

Research

Open Access

## Knee injury and Osteoarthritis Outcome Score (KOOS) – validation and comparison to the WOMAC in total knee replacement

Ewa M Roos\*<sup>1,2</sup> and Sören Toksvig-Larsen<sup>1</sup>

Address: <sup>1</sup>Department of Orthopedics, Lund University Hospital, SE-221 85 Lund, Sweden and <sup>2</sup>Center for Research and Development, Spenshult Hospital for Rheumatic Diseases, SE-313 92 Oskarström, Sweden

Email: Ewa M Roos\* - Ewa.Roos@ort.lu.se; Sören Toksvig-Larsen - Soren.Toksvig\_Larsen@ort.lu.se

\* Corresponding author

Published: 25 May 2003

Received: 5 March 2003

*Health and Quality of Life Outcomes* 2003, 1:17

Accepted: 25 May 2003

This article is available from: <http://www.hqlo.com/content/1/1/17>

© 2003 Roos and Toksvig-Larsen; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

### Abstract

**Background:** The Knee injury and Osteoarthritis Outcome Score (KOOS) is an extension of the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), the most commonly used outcome instrument for assessment of patient-relevant treatment effects in osteoarthritis. KOOS was developed for younger and/or more active patients with knee injury and knee osteoarthritis and has in previous studies on these groups been the more responsive instrument compared to the WOMAC. Some patients eligible for total knee replacement have expectations of more demanding physical functions than required for daily living. This encouraged us to study the use of the Knee injury and Osteoarthritis Outcome Score (KOOS) to assess the outcome of total knee replacement.

**Methods:** We studied the test-retest reliability, validity and responsiveness of the Swedish version LK 1.0 of the KOOS when used to prospectively evaluate the outcome of 105 patients (mean age 71.3, 66 women) after total knee replacement. The follow-up rates at 6 and 12 months were 92% and 86%, respectively.

**Results:** The intraclass correlation coefficients were over 0.75 for all subscales indicating sufficient test-retest reliability. Bland-Altman plots confirmed this finding. Over 90% of the patients regarded improvement in the subscales Pain, Symptoms, Activities of Daily Living, and knee-related Quality of Life to be extremely or very important when deciding to have their knee operated on indicating good content validity. The correlations found in comparison to the SF-36 indicated the KOOS measured expected constructs. The most responsive subscale was knee-related Quality of Life. The effect sizes of the five KOOS subscales at 12 months ranged from 1.08 to 3.54 and for the WOMAC from 1.65 to 2.56.

**Conclusion:** The Knee injury and Osteoarthritis Outcome Score (KOOS) is a valid, reliable, and responsive outcome measure in total joint replacement. In comparison to the WOMAC, the KOOS improved validity and may be at least as responsive as the WOMAC.

### Background

Patient-relevant outcome measures are now promoted in general health care, orthopaedics and rheumatology and

should be considered the primary outcome in clinical trials. Critical properties of an outcome measure include reliability, validity and responsiveness. The Outcome

Measures in Rheumatoid Arthritis Clinical Trials (OMER-ACT) group has suggested that the most important characteristic of an instrument may be its responsiveness [1]. Responsiveness of an assessment technique is defined as the sensitivity to change over time [2]. Responsiveness provides the basis for comparing measures with differing scales and can be measured using variables such as effect size [3], standardized response mean [4], and relative efficiency [5].

For assessment of treatment effects in patients with osteoarthritis (OA), the WOMAC is recommended and the most commonly used disease-specific outcome instrument [1]. The WOMAC was developed for elderly with osteoarthritis and assesses pain, stiffness and function of daily living in three separate subscales. However, there is an increasing interest in early treatment of OA to enhance the possibly to reverse or slow the disease process down. Since OA develops over decades, naturally the patients are younger and more active early in the disease process. To meet this need the Knee injury and Osteoarthritis Outcome Score (KOOS) was developed as an extension of the WOMAC for younger and/or more active patients with knee injury and/or knee osteoarthritis [6]. KOOS has in prior studies been proven to be more sensitive and responsive than the WOMAC in younger or more active patients [6,7]. Some patients eligible for total knee replacement have expectations of more demanding physical functions than required for daily living [8]. This encouraged us to study the use of the Knee injury and Osteoarthritis Outcome Score (KOOS) to assess the outcome of total knee replacement.

The objective of the present study was to study the usefulness of the KOOS in elderly patients with advanced osteoarthritis, eligible for total joint replacement. To do so we evaluated the relevance of the different subscales, the reliability, the construct validity and the responsiveness. In addition, we compared the responsiveness of the KOOS to the responsiveness of the WOMAC.

## Methods

### Patients

To recruit patients with osteoarthritis about to have primary total knee replacement (TKR), questionnaires were sent out to 125 consecutive patients on the waiting list at the Department of Orthopedics at Lund University Hospital in Lund, Sweden. Patients were recruited from December 1999 to April 2001. Of these 125 patients, 20 were excluded, ten underwent other operative procedures, eight were not operated on during the study period and two had rheumatoid arthritis. Thus preoperative data were available for 105 patients with knee osteoarthritis.

