Joint-Specific Multidimensional Assessment of Pain (J-MAP): Factor Structure, Reliability, Validity, and Responsiveness in Patients with Knee Osteoarthritis

KIMBERLY J. O'MALLEY, MARIA SUAREZ-ALMAZOR, JULIE ANIOL, PETER RICHARDSON, DAVID H. KUYKENDALL, J. BRUCE MOSELEY Jr, and NELDA P. WRAY

ABSTRACT. ***Objective***. To develop a reliable and valid instrument for measuring and monitoring joint-specific pain.

Methods. Developed using patient interviews, reviews of pain literature, and expert input from orthopedic surgeons, the final Joint-Specific Multidimensional Assessment of Pain (J-MAP) includes the 6-item Pain Sensory and the 4-item Pain Affect subscales. Scores on the J-MAP Pain Sensory and Affect subscales range from 0 to 100, with higher scores indicating more pain intensity and worse pain distastefulness, respectively. Following the assessment of the factor structure, patients' scores (n = 180) on the J-MAP subscales were converted to equal interval scores using Rasch analyses. A psychometric evaluation of the items and Rasch-calibrated scores was conducted and included an assessment of reliability, validity, and responsiveness for use with patients with radiographic knee osteoarthritis.

Results. Evidence from the factor analyses showed that the J-MAP Pain Sensory and Affect items made up 2 distinct factors. Internal consistency estimates for the J-MAP subscales exceeded 0.85. The J-MAP subscales showed evidence for validity and were shown to be internally and externally responsive, demonstrating greater responsiveness than the Arthritis Impact Measurement Scale or the Medical Outcome Study Short Form-36 pain subscales. Finally, evidence was found supporting the J-MAP subscales' ability to distinguish target joint pain from pain emanating from other musculoskeletal conditions.

Conclusion. The J-MAP is a reliable, valid, and responsive measure for assessing joint-specific pain at a single time point, or changes over time for one or a group of patients with knee osteoarthritis. With this initial evidence of its psychometric rigor, further testing of the measurement properties of the J-MAP in other joints and in other populations should be undertaken. (J Rheumatol 2003;30:534–43)

Key Indexing Terms:
JOINT PAIN OUTCOME MEASURE OSTEOARTHRITIS
VALIDITY EQUAL INTERVAL SCALING

Arthritis is a common and debilitating condition. Osteoarthritis (OA) affects about 12.1% of US adults or 20.7 million people[1]. In older adults, OA is one of the top conditions diagnosed by primary care physicians and is the primary cause of disability[2,3]. Further, with the increasing age of the population, the prevalence is expected to increase dramatically over the next 20 years[4-9].

The 2 most common complaints that patients with knee OA express include functional disability and pain. Although functional disability can exist with little pain, typically functional limitations and pain occur together. Indeed, the most common cause of functional limitation is joint pain. Several measures assessing joint pain currently exist. The most widely used include the Western Ontario and McMaster

From the Houston Center for Quality of Care and Utilization Studies, Houston VA Medical Center (VAMC), and Section of Health Services Research, Department of Medicine, Baylor College of Medicine, Houston, Texas, USA; International Survey Research, London, UK; and the Department of Orthopedic Surgery, Baylor College of Medicine.

Supported with resources and the use of facilities at the Houston Center for Quality of Care and Utilization Studies, Houston Veterans Affairs Medical Center, and by grant H-1639 from the US Department of Veterans Affairs.

K.J. O'Malley, PhD, Psychometrician, Assistant Professor; M. Suarez-Almazor, MD, PhD, Rheumatologist, Professor; J. Aniol, PhD, Clinical Psychologist, Instructor; P. Richardson, PhD, Statistical Programmer, Houston Center for Quality of Care and Utilization Studies, Houston VAMC, and Section of Health Services Research, Department of Medicine, Baylor College of Medicine; D.H. Kuykendall, PhD, Psychometrician, Director of Research and Project Director, International Survey Research, London, UK; J.B. Moseley Jr, MD, Clinical Associate Professor, Houston Center for Quality of Care and Utilization Studies, Houston VAMC, and Section of Health Services Research, Department of Medicine; N.P. Wray, MD, MPH, Senior Researcher, Chief of the Section of Health Services Research, Professor, Department of Orthopedic Surgery, Baylor College of Medicine.

Address reprint requests to Dr. K.J. O'Malley, VA Medical Center (152), 2002 Holcombe Boulevard, Houston, TX 77030.
E-mail: komalley@bcm.tmc.edu

Submitted February 11, 2002; revision accepted August 28, 2002.

University Osteoarthritis Index (WOMAC)[10], the McGill Pain Questionnaire[11], the Arthritis Impact Measurement Scale (AIMS2)[12], the Medical Outcome Study Short Form-36 Bodily Pain subscale (SF-36 BP)[13], various forms of the visual analog scales (VAS), and single-item 5-point Likert pain scales. From the clinical orthopedic literature, some common knee measures that incorporate an assessment of pain include the Marshall Knee Score[14], the Knee Society Clinical Rating Scale[15], and the Hospital for Special Surgery Knee Score[16].

An important clinical limitation of several of these measures including McGill, AIMS2, and SF-36 BP is that they do not assess pain in a specific joint and they focus on activities that require overall body movement, leading to confounding from pain in other areas. If an intervention is conducted to improve pain in a single joint (e.g., surgery or local injections), pain measures that assess joint pain in general (i.e., arthritis pain or overall body pain) confound the assessment of target joint pain with pain from other sources. Whereas the WOMAC and VAS measures have been used to assess pain in a specific knee joint, psychometric testing of these joint-specific measures has been minimal. In addition, the bulk of the psychometric research on the WOMAC has not been conducted on the version that asks about pain in one joint. Further, evidence supporting that these joint-specific measures are able to distinguish target joint pain from pain emanating from other musculoskeletal conditions has not been published. Regarding the WOMAC, studies have presented evidence that the WOMAC is *not* able to distinguish target joint pain from back pain or other joint pain[17,18].

Another clinical disadvantage of many of the current measures such as the WOMAC, AIMS2, SF-36 BP, VAS, the single-item scales, and the clinical measures from the orthopedic literature is that they do not assess both the sensory and affective components of pain. Past research has provided evidence that pain is multidimensional, with distinct sensory and affective dimensions[19-23]. Further, differential patient responses to pain dimensions can influence the way in which clinicians choose to intervene. For example, high pain intensity responses might indicate a need for pain medications, whereas high pain affect responses may indicate the need for a psychological intervention. Therefore, an instrument that measures only one dimension will limit clinicians' ability to adequately understand the complexity of the pain and treat the pain most appropriately[19,20,22-25]. Another practical limitation of some pain measures is their complexity. For example, many of the adjectives patients are asked to rate on the McGill, such as gnawing, lancinating, and gruelling, are difficult to understand and differentiate, especially for older adults who may suffer from cognitive difficulties.

