Osteoarthritis and Cartilage

ICRS

International Cartilage Repair Society

OARSI OSTEOARTHRITIS RESEARCH SOCIETY INTERNATIONAL

Acceptability, reliability, validity and responsiveness of the Turkish version of WOMAC osteoarthritis index

E. H. Tüzün P.T., Ph.D., Assist. Prof.†*, L. Eker M.D.‡, A. Aytar B.Sc.§, A. Daşkapan P.T., Ph.D., Assist. Prof.† and M. Bayramoğlu M.D., Assist. Prof.§
† *Department of Physical Therapy and Rehabilitation, Faculty of Health Sciences,*
Başkent University, Ankara, Turkey
‡ *Mother and Child Health Care and Family Planning General Directorate, Ministry of Health, Ankara, Turkey*
§ *Department of Physical Medicine and Rehabilitation, Başkent University Hospital, Ankara, Turkey*

Summary

Objective: To evaluate the acceptability, reliability, validity and responsiveness of the Turkish version of Western Ontario and McMaster Universities (WOMAC) osteoarthritis (OA) index in physiotherapy outpatient practice in Turkey.

Method: Data were obtained from 72 patients with OA of the knee. They were asked to answer two disease-specific questionnaires (WOMAC LK 3.1 and Lequesne-Algofunctional Index of Severity for the knee) and one generic instrument (Medical outcomes study SF-36 Survey-SF-36). Acceptability was assessed in terms of refusal rate, rates of missing responses, and administration time. Reliability was assessed using Cronbach's alpha. Content validity was assessed by examining the floor and ceiling effects, and skew of the distributions. Convergent and divergent validity was assessed by examining the Pearson's correlation coefficients. Responsiveness was determined by examining effect size (ES), standardized response means (SRM) and *P* values generated using Wilcoxon's test.

Results: The overall response rate was 100%. Alpha values for all WOMAC subscales exceeded the value of 0.70 at both baseline and follow-up assessments. Frequency distributions of scores were symmetrical. Subscales had negligible floor and ceiling effects. Both pain and physical function subscales were fairly correlated with the subscales measuring similar constructs of SF-36, whereas they were weakly correlated with other dimensions of SF-36. A good correlation was obtained between WOMAC total and Lequesne index. The pain and physical function subscales of WOMAC index were the most responsive subscales.

Conclusion: The Turkish WOMAC OA index is acceptable, valid, reliable and responsive for use in Turkish patients with knee OA.
© 2004 OsteoArthritis Research Society International. Published by Elsevier Ltd. All rights reserved.

Key words: Osteoarthritis, Reliability, Responsiveness, Turkish WOMAC, Validity.

Introduction

Evaluation has always been an important component of any health intervention program. The impact of chronic disease and therapeutic interventions must be evaluated from the patient's perspective in addition to more traditional measures of medical outcomes[1]. Patient-based outcome measures, addressing constructs such as health-related quality of life, subjective health status and functional status, are increasingly used as primary or secondary end-points in clinical trials. These instruments, if proved and valid, offer an important additional information source to clinicians in the care of their patients[2,3].

Various clinical instruments can be used to assess the patients with knee osteoarthritis (OA). The Western Ontario and McMaster Universities (WOMAC) OA index is the most widely used disease-specific instrument for assessment of patients with OA[4,5]. It is recommended by Outcome Measures in Rheumatology Clinical Trials (OMERACT) to measure the dimension of function in OA trials[6]. It has been argued that cross-cultural validation of patient-based outcome measures are needed to compare and contrast results, aggregate data from different countries, strengthen causal inference on effects of treatments, and to investigate the natural history of health conditions.

Although there was a linguistically validated Turkish version of WOMAC index, there is no available literature that suggests this version of WOMAC index is culturally equivalent to the original questionnaire, and has a similar level of validity and reliability in patients with knee OA. The aim of the present study was to evaluate the acceptability, reliability, validity and responsiveness of the Turkish version of WOMAC index in physiotherapy outpatient practice in Turkey.