### Questionnaires

All questionnaires were mailed to the patients and returned by mail in a pre-paid envelope. In addition to the KOOS, which includes the WOMAC, patients were also sent the SF-36 and questions regarding background data. The Swedish version LK 1.0 of the KOOS [9], including the Swedish version LK 1.0 of the WOMAC [10], and the Acute Swedish version of the SF-36 [11] were used. Literacy of the subjects was not assessed.

### KOOS

The Knee injury and Osteoarthritis Outcome Score (KOOS) is an extension of the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) [12]. KOOS was developed and is validated for several cohorts of younger and/or more active patients with knee injury and/or knee osteoarthritis [6,7,9]. KOOS is a 42-item self-administered self-explanatory questionnaire that covers five patient-relevant dimensions: Pain, Other Disease-Specific Symptoms, ADL Function, Sport and Recreation Function, and knee-related Quality of Life. The WOMAC pain questions are included in the subscale Pain, the WOMAC stiffness questions are included in the subscale Other Disease-Specific Symptoms and the WOMAC subscale Function is equivalent to the KOOS subscale ADL. The questionnaire, scoring manual and user's guide can be downloaded from <http://www.koos.nu>

### KOOS Score Calculation

The KOOS's five patient-relevant dimensions are scored separately: Pain (nine items); Symptoms (seven items); ADL Function (17 items); Sport and Recreation Function (five items); Quality of Life (four items). A Likert scale is used and all items have five possible answer options scored from 0 (No Problems) to 4 (Extreme Problems) and each of the five scores is calculated as the sum of the items included. Scores are transformed to a 0–100 scale, with zero representing extreme knee problems and 100 representing no knee problems as common in orthopaedic scales [13,14] and generic measures [15]. Scores between 0 and 100 represent the percentage of total possible score achieved. An aggregate score was not calculated since it was regarded desirable to analyze and interpret the five dimensions separately.

Since it was believed a priori that functions such as running, jumping, squatting, kneeling and pivoting were not applicable to all patients undergoing total knee replacement, a sixth answer option (not applicable) was given for the five items included in the subscale Sport and Recreation Function. If the box "not applicable" was marked the item was treated as missing data.

Missing data. If a mark was placed outside a box, the closest box was used. If two boxes were marked, that which

indicated the more severe problems was chosen. Missing data were treated as such; one or two missing values were substituted with the average value for that subscale. If more than two items were omitted, the response was considered invalid and no subscale score was calculated.

### **SF-36**

The SF-36 is a widely used measure of general health status which comprises eight subscales; Physical Functioning, Role-Physical, Bodily Pain, General Health, Vitality, Social Functioning, Role-Emotional and Mental Health [11,15]. The SF-36 is self-explanatory and takes about 10 minutes to complete. The SF-36 is scored from 0 to 100, 0 indicating extreme problems and 100 indicating no problems.

### **Background data**

In addition to demographic data, patients were asked to report co-morbid conditions. Patients were asked if they were currently treated by a doctor, or had been treated during the last year, for any of the following 11 conditions: Back problems, Lung disease, High blood pressure, Heart disease, Impaired circulation in the lower extremity, Neurologic disease, Diabetes, Cancer, Ulcer, Kidney disease, Impaired vision or eye disease.

### **Reliability**

To assess test-retest stability questionnaires were sent out one week apart on two separate occasions (pre-operatively and 6 month follow-up) for two different randomly selected subsets of patients. Wilcoxon signed rank test was used to determine if any significant changes occurred between the test and retest administration of the questionnaire. Intraclass correlation coefficients (ICC<sub>2,1</sub>) were calculated for all patients together and for the pre-operative and post-operative assessments separately. According to the method suggested by Bland and Altman the difference between the two assessments was plotted against the mean of the two assessments for each subject. 95% of differences were expected to be less than two standard deviations [16].

### **Validity**

Content validity was assessed at baseline by asking the patients to rate the importance of improvement in each of the five KOOS subscales on a 5-point Likert-scale as extremely important, very important, moderately important, somewhat important, or not important at all. For each subscale examples of included questions were given.

Convergent and divergent construct validity was determined by comparison of the pre-operative administrations of the KOOS and the SF-36. The SF-36 subscale Physical Functioning measures limitations of the ability to perform general physical activities, a corresponding

construct to what the ADL and Sport scales of the KOOS are intended to measure. SF-36 Bodily Pain measures pain/ache and disturbances in normal activities, a construct similar to knee pain which the KOOS Pain scale is designed to measure. We expected the highest correlations when comparing the scales that are supposed to measure the same or similar constructs. Further the eight subscales of SF-36 have been shown to produce valid indices of Physical Health and Mental Health [17]. Since the KOOS is designed to measure physical health rather than mental health we expected to observe higher correlations between the KOOS subscales and the SF-36 subscales of Physical Function, Bodily Pain, and Role Physical (convergent construct validity) than between KOOS subscales and the SF-36 subscales of Mental health, Vitality, Role Emotional, Social Functioning, and General Health (divergent construct validity). However based on previous methodological studies of the KOOS, we expected the correlations to the SF-36 subscale Role Physical to be lower than the correlations to the subscales Physical Function and Bodily Pain [6,9].

