Meniscal problems: to repair and to replace

Wal, R.J.P. van der

8

Translation and Validation of the Dutch Western Ontario Meniscal Evaluation Tool (WOMET).

Robert J.P. van der Wal
Bastiaan T.J. Heemskerk
Ewoud R.A. van Arkel
Ludwine B. Mokkink
Bregje J.W. Thomassen

J Knee Surg. 2017 May;30(4):314-322.

ABSTRACT

Purpose: The purpose of this study was to translate the Western Ontario Meniscal Evaluation Tool (WOMET) into Dutch and to determine validity, reliability and responsiveness of the Dutch version.

Methods: The WOMET was translated into Dutch according to a standardized forward-backward translation protocol. Eighty-six patients (51 male, 35 female, mean age 52.2 (standard deviation (SD) 11.4)) with isolated meniscal pathology were included. WOMET was completed three times (at baseline around 2 weeks and after 3 months from baseline). Knee injury and Osteoarthritis Outcome Score (KOOS), International Knee Documentation Committee (IKDC) Subjective Knee Form, IKDC Current Health Assessment Form and two anchor questions were also answered. Content validity, construct validity, reliability and responsiveness was determined.

Results: The Dutch WOMET showed good construct validity (good correlation with all other questionnaires and all hypotheses confirmed), content validity (floor and ceiling effects (<30%), internal consistency (Cronbach's alpha = 0.88) and test-retest reliability (intraclass correlation coefficient = 0.78). The Dutch WOMET was found responsive to change (88% confirmation of the predefined hypotheses). The smallest detectable change (SDC) and minimal important change (MIC) for the Dutch WOMET are 15.4 and 14.7, respectively.

Conclusions: The Dutch version of the WOMET is valid and reliable. It can be used as a disease specific tool to evaluate health related quality of life of Dutch patients with meniscal pathology.

INTRODUCTION

Meniscus injuries are the most common knee injury,[32] and can be classified into traumatic tears[37] or degenerative lesions.[31] The most common treatment for meniscal lesion is an arthroscopic partial meniscectomy, making this the most frequent orthopaedic intervention. Each year, almost half a million partial meniscectomies occur annually in the United States and this cost several thousand dollars each.[10,26] In the Netherlands about 30.000 arthroscopies for meniscal pathology were performed in 2000. In 2010 this has increased to about 42.000, of which 65% in the patients who were older than 45 years of age. This represents an increase in ten years' time of 46%. The percentage of meniscal repair is unknown.[7]

To evaluate an intervention, measures of impairment (such as pain, swelling and mechanical problems) as assessed after arthroscopic meniscal treatment, outcomes have to be determined. Surgeons can use objective measures such as range of motion or radiographic imaging to confirm meniscal healing after meniscal repair or the presence of a tear remnant after partial meniscectomy. However, to ensure the intervention satisfies the expectations of the patients, an instrument must measure the patient relevant outcomes, emphasizing health status, disability and function.[43] Therefore, patient reported outcomes measures (PROMs) can be used. Although this instrument does not measure objective outcomes, it can be an important instrument to measure the outcome of an intervention, because an objectively measured good result does not always ensure good patients satisfaction.[4] For instance, the result of a partial meniscectomy can be successful when judged with objective measures like wound healing, range of motion and swelling, but can be disappointing for the patient when there are still functional problems.[4,34]

Although arthroscopic meniscal treatment (partial meniscectomy and meniscal repair) is the most common orthopaedic procedure, there is no validated pathology-specific health-related quality of life outcome measurement for this type of injury in the Netherlands. Nowadays, various instruments have been used to assess the outcome of treatment of knee pathology: the Knee injury and Osteoarthritis Outcome Score (KOOS),[33] the International Knee Documentation Committee (IKDC) Subjective Knee Form,[19] the Lysholm score,[24] the Oxford Knee Score,[11] and the Cincinnati Knee Rating Scale.[29] The KOOS has been validated for patients with multiple knee problems, including meniscal lesions.[35] The Lysholm knee score and the IKDC subjective knee has been specially validated for patients with meniscal pathology.[5,8] However, these instruments do not all measure the outcomes for specific meniscal pathology or are not a specific health-related quality-of-life instrument. Kirkley et al. developed the Western Ontario Meniscal Evaluation Tool (WOMET), which is the first meniscal pathology-specific health-related quality of life instrument.[23] The WOMET is a valid, reliable, and responsiveness disease-specific outcome measure for

the assessment of health-related quality of life deficits in patients with meniscal pathology.[23]

To our knowledge, there is no Dutch version of the WOMET available at this moment. Considering the fact that the original version of the WOMET has already been validated in English[23] and recently is translated and validated in Finnish[36] and Turkish,[6] we hypothesized that the Dutch version of the WOMET is also a valid instrument to for the assessment of health-related quality of life deficits in Dutch patients with meniscal pathology. Therefore, the purpose of this study was to translate the WOMET into the Dutch language and determine the validation, reliability, and responsiveness for patients with traumatic or degenerative meniscal injury.

MATERIALS AND METHODS

Study Population

From July 2013 till June 2015, all consecutive patients who had a Magnetic Resonance Imaging (MRI) confirmed symptomatic meniscal tear were assessed for participation in this study. Patients were included at the orthopaedic outpatient clinic of the Medical Center Haaglanden by their orthopaedic consultant. Patients signed an informed consent after meeting all inclusion criteria. Inclusion criteria were: patients between 18 and 70 years old, a MRI confirmed symptomatic, isolated, traumatic or degenerative meniscal tear, no signs of osteoarthritis on plain radiographs, understanding of the Dutch language. Patients were excluded if they had concomitant ligament injuries or previous ligament injury with persistent knee instability, any previous knee operation, chondropathy higher than grade two on the Outerbridge scale[30] seen on MRI or during the operation, or inability to participate due to cognitive impairment. Because of the high prevalence of meniscal lesions among elderly people,[16] it was decided to exclude all patients over 70 years of age.