In addition to clinical and practical limitations of current instruments, the psychometric qualities demonstrated by currently used pain instruments vary. For example, VAS and single-item Likert scales are typically used as single-item scales; therefore they have poor reliability[21,26,27]. Most of the clinical measures from the orthopedic literature have undergone little, if any, psychometric testing. Moreover, none of the scales have been recalibrated so their scores have equal interval properties, resulting in equal score differences that may not be representative of equal differences in the measured trait[28,29]. Instruments without equal interval properties reduce the accuracy of information used to monitor patients' progress over time, since a 10 point score change from 20 to 30 may not mean the same as a 10 point score change from 50 to 60.

Given the shortcomings of current pain instruments, there is a need for a joint-specific pain measure that is easy to administer and psychometrically sound. Thus, we undertook the current study to develop the Joint-Specific Multidimensional Assessment of Pain (J-MAP). To be most useful for clinicians and researchers, the J-MAP needed to (1) assess both pain sensory and pain affect; (2) focus on pain from the target joint; (3) be easy to administer; (4) demonstrate adequate reliability, equal interval properties, validity, and responsiveness; and (5) distinguish target joint pain from pain from other conditions.

We describe the development and factor structure of the J-MAP, describe the Rasch analyses used to scale the J-MAP, evaluate the final J-MAP instrument's reliability, validity and responsiveness to change with patients suffering from knee joint pain, and assess whether the J-MAP is able to differentiate target knee pain and pain from other conditions.

MATERIALS AND METHODS

Data for this study were collected in a double-blind randomized placebo controlled trial comparing 3 treatments (arthroscopic debridement plus lavage, arthroscopic lavage alone, and placebo arthroscopy) for patients with OA of the knee. The local Veterans Affairs Institutional Review Board (Houston, Texas, USA) approved the study's protocol. Patients from Veterans Affairs clinics were invited to participate in the study if one of the investigators (JBM) determined the presence of OA according to the American College of Rheumatology definition of knee OA[30]. Study candidates were not older than 75 years of age, and in spite of maximum medical treatment for 6 months or more, reported at least moderate knee pain on average over the past week. Exclusion criteria included: (1) potential candidacy for total knee replacement; (2) high risk for anesthesia (rating of 3 or higher on the American Society of Anesthesiology Classification of Physical Status Scale)[31]; (3) functional limitation due to medical conditions other than knee OA; (4) high likelihood for dropping out of study (e.g., history of drug or alcohol abuse, or mental illness or homelessness); or (5) knee arthroscopy in the 2 years prior to the study. A more detailed description of the study protocol has been published and detailed results of the trial are presented elsewhere[32,33].

Data collection. Data collection occurred from October 1994 to October 2000. Patients completed self-report measures at baseline, 2 weeks, 6 weeks, 3 months, 6 months, 1 year, 18 months, and 2 years. Patient self-report measures used in this study included multiple measures of pain, function, satisfaction, physical health, and emotional health (see below).

Development of the J-MAP. The original J-MAP items, presented in Table 1,

Table 1. Original J-MAP items.

Item	Responses	Meaning
Pain intensity		
1. How much pain are you currently having in your left/right* knee?	1–7	Severe pain to no pain
2. At the present time (right now), how intense is your left/right knee pain?	0–10	No pain to bad as could be
3. In the past week, how intense was your worst left/right knee pain?	0–10	No pain to bad as could be
4. In the past week, ON THE AVERAGE, how intense was your left/right knee pain?	0–10	No pain to bad as could be
5. On about how many days have you had knee pain in the last week in your left/right knee?	0–7	No days to all days
6. On days when you've had knee pain in the past week, how many hours were you usually in pain in your left/right knee?**	0–24	No hours to all hours
Pain distaste		
1. How pleased are you with your current situation regarding pain?	1–5	Very displeased to very pleased
2. How much of a problem do you have with pain because of your left/right knee?	1–6	None to very severe
3. How satisfied are you with your current situation regarding pain?	1–5	Very satisfied to very dissatisfied
4. Compared to other people your age, do you rate your situation regarding pain as...	1–5	Very poor to excellent

Note. See scoring and conversion guide for instructions on calculating the total J-MAP scores (Tables 2 and 3). *Only "left" or "right" was included to indicate a patient's specific study knee. To assess pain in an alternative joint, substitute the joint name for "knee."**Item 6 was not included in the final J-MAP.

were developed based on interviews with patients in a preliminary feasibility trial, reviews of pain literature, and expert input from 2 orthopedic surgeons. Melzack and Wall's gate control theory[22] formed the theoretical underpinnings of the J-MAP. This theory posits that the report of pain is a result of integrating physiological and psychological factors. The J-MAP incorporates the sensory-discriminative and motivation-affective components described in the gate control theory, where the sensory component relates to the magnitude and spatiotemporal properties of the noxious stimulus, and the affective component relates to the emotional response triggered by the noxious stimulus[19,20,23,24,34].

Pain sensory items were created by adapting and expanding the pain intensity items from the Wisconsin Brief Pain Questionnaire and adding pain duration items[35]. The 6 original pain sensory items asked respondents to rate their target joint (knee in this study) based on current level of pain, intensity of pain (current, in past week, average), or duration of pain over the past week. The 4 original pain affect items were created by combining positive affect pain items such as how pleased and satisfied a person is with the pain and negative affect pain items such as how problematic the pain feels[23,36].

Other measures. Several measures of pain, satisfaction, physical health, and emotional health were used to evaluate the validity and responsiveness of the J-MAP. In addition to the J-MAP, participants completed 2 existing pain measures shown to be reliable and valid for use with OA patients, including the 4-item revised Arthritis Impact Measurement Scales (AIMS 2) pain subscale[12,37] and the 2-item bodily pain subscale from the SF-36[13]. In our study, these 2 pain scales had a median Cronbach's alpha (over 8 occasions) of 0.84. Patients also completed a 3-item knee-specific pain measure that assessed pain intensity in the alternate knee. This alternate knee measure was shown to have high internal consistency (median Cronbach's alpha of 0.96 over all occasions) and evidence of unidimensionality.