### **Responsiveness**

We expected that total knee replacement would induce a change in patients' perception of symptoms and function that could be measured by the questionnaires. Responsiveness was calculated as effect size, standardized response mean (SRM) and relative efficiency. Effect size is defined as mean score change divided by the standard deviation of the pre-operative score [3]. Effect sizes >0.8 are considered large [18]. Standardized response mean is defined as mean score change divided by the standard deviation of the change score [4]. Relative efficiency was computed by squaring the ratio of the z-statistics [5].

In part, the ability to respond to change can be assessed in terms of the proportion of patients at the floor (i.e. the worst score) or the ceiling (i.e. the best score) of each scale [19]. To assess the ability to respond to change the floor and ceiling effects were determined pre-operatively, at 6, and 12 months. For comparative reasons the WOMAC was examined in the same way.

## **Results**

### **Patients**

Of the 105 included patients, 39 were men and 66 (63%) were women, with a mean age of 71.3 years (range, 43–86). 22 patients had undergone a prior knee replacement of the other knee, and 5 patients had undergone a prior hip replacement. The patients self-reported on average 1.3 co-morbid diseases (median 1, range 1–11). In 56 cases (53%) the right knee was operated, in 43 (41%) the left, and in 6 (6%) both knees were operated simultaneously.

One patient died before the 6 month-follow up, and 7 did not return the questionnaires. Thus 6 month-follow up data were available for 97 (92%) patients. 12 month follow-up data were available for 90 (86%) of the included patients (two patients died between the 6 and 12 month follow-ups, and an additional five did not return the questionnaires).

#### **Missing baseline data**

KOOS. Few individual items were missing for the four subscales Pain, Symptoms, ADL and knee-related Quality of Life (126 items of 105 patients  $\times$  37 items = 3.2 %). A subscale score could be calculated for 103/105 patients for the subscale Pain, 105/105 for Symptoms, 104/105 for the subscale ADL, and 105/105 for the subscale knee-related Quality of Life. For the subscale Sport and Recreation Function 391 items of 105 patients  $\times$  5 items = 74% were noted as "not applicable" and thus treated as missing. A subscore could be calculated for 58/105 patients.

#### **Reliability**

Totally, 54 patients included in the study completed the KOOS twice within one to 23 days. Pre-operatively, test-retest data were available for 28 patients (mean number of days between the two assessments  $9.9 \pm 3.8$  days). At the six-month follow-up, test-retest data were available for another 26 patients (mean number of days between the two assessments  $10.2 \pm 5.6$  days). There were no statistically significant differences of the scores between the first and second assessments with the exception of the subscale Symptoms. When calculated for all 54 patients together, the patients reported on average more symptoms at the second test occasion (60/100 vs. 58/100 points,  $p = 0.04$ , Wilcoxon signed-rank test). The intraclass correlation coefficients (ICC<sub>2,1</sub>) were all over 0.75 when determined for all 54 patients together (Table 1). Bland and Altman plots for the five KOOS scales are given in Figure 1.

#### **Validity**

Content validity. Over 90% reported that improvement in the four subscales Pain, Symptoms, Activities of Daily Living, and knee-related Quality of Life was extremely or very important when deciding to have their knee operated on, Table 2. 51% reported that improvement in functions included in the subscale Sport and Recreation Function such as squatting, kneeling, jumping, turning/twisting and running was extremely or very important when deciding to have their knee operated on. The group reporting items related to Sport and Recreation Function being extremely or very important held more men (48% vs. 30%,  $p = 0.08$ ) but was similar with regard to age (71 vs. 70,  $p = 0.6$ ) and preoperative ADL function (41/100 vs. 40/100,  $p = 0.8$ ).

Following surgery, patients tended to start doing physical functions that they had not performed pre-operatively.

