Since achieving adequate pain relief is the ultimate goal of pain treatment, the PASS is a clinically relevant concept. Moreover, the patient driven PASS meets the growing need for measures of major improvement from the patients' perspective as opposed to measures of minimal important difference (Wells et al., 2001; Wolfe et al., 2005). A drawback of its use, however, is that it entails separate analyses for patients achieving a relevant change and patients achieving an acceptable state.

This study presents an investigation into meaningful changes in pain from the patient's perspective that combines the strengths of both the MCID and the PASS. The first objective of the study was to assess the magnitude of change on the VAS-PI that most closely represents patient-perceived satisfactory improvement (PPSI) in arthritis patients with localized musculoskeletal pain. The second objective was to investigate the stability of PPSI across groups of patients. Since patients' perceptions may also change over the course of treatment, the third objective was to examine whether PPSI corresponds with the change in pain that patients before treatment consider necessary for satisfactory improvement.

2. Methods

2.1. Patient selection and study design

Participants were recruited at the outpatient rheumatology clinic. All consecutive patients with localized musculoskeletal pain and an indication for a local corticosteroid injection were asked to participate. Patients were excluded if they were aged <16 years or unable to mark a visual analogue scale for pain intensity (VAS-PI). The study did not interfere with usual treatment.

Prior to the injection, patients indicated the average level of localized pain in the past week on a 100 mm, unmarked VAS-PI with endpoints "no pain" and "unbearable pain". Subsequently, patients marked the level of pain that would represent a satisfactory improvement on a separate VAS-PI. After 2 weeks, a follow-up questionnaire was mailed to the patients. After marking the VAS-PI for pain in the past week, patients judged the change in pain by answering the following question: "Compared to 2 weeks ago (before the local injection) the pain in the injected area is…" The response categories were "worse", "unchanged", "unsatisfactory improved", "satisfactory improved" and "good to very good improved".

2.2. Analyses

2.2.1. Statistics

Statistical analyses were performed using SPSS 11.0 for Windows. The valid use of parametric statistics was verified by testing for normal distribution of the variables (Kolmogorov–Smirnov test, normal distribution assumed when $P > 0.05$). When the assumption of normality was not met, non-parametric statistics were used. P values <0.05 were considered to indicate statistical significance. The mean and standard deviation (SD) were used for descriptive statistics unless otherwise specified.

2.2.2. Patients' judgments of change

To study meaningful changes in pain, an anchoring method based on the patient's judgment of change at follow-up was used. This categorical rating scale, however, has not been previously validated. Supporting evidence for its valid use as an external anchor would be an appreciable relationship between patients' ratings of change and actual changes on the VAS-PI (Guyatt et al., 2002a). To explore this relationship, the categorical ratings were compared with absolute change scores (VAS-PI follow-up − VAS-PI baseline) and percent change scores ((absolute change/VAS-PI baseline) × 100) by means of one-way analyses of variance (ANOVAs) followed by post hoc multiple comparisons (Bonferroni adjustment). Second, Spearman rank correlation coefficients of the categorical rating scale with absolute and percent change in pain on the VAS-PI were calculated. Correlations $\geqslant 0.5$ were considered indicative for the valid use of the rating scale (Cella et al., 2002; Guyatt et al., 2002b).

2.2.3. Patient-perceived satisfactory improvement on the VAS-PI

Patient-perceived satisfactory improvement (PPSI) was defined as the change in pain on the VAS-PI associated with a minimal rating of satisfactory improvement at follow-up. Ratings of "satisfactory improved" and "good to very good improved" were pooled to define satisfactory improved patients. Patients were considered unimproved when they rated themselves as worsened, unchanged or unsatisfactory improved. To evaluate the change in pain that was most closely associated with PPSI, receiver operating characteristic (ROC) curves were computed for both absolute and percent change scores (cf. Deyo and Centor, 1986; Ward et al., 2000). As opposed to the analyses of group means, as suggested by Jaeschke et al. (1989), ROC analysis offers the opportunity to study patient-perceived improvement at the individual level. An area under the ROC curve (AUC) $\geqslant 0.7$ was considered adequately accurate in classifying satisfactory improved patients (Grzybowski and Younger, 1997). The change score with the highest combination of sensitivity and specificity was selected as the optimal cut-off point for PPSI. The com-

parative accuracy of absolute and percent change scores was determined by comparing the areas under the curve (Hanley and McNeil, 1983).