Patient satisfaction with their general health and with their knee procedure was assessed using the 2-item General Health Satisfaction (GHS) and 4-item Satisfaction with Knee Procedure (SKiP) scales, respectively. Both satisfaction scales were developed for this study[38]. On the GHS patients reported how satisfied and pleased they were with their general health. On the SKiP, patients reported whether the operation was worthwhile, if they were helped by the procedure, if they thought the procedure was a waste of time, and if they would recommend the knee procedure to a family member if they needed care for the same problem. Evidence for the unidimensionality of these scales was demonstrated in confirmatory factor analytic models, and the median Cronbach's alpha over the assessment occasions for the GHS and SKiP, using data from this study, was 0.90 and 0.92, respectively.

Patients' physical health was evaluated using 2 physical functioning and 3 other physical health measures. Physical functioning was assessed with the 5-item AIMS2 Walking and Bending and the 10-item SF-36 physical functioning subscales[12,13,37]. The three other physical health measures included the SF-36 General Health, Role Physical, and Social Functioning subscales[13,63]. Patients completed 4 emotional health measures including 3 SF-36 subscales (Mental Health, Role Emotional, and Vitality), and the Perceived Stress Scale[13,40]. The median Cronbach's alpha for these physical and emotional health measures exceeded 0.75 in this study.

Analyses

Confirmatory factor analysis. Our first step in the analysis process was to evaluate whether data from the 10 J-MAP items indicated 1 latent variable or 2. To assess the underlying dimensions in our data, we initially fit item-level data to a model with 1 latent variable using Lisrel 8.14 for Windows[41]. We then fit a 2 latent variable model in which the pain sensory items indicated 1 latent variable and pain affect items indicated the other. We expected a significant correlation between the 2 latent variables, so we allowed them to correlate in the model. In addition, we expected shared format variance among items with the same number of response categories and between the 2 duration items. Therefore, we allowed the residual correlations among these items to be freely estimated in both the 1 and 2 latent variable models.

Fit of the data to the latent variable models was assessed using global and local fit indices. Global fit indices, which inform about how the model as a whole reproduces the relations among the variables[42,43], were assessed using 5 fit indices. The indices included the chi-square goodness-of-fit statistic, the Root Mean Square Error of Approximation (RMSEA), the Goodness-of-fit and Adjusted Goodness-of-fit indices, and the Comparative Fit Index[41,44-47]. Overall, we considered a nonsignificant chi-square test combined with RMSEA value below 0.08, and goodness-of-fit, adjusted Goodness-of-fit, and comparative fit indices 0.95 and higher to indicate evidence that our model adequately represented our data, or that the amount of misfit was small. Local fit indices, which reflect fit of specific features of the model, were the squared multiple correlations, modification indices, and significance tests of the structural parameters relating items to latent variables. Squared multiple correlations > 0.30, modification indices < 10, and significance tests with $p < 0.05$ were considered evidence of good local fit[42-45].

Rasch analyses. We intend to use the J-MAP to monitor patients' pain over time, so it was necessary that the scale have equal-interval properties. An equal-interval scale is one in which equal increases or decreases in score units correspond with equal changes in the outcome being measured[28,29].

We converted the raw J-MAP scores to equal-interval scales using the partial credit Rasch model, which is appropriate for items with multiple response categories[48], and calibrated the Pain Sensory and Pain Affect items and patient scores on the same measurement continuum using BIGSTEPS software[49]. We evaluated the appropriateness of the Rasch model by examining the outfit and infit, statistics that evaluate the fit between the scale items and the model with which they were calibrated[49,50]. Once the final items were selected, the J-MAP subscale scores were rescaled to range from 0-100, and these rescaled scores were used in all subsequent analyses involving total scores.

Reliability. The internal consistency of the final version of the J-MAP was evaluated using Cronbach's alpha coefficient at each of the 8 occasions of measure. The J-MAP was constructed for application in groups of patients and in individual patients. Since a scale applied to individuals requires higher reliability than the scale applied to groups, we considered internal consistency estimates ≥ 0.80 as adequate for group comparisons and estimates ≥ 0.90 adequate for individual application[27]. Since patients in our study received an intervention, their knee symptoms were expected to change over the course of the trial, leading to reproducibility (test-retest) estimates that would confound true symptom change with change due to random error. Therefore, we did not estimate reproducibility.

Validity. By comparing the J-MAP to other measures at baseline and the final occasion, we examined the convergent and discriminant validity of the J-MAP[51,52]. We expected high correlations between scales measuring the same construct, low correlations between scales measuring different constructs, and moderate correlations between scales measuring related constructs as evidence of construct validity. Specifically, for the Pain Sensory subscale, we expected high correlations (in absolute value) with other measures of pain, and moderate correlations with satisfaction and other measures of physical health, since pain, satisfaction, and physical health have been shown to have moderate correlations in other studies[17,53]. We expected the lowest correlations between the J-MAP Pain Sensory subscale and the emotional health measures. For the Pain Affect subscale, we expected high correlations (in absolute value) with other measures of pain and satisfaction and moderate correlations with other measures of physical and emotional health measures.

We defined our categories of correlations based on the suggested values of Cohen[54]: a correlation of 0.50 represents a large effect size, since it indicates that 25% of the variance in one variable is linearly associated with variance in the other. A correlation of 0.30 represents a medium effect size, since it indicates that the variables share 9% of their variance. Similarly, a correlation of 0.10 represents a small effect size, since it indicates that the variables share 1% of their variance. Given Cohen's recommended values, we categorized correlations ranging from 0 to 0.20 as low, 0.21–0.40 as moderate, and ≥ 0.41 as high.

Responsiveness. Responsiveness is the ability of a scale to detect change, where change is defined as the minimal amount of change considered important by patients[55,56]. Researchers have defined internal responsiveness as the within-person scale score changes over a prespecified timeframe during which a treatment of known effectiveness is administered, and external responsiveness as the extent to which score changes on a measure relate to corresponding changes in a reference measure of health status[56].

Before evaluating the responsiveness of the J-MAP subscales, we estimated the minimal clinically important difference (MCID) for the two J-MAP subscales, where the MCID is defined as "the smallest difference in measured health status that signifies an important rather than trivial difference in patient symptoms"[57,58]. Researchers have calculated the MCID using both empirical and patient-based methods[57-62]. We determined the MCID for this study by taking the average of an empirical and patient-based MCID. The empirical MCID was calculated at each occasion as the standard error of measurement (SEM), obtained by multiplying the standard deviation of the measure times the square root of the reliability coefficient subtracted by 1[61,62]. We used the internal consistency estimate as the reliability coefficient. We then used the average empirical MCID over the 8 occasions as the empirical MCID. The anchor-based MCID was calculated as the average score change for patients who reported that they were somewhat better on the SF-36 Reported Health Transition question, which asks, "Compared to one year ago, how would you rate your health in general now?"[63]. We averaged the patient-based MCID from Year 1 and Year 2 and used this average as the patient-based MCID. Finally, we took the average of the empirical MCID and the patient-based MCID as the final MCID for this study.