Method

SUBJECTS

Seventy-two consecutive knee OA patients on the waiting list at the Department of Physical Medicine and Rehabilitation at Başkent University Hospital in Ankara, Turkey, were asked to be study subjects. Patients were recruited

*Address correspondence and reprint requests to: Emine Handan Tüzün, P.T., Ph.D., Assist. Prof., Başkent University, Faculty of Health Sciences, Department of Physical Therapy and Rehabilitation, Eskisehir Yolu 20.Km, Baglica, Ankara, Turkey. Tel: 90-312-2341010/1633; Fax: 90-312-2341054.; E-mail: htuzun@baskent.edu.tr

from February to June 2003. Inclusion criteria for the study were age over 40 years and primary knee OA. All patients fulfilled clinical and radiological criteria of the American College of Rheumatology for primary knee OA[7]. Disease severity was graded on the basis of the Kellegren and Lawrence radiographic system[8]. Exclusion criteria included history or active presence of other rheumatic diseases potentially responsible for a secondary OA, severe articular inflammation as confirmed by physical examination and laboratory parameters (ESR > 40 mm/h and serum rheumatoid factor titer > 1:40), traumatic knee lesions, intra-articular or systemic use of corticosteroids within the previous 3 months, and use of analgesics and non-steroidal anti-inflammatory drugs during the previous 2 months. Patients who had advanced cardiac diseases or peripheral vascular diseases were also excluded from the study.

Subjects were randomly allocated to two groups. The first group ($n = 36$) received isokinetic quadriceps exercise program. The second group ($n = 36$) received isotonic quadriceps exercise program. Patients in both intervention groups also received transcutaneous electrical nerve stimulation and hot packs application. Treatment was administered five times a week for 2 weeks in both groups. In addition to demographic data, patients were asked to report co-morbid conditions such as chronic respiratory disease, hypertension, gastrointestinal disease and endocrine disease. The full details of this randomized trial of isotonic vs isokinetic quadriceps strengthening were presented elsewhere[9].

The ethics committee of Başkent University approved the study. Each patient was informed about the study and they gave their written informed consent to participate.

INSTRUMENTS

Patients were assessed for functional ability and general health status, using two disease-specific questionnaires (WOMAC index and Lequesne-Algofunctional Index of Severity for the knee) and one general health questionnaire (Medical Outcomes Study Short Form (SF-36). Disease-specific questionnaires were administered at baseline (Time 1) and on the third day after treatment period (Time 2), whereas the SF-36 was administered at baseline. The WOMAC index and the SF-36 were filled out by patients in the same room in the clinic. The third investigator (AA) filled out the Lequesne index using interview technique. Time taken to fill out WOMAC index and SF-36 was recorded.

The WOMAC index is a disease-specific, self-administered questionnaire developed to study patients with hip or knee OA. It consists of 24 questions, grouped into 3 subscales (pain, stiffness, and physical function)[10]. In this study, after obtaining permission from the copyright owner (Dr Nicholas Bellamy), we used linguistically validated Turkish version of WOMAC LK 3.1 scale. In Likert (LK) scale, there are five alternative answers to every question (0 = none, 1 = mild, 2 = moderate, 3 = severe, 4 = extreme). The maximum score in LK scale is 20 points for pain, 8 points for stiffness, and 68 points for physical function. Higher scores indicate more or worse symptoms, maximal limitations, and poor health. In this study, we used a normalization procedure to correct for differences in scale length. In order to normalize the LK scale on a scale of 0−10, the following correction factors were used where S = sum of raw scores of items in dimension: Pain normalization = $S \times 0.50$; Stiffness normalization = $S \times 1.25$;

and Physical function normalization = $S \times 0.147$[11]. After subscale values were normalized, they were summated to provide a single value in which the three component subscales were equally weighted (WOMAC total score)[11].