2.2.4. Consistency of PPSI over groups of patients

The consistency of PPSI across baseline demographic and clinical variables was investigated using the data of patients who rated their pain as satisfactory or good to very good improved. Dependency of absolute change on baseline VAS-PI scores was determined by linear regression analysis. The consistency of absolute change over age and disease duration was assessed using Pearson correlation coefficients and the differences between men and women and between the five primary diagnoses were investigated using an independent t test and a one-way ANOVA. Since percent change in these patients was not normally distributed, the non-parametric Spearman rank correlation coefficient, Mann–Whitney U test and Kruskal–Wallis H tests were used to assess the stability of percent change scores.

2.2.5. Consistency of PPSI over the course of treatment

To assess whether patient's perceptions of satisfactory improvement had changed over the course of treatment, the mean actual change scores of improved patients were compared with the mean change scores patients initially judged necessary to be satisfied. The agreement between actual and initially defined change scores of satisfactory improved patients was calculated using the intraclass correlation coefficient (ICC). ICCs were considered excellent when >0.75, fair to good when $\geqslant 0.40 \leqslant 0.75$ and poor when <0.40 (Fleiss, 1986). Since the ICC does not provide information on the magnitude of within-person differences, a Bland–Altman plot of the difference against the average of the actual and initially defined change scores was constructed (Bland and Altman, 1986).

3. Results

3.1. Patient characteristics

Between May and December 2004, 200 patients agreed to participate in the study and completed the baseline questionnaire. Despite sending reminders, six follow-up questionnaires were not returned. Thirteen follow-up questionnaires were not interpretable. Descriptive baseline characteristics of the included patients are listed in Table 1.

3.2. Patients' judgments of change

ANOVAs showed that absolute and percent change scores on the VAS-PI were significantly different – in the expected direction – between groups based on the patients' ratings of change (Table 2). Both absolute and percent change scores were significantly different between satisfactory improved patients and worsened, unchanged, or unsatisfactory improved patients. The association between patient-perceived ratings of change and actual change scores was supported by moderate (Spearman $r = -0.51$, $P < 0.001$) to good (Spearman $r = -0.70$, $P < 0.001$) correlation for absolute and percent change, respectively.

Table 1
Baseline demographic and clinical characteristics ($N = 181$)

Age in years (mean ± SD)	59.5 ± 14.7
Gender (% female/male)	70.7/29.3
Primary diagnosis	
Rheumatoid arthritis (%)	37.0
Osteoarthritis (%)	17.7
Psoriatic arthritis (%)	8.8
Tendinitis/bursitis (%)	8.3
Other (%)	28.2
Disease duration (median, range)	5, 0–52
VAS-PI (mean ± SD)	58.6 ± 24.0

There were no significant differences in baseline VAS-PI scores between patients based on primary diagnosis (ANOVA) and baseline pain was not related to age (Pearson r). Women tended to report more pain then men, although this difference was not significant (60.9 ± 24.1 vs. 53.2 ± 23.1 mm, $t(179) = -1.96$, $P = 0.052$). Patients with longer disease duration reported more pain (Spearman $r = 0.14$, $P < 0.05$).

3.3. Patient-perceived satisfactory improvement on the VAS-PI

Fig. 1 presents the ROC curves for absolute and percent change on the VAS-PI at 2-week follow-up, associated with patients' ratings of satisfactory or good to very good improvement. Both absolute and percent change scores had good diagnostic power in identifying satisfactory improved patients, with AUCs of 0.80 (95% CI: 0.73–0.85, $P < 0.0001$) and 0.86 (95% CI: 0.80–0.91, $P < 0.0001$), respectively. The better overall accuracy of PPSI expressed as a percent change score was represented by a significantly higher AUC for percent change scores ($P < 0.05$). The optimal cut-off point for an absolute change in pain was −30 mm, corresponding to a sensitivity of 0.68 (95% CI: 0.58–0.76) and specificity of 0.84 (95% CI: 0.73–0.92). The best cut-off for a

Table 2
Group level analysis of mean change in pain on the VAS-PI (in mm) at follow-up associated with categories of patient-perceived rating of change ($N = 181$)