Study Design

All patients were asked to complete three sets of questionnaires at three time points: T0 (baseline), T1 (around 2 weeks after baseline and before start of intervention) and T2 (at least at 12 weeks after start of intervention; surgery or start of conservative treatment). All sets contained the Dutch WOMET, KOOS, IKDC Current Health Assessment form and IKDC Subjective Knee Form. At T1 an anchor question (one on remembrance of questions) and at T2 two anchor questions (one on remembrance and one about the effect of treatment) were asked as well. Patients received their first questionnaire at the outpatient department or by post to fill in at home. The second and third questionnaire was send by post to fill in at home. For this study the KOOS and IKDC Subjective Knee Form were used because

these questionnaires were the second and third best instruments to measure symptoms in patients with meniscal lesions after the WOMET.[39]

Outcome Measures

WOMET

The WOMET consists of 16 questions and is specifically developed to evaluate health-related quality of life in patients with meniscal pathology.[23] The score may be reported as a total overall score and a total score per subscale (symptoms; sport, recreation, work, and lifestyle; emotions). The original questionnaire contains a visual analogue scale (VAS), however in this study we have chosen for numeric rating scale (NRS) from 0-10, with 0 indicating no problems at all, and 10 indicating the worst problems. We divided the WOMET in three different subscales to compare the WOMET with the other questionnaires: pain, function and quality of life (QoL). Questions 1,3,4,5,6,7,9,11 and 12 represented the subscale 'function', questions 2,8,13 represented the subscale 'pain' and the remaining questions 10,14,15,16 represented the subscale 'quality of life'.

The WOMET was translated according to a forward-backward translation protocol.[17,27,28] Two independent bilingual translators, one with a medical background and one without a medical background, and a translation agency created a Dutch version of the WOMET. Any discrepancies between the translators were solved by consensus between all translators. Two other independent bilingual translators, who were blinded for the original version of the WOMET, then translated the Dutch version back to English. Next to that, the forward-backward translated WOMET was compared with the original WOMET, to see if any changes occurred during the translation process. The Dutch version of the WOMET was presented to a focus group, consisting of seven patients with meniscal pathology and one independent orthopaedic surgeon, for feedback on the clarity, content and relevance of the questions. The feedback was used to improve the Dutch version of the WOMET and the final version was composed (see attachment). This final version was pre-tested on 25 patients with isolated meniscal injury to check interpretation, cultural relevance of translation and ease of comprehension.

KOOS

The KOOS is a Swedish questionnaire developed by Roos et al.[33] The KOOS is developed with the purpose of evaluating short-term and long-term symptoms and function in subjects with knee injury and osteoarthritis. The questionnaire consists of 42 questions divided into five subdomains: symptoms (7), pain (9), Activities of Daily Living (ADL) (17), function in Sport and Recreation (S&R) (5) and Quality of Life (QoL) (4). All answers are multiple-choice with a 5-Likert scale from 0 to 4. The score is reported as a total score per domain. The KOOS was translated and validated for Dutch patients with osteoarthritis.[12]

IKDC Subjective Knee Form

The IKDC Subjective Knee Form (IKDC-knee) is developed by Irrgang et al.[21] The purpose of this questionnaire is to evaluate the symptoms and limitations caused by knee injuries during daily activities and sports. We divided the 18 questions into two subscales: pain (2) and function (16). Therefore, the score was calculated as a total overall score and a total score per subscale. The total overall score was calculated with the following formula: IKDC-knee total score = ((total rough score – 18) / 87) * 100. The total function score was calculated with the following formula: IKDC-knee total function score = ((total rough function score – 16) / 67) * 100

The total pain score was calculated with the following formula: IKDC-knee total pain score = ((total rough pain score – 2) / 20) * 100. The Dutch version of the IKDC Subjective Knee Form is validated in patients with a variety of knee-related problems.[18]

IKDC Current Health Assessment Form

The International Knee Documentation Committee (IKDC) Current Health Assessment Form (IKDC-health) is the Short-Form 36 (SF-36) and contains 36 questions. It measures health on eight multi-item dimensions: physical function, social function, physical problems, emotional problems, mental health, vitality, pain, general health perception and health change. In this study, the total overall score was used and was calculated with the following formula: IKDC-health = ((total rough score – 35) / 110) * 100. The total score was defined as a QoL score. The percentage of normal score was used, with 0% represents the worst possible score and 100% represents the best possible score. The SF-36 has shown to be reliable and valid in the Dutch general population.[1]

For all total scores the percentage of normal score was used, with 0% represents the worst possible score and 100% represents the best possible score. Missing values were calculated according to the scoring instructions of the questionnaire.

Anchor questions

In this study two different anchor questions were asked. The first anchor question was asked to find out, to what extent the patient could remember their answers on the previous questionnaires. The patients filled in a self-reported 3-Likert scale containing the following answers: I can remember every answer, I can partly remember every answer, I cannot remember any answer. The second anchor question was used to find out if the patient's complaints had been improved or worsened since completing the first questionnaires. We used a self-reported 7-Likert scale containing the following answers: completely recovered, much improved, slightly improved, unchanged, slightly worse, much worse, and worse than ever.

Validity

A questionnaire is valid if it measures the construct it is supposed to measure. Validity is divided into several domains: construct validity, content validity and criterion validity.[27,28] Because the Dutch version of the WOMET is the first meniscal pathology-specific health-related quality-of-life instrument, there is no golden standard for the criterion validity. Therefore, for the validation of the Dutch version of the WOMET only the construct and content validity were determined.

Construct validity

Construct validity refers to the degree to which the scores of a health related questionnaire are consistent with hypotheses based on the assumption that the questionnaire validly measures the construct to be measured.[27,28] The following measures were used to set up the hypotheses: Numeric Rating Scale (NRS) for pain, IKDC-knee subscale pain and function, KOOS subscales symptoms, pain, ADL, S&R and QoL and IKDC-health.