Once the MCID was defined, we evaluated internal responsiveness of the J-MAP by comparing all patient scores on both subscales at baseline to their scores at 6 weeks after treatment, a time when we expected the most improvement. Since no pain differences were found between the 3 treatment groups at baseline or 6 weeks[33], we combined the 3 treatment groups in the responsiveness analyses. We examined the 95% confidence interval (CI) around the 2 J-MAP subscale mean differences at 6 weeks and evaluated whether the intervals excluded and exceeded the MCID. Since the MCID signified clinically significant change and the interval around the mean difference signified the score range in which we were 95% confident that the true mean difference fell, we considered an interval above the MCID not overlapping with the MCID as evidence that change was clinically significant. We therefore considered CI greater than the MCID as evidence that the scores on the J-MAP were responsive to patients' clinical changes.

After gathering evidence for the internal responsiveness of the J-MAP, we compared the internal responsiveness of the J-MAP to the internal responsiveness of the AIMS2 and SF-36 pain subscales. Using paired samples t tests comparing baseline and 6-week scores, we calculated the effect sizes of the score changes and compared these effect sizes.

We evaluated external responsiveness using the linear regression model proposed by Husted, *et al*[56]. We regressed the difference in patient responses to the SF-36 health transitional index from baseline to 6 weeks on the difference in the J-MAP subscales over the same period[63]. Since the regression coefficient represented the average increase in the health transition index associated with a one-unit change in the J-MAP subscales, the significance of the regression coefficients served as the test of external responsiveness for the subscales. In addition, we examined the squared multiple correlations, or R^2 values, as measures of the effect size of the external responsiveness.

Discrimination of target joint pain. To evaluate the extent to which the J-MAP could discriminate target joint pain from other joint pain or pain emanating from other musculoskeletal conditions, we conducted 2 sets of analyses. (1) We compared pain changes in the study knee, using the J-MAP from baseline to Year 1, to pain changes in the alternate knee from baseline to Year 1. Since the study intervention was only performed in one knee, significant pain changes in the treated knee as measured by the J-MAP combined with no significant pain changes in the alternate knee as measured by the alternate knee pain measure would provide evidence that the J-MAP could distinguish target knee pain from alternate knee pain. We tested for the significance of the pain changes in the 2 knees using paired samples t tests. (2) We compared target knee pain changes using the J-MAP subscales for patients who did and did not report other musculoskeletal conditions such as low back pain, hip arthritis, and sciatica. We hypothesized that if the J-MAP subscales did distinguish between the target knee pain and pain from other conditions, changes in J-MAP scores would not differ significantly for patients with and without other conditions. We tested our hypothesis using 6 general linear models. In 3 of the models, the dependent variable was the J-MAP Pain Sensory score at Year 1, the independent variable was the dichotomous variable indicating the presence or absence of other orthopedic conditions (low back pain, hip arthritis, and sciatica), and the covariate was the J-MAP Pain Sensory score at baseline. In the other 3 models, we substituted the J-MAP Pain Affect scores for the J-MAP Pain Sensory scores.

RESULTS

Participants. The average age of the 180 study participants was 56 ± 11 years, 93% were male, 62% were white, 29% were black, and 57% were married. Regarding OA severity, 29% of the patients had mild, 46% had moderate, and 25% had severe radiographic knee OA.

To assess how similar these patients' functional status was to the general population, we compared our study patients' SF-36 subscale means to the normative means published for the general population[63]. Results indicated that study patients reported statistically significantly worse bodily pain (38.0 ± 17.5 vs 75.2 ± 23.5), general health (63.0 ± 19.6 vs 72.0 ± 20.3), physical function (44.5 ± 22.5 vs 84.3 ± 23.3), role-emotional (64.4 ± 43.3 vs 81.3 ± 33.0), role-physical (33.9 ± 39.1 vs 81.0 ± 34.0), social functioning (64.4 ± 25.0 vs 83.3 ± 22.7), and vitality (55.0 ± 20.0 vs 60.9 ± 21.0) than the general population ($p < 0.01$). Study participants reported similar mental health compared to the general population (75.5 ± 18.8 vs 74.7 ± 18.0) ($p > 0.05$).

Missing data. Of the 180 study participants, 165 (92%) had complete J-MAP data at the end of the trial. We assessed whether participants with complete and incomplete data at the end of the trial differed on their J-MAP subscales, other pain scales, and demographic variables at the start of the trial using t tests for continuous variables and chi-square tests for categorical data. Results indicated that participants with complete J-MAP Pain Sensory data did not differ from those with incomplete data on the J-MAP Pain Sensory measure ($p = 0.42$), J-MAP Pain Affect measure ($p = 0.39$), the AIMS2 pain subscale ($p = 0.65$), and the SF-36 pain subscale ($p = 0.97$) at baseline. In addition, those who were and were not missing J-MAP Pain Sensory data at the end of the trial did not differ on age ($p = 0.36$), general health ($p = 0.94$), mental health ($p = 0.85$), or radiographic knee severity ($p = 0.61$). Results indicated that participants with complete J-MAP Pain Affect data did not differ from those with incomplete data on the J-MAP Pain Sensory measure ($p = 0.50$), J-MAP Pain Affect measure ($p = 0.24$), the AIMS 2 pain subscale ($p = 0.69$), or the SF-36 pain subscale ($p = 0.89$) at baseline. In addition, those who were and were not missing J-MAP pain intensity data at the end of the trial did not differ on age ($p = 0.47$), general health ($p = 0.81$), mental health ($p = 0.62$), or radiographic knee severity ($p = 0.69$).

Confirmatory factor analyses. Results from the confirmatory factor analyses with one latent variable provided evidence of less than adequate global fit [chi-square (degrees of freedom, df, 29) = 330.27, $p = 0.00$, RMSEA = 0.25, Goodness-of-fit index = 0.72, Adjusted Goodness-of-fit index = 0.47, and Comparative Fit index = 0.78]. Local fit indices indicated that the single latent variable model did not account for the correlations among the affect items well. Global fit indices for the 2 latent variable model with the 6 intensity items comprising one latent variable and the 4 affect items comprising the second latent variable indicated much better fit [chi-square (df 28) = 35.7, $p = 0.15$, RMSEA = 0.04, goodness-of-fit index = 0.96, adjusted goodness-of-fit index = 0.92, and Comparative Fit Index = 0.99]. Local fit indices also provided evidence of good local fit. Specifically, all squared multiple correlations exceeded 0.30, with one exception (0.17 for Item 6). Modification indices all fell below 8.0 and all structural parameters were statistically significant, with $p < 0.01$. The correlation between the 2 latent variables was 0.57.