	Absolute change (mean ± SD)[a]	Percent change (mean ± SD)[b]
Worsened ($n = 3$)	$16.3 \pm 21.0_{a}$	$35.5 \pm 42.8_{a}$
Unchanged ($n = 17$)	$-2.8 \pm 18.9_{a}$	$-6.7 \pm 38.2_{a,b}$
Unsatisfactory improved ($n = 49$)	$-16.3 \pm 19.8_{a}$	$-22.7 \pm 31.3_{b}$
Satisfactory improved ($n = 76$)	$-37.2 \pm 25.4_{b}$	$-56.1 \pm 34.6_{c}$
Good to very good improved ($n = 35$)	$-43.5 \pm 23.5_{b}$	$-85.6 \pm 15.5_{d}$

[a] One-way ANOVA: $F(4, 175) = 18.0$, $P < 0.001$.
[b] One-way ANOVA: $F(4, 175) = 34.5$, $P < 0.001$. Means in the same column that do not have the same subscript differ at $P < 0.05$ (Bonferroni adjustment).

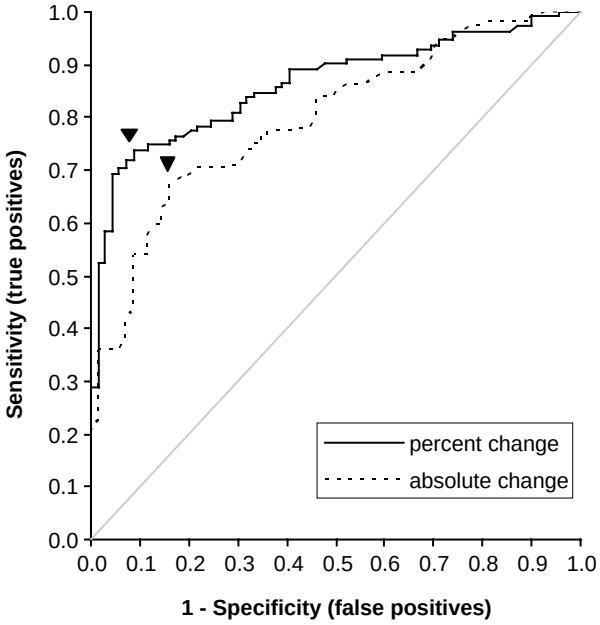

Fig. 1. Receiver operating characteristic curves for absolute and percent change in pain on the VAS-PI at 2-week follow-up associated with PPSI ($N = 181$). ▼ = Optimal cut-off point: absolute change = −30 mm (sensitivity 0.68, specificity 0.84); percent change = −54.6% (sensitivity 0.74, specificity 0.91). AUC absolute change = 0.80 (95% CI: 0.73–0.85); AUC percent change = 0.86 (95% CI: 0.80–0.91); $P < 0.001$ for difference between AUCs.

percent change from baseline was −54.6%, with a sensitivity of 0.74 (95% CI: 0.65–0.82) and a specificity of 0.91 (95% CI: 0.82–0.97).

3.4. Consistency of PPSI over groups of patients

The results from the ROC analyses indicated that percent change scores performed better in identifying satisfactory improved patients than absolute change scores. This dependency of PPSI on baseline pain was confirmed by analysis of the change scores of satisfactory and good to very good improved patients ($n = 111$). The relation between absolute change in pain and baseline pain is illustrated in Fig. 2. Patients with high baseline pain required greater absolute reductions in pain to reach a satisfactory improvement ($r^2 = 0.58$, $P < 0.001$). The magnitude of both absolute and percent change in pain was not related to age or disease duration and did not vary between groups based on primary diagnosis. However, absolute change scores in female patients were significantly larger than those in male patients (-42.0 ± 25.4 vs. -31.2 ± 22.1 mm, $t(109) = 2.03$, $P < 0.05$). Percent change scores did not significantly differ between men and women.