Correlation was classified in: very high correlation (0.90 to 1.00), high correlation (0.70 to 0.90), moderate correlation (0.50 to 0.70), low correlation (0.30 to 0.50) and negligible correlation (0.00 to 0.30).[20]

Nine hypotheses were drawn up. WOMET subscale pain has at least a moderate positive correlation with KOOS and IKDC-knee pain subscales and at least a moderate negative correlation with the NRS for pain. WOMET subscale function has at least a moderate positive correlation with KOOS subscales symptoms, ADL and S&R, and with IKDC-knee subscale function. WOMET subscale QoL has at least a moderate positive correlation with KOOS subscale QoL and at least a low positive correlation with the IKDC-health. The IKDC-knee was specifically designed to assess overall knee problems and the KOOS to measure post traumatic osteoarthritis. Therefore, we predicted that the correlation between the WOMET and the IKDC-knee would be slightly stronger (r = 0.7) compared to the correlation between the WOMET and the KOOS (r = 0.5). Next to that, the IKDC-health was used to measure the overall conditions of the patient. That is why, we predicted a weaker (r = 0.4) correlation between the WOMET and the IKDC-health.

Content validity

Content validity examines the degree to which the content of a health related questionnaire is an adequate reflection of the construct to be measured.[27,28] For validation of the Dutch version of the WOMET, the floor and ceiling effects were calculated to determine the validity of its content. The floor and ceiling effect give insight of the variance in scores, that will not be measured anymore above or below a certain level. If many patients have the minimal or maximal score the question might be less relevant and patients cannot improve or deteriorate over time. Floor (minimal score) and ceiling (maximal score) effects at baseline were evaluated, because they could influence the content validity and respon-

siveness.[40] The floor and ceiling effects at baseline (T0 and T1 together) were determined for the overall WOMET score, for the three subscales of the WOMET and for the sixteen questions separately. The floor and ceiling effect was assessed by calculating the percentage of patients with a minimum or maximum score and was consideredacceptable if less than 30% of the patients had a minimum or maximum score.

Reliability

A questionnaire is reliable when a patient gets the same score on repeated admissions of the measurement. For reliability of the Dutch WOMET test-retest reliability and measurement error was calculated.

Internal Consistency

Internal consistency was determined in 86 patients for the overall WOMET score and for the three domains. Good internal consistency exists when Cronbach's alpha is >0.7.[41]

Test-retest reliability

The test-retest reliability, is the proportion of the total variance in the measurements which is due to 'true' differences between patients over time.[27,28] To asses test-retest reliability, first, patients with no significant change in QoL scores between T0 and T1 were selected. This was tested with the comparison of the KOOS subscale QoL and the IKDC-health. Second, the Intraclass Correlation Coefficient (ICC) was used to determine test-retest reliability in the selected patients. The 95% limits of agreement for the differences for the overall WOMET were determined as well.

Measurement error

Measurement error is the systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured.[27,28] The measurement error was expressed by the Standard Error of the Measurement (SEM).[27] The precision of the WOMET was expressed in Standard Error of Measurement (SEM), which was be calculated by repeated measures in one participant: $SEM = SD_{baseline} * \sqrt{(1 - Cronbach\ \alpha)}$.[25,42] After calculating the SEM, it was used to determine the SDC: SDC = 1.96 x √2 x SEM. For our study, we use the standard deviation (SD) of the WOMET score at T0 and the Cronbach's alpha. The smallest detectable change (SDC) represents the within-person change due to real change in one individual and without the measurement error.[3,13] A low SDC reflects to no real change and represents a high reliability.

Responsiveness

The responsiveness is the ability of a questionnaire to detect changes over time in the construct to be measured.[27,28] Because of lack of a gold standard, the second best op-

tion was to compare changes on the WOMET with changes on other questionnaires or subscales that measure slightly different constructs.[28] This was assessed by testing eight predefined hypotheses about the expected direction and magnitude of the correlation coefficients between the change scores of the questionnaires. To evaluate the responsiveness the changes on the WOMET scores (T0 versus T2) and subscales were compared with the other subscales. The following eight hypothesis, similar to the hypotheses for the construct validity, were drawn up. Changes on the WOMET subscale pain has at least a moderate positive correlation with changes on the KOOS subscale pain and at least a high positive correlation with changes on the IKDC-knee subscale pain . Changes on the WOMET subscale function has at least a moderate positive correlation with changes on the KOOS subscales symptoms, ADL and S&R, and at least a high positive correlation with changes on the IKDC-knee subscale function. Changes on the WOMET subscale QoL has at least a moderate positive with changes on the KOOS subscale QoL and at least a moderate positive correlation with changes on the IKDC-health. We considered the responsiveness of the WOMET to be good if at least 75% of the hypotheses were confirmed.[40]

Interpretability

Interpretability is the degree to which one can assign qualitative meaning to an instrument's quantitative scores or change in scores.[28] According to the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) guidelines it is important to determine the minimal important change (MIC).[28] The MIC is defined as the smallest measured change score that patients perceive to be important. An instrument is useful if the SDC is smaller than the MIC.[15] The change scores on the questionnaires were calculated by subtracting each patient's T2 score from the T0 (baseline) score, and were then used to determine MIC using an anchor-based mean change score technique.[9] The anchor scores were used to categorize patients into seven subgroups, varying from completely recovered to worse than ever. Change scores were calculated in each of the seven subgroups. The MIC was defined as the mean change score in the subcategory of patients who were "slightly improved" according to the anchor scores.[14]

Statistical Analysis

All statistical analysis were performed with the use of SPSS (version 22.0). The questionnaire scores at T0, T1 and T2 were checked for normality with the one-sample Kolmogorov-Smirnov test. The hypotheses of the construct validity were tested at T0, T1 and T2 with the Pearson's correlation coefficient. The paired sample t-test (for normal distribution) or the Wilcoxon signed-rank test (for ordinal distribution) was used for the comparison of the KOOS subscale QoL and the IKDC-health between T0 and T1. After that, the test-retest reliability was calculated with the ICC using a two way random model. The Pearson's correlation coefficient was used to make a comparison between the change scores on the

WOMET subscale and the KOOS subscale, the IKDC-health and the IKDC-knee subscale. All reported p values were two-tailed with an α of 0.05 indicating significance.