Lequesne-Algofunctional Index of Severity for the knee is a disease-specific questionnaire. It consists of 10 questions grouped into three subscales (pain or discomfort, activities of daily living and maximum distance walked). Although the metric properties of the Lequesne index have been established[12,13], separate subscales have not been validated for independent applications[14]. Total score ranges from 0 to 24, and higher scores indicate more severe handicap.

In this study, after we obtaining permission from the copyright owner (QualityMetric Incorporated), Turkish version of the SF-36 was used to assess the general health status of subjects. The SF-36 includes eight multi-item scales containing two to 10 items each plus a single item to compare the current health with a person's health one year ago (health transition)[15]. The scales cover the dimensions of physical functioning (PF), role physical (RP), bodily pain (P), general health (GH), vitality (V), social functioning (SF), role emotional (RE), and mental health (MH). All items pertaining to each scale (excluding health transition) are summed and transformed to form a scale from 0 to 100, where a higher score indicates a better state of health or well-being.

ANALYSIS

In this study, we used the Statistical Package for the Social Sciences (SPSS for Windows 9.0) for statistical analyses. We checked the missing values for each questionnaire prior to further analysis. To evaluate the association between the WOMAC index subscales and patient characteristics, such as age, marital status, level of education, body-mass index (BMI), duration of OA and radiographic severity, correlation coefficients were calculated. Acceptability was assessed in terms of refusal rate, rates of missing responses and administration time.

Internal consistency reliability was assessed by Cronbach's alpha coefficient at Time 1 and Time 2. An alpha value of 0.70 or higher was considered as acceptable reliability for group comparisons[16,17].

Content validity was assessed at Time 1 by examining the floor and ceiling effects, and skew of the distribution in each subscale. We hypothesized that the skewness statistics range from -1 to $+1$[18], and floor and ceiling effects are less than 20%[19].

Convergent and divergent construct validity was assessed at Time 1 by examining the Pearson's correlation coefficients of the WOMAC index subscales compared to the subscales of SF-36. We also examined the correlation coefficient between WOMAC total and Lequesne index. Based on the literature, we hypothesized that WOMAC pain and physical function (and lesser extent stiffness) scores are highly correlated with similar subscale (physical function and bodily pain) scores derived from SF-36[5]. We further hypothesized that WOMAC total is highly correlated with Lequesne index (convergent validity)[20]. Additionally, we hypothesized that WOMAC index subscales are inadequately or weakly correlated with theoretically unrelated subscale (mental component subscales) scores derived from SF-36 (divergent validity). Pearson's correlation coefficients were interpreted as either excellent relationship $|r| \geq 0.91$; good $0.90 \geq |r| \geq 0.71$; fair $0.70 \geq |r| \geq 0.51$; weak $0.50 \geq |r| \geq 0.31$; little or none $|r| \leq 0.3$. A P value of 0.01 was taken as the level of significance[21].

Table I
Baseline characteristics of study population

Characteristics	Isotonic exercise group ($N = 36$)	Isokinetic exercise group ($N = 36$)	Total ($N = 72$)
Sex (female), *n* (%)	31 (86.1)	31 (86.1)	62 (86.1)
Age, mean (SD), years	61.9 (8.9)	60.2 (9.2)	61.0 (9.0)
Marital status, *n* (%)			
Married	26 (72.2)	22 (61.1)	48 (66.7)
Single	10 (27.8)	14 (38.9)	24 (33.3)
BMI, mean (SD), kg/m^2	29.5 (5.3)	30.2 (4.6)	29.8 (4.9)
Education level, *n* (%)			
Primary level	11 (30.6)	5 (13.9)	16 (22.2)
Secondary level	4 (11.1)	4 (11.1)	8 (11.1)
High school level	11 (30.6)	16 (44.4)	27 (37.5)
University	10 (27.8)	11 (30.6)	21 (29.2)
Duration of OA, mean (SD), years	6.1 (6.1)	4.7 (4.5)	5.4 (5.3)
Radiological grade, *n* (%)			
Grade 1	9 (25.0)	13 (36.1)	22 (30.6)
Grade 2	18 (50.0)	12 (33.3)	30 (41.7)
Grade 3	8 (22.2)	8 (22.2)	16 (22.2)
Grade 4	1 (2.8)	3 (8.3)	4 (5.6)
Number of co-morbid conditions, *n* (%)			
0	10 (27.8)	13 (36.1)	23 (31.9)
1	21 (58.3)	10 (27.8)	31 (43.1)
2 or more	5 (13.9)	13 (36.1)	18 (25.0)