3.5. Consistency of PPSI over the course of treatment

The absolute change in pain that patients ($N = 181$) initially considered necessary to achieve a satisfactory

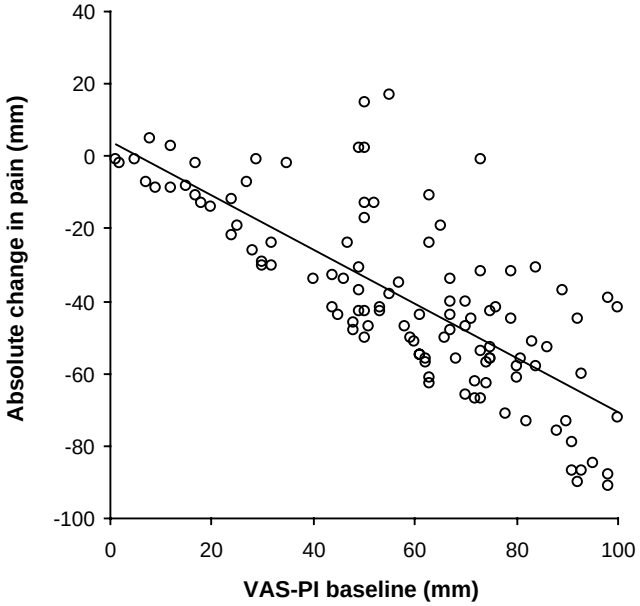

Fig. 2. Scatter plot of absolute change in pain in satisfactory or good to very good improved patients related to baseline pain intensity ($n = 111$). The straight line represents the linear regression line through the data points ($r^2 = 0.58$, $P < 0.001$), demonstrating the dependency of PPSI on baseline pain.

improvement was -32.0 ± 19.7 mm, corresponding to a percent change of -54.7 ± 27.8%. The actual change scores of satisfactory improved patients ($n = 76$) were adequately correlated with the initially defined satisfactory change scores (ICC = 0.61). However, Bland–Altman analysis (Fig. 3) showed that actual change scores

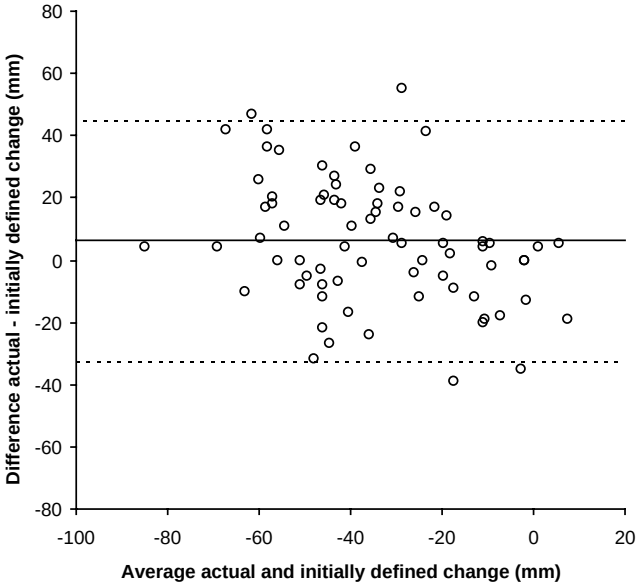

Fig. 3. Agreement between actual and initially defined change for satisfactory improvement within satisfactory improved patients ($n = 76$). The dashed lines represent the 95% limits of agreement (−32.9 and 45.1 mm). The horizontal solid line represents the mean difference between both change scores (6.1 mm).

were systematically larger than initially defined change scores (mean difference: 6.1 ± 19.9 mm; paired t test, $P < 0.05$). As expected, this difference was weakly related to the magnitude of actual change in pain (Pearson $r = 0.23$, $P < 0.05$), indicating that patients with relatively high actual change scores had improved more than they initially considered necessary for satisfactory improvement. The difference between initially defined and actual change in pain was highly variable, as represented by the relatively wide limits of agreement (± 39.0 mm) of the Bland–Altman plot. Since systematic bias was only moderate, the predominant source of error, however, was due to random variation instead of a systematic difference between actual and initially defined satisfactory improvement.

4. Discussion

This study presents the PPSI as a new outcome for individual, within-person improvement in pain intensity. Defining meaningful improvement is becoming increasingly important in interpreting the effectiveness of the treatment of pain. However, currently used data driven constructs for identifying clinical improvements such as the MCID do not satisfy the need for information on relevant changes from the patient's perspective. PPSI is assessed using patients' judgments of satisfactory change as the only criterion and prevents arbitrarily chosen cut-off points on the external anchor. Moreover, it gives a better representation of relevant change since patients tend to judge changes based on the acceptability of their present state. As such, it allows for a patient-centred approach in determining thresholds for true meaningful change, which combines the strengths of both the MCID and the PASS.