Rasch analyses. Using Rasch analyses, we examined the standardized fit statistics, effective measurement range, and reliability of the J-MAP Pain subscales to evaluate whether item deletions were necessary and to rescale the total scores. We decided to retain all items except Item 6 of the Pain Sensory subscale. The infit and outfit statistics for Item 6 were greater than 4.0. In addition, Item 6 has the lowest squared multiple correlation in the confirmatory factor analysis. After eliminating Item 6, we reexamined the standardized fit statistics, effective measurement range, and reliability of the J-MAP Pain Sensory subscale without Item 6 and found them improved. The final version of J-MAP includes all items shown in Table 1 except Item 6, and the score conversion tables are presented in Tables 2 and 3. The final J-MAP Pain Sensory and Pain Affect subscales ranged from 0 to 100, with higher scores indicating more pain and worse pain affect, respectively.

Reliability. The median internal consistency coefficients for the J-MAP subscales were 0.90 for the Pain Sensory subscale (range 0.85–0.94) and 0.86 for the Pain Affect subscale (range 0.82–0.90). These estimates both met the criterion for group comparisons; however, the 0.86 coefficient for the Pain Affect subscale fell just short of our preset criterion for individual comparisons.

Validity. The pattern of correlations between the J-MAP Pain Sensory and Pain Affect subscales and other measures of pain, satisfaction, physical health, and emotional health at baseline and the final occasion (Table 4) provided evidence for convergent and discriminant validity of the J-MAP subscales. As predicted, scores on the J-MAP Pain Sensory subscale were highly correlated (0.49–0.64) with the other pain measures. The Pain Sensory subscale correlated moderately and highly with the satisfaction and other physical health measures at baseline (0.36–0.43) and at the final occasion (0.28–0.56), with one exception (SF-36 Role Physical). Low to moderate correlations were found between the J-MAP Pain Sensory subscale and measures of emotional health on both occasions (0.17–0.35).

Correlations between the J-MAP Pain Affect subscale and the other measures of pain and satisfaction were high at both occasions (0.56–0.79). The J-MAP Pain Affect subscale was moderately and highly correlated with other measures of physical health at both occasions (0.24–0.62) and moderately correlated with emotional health measures

Table 2. J-MAP Pain Sensory scoring and conversion table. To calculate the final J-MAP Pain Sensory score, complete the following: Step 1: Reverse score Items 1 and 3. Step 2: Calculate the total raw score as the sum of Item 1 (reversed), Item 2, Item 3 (reversed), Item 4, and Item 5. Step 3: Convert the raw score to the scaled score using the table.

Raw Score	Scaled Score	Raw Score	Scaled Score	Raw Score	Scaled Score
1	0	16	39	31	52
2	7	17	40	32	53
3	15	18	41	33	54
4	20	19	41	34	56
5	23	20	42	35	57
6	26	21	43	36	59
7	28	22	44	37	62
8	30	23	45	38	65
9	32	24	45	39	68
10	33	25	46	40	73
11	34	26	47	41	78
12	35	27	48	42	83
13	36	28	49	43	92
14	37	29	50	44	100
15	38	30	51		

Table 3. J-MAP Pain Affect scoring and conversion table. To calculate the final J-MAP Pain Affect score, complete the following: Step 1: Reverse score Items 1 and 4. Step 2: Calculate the total raw score as the sum of Item 1 (reversed), Item 2, Item 3, and Item 4 (reversed). Step 3: Convert the raw score to the scaled score using the table.

Raw Score	Scaled Score	Raw Score	Scaled Score
4	0	14	54
5	0	15	58
6	6	16	63
7	14	17	69
8	21	18	75
9	28	19	82
10	36	20	92
11	42	21	100
12	47		
13	51		

at both occasions (0.23–0.34), with 2 exceptions (Perceived Stress Scale and SF-36 Role Emotional).

Overall, the patterns of correlations between the J-MAP subscales and the other measures conformed to most of our predictions except that correlations were slightly higher than expected with several of the satisfaction, physical health, and emotional health measures. Nonetheless, given that none of the correlations were dramatically different from our expectations and the pattern of results corresponded with our hypotheses, we considered the evidence for the convergent and discriminant validity of the J-MAP Pain Sensory and Pain Affect subscales to be strong.

Responsiveness. The average empirical MCID (over the 8 occasions) for the J-MAP Pain Sensory subscale calculated using the standard error of measurement (SEM) method was 5.0 (range 4.5–5.3). For the Pain Affect subscale, the average empirical MCID was 8.0 (range 7.4-8.5). The average patient-based MCID for the J-MAP Pain Sensory and Pain Affect subscales calculated from Years 1 and 2 were 8.6 and 12.4, respectively. Therefore, the final MCID values for the J-MAP Pain Sensory and Pain Affect subscales, calculated as the average of the MCID from the 2 methods, were 6.8 and 10.2, respectively.

Results from the evaluation of internal responsiveness showed that the CI around the mean score changes on the J-MAP subscales at 6 weeks were sensitive enough to detect knee pain differences after treatment. Specifically, the 95% CI around the Pain Sensory subscale mean change of 11.3 scale points (8.7, 14.0) excluded and exceeded the 6.8 MCID. The 95% CI around the Pain Affect subscale mean change of 13.4 scale points (10.3, 16.5) also excluded and exceeded the MCID of 10.2.

Comparison of the internal responsiveness effect sizes of the J-MAP, the AIMS 2 pain scale, and the SF-36 pain subscale revealed that the J-MAP subscales were the most internally responsive. Specifically, the effect sizes comparing baseline and 6-week scores were 0.65 and 0.66 for the J-MAP Pain Sensory and Pain Affect subscales, respectively. In contrast, the effect sizes for the AIMS2 and SF-36 pain subscales were 0.45 and 0.48, respectively. The J-MAP subscales therefore showed superior internal responsiveness compared to the other 2, more general, pain measures.

Evidence for external responsiveness of the J-MAP was found from the linear regression models. Specifically, the regression coefficients (and standard errors) for the J-MAP Pain Sensory and Affect subscales were 0.02 (0.004) and 0.01 (0.004), $p < 0.001$ for both. These coefficients showed that a one-unit change in the J-MAP subscales (with scores ranging from 1 to 100) represented a 0.02 and 0.01 change in the health transitional index, respectively. The R^2 values

Table 4. Correlations between the J-MAP and measures of pain, satisfaction, physical health, and emotional health: evidence for convergent and discriminant validity. All correlations are presented as absolute values.