Responsiveness was determined by examining the *P* values generated using Wilcoxon's test comparing Time 1 scores with Time 2 scores. We also examined the effect size (ES) and standardized response means (SRM). The ES is equal to the mean change in score divided by the standard deviation of individuals' baseline (Time 1) score. The SRM is equal to the mean change in score divided by the standard deviation of individuals' changes in score. A value of 0.2 to 0.5 was regarded as "small", 0.5 to 0.8 as "medium", and those above 0.8 as "large"[22].

Results

As shown in Table I, the majority of patients were women (86%). The mean age of 72 patients was 61.0 ± 9 years (range $41-80$). The mean duration of OA was 5.4 ± 5.3 years (range 6 months-20 years). Thirty-two patients (44.4%) were obese (BMI > 30.0 kg/m^2), another 27 (37.5%) were overweight (BMI $= 25.0-29.9$ kg/m^2). Forty-nine patients (68.1%) reported one or more co-morbid conditions. Distribution of radiographic severity showed that majority of patients had a radiological severity of first and second degree on Kellegren's scale (30.6% and 41.7%, respectively).

Table II
Correlation coefficients between the WOMAC scores and independent variables

Independent variables	WOMAC subscales			
	Pain	Stiffness	Physical function	Total
Pearson				
Age	−0.03	−0.05	0.07	0.04
BMI	0.08	0.16	0.17	0.17
Duration of disease	−0.14	0.11	0.03	0.01
Spearman				
Marital status	0.17	0.08	0.09	0.11
Education level	−0.14	0.03	−0.19	−0.17
Radiological grade	−0.09	−0.10	−0.07	−0.06
All *P* values > 0.05.				

As shown in Table II, there was no significant correlation between the WOMAC subscale scores and the patient's characteristics (age, marital status, education level, BMI, duration of disease and radiological grade).

DATA QUALITY AND ACCEPTABILITY

All eligible subjects accepted to take part in the study. Missing data for items were low (3.2%). Ninety-seven percent of patients completed all 24 items (100% complete data), 1.6% of patients missed out two physical function items and 1.6% of patients missed out one physical function item. We substituted the average value for the subscale in lieu of missing item values as recommended by Dr Bellamy[11]. Therefore, WOMAC index scores were computed for 72 patients (100%). Average time to complete the WOMAC index was 3.5 min, while it was 7.8 min for SF-36.

RELIABILITY

Cronbach's alpha coefficients for the three subscales at Time 1 and Time 2 were as follows: Pain: 0.75, 0.81; Stiffness: 0.71, 0.76; Physical function: 0.94, 0.96.

CONTENT VALIDITY

As shown in Table III, frequency distribution of scores on each subscale of WOMAC index was quite symmetrical and

Table III
Content validity of WOMAC

WOMAC subscales	Content validity				
	N	Mean (SD)	Floor *n* (%)	Ceiling *n* (%)	Skewness
Pain	72	4.1 (1.9)	1 (1.4)	0	−0.13
Stiffness	72	3.7 (2.5)	14 (19.4)	0	−0.18
Physical function	72	4.6 (2.0)	1 (1.4)	0	−0.03
Total	72	12.4 (5.5)	0	0	0.01

Table IV
Convergent and divergent validity (Pearson's correlation coefficient)