In the present study, the threshold for PPSI in musculoskeletal pain was best represented by a decrease on the VAS-PI of at least 55% or 30 mm. This threshold is characterized by a high sensitivity and specificity, supporting the responsiveness of the VAS-PI in measuring musculoskeletal pain. The magnitude of change required to achieve PPSI is considerably larger than most current definitions of meaningful change, such as the proposed 30% improvement criterion (Dworkin et al., 2005). In fact, it corresponds more closely with the formally most often used 50% pain relief threshold (e.g., Scott et al., 1990; Moore et al., 1996). Although this higher threshold is neither supported by empirical research (Seres, 1999) nor has its importance to patients been established (Farrar, 2000), a 50% reduction in pain does have clinical intuitive appeal as a threshold for satisfactory improvement (Moore et al., 1996). The difference in magnitude between PPSI and the previously established MCIDs on the VAS-PI could have several possible explanations. The difference could be related to the specific clinical setting of this study, the patients' demographic characteristics or diverging patient expectations. A more likely explanation, however, is that patients are not as easily satisfied with an improvement in pain as investigators are. Changes in pain may need to exceed the cut-offs defined by investigators to be considered satisfactory by patients. Evidence supporting this assumption is that the 55% or 30 mm cut-off for satisfactory improvement is in close accordance with recent studies examining patient-perceived, relevant improvements on the VAS-PI. Concepts defined as "adequate pain treatment" (Lee et al., 2003), "important improvement or recovery" (Giraudeau et al., 2004) or "considerable improvement" (Kvien et al., 2004) show similar cut-off points. As such, the cut-off for satisfactory improvement seems to answer the growing need for definite, relevant response criteria as opposed to minimal detectable responses (Felson and Anderson, 2001; Kelly, 2001b; Tugwell et al., 2001; Wolfe et al., 2005).

The present study confirms that patient-perceived improvement is not uniformly distributed over the range of the VAS. Whereas initially important improvements were considered to be absolute values (Kelly, 2001a; Todd, 2001), more recently it was shown that the magnitude of a MCID increases as baseline pain intensity increases (Bird and Dickson, 2001; Farrar et al., 2001; Jensen et al., 2003; Tubach et al., 2005a). This dependency on baseline pain status also applies to satisfactory improvements. Patients with high baseline pain need larger reductions in pain to consider themselves satisfactory improved. The ROC analyses also indicate that the diagnostic accuracy of the VAS-PI in discriminating between satisfactory and not-satisfactory improved patients increases when change scores are expressed as a percent change from baseline. Like MCID, PPSI is thus best represented as a percent change from baseline.

The magnitude of a satisfactory improvement proves to be consistent across groups of patients, except for gender. The lower absolute value of PPSI in men can be partly explained by their lower baseline pain scores, since percent change scores were more consistent between men and women. PPSI is also consistent over the course of treatment. Retrospective judgements of satisfactory improvement are adequately correlated with the change in pain patients at baseline consider satisfactory. The relatively wide limits of agreement between actual and initially defined satisfactory change scores may be due to the inherent problem of high measurement error associated with the use of visual analogue scales (DeLoach et al., 1998; Lassere et al., 2001; Kropmans et al., 2002).

The results of the study support the valid use of the 5-point categorical rating scale as an anchor to assess PPSI. The rating scale allows for a clear distinction between satisfactory and unsatisfactory improved patients. Moreover, the categorical rating

scale correlates adequately with the absolute change on the VAS-PI and good with percent change from baseline. However, the assumption that the categorical rating scale is also a reliable standard for measuring change could be a concern. The design of the study did not allow for an assessment of the reliability of this scale. This is a common problem for global rating scales, since internal consistency (Cronbach's α) cannot be computed for a single-item scale and test–retest studies are often complicated or impractical (Norman et al., 1997). Future studies are required using this scale on successive occasions in patients with a stable VAS-PI after the first follow-up, in order to assess the test–retest reliability of the categorical rating scale.