	Initial Occasion		Final Occasion	
	Pain Sensory	Pain Affect	Pain Sensory	Pain Affect
J-MAP measures				
Pain sensory	1.0	0.56**	1.0	0.68**
Pain affect	0.56**	1.0	0.68**	1.0
Pain measures				
AIMS2 pain	0.49**	0.56**	0.64**	0.68**
SF-36 bodily pain	0.63**	0.59**	0.59**	0.65**
Satisfaction measures				
Satisfaction with knee procedure†	N/A	N/A	0.55**	0.62**
General health satisfaction	0.36**	0.64**	0.52**	0.79**
Physical health measures				
AIMS2 walking and bending	0.43**	0.53**	0.56**	0.62**
SF-36 physical functioning	0.40**	0.44**	0.48**	0.53**
SF-36 general health	0.36**	0.34**	0.28**	0.48**
SF-36 role physical	0.15	0.24*	0.35**	0.45**
SF-36 social functioning	0.42**	0.44**	0.39**	0.40**
Emotional health measures				
SF-36 mental health	0.33**	0.33**	0.35**	0.34**
Perceived stress scale	0.32**	0.19	0.27*	0.28**
SF-36 role emotional	0.17	0.08	0.29*	0.23*
SF-36 vitality	0.31**	0.29**	0.23*	0.34**

J-MAP:Joint-Specific Multidimensional Assessment of Pain, AIMS2: Arthritis Impact Measurement Scale Revised; SF-36: Medical Outcomes Trust Health Survey. †Satisfaction with knee procedure was not assessed preoperatively, so correlations are only presented at the final occasion. *$p < 0.01$, **$p < 0.001$.

for the models were 0.08 and 0.07, respectively, suggesting that the proportion of variance for the health transitional index explained by the J-MAP subscale changes was just less than 10%. The significance of these coefficients and the R^2 values provided evidence of the external responsiveness of the J-MAP subscales.

Discrimination of target joint pain. The analysis comparing pain changes for the target knee using the J-MAP and pain changes for the alternate knee using the 3-item pain sensory measure provided evidence that the J-MAP was able to discriminate target knee pain from alternate knee pain. Specifically, the paired samples t tests comparing patients' study knee pain from baseline to Year 1 using the J-MAP Pain Intensity and Pain Affect were 6.7 and 7.2 ($p < 0.001$). The paired samples t test comparing patients' alternate knee pain from baseline to Year 1 was –1.4 ($p = 0.17$). In fact, patients reported an average *increase* of pain in the alternate knee over the year (mean increase of 2.5 points), whereas they reported an average decrease in pain for the J-MAP subscales. These results indicate that the J-MAP indeed differentiated target knee pain from alternate knee pain.

Results of analyses comparing knee pain changes (using the J-MAP) for patients with and without other musculoskeletal conditions provided support that the J-MAP discriminated target joint pain and pain emanating from other conditions. Specifically, the decrease in J-MAP Pain Sensory scores for patients with and without low back pain and sciatica was not statistically different. In addition, the decrease in J-MAP Pain Affect scores for patients with and without low back pain, hip arthritis, and sciatica was not statistically different. The only exception was that decreases in the J-MAP Pain Sensory scores were significantly greater for patients with hip arthritis than for patients without hip arthritis ($p = 0.03$). Given that we conducted 6 statistical tests without correcting for multiple comparisons and that the p value for this single significant test was close to 0.05, we did not find this exception to indicate strong evidence against our hypothesis that the J-MAP was able to discriminate target joint pain and pain from other musculoskeletal conditions.

DISCUSSION

Our results indicate that the J-MAP is a reliable, valid, and responsive outcome measure for assessing and monitoring knee pain in patients with OA. The J-MAP offers several important advantages over other pain measures including that (1) it assesses pain in the study joint and is able to discriminate study joint pain and pain from other conditions; (2) it has been scaled to have equal-interval properties and can therefore more optimally be used to evaluate patients' changes in joint pain; (3) it assesses both sensory and affective dimensions of pain; (4) it has undergone thorough psychometric testing, confirming its excellent measurement properties; and (5) it is more responsive than more generic pain measures.

By focusing on the target joint, the J-MAP subscales are

more likely than more generic pain measures to isolate target joint pain. Our finding that the J-MAP subscales distinguished pain in the study knee from pain in the other knee and from pain due to other conditions supports this advantage. Although other joint-specific measures of knee pain are available, including visual analog measures, the WOMAC, and several pain measures from the orthopedic literature, the psychometric testing of these measures has been minimal and no evidence has been presented that these other joint-specific measures can discriminate target joint pain and pain from other conditions.

The equal-interval property of J-MAP serves as one of its most important advantages over other measures, since change scores on the J-MAP represent meaningful symptom changes across the measurement range continuum. Reduction or relief of joint pain is often the primary goal of therapy. Since reduction and relief imply change, meaningful change scores are a critical characteristic of a joint pain measure.

J-MAP assessment of sensory and affective pain dimensions promotes a richer understanding of a patient's pain compared to one-dimensional measures. Clinicians and researchers can use the J-MAP to understand the intensity of patient knee pain and the extent to which this pain poses a problem for patients. This more detailed information can assist clinicians and researchers as they tailor interventions to individual needs of patients.

Regarding the J-MAP's psychometric properties, internal consistency estimates suggest that the J-MAP is as reliable or more reliable than most pain measures[10,64]. Evidence of the validity of J-MAP is strong and comparable to that of other pain measures, suggesting that J-MAP is a valid tool for assessing knee pain in patients with OA. Evidence of responsiveness in the form of confidence intervals around the J-MAP change scores from baseline to 6 weeks that excluded the MCID and a significant relation between changes in the J-MAP and changes in an external health index is commensurate with that of other pain measures such as the AIMS2, SF-36 Bodily Pain Subscale, and the WOMAC[64-67]. Moreover, J-MAP is sensitive enough to detect meaningful pain changes over time; that responsiveness of J-MAP subscales was greater than that of AIMS2 and SF-36 pain subscales provides evidence that the J-MAP is more responsive to change than more generic measures.