	WOMAC subscales			
	Pain	Stiffness	Physical function	Total
SF-36 subscales				
Physical functioning	−0.55	−0.38	−0.72	−0.63
Role physical	−0.36	−0.47	−0.48	−0.52
Bodily pain	−0.51	−0.40	−0.54	−0.55
General health	−0.34	−0.34	−0.49	−0.45
Vitality	−0.43	−0.35	−0.43	−0.47
Role emotional	−0.28	−0.34	−0.40	−0.40
Social functioning	−0.34	−0.36	−0.49	−0.46
Mental health	−0.40	−0.31	−0.37	−0.42
Lequesne index	0.64	0.50	0.68	0.70

All *P* values < 0.01 (two tailed).

not notably skewed (range −0.13 to +0.01). The WOMAC subscales had negligible floor and ceiling effects.

CONVERGENT AND DIVERGENT VALIDITY

Table IV shows Pearson's correlation coefficients between WOMAC subscales and SF-36 subscales, and Lequesne index. Both pain and physical function subscales of WOMAC index showed statistically significant correlation with the physical functioning and bodily pain subscales of SF-36. Although correlation coefficients are significant at 0.01 levels, these two subscales had weak correlations with the other dimensions of SF-36. There was a fairly good correlation, at 0.01 significant level, between the scores on WOMAC total and Lequesne index.

RESPONSIVENESS

As shown in Table V, *P* values generated using Wilcoxon's test for isokinetic quadriceps exercise group showed that there were significant improvements at the end of treatment period in all WOMAC subscales (all *P* values < 0.01). Likewise, there were significant improvements for all subscales, except stiffness, in Time 2 compared to Time 1 for isotonic quadriceps exercise group.

In terms of SRM, the patients in group 1 showed large effect in the pain, physical function, and total scores of WOMAC index, whereas the patients in group 2 showed medium effect. In terms of stiffness, the size of the effect was medium in group 1, whereas it was small in group 2. In both groups, the values of ES and SRM for Lequesne index were greater than those calculated for WOMAC total.

Discussion

The use of measurement tools in physical therapy has increased dramatically since early 1900s. For clinicians, valid measurement tools provide important information to support effective clinical interpretation. These instruments helped to define the nature and scope of clinical problems, provide standards against which to monitor progress, and serve to summarize clinically important changes that occur as a result of the therapy process. For any test or measure of health status to be useful it must be reliable, valid, responsive and acceptable. The WOMAC index has been designed specifically to evaluate patients with OA of the knee or hip. The English-language version of the original index has been shown to be a reliable, valid, responsive and acceptable outcome measure[10,23]. It has been translated into over 50 different languages including Turkish[11]. However, to enable comparison between evaluations made in different countries, this questionnaire needs not only to be translated and linguistically validated, but also the demonstration of its psychometric properties (reliability, validity and sensitivity to change). Here, we present data on metric properties of the Turkish version of WOMAC index.

In this study, we found that there was no significant correlation between WOMAC scores and age of patient. Our result confirms earlier observations that there was no association between age and severity of pain and physical function in OA[24−26]. We found that BMI was not associated with WOMAC scores. This result is consistent with the observation made by Dr Salaffi and co-workers[20]. While many investigators suggested that formal education level appear to be an important factor in self-reported pain severity in knee OA[20,25,26], we were unable to show a significant correlation between the educational level and WOMAC scores. Although this finding shows that patient's

Table V
Changes in the WOMAC and Lequesne index scores for subjects in isokinetic quadriceps exercise program (group 1) and isotonic quadriceps exercise program (group 2)