RESULTS

Forward and backward translation of the WOMET revealed no problems or language difficulties. A total of 296 consecutive patients with meniscal pathology were eligible for this study. After first exclusion, 152 patients started to fill in questionnaires. A total of 86 patients completed all questionnaires at the three time moments. Reasons for exclusion are shown in Figure 1. The demographic data of all 86 patients included in the study is noted in Table 1. Median time (interquartile range (IQR)) of the first time interval (T0-T1) and second time interval (T0-T2) was 16 (13 – 22) days and 105 (91 - 209) days, respectively. The mean scores of the questionnaires at T0, T1 and T2 are noted in Table 2. Just one of the patients (1.2%) could remember every answer to the previous questionnaires on T1, while two patients (2.3%) did on T2. Most of the patients partly remembered their answers to previous questionnaires on T1 and T2, 79.8% and 58.1%, respectively. The rest of the patients could not remember any of their previous given answers.

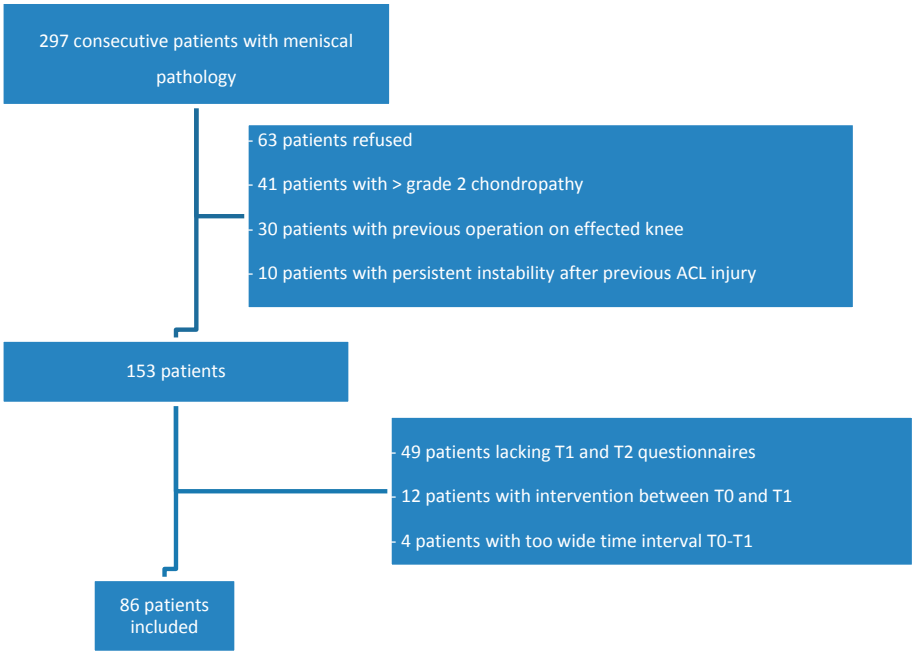

Figure 1. Flow chart of inclusion.

Table 2. Mean and SD of questionnaires at T0, T1, T2

	T0	T1	ICC	T2
WOMET				
Total score	40.8 (15.8)	43.6 (17.7)	0.78	63.0 (23.8)
Function	46.4 (17.4)	48.6 (18.5)	0.73	67.7 (22.5)
Pain	32.1 (18.5)	34.9 (19.6)	0.77	59.8 (27. 6)
QoL	34.8 (19.1)	38.1 (22.5)	0.75	55.3 (29.3)
KOOS	T0	T1		T2
Symptoms	60.5 (16.6)	62.0 (19.1)	0.72	75.0 (17.3)
Pain	48.9 (19.0)	52.9 (20.1)	0.82	71.8 (21.1)
ADL	57.1 (20.3)	60.1 (21.4)	0.81	76.8 (22.2)
S&R	23.4 (20.6)	28.5 (22.2)	0.73	48.8 (30.3)
QoL	35.7 (15.9)	37.5 (16.3)	0.64	54.3 (23.6)
IKDC-health	T0	T1		T2
Total score	62.6 (14.9)	63.1 (15.4)	0.85	70.3 (16.8)
IKDC-knee	T0	T1		T2
Total score	42.6 (13.8)	45.8 (15.3)	0.75	61.5 (19.5)
Function	46.0 (14.1)	44.8 (14.0)	0.71	61.7 (18. 6)
Pain	31.5 (18.8)	38.1 (21.2)	0.78	60.3 (26.9)

ICC = intraclass correlation coefficient. QoL = quality of life, ADL = activity in daily living, S&R = sports and recreation.

Validity

Construct validity

The WOMET subscales shows good correlation with the NRS pain score, IKDC Subjective Knee Form subscale pain and function, all the KOOS subscales, and with the IKDC Current Health Assessment (Table 3). As predicted, the correlation between the WOMET subscales and the IKDC Subjective Knee Form subscales were generally stronger compared to the correlation between the WOMET subscales and the KOOS subscales. Next to that, the weakest correlation was between the WOMET and the IKDC Current Health Assessment. All hypotheses were confirmed.

Content validity

For the total WOMET score there was no floor and ceiling effect. WOMET subscale pain, function and quality of life showed acceptable floor and ceiling effect as well. Next to that, every question of the WOMET was analysed for floor and ceiling effects. An acceptable but high floor effect was found for question four, about 'numbness', and for question seven, about 'swelling', 25.6% and 27.9% respectively.