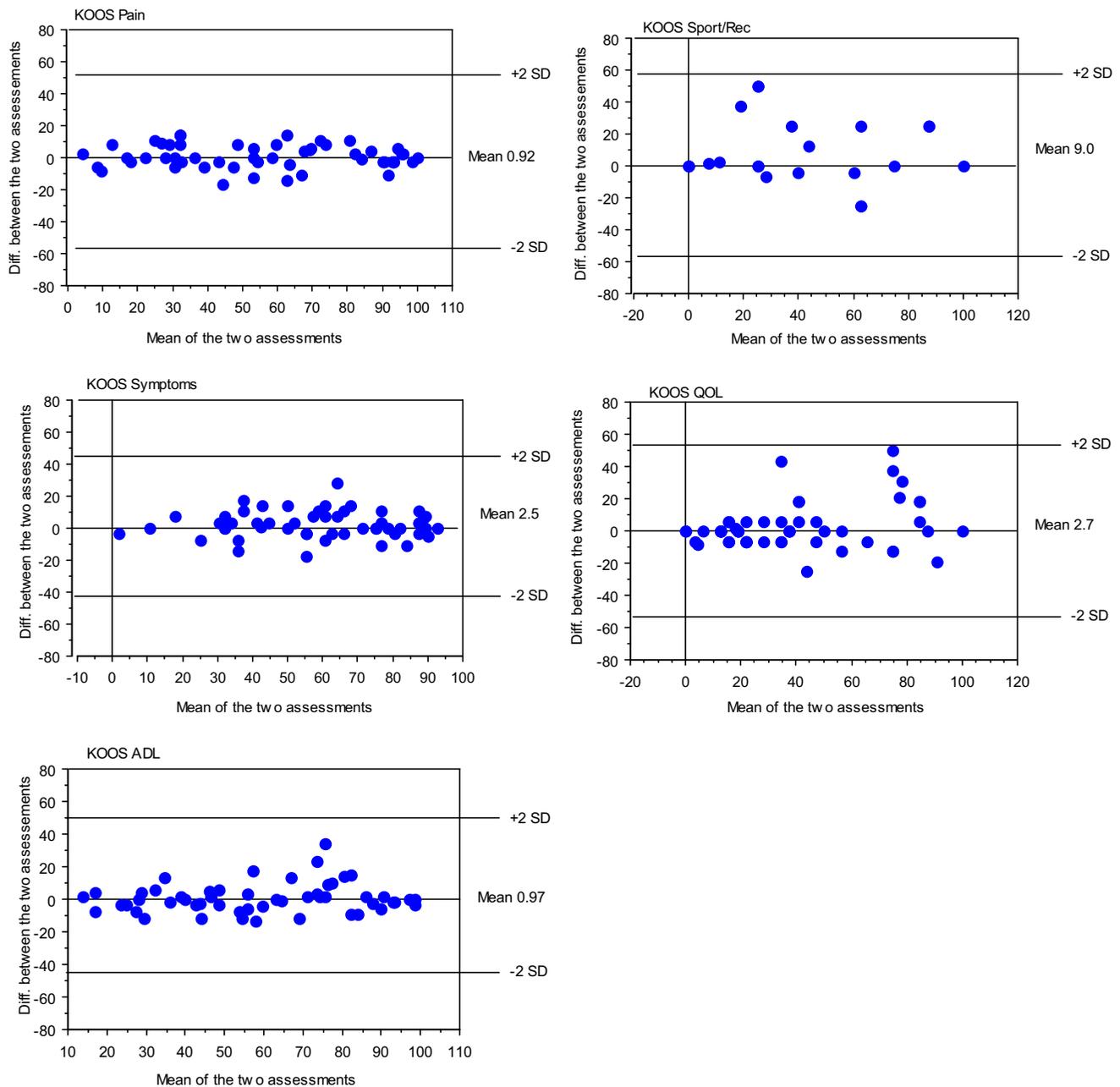
Pre-operatively, 27% rated their degree of difficulty with squatting, 17% with running, 12% with jumping, 42% with twisting/pivoting, and 34% with kneeling. The others reported not performing the function. At six months the percentages of patients reporting doing the functions had increased to 40%, 28%, 23%, and 46% for squatting, running, jumping, and twisting/pivoting. The percentage reporting kneeling had decreased to 26%. These trends were confirmed at the 12 month follow-up.

Construct validity. As expected, high correlations occurred between the SF-36 scales and the KOOS scales that are intended to measure similar constructs (bodily pain vs. pain,  $r_S = 0.62$ ; physical function vs. activities of daily living,  $r_S = 0.48$ ). Generally, higher correlations were seen when comparing KOOS scales to SF-36 scales with a high ability to measure physical health (convergent construct validity), and lower correlations were seen when comparing KOOS scales to SF-36 scales with a high ability to measure mental health (divergent construct validity). The correlations of the KOOS scales to the SF-36 subscale Role Physical were lower compared to the other SF-36 subscales with a high ability to measure physical health (Table 3).

#### **Responsiveness**

A significant improvement ( $p < 0.001$ ) was seen post-operatively in all subscales (Table 4). The most responsive subscale was knee-related quality of life (QOL) with an effect size of 2.86 at 6 months and 3.54 at 12 months. The second most responsive subscale was Pain with effect sizes of 2.28 and 2.55 at 6 and 12 months, respectively. The subscale sport and recreation function (Sport/Rec) was the least responsive subscale with effect sizes of 1.18 and 1.08 at 6 and 12 months, respectively. It should be noted that the effect size calculation for the subscale Sport/Rec are based on 29 and 27 patients only. Generally the effect sizes were larger at 12 months, implying improvement occurring between 6 and 12 months (Table 5). The calculation of SRM generally yielded somewhat smaller numbers but did not change the interpretation of the data (Table 5).

Floor and ceiling effects. Pre-operatively, no notable ceiling effects were found. At 6 months, 15% reported best possible pain score and 16% reported best possible sport and recreation score making detection of further improvement impossible. The ceiling effects for the other subscales were lower. At 12 months, 22% reported best possible pain score and 17% reported best possible quality of life score (Table 6).