	Mean (SD)			*P* value†	ES	SRM
	Time 1	Time 2	Difference			
Group 1						
WOMAC subscales						
Pain	4.5 (1.8)	2.9 (1.9)	1.6 (1.6)	0.001	0.88	1.00
Stiffness	3.9 (2.5)	2.8 (2.2)	1.1 (2.1)	0.009	0.44	0.52
Physical function	4.7 (1.9)	3.2 (2.1)	1.5 (1.6)	0.001	0.79	0.94
Total	13.1 (5.4)	8.9 (5.9)	4.2 (4.1)	0.001	0.78	1.02
Lequesne index	10.4 (4.1)	7.0 (4.8)	3.4 (2.9)	0.001	0.83	1.17
Group 2						
WOMAC subscales						
Pain	3.8 (1.9)	2.4 (1.8)	1.4 (1.8)	0.001	0.74	0.78
Stiffness	3.6 (2.5)	2.8 (2.2)	0.8 (2.8)	0.107	0.32	0.29
Physical function	4.4 (2.2)	3.3 (2.1)	1.1 (1.6)	0.001	0.50	0.69
Total	11.8 (5.5)	8.5 (5.3)	3.3 (4.7)	0.001	0.60	0.70
Lequesne index	10.1 (3.9)	6.8 (3.2)	3.3 (3.6)	0.001	0.85	0.92

ES = effect size, SRM = standardized response mean.
†Wilcoxon's test.

educational level, a potential bias source, does not affect the response to items in the questionnaire, further studies are needed to confirm our finding. In this study, we did not find a significant correlation between the duration of OA and WOMAC scores. This finding is consistent with the finding of Dr Creamer and co-workers[26]. The findings of the present study showed that functional disability and pain severity were not correlated with radiographic damage. This finding confirms earlier observations on the association between the radiographic severity and functional disability and pain[27–29].

In this study, we found that the non-response rate and the rates of missing responses were quite low. The time taken to fill out WOMAC index was lesser than in SF-36. These results showed that the Turkish version of WOMAC index is an acceptable instrument to patients.

In this study, we found that Cronbach's alpha values for all subscales of WOMAC index exceed the suggested cutoff value of 0.70, revealing an acceptable level of reliability for group comparisons[30]. The alpha value of WOMAC stiffness subscale was lower than all other subscale values. This was an expected result, since the alpha value is dependent on the number of items. The WOMAC stiffness subscale is derived from only two questions, while the pain and physical function subscales contain 5 and 17 items, respectively. Similar findings have been reported for the Italian version[20] and the Spanish version[31].

The results of present study showed that the Turkish version of WOMAC index has a minimal floor and ceiling effects, as well as good frequency distribution of scores on each subscale. This finding is consistent with such previous studies[32,33] and indicates that the Turkish version of WOMAC index has a good content validity. Likewise, correlation coefficients between the Turkish version of WOMAC index and both SF-36 and Lequesne index demonstrated good convergent and divergent validity. There were linear relationships between the pain subscale of WOMAC index and the physical functioning subscale of SF-36, and between the physical function subscale of WOMAC index and the bodily pain of SF-36. This finding is not surprising, since physical function in OA is closely related with the pain[34].

The self-report questionnaires are widely used for the assessment of treatment outcomes in evaluation studies. In these studies assessments are performed at various time points and outcomes are presented as differences between the baseline and follow-up measurements. There is growing recognition that assessing an intervention's effect should not only focus at the statistical significance of the differences in health outcome but also at the baseline and follow-up measurements. If a P value is interpreted as statistically significant, rejecting the null hypothesis does not imply an effect of important magnitude; likewise, a non-significant P value does not indicate a trivial result[35,36]. The results of present study showed that after 2 weeks treatment period, both treatment protocols are effective in reducing pain and in increasing physical function. The pain and physical function subscales of WOMAC index were the most responsive subscales, with the highest ES and SRM resulting in comparably smaller sample sizes required for future studies. ES values for stiffness subscale were lower than in the other subscales. This finding is consistent with such previous studies[37,38]. One possible explanation for this result may be that the reliability can spuriously affect the ES. As the internal consistency decreases, within-group variability increases, which means that between-group ES decreases[39]. In a study conducted among ambulatory patients undergoing hip or knee arthroplasty, it had been found that WOMAC global index was more responsive than Lequesne index[40]. On the contrary, our results showed that Lequesne-Algofunctional Index of Severity for the knee was more responsive than WOMAC total. The differences between the results of two studies may be the result of differences between treatment methods as well as the differences in characteristics of study populations.