Table 3. Content validity. Correlations between WOMET subscales (pain, function and quality of life) and IKDC-knee, IKDC-health and KOOS subscales at T0, T1 and T2.

	WOMET Pain		
	T0	T1	T2
NRS			
Pain	-0.61*	-0.72*	-0.83*
IKDC knee			
Pain	0.66*	0.78*	0.86*
KOOS			
Pain	0.52*	0.73*	0.81*
	WOMET Function		
IKDC-knee	T0	T1	T2
Function	0.68*	0.78*	0.77*
KOOS	T0	T1	T2
Symptoms	0.69*	0.78*	0.80*
ADL	0.62*	0.75*	0.83*
S&R	0.59*	0.61*	0.71*
	WOMET QoL		
KOOS	T0	T1	T2
QoL	0.64*	0.70*	0.79*
IKDC-health	T0	T1	T2
	0.50*	0.51*	0.66*

ADL = activity in daily living, S&R = sports and recreation, QoL = quality of life.
*: significant (<0.05)

Reliability

Internal consistency

Internal consistency of the overall WOMET score was good (Cronbach's alpha = 0.88).

Test-retest reliability

There was no significant difference between the KOOS subscale QoL and IKDC Current Health Assessment at T0 and T1. Therefore, all patients were used to measure the test-retest reliability. The test-retest reliability of the overall WOMET score and the three domains were all found to be good (Table 2). The 95% limits of agreement for the differences for the overall WOMET score are shown in Figure 2.

Measurement error

SEM and SDC for the overall WOMET score was 5.5 and 15.4 respectively.

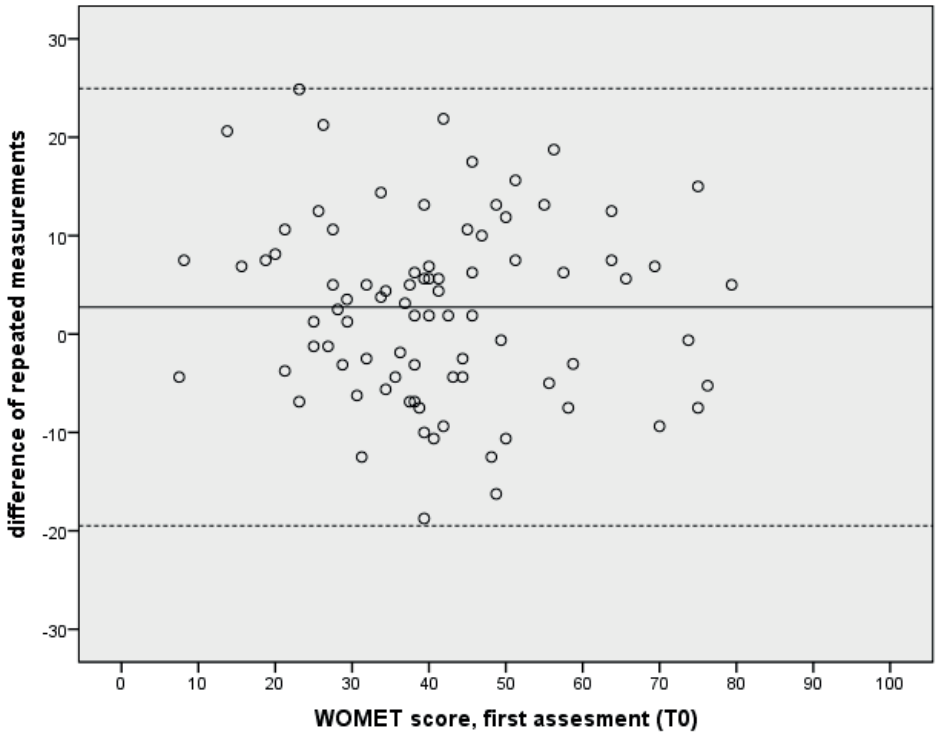

Figure 2. Scatter plot illustrating test-retest reliability, difference in overall WOMET score for each patient (n=85) between T0 and T1.

Responsiveness

The results of the responsiveness analyses is shown in Table 4. The changes on the WOMET subscales showed good correlation with the IKDC Subjective Knee Form subscales pain and function, all KOOS subscales, and with the IKDC Current Health Assessment (Table 4). Seven out of eight hypotheses (88%) were confirmed.

Interpretability

The MIC for the overall WOMET score was 14.7 (Table 5).

DISCUSSION

This study showed that the Dutch version of the WOMET has good construct validity, content validity, test-retest reliability and responsiveness for measuring meniscal pathology-specific health-related quality of life in patients with traumatic or degenerative meniscal tears treated conservatively or operatively.

Table 4. Responsiveness. Mean changes in subscales of WOMET, IKDC-knee, KOOS and IKDC-health between T0 and T2. Correlation between changes on the WOMET subscales and the subscales of KOOS, IKDC-knee and the IKDC-health.

	Mean change (SD)	Correlation (Predefined correlation)
WOMET pain	27.7 (28.1)	
IKDC-knee pain*	28.8 (25.0)	0.78 (0.7)
KOOS Pain*	22.8 (19.4)	0.62 (0.5)
WOMET function	21.3 (22.5)	
IKDC-knee function#	15.6 (17.4)	0.66 (0.7)
KOOS Symptoms#	14.6 (17.3)	0.65 (0.5)
KOOS ADL#	19.7 (19.6)	0.66 (0.5)
KOOS S&R#	25.6 (26.3)	0.64 (0.5)
WOMET QoL	20.5 (25.7)	
KOOS QoL$	18.6 (23.7)	0.80 (0.5)
IKDC-health$	7.7 (11.9)	0.51 (0.4)
Hypotheses confirmed		7/8 (88%)

ADL = activity in daily living, S&R = sports and recreation, QoL = quality of life.
* = correlation with WOMET pain, # = correlation with WOMET function, $ =correlation with WOMET QoL.