**Figure 1**  
Bland-Altman plots for the five KOOS subscales

**Comparison of the KOOS to the WOMAC Validity**

All WOMAC subscales and the corresponding KOOS subscales Pain, Symptoms and ADL were rated as extremely or very important by over 90% of the patients (Table 2). 91% of the patients rated the KOOS subscale knee-related quality of life as extremely or very important, indicating

items such as awareness, life style modifications, and confidence being just as important as questions related to daily living. The WOMAC does not assess this dimension of disease. The KOOS subscale Sport and Recreation Function was considered as extremely or very important by 51%, indicating functions such as squatting, running,

**Table 1: Intraclass correlation coefficients**

KOOS subscale	Intraclass correlation coefficients		
	Pre-op, n = 28*	Post-op, n = 26†	Total group, n = 54
Pain	0.90	0.94	0.97
Symptoms	0.92	0.85	0.94
ADL	0.89	0.82	0.93
Sport/Rec	0.58	0.74	0.78
QOL	0.89	0.73	0.88

\* n = 8 for the subscale Sport and Recreation pre-operatively † n = 13 for the subscale Sport and Recreation post-operatively

**Table 2: Content validity. The percentage of patients reporting the importance of the five different koos subscales when deciding to have the operation**

	Pain	Symptoms	ADL	Sport/Rec	QOL
Extremely or very important	95%	92%	93%	51%	91%
Somewhat or not important at all	5%	8%	7%	49%	9%

**Table 3: Construct validity. Spearman's correlation coefficients ( $r_s$ ) determined when comparing KOOS' five subscales to the SF-36 eight different subscales. N = 103-105 with the exception of Sport/Rec where n = 58.**

SF-36 subscale	Physical Health†	Mental Health†	KOOS Pain	KOOS Symptoms	KOOS ADL	KOOS Sport/Rec	KOOS QOL
Physical Function	Strong	Weak	0.19	0.20*	0.48***	0.11	0.34***
Role Physical	Strong	Weak	0.16	0.06	0.26**	-0.06	0.10
Bodily Pain	Strong	Weak	0.62***	0.37***	0.68***	0.34*	0.60***
General Health	Moderate	Moderate	0.16	0.08	0.29**	-0.1	0.07
Vitality	Moderate	Moderate	0.38***	0.30**	0.50***	0.24	0.33***
Social Functioning	Moderate	Strong	0.26**	0.14	0.39***	0.12	0.40***
Role Emotional	Weak	Strong	0.12	0.09	0.36***	0.24	0.06
Mental Health	Weak	Strong	0.14	0.05	0.28**	0.09	0.10

† SF-36 subscales ability to measure physical health vs. mental health [17]

**Table 4: Mean (SD) of the KOOS and WOMAC at baseline and follow-ups at 6 and 12 months. 0-100 worst to best scale**

	Baseline	6 months	12 months
<b>KOOS</b>			
Pain	38(18)	79(20)	83(16)
Symptoms	47(20)	72(18)	84(16)
ADL	41(16)	77(17)	82(16)
Sport/Rec	16(22)	48(33)	46(30)
QOL	19(14)	59(25)	69(24)
<b>WOMAC</b>			
Pain	42(19)	83(17)	86(15)
Stiffness	40(22)	70(20)	77(19)
Function	41(16)	77(17)	82(16)

**Table 5: Effect sizes and Standardized Response Mean (SRM) 6 and 12 months post-operatively**

Subscale	Effect Size 6 months	Effect Size 12 months	SRM 6 months	SRM 12 months
KOOS Pain	2.28	2.55	1.67	2.12
WOMAC Pain	2.08	2.27	1.77	2.02
SF-36 Bodily Pain	2.10	2.28	1.31	1.51
KOOS Symptoms	1.24	1.59	0.99	1.25
WOMAC Stiffness	1.30	1.65	0.92	1.20
KOOS ADL	2.25	2.56	1.70	1.90
KOOS Sport/Rec*	1.18	1.08	0.81	0.88
WOMAC Function	2.25	2.56	1.70	1.90
SF-36 physical function	2.01	2.23	1.39	1.54
KOOS QOL	2.86	3.54	1.60	1.99

\*For the subscale Sport/Rec, 29 subjects and 27 subjects, respectively, were used for the calculations at 6 and 12 months.

jumping, turning/twisting and kneeling being of great importance to every second patient undergoing total knee replacement. These functions are not assessed by the WOMAC.

#### Responsiveness

The KOOS subscale knee-related QOL had the highest effect size of all subscales of the KOOS and the WOMAC (Table 5). The relative efficiency when comparing corresponding subscales of the KOOS and the WOMAC (pain vs. pain, symptoms vs. stiffness) at the 6 and 12 month follow-up ranged from 1.0 to 1.04 indicating corresponding subscales of both measures being equally responsive. The KOOS subscale ADL is equivalent to the WOMAC subscale Function. No comparison to the KOOS subscales Sport and Recreation Function and knee-related Quality of Life were made since the WOMAC does not assess corresponding constructs.

The ceiling effects at 6 and 12 months of the WOMAC subscales Pain and Stiffness were higher than for the corresponding subscales of the KOOS, indicating KOOS having a better ability than WOMAC to detect future improvement post-operatively (Table 6).

#### Discussion

We have shown that KOOS is a useful, reliable, valid and responsive instrument for assessment of patient-relevant outcomes in elderly subjects with advanced osteoarthritis. The KOOS may have some advantages and some disadvantages compared to the WOMAC.