Another concern is the exclusive focus on improvements in pain. Since only three patients indicated an increase in pain, clinically meaningful deteriorations on the VAS-PI could not be calculated. The magnitude of the change that patients perceive as meaningful may differ between improvements and deteriorations (Hays and Woolley, 2000; Cella et al., 2002; Guyatt et al., 2002a). The goal of this investigation and most clinical studies, however, was to study important improvement since this is the result that clinicians and researchers are usually most interested in.

Moreover, the correlation between actual changes in satisfactory improved patients and initially defined changes may have been influenced by a testing effect, i.e., patients may have recalled the position they originally marked on the VAS-PI that would constitute satisfactory improvement.

A final issue concerns the generalizability of the findings. In the current sample, only patients who were treated with a corticosteroid injection were included. These injections are usually administered to patients who experience an exacerbation of pain. The relatively acute nature of their pain may have influenced patients' ratings of their pain and improvement. To determine the generalizability of the study, the findings should be confirmed in different clinical settings. Moreover, the magnitude of PPSI may very well differ for other outcome domains, such as physical functioning, global health status or quality of life. Since the procedures for assessing PPSI can be applied to all patient-reported outcomes, meaningful improvements from the patient's perspective can also be determined for these outcome domains.

In conclusion, PPSI is a clinically relevant and stable concept and can be used to assess true meaningful change in pain from the patient's perspective. Its straightforward character and analyses allows for the unambiguous assessment of satisfactory improved patients. The application of this measure in future clinical studies could lead to new standards for defining clinically meaningful improvement in pain and other outcome domains.

Acknowledgements

The authors thank the respondents who participated in this study and the rheumatologists of Medisch Spectrum Twente for their help in recruiting patients.

References

Aseltine Jr RH, Carlson KJ, Fowler Jr FJ, Barry MJ. Comparing prospective and retrospective measures of treatment outcomes. Med Care 1995;33:AS67–76.

Beaton DE, Bombardier C, Katz JN, Wright JG. A taxonomy for responsiveness. J Clin Epidemiol 2001;54:1204–17.

Benedetti F, Vighetti S, Amanzio M, Casadio C, Oliaro A, Bergamasco B, et al. Dose–response relationship of opioids in nociceptive and neuropathic postoperative pain. Pain 1998;74:205–11.

Bird SB, Dickson EW. Clinically significant changes in pain along the visual analog scale. Ann Emerg Med 2001;38:639–43.

Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–10.

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

Cepeda MS, Africano JM, Polo R, Alcala R, Carr DB. What decline in pain intensity is meaningful to patients with acute pain? Pain 2003;105:151–7.

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

DeLoach LJ, Higgins MS, Caplan AB, Stiff JL. The visual analog scale in the immediate postoperative period: intrasubject variability and correlation with a numeric scale. Anesth Analg 1998;86:102–6.

Deyo RA, Centor RM. Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. J Chronic Dis 1986;39:897–906.

Deyo RA, Patrick DL. The significance of treatment effects: the clinical perspective. Med Care 1995;33:AS286–91.

Dhanani S, Quenneville J, Perron M, Abdolell M, Feldman BM. Minimal difference in pain associated with change in quality of life in children with rheumatic disease. Arthritis Rheum 2002;47:501–5.

Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 2005;113:9–19.

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 2001;94:149–58.

Farrar JT. What is clinically meaningful: outcome measures in pain clinical trials. Clin J Pain 2000;16:S106–12.

Feine JS, Lavigne GJ, Dao TT, Morin C, Lund JP. Memories of chronic pain and perceptions of relief. Pain 1998;77:137–41.

Felson DT, Anderson JJ. A review of evidence on the discriminant validity of outcome measures in rheumatoid arthritis. J Rheumatol 2001;28:422–6.

Fischer D, Stewart AL, Bloch DA, Lorig K, Laurent D, Holman H. Capturing the patient's view of change as a clinical outcome measure. J Am Med Assoc 1999;282:1157–62.

Fleiss JL. The design and analysis of clinical experiments. New York: Wiley; 1986.

Giraudeau B, Rozenberg S, Valat JP. Assessment of the clinically relevant change in pain for patients with sciatica. Ann Rheum Dis 2004;63:1180–1.

Grzybowski M, Younger JG. Statistical methodology: III. Receiver operating characteristic (ROC) curves. Acad Emerg Med 1997;4:818–26.

Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR. Methods to explain the clinical significance of health status measures. Mayo Clin Proc 2002a;77:371–83.

Guyatt GH, Norman GR, Juniper EF, Griffith LE. A critical look at transition ratings. J Clin Epidemiol 2002b;55:900–8.

Haas M, Nyiendo J, Aickin M. One-year trend in pain and disability relief recall in acute and chronic ambulatory low back pain patients. Pain 2002;95:83–91.

Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;148:839–43.

Hays RD, Woolley JM. The concept of clinically meaningful difference in health-related quality-of-life research. How meaningful is it? Pharmacoeconomics 2000;18:419–23.

Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials 1989;10:407–15.

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 2003;4:407–14.

Kelly AM. The minimum clinically significant difference in visual analogue scale pain score does not differ with severity of pain. Emerg Med J 2001a;18:205–7.

Kelly AM. Setting the benchmark for research in the management of acute pain in emergency departments. Emerg Med (Fremantle) 2001b;13:57–60.

Kropmans TJ, Dijkstra PU, Stegenga B, Stewart R, de Bont LG. Repeated assessment of temporomandibular joint pain: reasoned decision-making with use of unidimensional and multidimensional pain scales. Clin J Pain 2002;18:107–15.

Kvien TK, Dougados M, Mowinckel P, Skomsvoll JF, Mikkelsen K. Values for Minimal Clinically Important Improvement (MCII) in patients with Ankylosing Spondylitis (Poster). Annual Scientific Meeting of the American College of Rheumatology, San Antonio, Texas, 2004.

Lassere MN, van der Heijde D, Johnson KR, Boers M, Edmonds J. Reliability of measures of disease activity and disease damage in rheumatoid arthritis: implications for smallest detectable difference, minimal clinically important difference, and analysis of treatment effects in randomized controlled trials. J Rheumatol 2001;28: 892–903.

Lee JS, Hobden E, Stiell IG, Wells GA. Clinically important change in the visual analog scale after adequate pain control. Acad Emerg Med 2003;10:1128–30.

Moore A, McQuay H, Gavaghan D. Deriving dichotomous outcome measures from continuous data in randomised controlled trials of analgesics. Pain 1996;66:229–37.

Norman GR, Stratford P, Regehr G. Methodological problems in the retrospective computation of responsiveness to change: the lesson of Cronbach. J Clin Epidemiol 1997;50:869–79.

Salaffi F, Stancati A, Silvestri CA, Ciapetti A, Grassi W. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. Eur J Pain 2004;8:283–91.

Seres JL. The fallacy of using 50% pain relief as the standard for satisfactory pain treatment outcome. Pain Forum 1999;8:183–8.

Scott DL, Dacre JE, Greenwood A, Treasure L, Huskisson EC. Can we develop simple response criteria for slow acting antirheumatic drugs? Ann Rheum Dis 1990;49:196–8.

Stratford PW, Binkley JM, Riddle DL, Guyatt GH. Sensitivity to change of the Roland–Morris Back Pain Questionnaire: part 1. Phys Ther 1998;78:1186–96.

Todd KH. Patient-oriented outcome measures: the promise of definition. Ann Emerg Med 2001;38:672–4.

Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, et al. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. Ann Rheum Dis 2005a;64:29–33.

Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, et al. Evaluation of clinically relevant states in patient reported outcomes in knee and hip osteoarthritis: the patient acceptable symptom state. Ann Rheum Dis 2005b;64:34–7.

Tugwell P, Boers M, Brooks PM, Simon L, Strand CV. OMERACT 5: International consensus conference on outcome measures in rheumatology. J Rheumatol 2001;28:395–7.

van der Heijde DM, van 't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA, et al. Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. Ann Rheum Dis 1990;49:916–20.

Ward MM, Marx AS, Barry NN. Identification of clinically important changes in health status using receiver operating characteristic curves. J Clin Epidemiol 2000;53:279–84.

Wells G, Anderson J, Beaton D, Bellamy N, Boers M, Bombardier C, et al. Minimal clinically important difference module: summary, recommendations, and research agenda. J Rheumatol 2001;28: 452–4.

Wolfe F, Michaud K, Strand V. Expanding the definition of clinical differences: from minimally clinically important differences to really important differences. Analyses in 8931 patients with rheumatoid arthritis. J Rheumatol 2005;32:583–9.