Table 5. Interpretability. Mean change score for overall WOMET according to the anchor question on improvement. The 'slightly improved' group was used to determine the minimal important change.

	n	Δ overall WOMET (SD)
Completely improved	11	46.5 (17.3)
Much improved	32	33.7 (14.9)
Slightly improved	**22**	**14.7 (15.2)**
Unchanged	10	4.9 (12.8)
Slightly worse	8	-0.5 (14.3)
Much worse	2	-32.8 (19.0)
Worse than ever	0	N/A

Δ = mean change, SD = standard deviation. N/A = not applicable.

Sihvonen et al.[36] determined criterion validity using the Lysholm knee score. In our opinion Lysholm knee score is not suitable to use as a golden standard. Lysholm knee score was validated for patients with meniscal injury, but determination of quality of life deficits was not assessed.[5] So, in absence of a gold standard for evaluation meniscal pathology-specific health-related quality of life, construct validity was determined. Similar to the findings of Kirkley et al.[23] and Sihvonen et al.[36] we found good construct validity. All WOMET subscales showed good correlation with the subscales of the other questionnaires leading to confirmation of all our hypotheses. As we predicted the weakest correlation was found with IKDC-health, this is comparable to other WOMET validation studies.[6,36]

This confirms that the IKDC-Health measures additional aspects of the physical health and provides more comprehensive, but less specific, information about the patients' overall health compared to condition-specific questionnaires.[22]

The overall WOMET score, the three subscales of the WOMET score and all the sixteen questions separately had acceptable floor and ceiling effects (<30%). However, for the questions about 'numbness' and 'swelling' quite high floor effects were found. This was similar to the findings published by Sihvonen et al.[36] and Celik et al.[6] We agree with Celik et al, that 'numbness' is one of the rare symptoms of meniscal pathology, which can be an explanation for the high floor effect of this item. Compared to 'numbness', 'swelling' is a less rare symptom, but more often found in patients with isolated meniscal tears in combination with osteoarthritis.[36] Our study population consisted of patients with isolated meniscal tears without radiological signs of osteoarthritis, which may be a reason for the high floor effect of the question about 'swelling'. Taken together, these findings suggests that questions about 'numbness' and 'swelling' are less relevant for evaluation of patients with isolated, traumatic or degenerative meniscal tears without osteoarthritis.

In determining reliability, we found a high ICC (0.78), which equals the original study[23] and was comparable to the Turkish and Finnish validation studies.[6,36] In addition we determined measurement error in terms of SEM and SDC. This SDC means that if you want to determine a treatment effect, you need to find a difference of at least 15.4 points in an individual patient to make sure that the difference is not due to random error. We also found acceptable internal consistency (Cronbach's alfa = 0.88) for the overall WOMET score, which was similar to previous studies (α= 0.92 for the original WOMET, α=0.91 for the Finnish WOMET and α= 0.89 for the Turkish WOMET).[6,23,36] This means that all items of the WOMET reflect the same phenomenon.

As mentioned earlier, because of lack of a gold standard, the second best option to define responsiveness was to compare changes on the WOMET with changes on other questionnaires or subscales that measure slightly different constructs.[28] Our good responsiveness could not be compared with previous studies, in which responsiveness was only expressed by a calculated standardized response mean.[23,26]

To the best of our knowledge, there is no previous data on MIC for the WOMET score. We are the first who determined the MIC of the WOMET, which is a strength of this study. A score increased with 14.7 points on the WOMET score was considered clinically relevant. An instrument is more useful if the SDC is smaller than the MIC.[15] In our study, the SDC that was slightly larger than the MIC. This means that if an individual patient has a change score as large as the MIC, we cannot be 95% sure that this change is not due to measurement error. However, as the differences between the SDC and the MIC were rather small, we think that the WOMET is suitable for use in clinical practice and research.

Another strength of this study was that the participants were representative of patients with meniscal injury. Our study population consisted of young and old patients with differ-

ent meniscal injury and meniscal treatment; acute and chronic meniscal injury, traumatic and degenerative meniscal injury, treated operatively (partial meniscectomy and meniscal repair) or conservatively.

Compared to the original WOMET instead of VAS, NRS was used on recommendation of the focus group patients. For our patient population the NRS appeared to be easier to understand and answer, shorter to complete and preferable above VAS, also reported previously.[2] Another distinction made compared to the original WOMET was the distribution of questions in the different domains or subscales. We divided the questions in the subscales: pain, function and quality of life, according to the International Classification of Functioning, Disability and Health (ICF).[44]

There are some limitations of this study. Firstly, time-interval between T0 and T1 was relatively long. Test-retest time-interval of two weeks is considered appropriate for the evaluation of PROMs instruments.[38] Secondly, we had to exclude 67 patients because of a too wide time interval, incomplete amount of questionnaires or an intervention which was started in the first time interval. A more strict control on returning questionnaires would probably had increased patient inclusion, at least to have more data to analyse test-retest reliability. Thirdly, defining hypotheses remains arbitrary and there is no consensus about the number of hypotheses which should be confirmed.

CONCLUSION

The Dutch version of the WOMET seems valid and reliable. It can be used as a disease specific tool to evaluate health related quality of life of Dutch patients with meniscal pathology.

ACKNOWLEDGEMENT

This study was funded by a grant from the Research Fund of Medical Center Haaglanden (Wetenschaps-beurs 2014).