#### Validation of the KOOS in different populations

KOOS has been proven reliable, valid and responsive in operative treatment of knee injury such as arthroscopy [9] and reconstruction of the anterior cruciate ligament [6]. However, validation is an ongoing process, and to fully validate an outcome instrument it has to perform as ex-

pected over time in different settings [20]. Previously, we have shown that it was possible to adapt the KOOS to assess patient-relevant outcomes related to other joints, the foot and ankle, in patients of similar age and activity level to the knee patients the instrument was initially developed for [21]. In the current study, we have shown that items found relevant for younger or more active individuals with osteoarthritis match the expectations of older OA patients and thus improve validity and make possible greater responsiveness also for older OA patients.

#### Reliability

The test-retest reliability in the current study with intraclass correlation coefficients ranging from 0.78 to 0.97 are comparative to previous methodological studies of the KOOS [6,9] and the adaptation of the KOOS to foot- and ankle related problems [21] where the intraclass correlation coefficients have ranged from 0.70 (ADL subscale of the foot- and ankle adaptation) to 0.93 (Symptom subscale for the KOOS when used in patients with anterior cruciate ligament injury). A statistically significant difference of 2/100 points in mean score between the first and second assessment was found for the subscale Symptoms. This difference is far smaller than the clinically significant difference which is thought to be in the magnitude of 10 points [22], and also far smaller than previously detected changes over time for the KOOS subscale symptoms. Three months after meniscectomy, in a middle-aged subgroup with open lesion of the cartilage or exposed bone, the average postoperative improvement in the subscale Symptoms was 11 points [23]. In a double-blind placebo-controlled trial on the effects of a nutritional supplement in adults with osteoarthritis, the mean improvement at six weeks in the subscale Symptoms for the treatment group was 9 points compared to 1 point for the control group [24].

**Table 6: Ceiling and floor effects of the KOOS and WOMAC. Percentage of patients reporting best possible score (ceiling effect)/ worst possible score (floor effect).**

	Pain	Symptoms	KOOS ADL	Sport/Rec	QOL	Pain	WOMAC Stiffness	Function
Pre-op	1/0	0/1	0/0	0/48	0/14	1/2	2/7	0/0
6 mo	15/0	3/0	8/0	16/16	11/1	27/0	15/0	8/0
12 mo	22/0	12/0	11/0	9/12	17/0	30/0	27/0	11/0

**Validity**

Determination of content validity, or the extent to which measures represent functions or items of relevance given the purpose and matter at issue [25], should be a minimum prerequisite for acceptance of a measure [26]. We have asked the patients to rate how relevant or important they find each subscale of the KOOS as a method to determine content validity. Another possibility would have been to ask the patients to rate each of the 42 items separately. It is however confusing to rate both relevance and degree of difficulty with an item at the same time. Since the KOOS is an already existing measure and we in the current study have studied the measurement properties in another population, with problems from the same joint as the measure was developed for, we found it appropriate to rate the relevance of each subscale instead of each item to ensure content validity. By doing so we found that physical functions such as squatting, running, jumping, twisting/pivoting and kneeling are extremely or very important to every second patient undergoing total knee replacement due to osteoarthritis. This was a surprising finding that might reflect that the aging population today has higher demands of physical activity compared to some decades ago. The mean age of the patients, 71.3 years, the sex distribution (63% women), the number of reported co-morbidities (mean 1.3, range 1–11) and baseline WOMAC scores seem comparable to other knee replacement populations studied in the UK, US and Australia [27] (E Lingard, Kinemax Outcomes Study, personal communication, Januari 2003) but to generalize the finding to other populations with varying physical exercise habits, further studies are needed.

Construct validity was determined in comparison to the eight subscales of the SF-36. The pattern hypothesized a priori, with higher correlations to subscales with a high ability to measure physical health and lower correlations to subscales with a high ability to measure mental health, was found confirming the KOOS measuring the suggested constructs. As found in previous methodological studies of the KOOS [6,9] the correlations to the subscale Role Physical were lower than the correlations to the other subscales with a strong ability to measure physical health.

This could reflect the previously observed different psychometric properties of the subscale Role Physical as opposed to Physical Function and Bodily Pain in orthopaedic patients [28,29]. The patterns are comparable with some exceptions reflecting the different populations for which the KOOS has been studied. The strongest correlation of the KOOS subscale Sport and Recreation Function to the SF-36 subscale Physical Function was found in patients undergoing anterior cruciate ligament reconstruction [6] or arthroscopy [9], while in the current study on patients having total knee replacement, the strongest correlation of Sport and Recreation Function was to the SF-36 subscale Bodily Pain.

**Responsiveness**

Responsiveness is defined as the sensitivity of an assessment technique to change over time [2]. A high effect size or standardized response mean indicate fewer patients needed to demonstrate a statistical difference and a diminished risk for type II error. Responsiveness can be measured using variables such as effect size [3], standardized response mean [4] or relative efficiency [25]. Since no gold standard for measuring responsiveness exists, we choose to calculate all three measures. The standardized response mean yielded somewhat smaller numbers than the effect size calculations but did not change the interpretation of the results.