In conclusion, the Turkish WOMAC OA index is acceptable, valid, reliable and responsive for use in Turkish patients with OA of the knee in physiotherapy outpatient practice.

References

1. Slevin ML, Plant H, Lynch D, Drinkwater J, Gregory WM. Who should measure quality of life, the doctor or the patient? Br J Cancer 1988;57:109–12.
2. Tarlov AR, Ware JE Jr, Greenfield S, Nelson EC, Perrin E, Zubkoff M. The Medical Outcomes Study. An application of methods for monitoring the results of medical care. JAMA 1989;262:925–30.
3. Anonymous. Recognising disability. Lancet 1991;338: 154–5.
4. Anderson JG, Wixson RL, Tsai D, Stulberg SD, Chang RW. Functional outcome and patient satisfaction in total knee patients over the age of 75. J Arthroplasty 1996;11:831–40.
5. Brazier JE, Harper R, Munro J, Walters SJ, Snaith ML. Generic and condition-specific outcome measures for people with osteoarthritis of the knee. Rheumatology (Oxford) 1999;38:870–7.
6. Bellamy N, Kirwan J, Boers M, Brooks P, Strand V, Tugwell P, *et al.* Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand osteoarthritis. Consensus development at OMERACT III. J Rheumatol 1997;24: 799–802.
7. Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, *et al.* Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. Arthritis Rheum 1986;29: 1039–49.
8. Altman R, Brandt K, Hochberg M, Moskowitz R, Bellamy N, Bloch DA, *et al.* Design and conduct of clinical trials in patients with osteoarthritis: recommendations from a task force of the Osteoarthritis Research Society. Results from a workshop. Osteoarthritis Cartilage 1996;4:217–43.
9. Tuzun EH, Aytar A, Eker L, Daskapan A. Effectiveness of two different physical therapy programmes in the treatment of knee osteoarthritis. Pain Clin 2004;16: 379–87.
10. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15:1833–40.
11. Bellamy N. WOMAC. Osteoarthritis Index. User Guide V 2002.
12. Lequesne MG, Mery C, Samson M, Gerard P. Indexes of severity for osteoarthritis of the hip and knee. Validation—value in comparison with other assessment tests. Scand J Rheumatol 1987;65(Suppl):85–9.

13. Lequesne MG, Samson M. Indices of severity in osteoarthritis for weight bearing joints. J Rheumatol 1991;18(Suppl 27):16−8.

14. Lequesne MG, Mery C, Samson M, Marty M. Comparison between the WOMAC and the Lequesne indices in patients with knee and hip osteoarthritis. Osteoarthritis Cartilage 1998;6:441−2.

15. Ware JE Jr, Gandek B. Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project. J Clin Epidemiol 1998;51:903−12.

16. Nunnally JC, Bernstein IH. Psychometric Theory. 3rd edn. New York: McGraw-Hill 1994.

17. DeVellis RF. Scale Development: Theory and Applications. Newbury Park: Sage Publications, 1991.

18. Holmes W, Bix B, Shea J. SF-20 score and item distributions in a human immunodeficiency virus-seropositive sample. Med Care 1996;34:562−9.

19. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? Qual Life Res 1995;4:293−307.

20. Salaffi F, Leardini G, Canesi B, Mannoni A, Fioravanti A, Caporali R, et al. Reliability and validity of the Western Ontario and McMaster Universities (WOMAC) Osteoarthritis Index in Italian patients with osteoarthritis of the knee. Osteoarthritis Cartilage 2003;11:551−60.