Appendix: Nederlandse WOMET vragenlijst (Dutch version of the WOMET questionnaire)

REFERENCES

1. Aaronson NK, Muller M, Cohen PD, Essink-Bot ML, Fekkes M, Sanderman R, Sprangers MA, te Velde A, Verrips E (1998) Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. J Clin Epidemiol 51:1055-1068.
2. Akad K, Solmaz D, Sari I, Onen F, Akkoc N, Akar S (2013) Performance of response scales of activity and functional measures of ankylosing spondylitis: numerical rating scale versus visual analog scale. Rheumatol Int 33:2617-2623.
3. Beckerman H, Roebroeck ME, Lankhorst GJ, Becher JG, Bezemer PD, Verbeek AL (2001) Smallest real difference, a link between reproducibility and responsiveness. Qual Life Res 10:571-578
4. Black N (2013) Patient reported outcome measures could help transform healthcare. BMJ 346:f167.
5. Briggs KK, Kocher MS, Rodkey WG, Steadman JR (2006) Reliability, validity, and responsiveness of the Lysholm knee score and Tegner activity scale for patients with meniscal injury of the knee. J Bone Jt Surg 88A:698-705.
6. Celik D, Demirel M, Kuş G, Erdil M, Razak Özdiçler A (2015) Translation, cross-cultural adaptation, reliability and validity of the Turkish version of the Western Ontario Meniscal Evaluation Tool (WOMET). Knee Surg Sports Traumatol Arthrosc 23:816-825.
7. Centraal Bureau voor de Statistiek (2014) In: Operaties in het ziekenhuis; soort opname, leeftijd en geslacht, 1995-2010, Available via DIALOG. http://statline.cbs.nl/StatWeb/publication/?VW=T&DM=SLNL&PA=80386NED&LA=NL Accessed 5 Feb 2014 (in Dutch).
8. Crawford K, Briggs, KK, Rodkey WG, Steadman JR (2007) Reliability, validity, and responsiveness of the IKDC score for meniscal injuries of the knee. Arthroscopy 23:839-844.
9. Crosby RD, Kolotkin RL, Williams GR (2003) Defining clinically meaningful change in health-related quality of life. J Clin Epidemiol 56:395–407.
10. Cullen KA, Hal, MJ, Golosinskiy A (2009) Ambulatory surgery in the United States. Natl Health Stat Report 11:1-25.
11. Dawson J, Fitzpatrick R, Murray D, Carr A (1998) Questionnaire on the perceptions of patients about total knee replacement. J Bone Jt Surg 80B:63-69.
12. De Groot IB, Favejee MM, Reijman M, Verhaar JA, Terwee CB (2008) The Dutch version of the Knee Injury and Osteoarthritis Outcome Score: a validation study. Health Qual Life Outcomes. Doi:10.1186/1477-7525-6-16.
13. De Vet HC, Bouter LM, Bezemer PD, Beurskens AJ (2001) Reproducibility and responsiveness of evaluative outcome measures. Theoretical considerations illustrated by an empirical example. Int J Technol Assess Health Care 17:479-487.
14. De Vet HC, Terwee CB, Mokkink LB, Knol DL (2011) Measurement in Medicine. Cambridge University Press, Cambridge.
15. Eliasziw M, Young SL, Woodbury MG, Fryday-Field K (1994) Statistical methodology for the concurrent assessment of interrater and intrarater reliability: using goniometric measurements as an example. Phys Ther 74:777-788.
16. Englund M, Guermazi A. Gale D, Hunter DJ, Aliabadi P, Clancy M, Felson DT (2008) Incidental meniscal findings on knee MRI in middle-aged and elderly persons. N Engl J Med 359:1108-1115.
17. Guillemin F (1995) Cross-cultural adaptation and validation of health status measures. Scand J Rheumatol 24:61-63.
18. Haverkamp D, Sierevelt IN, Breugem SJ, Lohuis K, Blankevoort L, van Dijk CN (2006) Translation and validation of the Dutch version of the International Knee Documentation Committee Subjective Knee Form. Am J Sports Med 34:1680-1684.

19. Hefti F, Müller W, Jacob R, Stäubli HU (1993) Evaluation of knee ligament injuries with the IKDC form. Arthrosc 1:226-234.
20. Hinkle DE, Wiersma W, Jurs SG (2003) Applied Statistics for the Behavioral Sciences, 5th edn. Houghton Mifflin, Boston.
21. Irrgang JJ, Anderson AF, Boland AL (2001) Development and validation of the International Knee Documentation Committee Subjective Knee Form. Am J Sports Med 29:600-613.
22. Irrgang JJ, Marx RG (2007) Clinical outcomes in sport and exercise physical therapies. In: Kolt GS, Synder-Mackler L (eds) Physical therapiues in sports and exercise, 2nd edn. Elsevier, Edinburgh.
23. Kirkley A, Griffin S, Whelan D (2007) The development and validation of a quality of life-measurement tool for patients with meniscal pathology: The Western Ontario Meniscal Evaluation Tool (WOMET). Clin J Sport Med 17:349-356.
24. Lysholm J, Gillquist J (1982) Evaluation of knee ligament surgery results with special emphasis on use of a scoring scale. Am J Sports Med 10:150-154.
25. McHorney CA, Tarlov AR (1995) Individual-patient monitoring in clinical practice: are available health status surveys adequate? Qual Life Res 4:293-307.
26. Meredith DS, Losina E, Mohamed NN, Wright J, Katz JN (2005) Factors predicting functional and radiographic outcomes after arthroscopic partial meniscectomy: a review of the literature. Arthrosc 21:211-223.
27. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL (2006) Protocol of the COSMIN study: Consensus-based Standards for the selection of health Measurement Instruments. BCM Med Res Methodol. Doi:10.1186/1471-2288-6-2.
28. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HC (2010) The COSMIN study reached international consensus on taxonomy terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 63:737-745.
29. Noyes FR (1987) The Noyes knee rating system. Cincinnati Sportsmedicine Research and Education Foundation. Cincinnati.
30. Outerbridge RE (1961) The etiology of chondromalacia patellae. J Bone Jt Surg 43B:752-757.
31. Poehling GG, Ruch DS, Chabon SJ (1990) The landscape of meniscal injuries. Clin Sports Med 9:539-549.
32. Rodkey WG, Stone KR, Steadman JR (1993) Replacement of the irreparably injured meniscus. Sports Med Arthrosc 1:168-176.
33. Roos EM, Lohmander LS (2003) Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. Health Qual Life Outcomes. Doi:10.1186/1477-7525-1-64.
34. Roos EM, Roos HP, Ryd L, Lohmander LS (2000) Substantial disability 3 months after arthroscopic partial meniscectomy: A prospective study of patient-relevant outcomes. Arthrosc.16:619-626.
35. Salavati M, Mazaheri M, Negahban H, Sohani SM, Ebrahimian MR, Abrahimi I, Kazemnejad A, Salavati M (2008) Validation of a Persian-version of Knee Injury and Osteoarthritis Outcome Score (KOOS) in Iranians with knee injuries. Osteoarthr Cartil 16:1178-1182.
36. Sihvonen R, Järvelä T, Aho H, Järvinen TL (2012) Validation of the Western Ontario Meniscal Evaluation Tool (WOMET) for patients with a degenerative meniscal tear: a meniscal pathology-specific quality-of-life index. J Bone Jt Surg Am. Doi:10.2106/JBJS.K.00804.
37. Smillie IS (1968) The current pattern of the pathology of meniscus tears. Proc R Soc Med 61:44-45.
38. Streiner DL, Norman GR (2008) Health measurement scales. A practical guide to their development and use, 4 edn. Oxford University Press, New York.