**Missing baseline data**

A subscale score could be calculated pre-operatively for all subscales for 97% of the patients, a surprisingly high number in this elderly population. In a previous study using the WOMAC to assess total knee replacement completion rates over 90% were found [27], and in a previous validation study of the KOOS in a younger population the completion rates was almost 100% (for 153 patients one subscale score could not be calculated due to missing data) [9].

**Comparison of the KOOS to the WOMAC**

The WOMAC is recommended to use for evaluation of treatment effects in trials including elderly with knee osteoarthritis [1] and in total knee replacement [27]. We

found the effect sizes of the WOMAC in the current study being sufficient to enable convenient sample sizes in clinical studies of total knee replacement and thus confirm the latter recommendation. However, using the KOOS instead of the WOMAC may during some circumstances be considered advantageous.

The inclusion of the subscale Sport and Recreation Function may be considered an advantage. Although Sport and Recreation Function is not relevant to all patients, the subscale improves validity by assessing functions considered extremely or very important by every second patient undergoing total knee replacement. Thus by adding the subscale Sport and Recreation Function, assessment of functional improvement being undetected by other commonly used disease-specific instruments is possible. We found that patients following total knee replacement start doing more demanding physical functions than they did prior to the operation. In the current study the patients were given the option to rate the items considered as more demanding activities as "not applicable". To make possible measuring improvement over time, and minimize data loss in clinical studies, we suggest that in future studies the five answer options normally included in the KOOS questionnaire (ranging from "no difficulty" to "extreme difficulty") should be used without the addition of the option "not applicable". It is our experience that patients will choose to answer "extreme difficulty" with e.g. squatting if they want to squat but are not able to. If they are not interested in squatting however, they commonly choose the answer option "no difficulty", since this lack of function does not present a problem to them. To ask the patients to rate difficulty, activity limitations according to International Classification of Functioning, Disability and Health (ICF) [30] instead of ability (impairment), as common in orthopaedic rating scales, provides a possibility for the patient to individualize the importance of each item.

Another advantage of the KOOS compared to the WOMAC is the inclusion of the subscale knee-related Quality of Life. Knee-related Quality of Life was reported to be extremely or very important by over 90% of the patients, was the most responsive subscale both at 6 and 12 months, and had an effect size of 3.54 at 12 months post-operatively. In addition, it was the subscale that best showed the improvement occurring between 6 and 12 months. This latter finding could indicate that improvement in pain is faster than adaptation to the new situation and improvement in items such as trust in knee, awareness of knee and life-style changes because of the knee-problems. Corresponding findings have been seen with the use of a generic measure, the SF-36, in total hip replacement [31].

A disadvantage of the KOOS compared to the WOMAC is the increased number of items, 42 compared to 24, resulting in a larger burden for the patient. This might be an issue if multiple instruments are administered at the same time.

#### **Future applications**

Total joint replacement is a very successful treatment of severe osteoarthritis. The reduction in pain is immediate and over 90 % of the patients report being satisfied with the procedure [32]. However, patients report expectations of functional improvement to be just as frequent and important as expectations of pain relief [8]. Physical function do not necessary improve because of pain reduction, implying a need for rehabilitation, exercise and physical therapy aiming at restoring physical function. Few studies on rehabilitation, exercise and physical therapy after total joint replacement are found in the literature but it seem possible to improve physical activity by exercise programs carried out at home or in groups [33]. The evidence is however not strong, presumably because of poorly designed studies with small sample sizes. An alternative explanation is the use of outcome measures with poor validity and responsiveness concerning physical function. In studies where physical function is the primary outcome measure it may be an advantage to use the KOOS compared to the WOMAC.

#### **Conclusions**

The KOOS is a useful, reliable, valid and responsive instrument for assessment of patient-relevant outcomes in elderly subjects with advanced osteoarthritis. Compared to the WOMAC, the KOOS could be advantageous when assessing younger groups, groups with high expectations of physical activity, interventions with smaller effects or interventions where physical function is the primary outcome, and when assessing long-term outcome.

#### **Authors' contributions**

ER and STL designed the study and coordinated the data collection. ER analyzed the data and drafted the manuscript. Both authors read and approved the final manuscript.

#### **Acknowledgements**

We would like to acknowledge Mrs Lena M Hansson for excellent help with data collection.

#### **References**

1. Bellamy N, Kirwan J, Boers M, Brooks P, Strand V, Tugwell P, Altman R, Brandt K, Dougados M and Lequesne M **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
2. Bellamy N **Musculoskeletal Clinical Metrology** London: Kluwer Academic Publishers 1993,
3. Kazis LE, Anderson JJ and Meenan RF **Effect sizes for interpreting changes in health status** *Med Care* 1989, **27**:S178-89