21. Martin DP, Engelberg R, Agel J, Swiontkowski MF. Comparison of the Musculoskeletal Function Assessment questionnaire with the Short Form-36, the Western Ontario and McMaster Universities Osteoarthritis Index, and the Sickness Impact Profile health-status measures. J Bone Joint Surg Am 1997;79: 1323−35.

22. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd edn. New York: Academic Press.

23. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J. Validation study of WOMAC: a health status instrument for measuring clinically-important patient-relevant outcomes following total hip or knee arthroplasty in osteoarthritis. J Orthop Rheumatol 1988;1:95−108.

24. Summers MN, Haley WE, Reveille JD, Alarcon GS. Radiographic assessment and psychologic variables as predictors of pain and functional impairment in osteoarthritis of the knee or hip. Arthritis Rheum 1988; 31:204−9.

25. 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.

26. Creamer P, Lethbridge-Cejku M, Hochberg MC. Factors associated with functional impairment in symptomatic knee osteoarthritis. Rheumatology 2000;39: 490−6.

27. Salaffi F, Cavalieri F, Nolli M, Ferraccioli G. Analysis of disability in knee osteoarthritis. Relationship with age and psychological variables but not with radiographic score. J Rheumatol 1991;18:1581−6.

28. McAlindon TE, Cooper C, Kirwan JR, Dieppe PA. Determinants of disability in osteoarthritis of the knee. Ann Rheum Dis 1993;52:258−62.

29. Odding E, Valkenburg HA, Algra D, Vandenouweland FA, Grobbee DE, Hofman A. Associations of radiological osteoarthritis of the hip and knee with locomotor disability in the Rotterdam Study. Ann Rheum Dis 1998;57:203−8.

30. Bland JM, Altman DG. Statistics notes: Cronbach's alpha. BMJ 1997;314:572.

31. Escobar A, Quintana JM, Bilbao A, Azkarate J, Guenaga JI. Validation of the Spanish version of the WOMAC questionnaire for patients with hip or knee osteoarthritis. Western Ontario and McMaster Universities Osteoarthritis Index. Clin Rheumatol 2002;21: 466−71.

32. Thumboo J, Chew LH, Soh CH. Validation of the Western Ontario and McMaster University osteoarthritis index in Asians with osteoarthritis in Singapore. Osteoarthritis Cartilage 2001;9:440−6.

33. Soderman P, Malchau H. Validity and reliability of Swedish WOMAC osteoarthritis index: a self-administered disease-specific questionnaire (WOMAC) versus generic instruments (SF-36 and NHP). Acta Orthop Scand 2000;71:39−46.

34. Guccione AA, Felson DT, Anderson JJ. Defining arthritis and measuring functional status in elders: methodological issues in the study of disease and physical disability. Am J Public Health 1990;80:945−9.

35. Rosenthal R, Rubin DB. The counter null value of an effect size: a new statistic. Psychol Sci 1994;5: 329−34.

36. Bartko JJ, Pulver AE, Carpenter WT. The power of analysis: statistical perspectives. Part 2. Psychiatry Res 1988;23:301−9.

37. Roos EM, Toksvig-Larsen S. Knee injury and Osteoarthritis Outcome Score (KOOS)—validation and comparison to the WOMAC in total knee replacement. Health Qual Life Outcomes 2003;1:17.

38. Angst F, Aeschlimann A, Stucki G. Smallest detectable and minimal clinically important differences of rehabilitation intervention with their implications for required sample sizes using WOMAC and SF-36 quality of life measurement instruments in patients with osteoarthritis of the lower extremities. Arthritis Rheum 2001;45:384−91.

39. Leon AC, Marzuk PM, Portera L. More reliable outcome measures can reduce sample size requirements. Arch Gen Psychiatry 1995;52:867−71.

40. Theiler R, Sangha O, Schaeren S, Michel BA, Tyndall A, Dick W, et al. Superior responsiveness of the pain and function sections of the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) as compared to the Lequesne-Algofunctional Index in patients with osteoarthritis of the lower extremities. Osteoarthritis Cartilage 1999;7:515−9.