Two aspects of the findings deserve further discussion: (1) The internal consistency estimates for the 2 J-MAP subscales met recommended values for group comparisons, and we recommend using the current version of the J-MAP for this purpose. The internal consistency estimate for the Pain Sensory subscale reached the recommended internal consistency value for individual monitoring purposes; however, the estimate for the Pain Affect subscale was slightly lower than the recommended value. Researchers and clinicians need highly reliable pain measures, to monitor joint pain in individual patients, that have reliability coefficients at or above 0.90[27]. One way to improve the internal consistency of the Pain Affect subscale is to add items[27,52], therefore further research evaluating the addition of a fifth and possibly sixth pain affect item would be beneficial. (2) The validity coefficients between the J-MAP subscales and a few of the satisfaction, physical health, and emotional health measures were higher than predicted according to Cohen's classification scheme[54]. A possible explanation for these higher coefficients derives from the conceptualization of pain according to Melzack and Wall's gate control theory[22]. Pain perception integrates physiological and psychological factors, thereby resulting in a latent variable that should theoretically relate to the physiological and psychological variables we used to assess convergent and discriminant validity[23,34]. In addition, many of the measures, such as the SF-36 Social Functioning and Role Emotional measures, conceptually overlapped when we classified them as physical and emotional health measures.

Limitations. The J-MAP was developed to assess pain in any joint in a variety of populations. Our analysis of its psychometric properties was based on a population of 180 veterans with knee OA who used the J-MAP to rate knee pain. Whether the psychometric properties can be replicated in other populations, such as with younger female patients, is unknown without further testing. Although we expect that the psychometric properties will hold up in other populations, this question warrants empirical study. In addition, the trial from which these data are used did not include patients who had minor sports related injuries or those with more severe arthritis requiring joint replacement, therefore there is potential for ceiling and floor effects for this scale in these populations. Further, whether the J-MAP can be used to assess pain in other joints such as elbows or shoulders has also not been confirmed in this study. Therefore, further assessment of the J-MAP in other joints and populations is needed. If the J-MAP demonstrates strong psychometric properties in other joints and with other populations, it will offer a single, versatile tool with potential to save time, facilitate comparisons across target joints, and allow application in multiple diseases (arthritis, sport injuries, etc.).

Another limitation of our study is that we did not include data from the WOMAC, currently the most common instrument for patients with knee OA. When we started data collection in 1994, the WOMAC was a relatively new measure with few studies available reporting its use. A comparison of the J-MAP with the WOMAC should be conducted.

When we compared the responsiveness of the J-MAP to other measures, our analyses were somewhat limited because we did not include joint-specific measures. Disease-specific measures are expected to be more responsive than more generic measures; therefore our finding that the J-MAP was more responsive than the SF-36 and AIMS2

pain subscales was not surprising. Future research comparing J-MAP's responsiveness to that of other joint-specific measures is needed.

The J-MAP is a symptom-specific instrument designed for measuring target joint pain. J-MAP scores are limited because they do not inform about how joint pain affects function or quality of life. Therefore, we concur with the recommendations of other researchers that the J-MAP be used in conjunction with some of the more generic functional measures to understand how the joint pain affects patients' functional ability and quality of life[57,62].

We have introduced the Joint-Specific Multidimensional Assessment of Pain, a new measure for documenting and monitoring knee-specific pain. Its clinical and psychometric properties make it a versatile measure for research and clinical applications that offers advantages over existing measures. Although further testing in other populations and joints is needed, evidence from our study suggests that the J-MAP is a multidimensional, joint-specific pain assessment instrument that is reliable, valid, and responsive for assessing and monitoring knee joint pain.

REFERENCES

1. McKinney R, Andersen RE. Exercise benefits patients with osteoarthritis. Physician Sports Med 2000;28:71-2.
2. Creamer P, Hochberg MC. Osteoarthritis. Lancet 1997;350:503-9.
3. Symmons DP. Knee pain in older adults: the latest musculoskeletal "epidemic". Ann Rheum Dis 2001;60:91-7.
4. National Institute of Arthritis and Musculoskeletal and Skin Diseases (2001). National Institutes of Health News Release. Available from: http://www.nih.gov/news/pr/jan2001/niams-05.htm [Accessed September 24, 2002].
5. Lawrence RC, Helmick CG, Arnett FC, et al. Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States. Arthritis Rheum 1998;41:778-99.
6. National Arthritis Data Workgroup. Arthritis prevalence and activity limitation — United States, 1990. MMWR Morb Mortal Wkly Rep 1994;43:433-8.
7. Felson DT, Lawrence RC, Dieppe PA, et al. Osteoarthritis: new insights. I. The disease and its risk factors. Ann Intern Med 2000;133:635-46.
8. Felson DT, Lawrence RC, Hochberg MC, et al. Osteoarthritis: new insights. II. Treatment approaches. Ann Intern Med 2000; 133:726-37.
9. Centers for Disease Control. Prevalence and impact of arthritis by race and ethnicity — United States, 1989-1991. MMWR Morb Mortal Wkly Rep 1996;45:373-8.
10. 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;15:1833-40.
11. Melzack R. McGill Pain Questionnaire. In: Turk DC, Melzack R, editors. Handbook of pain assessment. New York: Guilford; 1975:154-61;165-6.
12. Meenan RF, Mason JH, Anderson JJ, Guccione AA, Kazis LE. AIMS2: The content and properties of a revised and expanded Arthritis Impact Measurement Scales health status questionnaire. Arthritis Rheum 1992;35:1-10.
13. Ware J, Sherbourne C. The MOS 36-item Short Form Health Survey (SF-36). Med Care 1992;30:473-83.
14. Marshall JL, Fetto JF, Botero PM. Knee ligament injuries: a standardized evaluation method. Clin Orthop 1977;123:115-29.
15. Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the Knee Society Clinical Rating System. Clin Orthop 1989;248:13-14.
16. Ranawat CS, Insall J, Shine J. Duo-condylar knee arthroplasty: Hospital for Special Surgery design. Clin Orthop 1976;120:76-82.
17. McGrory BJ, Harris WH. Can the Western Ontario and McMaster Universities (WOMAC) osteoarthritis index be used to evaluate different hip joints in the same patient? J Arthroplasty 1996; 11:841-4.
18. Wolfe F. Determinants of WOMAC function, pain, and stiffness scores: evidence for the role of low back pain, symptom counts, fatigue and depression in osteoarthritis, rheumatoid arthritis and fibromyalgia. Rheumatology (Oxford) 1999;38:355-61.
19. Price DD, McGrath PA, Rafii A, Buckingham B. The validation of visual analog scales as ratio scale measures of chronic and experimental pain. Pain 1983;17:45-56.
20. Price DD, Rafii A, Watkins LR, Buckingham B. A psychophysical analysis of acupuncture analgesia. Pain 1984;19:27-42.
21. Jensen MP, Karoly P. Self-report scales and procedures for assessing pain in adults. In: Turk DC, Melzack R, editors. Handbook of pain assessment. New York: Guilford Press; 1992:135-50.
22. Melzack R, Wall PD. The challenge of pain. New York: Basic Books; 1983.
23. Price DD, Harkins SW, Baker C. Sensory-affective relationships among different types of clinical and experimental pain. Pain 1987;28:291-9.
24. Gracely RH. Pain psychophysics. In: Chapman CR, Loeser JD, editors. Advances in pain research and therapy. New York: Raven Press; 1989:211-29.
25. Turk DC, Meichenbaum D, Genest M. Pain and behavioral medicine: a cognitive-behavioral perspective. New York: Guilford Press; 1983.
26. 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.
27. Nunnally J, Bernstein IH. Psychometric theory. 3rd ed. New York: McGraw-Hill; 1994.
28. Hays RD, Morales LS, Reise SP. Item response theory and health outcomes measurement in the 21st century. Med Care 2000;38 Suppl 9:II28-42.
29. Mungus D, Reed BR. Application of item response theory for development of a global functioning measure of dementia with linear measurement properties. Stat Med 2000;19:1631-44.
30. Altman R. Criteria for the classification of clinical osteoarthritis. J Rheumatol 1991;18:10-8.
31. Miller RD. Anesthesia. 3rd ed. New York: Churchill Livingston; 1990.
32. Moseley B. Arthroscopic treatment of osteoarthritis of the knee: one-year evaluation of the placebo effect. Orthop Trans 1995;19:382.
33. Moseley JB, O'Malley K, Petersen N, et al. A randomized, placebo-controlled trial of arthroscopic surgery for osteoarthritis of the knee. N Engl J Med 2002;347:81-8.
34. Melzack R, Casey KL. Sensory, motivational, and central control determinants of pain. In: Kenshalo D, editor. The skin senses. Springfield: Thomas; 1968:423-39.
35. Daut R, Cleeland C, Flanery R. Development of the Wisconsin Brief Pain Questionnaire to assess pain in cancer and other disease. Pain 1983;17:197-210.
36. Turk DC, Melzack R. The measurement of pain and the assessment of people experiencing pain. In: Turk DC, Melzack R, editors. Handbook of pain assessment. New York: Guilford Press; 1992:3-12.

37. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. Med Care 1990;28:632-42.
38. O'Malley K, Greisinger AJ, Petersen NJ, Menke T, Moseley JB, Wray NP. The Knee Society Clinical Rating Scale: A validation study. Clin Orthop 2002:(in press).
39. Stucki G, Liang MH, Fossel AH, Katz JN. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. J Clin Epidemiol 1995;48:1369-78.
40. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983;24:385-96.
41. Joreskog KG, Sorbom D. LISREL 8: User's reference guide. Chicago: Scientific Software, International; 1996.
42. Bollen KA. Structural equations with latent variables. New York: John Wiley & Sons; 1989.
43. Francis DJ. An introduction to structural equation models. J Clin Exp Neuropsychol 1988;10:623-39.
44. Bollen KA, Long JS. Introduction. In: Bollen KA, Long JS, editors. Testing structural equation models. Newbury Park, CA: Sage Publications; 1993:1-9.
45. Browne MW, Cudeck R. Alternative ways of assessing model fit. In: Bollen KA, Long JS, editors. Testing structural equation models. Newbury Park, CA: Sage Publications; 1993:136-62.
46. Kline RB. Principles and practice of structural equation modeling. New York: The Guilford Press; 1988.
47. Bentler PM. Comparative fit indexes in structural models. Psychol Bull 1990;107:256-9.
48. Andrich DA. A rating formulation for ordered response categories. Psychometrika 1978;43:561-73.
49. Linacre JM, Wright BD. A user's guide to BIGSTEPS: Rasch-model computer program, version 2.7. Chicago: Mesa Press; 1997.
50. Cook KF, Ashton CM, Byrne MM, et al. A psychometric analysis of the measurement level of the rating scale, time trade-off, and standard gamble. Soc Sci Med 2001;53:1275-85.
51. Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychol Bull 1959;56:81-105.
52. Crocker L, Algina J. Introduction to classical and modern test theory. Fort Worth, TX: Harcourt Brace Jovanovich College Publishers; 1986.
53. Creamer P, Lethbridge-Cejku M, Hochberg MC. Determinants of pain severity in knee osteoarthritis: effect of demographic and psychosocial variables using 3 pain measures. J Rheumatol 1999;26:1785-92.
54. Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale, NJ: Erlbaum; 1977.
55. Medical Outcomes Trust. Scientific Advisory Committee instrument review criteria Internet publication 1995 Bulletin Vol 3 No 4. [Accessed October 17, 2002]. Available from: http://195.101.204.50:443/public/34sacrev.htm
56. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: A critical review and recommendations. J Clin Epidemiol 2000;53:459-68.
57. Kosinski M, Zhao SZ, Dedhiya S, Osterhaus JT, Ware JE. Determining minimally important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. Arthritis Rheum 2000;43:1478-87.
58. Jaeschke R, Singer J, Guyatt GH. Measurement of health status: Ascertaining the minimal clinically important differences. Control Clin Trials 1989;10:407-15.
59. Redelmeier DA, Guyatt GH, Goldstein RS. Assessing the minimal important difference in symptoms: A comparison of two techniques. J Clin Epidemiol 1996;49:1215-9.
60. Guyatt GH, Kirshner B, Jaeschke R. Measuring health status: What are the necessary measurement properties? J Clin Epidemiol 1992;45:1341-5.
61. 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 1999;37:469-78.
62. 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 1999;52:861-73.
63. Ware JE, Snow KK, Kosinsky M, Gandek B. SF-36 Health survey manual and interpretation guide. Boston, MA: The Health Institute, New England Medical Center; 1993.
64. Roos EM, Klassbo M, Lohmander LS. WOMAC Osteoarthritis Index: Reliability, validity, and responsiveness in patients with arthroscopically assessed osteoarthritis. Scand J Rheumatol 1999;28:210-5.
65. Brazier JE, Harper R, Munro J, Walter SJ, Snaith ML. Generic and condition-specific outcome measures for people with osteoarthritis of the knee. Rheumatology (Oxford) 1999;38:870-7.
66. Chatman AB, Hyams SP, Neel JM, et al. The Patient-Specific Functional Scale: Measurement properties in patients with knee dysfunction. Physical Ther 1997;77:820-9.
67. Ehrich EW, Davies GM, Watson DJ, Bolognese JA, Seidenberg BC, Bellamy N. Minimal perceptible clinical improvement with the Western Ontario and McMaster University osteoarthritis index questionnaire and global assessments in patients with osteoarthritis. J Rheumatol 2000;27:2635-41.