4. Liang MH, Fossel AH and Larson MG **Comparisons of five health status instruments for orthopedic evaluation** *Med Care* 1990, **28**:632-42
5. Liang MH, Larson MG, Cullen KE and Schwartz JA **Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research** *Arthritis Rheum* 1985, **28**:542-7
6. Roos EM, Roos HP, Lohmander LS, Ekdahl C and Beynnon BD **Knee Injury and Osteoarthritis Outcome Score (KOOS) – development of a self-administered outcome measure** *J Orthop Sports Phys Ther* 1998, **28**:88-96
7. Roos EM, Roos HP and Lohmander LS **WOMAC Osteoarthritis Index – additional dimensions for use in subjects with post-traumatic osteoarthritis of the knee. Western Ontario and McMaster Universities Osteoarthritis Cartilage** 1999, **7**:216-21
8. Roos EM, Nilsson AK and Toksvig-Larsen S **Patients' expectations suggest additional outcomes in total knee replacement** *In: Association of Rheumatology Health Professionals New Orleans* 2002,
9. Roos EM, Roos HP, Ekdahl C and Lohmander LS **Knee injury and Osteoarthritis Outcome Score (KOOS) – validation of a Swedish version** *Scand J Med Sci Sports* 1998, **8**:439-48
10. Roos EM, Klassbo M and Lohmander LS **WOMAC osteoarthritis index. Reliability, validity, and responsiveness in patients with arthroscopically assessed osteoarthritis. Western Ontario and McMaster Universities** *Scand J Rheumatol* 1999, **28**:210-5
11. Sullivan M and Karlsson J **SF-36 Hälsoenkät: Swedish Manual and Interpretation Guide** Gothenburg, Sweden: Health Care Unit, Sahlgrenska Hospital 1994,
12. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J and 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
13. Barber SD, Noyes FR, Mangine RE, McCloskey JW and Hartman W **Quantitative assessment of functional limitations in normal and anterior cruciate ligament-deficient knees** *Clin Orthop* 1990, **204**:14
14. Tegner Y and Lysholm J **Rating systems in the evaluation of knee ligament injuries** *Clin Orthop* 1985, **43**:9
15. Ware JE Jr and Sherbourne CD **The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection** *Med Care* 1992, **30**:473-83
16. Bland JM and Altman DG **Statistical methods for assessing agreement between two methods of clinical measurement** *Lancet* 1986, **1**:307-10
17. Ware JE Jr, Snow K, Kosinski M and Gandek B **SF-36 Health Survey Manual and Interpretation Guide** Boston, MA: The Health Institute, New England Medical Center 1993,
18. Cohen J **Statistical power analysis for the behavioural sciences** New York: Academic Press 1977,
19. Fortin PR, Stucki G and Katz JN **Measuring relevant change: an emerging challenge in rheumatologic clinical trials** *Arthritis Rheum* 1995, **38**:1027-30
20. Kirkley A, Griffin S, McLintock H and Ng L **The development and evaluation of a disease-specific quality of life measurement tool for shoulder instability. The Western Ontario Shoulder Instability Index (WOSI)** *Am J Sports Med* 1998, **26**:764-72
21. Roos EM, Brandsson S and Karlsson J **Validation of the foot and ankle outcome score for ankle ligament reconstruction** *Foot Ankle Int* 2001, **22**:788-94
22. Ehrich EW, Davies GM, Watson DJ, Bolognese JA, Seidenberg BC and Bellamy N **Minimal perceptible clinical improvement with the Western Ontario and McMaster Universities osteoarthritis index questionnaire and global assessments in patients with osteoarthritis** *J Rheumatol* 2000, **27**:2635-41
23. Roos EM, Roos HP, Ryd L and Lohmander LS **Substantial disability 3 months after arthroscopic partial meniscectomy: A prospective study of patient-relevant outcomes** *Arthroscopy* 2000, **16**:619-26
24. Colker CM, Swain M, Lynch L and Gingerich DA **Effects of a milk-based bioactive micronutrient beverage on pain symptoms and activity of adults with osteoarthritis: a double-blind, placebo-controlled clinical evaluation** *Nutrition* 2002, **18**:388-92
25. Johnston MV, Keith RA and Hinderer SR **Measurement standards for interdisciplinary medical rehabilitation** *Arch Phys Med Rehabil* 1992, **73**:S3-23
26. Streiner DL and Norman G **Health measurement scales. A practical guide to their development and use** Second edn. Oxford: Oxford University Press 1995,
27. Brazier JE, Harper R, Munro J, Walters SJ and Snaith ML **Generic and condition-specific outcome measures for people with osteoarthritis of the knee** *Rheumatology (Oxford)* 1999, **38**:870-7
28. Shields RK, Enloe LJ and Leo KC **Health related quality of life in patients with total hip or knee replacement** *Arch Phys Med Rehabil* 1999, **80**:572-9
29. Martin DP, Engelberg R, Agel J and 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
30. WHO **International classification of functioning, disability and health** *In: Book International classification of functioning, disability and health* City: WHO 2001, **2003**:
31. Nilsson AK and Lohmander LS **Age and waiting time as predictors of outcome after total hip replacement for osteoarthritis** *Rheumatology (Oxford)* 2002, **41**:1261-7
32. Robertsson O, Dunbar M, Pehrsson T, Knutson K and Lidgren L **Patient satisfaction after knee arthroplasty: a report on 27,372 knees operated on between 1981 and 1995 in Sweden** *Acta Orthop Scand* 2000, **71**:262-7
33. Roos EM **Effectiveness and practice variation of rehabilitation after joint replacement** *Curr Opin Rheumatol* 2003, **15**:160-2

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