39. Tanner SM, Dainty KN, Marx RG, Kirkley A (2007) Knee-specific quality-of-life instruments: which ones measure symptoms and disabilities most important to patients? Am J Sports Med 35:1450-1458.
40. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC (2007) Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 60:34-42.
41. Terwee CB, Dekker FW, Wiersinga WM, Prummel MF, Bossuyt PM (2003) On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. Qual Life Res 12:349-362.
42. Wells G, Beaton D, Shea B, Boers M, Simon L, Strand V, Brooks P, Tugwell P (2001) Minimal clinically important differences: review of methods. J Rheumatol 28:406-412.
43. World Health Organization (1980) International classification of impairments, disabilities, and handicaps. Geneva.
44. World Health Organization (2013) How to use the ICF: A practical manual for using the International Classification of Functioning, Disability and Health (ICF). Geneva.

APPENDIX

WOMET vragenlijst

Instructie:

De vragenlijst bestaat uit 16 vragen. De vragen gaan over de gevolgen van meniscusproblemen en hebben betrekking op fysieke problemen, emotie en het algemeen dagelijks functioneren. U kunt bij iedere vraag uw antwoord weergeven op een schaal van 0 tot 10. Hierbij geeft u aan geen last te hebben bij 0, en heel erg veel last te hebben bij 10.

1. Hoeveel last heeft u van het gevoel dat u door uw knie zakt of dat uw knie instabiel is?

	0	1	2	3	4	5	6	7	8	9	10	
totaal geen	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	heel erg veel

2. Hoeveel last heeft u van pijn of irritatie in uw knie na activiteit?

	0	1	2	3	4	5	6	7	8	9	10	
totaal geen	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	heel erg veel

3. Hoeveel last heeft u van het verlies aan beweeglijkheid van uw knie?

	0	1	2	3	4	5	6	7	8	9	10	
totaal geen	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	heel erg veel

4. Hoeveel last heeft u van een verminderd gevoel in of rondom uw knie?

	0	1	2	3	4	5	6	7	8	9	10	
totaal geen	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	heel erg veel

5. Hoeveel last heeft u van stijfheid van uw knie als u ‘s morgens opstaat of als u opstaat nadat u lang gezeten heeft?

	0	1	2	3	4	5	6	7	8	9	10	
totaal geen	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	heel erg veel

6. Hoeveel last heeft u van zwakte in uw knie?

	0	1	2	3	4	5	6	7	8	9	10	
totaal geen	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	heel erg veel

7. Hoeveel last heeft u van zwelling van uw knie?

	0	1	2	3	4	5	6	7	8	9	10	
totaal geen	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	heel erg veel

8. Hoeveel last heeft u van pijnscheuten in uw knie nadat u deze heeft belast?

	0	1	2	3	4	5	6	7	8	9	10	
totaal geen	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	heel erg veel

9. Hoeveel last heeft u van kraken, knakken of het gevoel iets te voelen wegschieten in uw knie?

	0	1	2	3	4	5	6	7	8	9	10	
totaal geen	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	heel erg veel

10. Hoe bang bent u uw knie weer te blesseren als u opnieuw gaat sporten of werken?

	0	1	2	3	4	5	6	7	8	9	10	
totaal niet	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	heel erg

11. Hoeveel wordt u beperkt in uw huidige activiteiten ten opzichte van de activiteiten van voor uw blessure?

	0	1	2	3	4	5	6	7	8	9	10	
totaal niet	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	heel erg

12. In hoeverre kunt u door uw knie minder goed uw sport beoefenen en/of uw werk doen? (als ze allebei slechter gaan, scoor de slechtste van de twee)

	0	1	2	3	4	5	6	7	8	9	10	
totaal niet	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	heel erg

13. Hoeveel last heeft u met hurken?

	0	1	2	3	4	5	6	7	8	9	10	
totaal niet	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	heel erg

14. Hoe vaak denkt u aan uw knie?

	0	1	2	3	4	5	6	7	8	9	10	
totaal niet	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	altijd

15. Hoe bezorgd bent u over hoe het verder zal gaan met uw knie?

	0	1	2	3	4	5	6	7	8	9	10	
totaal niet	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	heel erg

16. In hoeverre voelt u zich gefrustreerd of ontmoedigd vanwege uw knie?

	0	1	2	3	4	5	6	7	8	9	10	
totaal niet	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	heel erg