Reliability of the Dutch translation of the Kujala Patellofemoral Score Questionnaire

Ummels PEJ [1,2,3], Lenssen AF [1,4], Barendrecht M [3], Beurskens AJHM [5,6]

[1]Zuyd University of Applied Sciences, Faculty of Health, Department of Physiotherapy, Heerlen, The Netherlands
[2]Paramedic Centre Sittard Zuid, Sittard, The Netherlands
[3]Avans Plus Improving Professionals, Breda, The Netherlands
[4]Maastricht University Medical Center, Maastricht, The Netherlands
[5]Zuyd University of Applied Sciences, Faculty of Health, Research Centre Autonomy and Participation of patients with a chronic illness, Heerlen, The Netherlands
[6]Maastricht University, Caphri School, Maastricht, The Netherlands

Objective: There are no Dutch-language disease-specific questionnaires for patients with patellofemoral pain syndrome (PFPS) available that could help Dutch physiotherapists to assess and monitor these symptoms and functional limitations. The aim of this study was to translate the original disease-specific Kujala Patellofemoral Score (KPS) into Dutch and evaluate its reliability.

Methods: The questionnaire was translated from English into Dutch in accordance with internationally recommended guidelines. Reliability was determined in 50 stable subjects with an interval of one week. The patient inclusion criteria were age between 14 and 60 years, and knowledge of the Dutch language. The presence of at least three of the following symptoms: pain while taking the stairs, pain when squatting, pain when running, pain when cycling, pain when sitting with knees flexed for a prolonged period, grinding of the patella and a positive clinical patella test. The internal consistency, test-retest reliability, measurement error, and Limits of Agreement (LoA) were calculated.

Results: Internal consistency was 0.78 for the first assessment and 0.80 for the second assessment. The intraclass correlation coefficient ($ICC_{agreement}$) between the first and second assessments was 0.98 (95% Confidence interval (CI) 0.96-0.98). The mean difference between the first and second measurements was 0.64. The Bland and Altman plot showed that the LoA were -10.37 and 11.65. The Standard error measurement (SEM) was 0.78, and the smallest detectable change (SDC) was ±11.01.
Conclusions: The results of the present study indicated that the translated Dutch version of the KPS questionnaire is equivalent to its original English-language version, has good internal consistency and test-retest reliability Trial registration NTR (TC = 3258).

Key Words: *Anterior Knee Pain Scale (AKPS), Kujala Patellofemoral Score, Patellofemoral pain syndrome (PFPS), questionnaire, test-retest reliability*

INTRODUCTION

Patellofemoral pain syndrome (PFPS) is one of the most common knee problems in clinical settings. A retrospective case-control study of 2002 running injuries showed that PFPS is the most common overuse injury which was seen in 331 (16.5%) of the patients.[1] In total, 50% of the athletes reported an association between their knee complaint and their sport activity.[2] In female athletes, Myer et al found a point prevalence of PFPS of 16.3 per 100. The cumulative incidence risk rate for the development of new unilateral PFPS was 9.66 per 100 athletes pro season and 1.09 per 1000 athletic exposures.[3]

PFPS is defined as pain around the patella that occurs during or after high-loaded flexion and extension of the knee. The predominant symptom is pain, and the condition generally progresses to the point of functional impairment. Activities that are painful with PFPS are ascending and descending stairs, squatting, running, cycling, and sitting with knees flexed for a prolonged period. Symptoms usually start during adolescence when participation in sports is high and can occur over a prolonged period.[4]

There is no agreement concerning the pathology and etiology of PFPS. Hewitt et al reported that altered or decreased neuromuscular control during the execution of sports movements, which result in excessive resultant out of plane knee joint motion and load, appear to increase acute knee injury in female athletes and may contribute to the development of PFPS.[5] Clinicians suspect that PFPS results from abnormal patella tracking that leads to excessive compressive stress to the patellar facets. Factors that may contribute to abnormal patella tracking include quadriceps weakness, quadriceps muscle imbalances, excessive knee soft tissue tightness, an increased quadriceps angle (Q angle), hip weakness, and altered foot kinematics.[6] Based on this clinical theory, the goals of treatment interventions for PFPS is to reduce pain, to improve patella tracking, and to reduce abnormal stress to the patellofemoral joint structures.[6,7,8,]

Extensive diagnostic investigations do not reveal specific pathology. Cook et al [9] investigated the diagnostic accuracy in patients with anterior knee pain and the association with disability of selected functional findings or the association with physical examination tests for PFPS. They found that the most accurate single diagnostic test is pain encountered during resisted muscle contraction of the knee: positive predictive value (PPV) = 82%, positive likelihood ratio (LR+) = 2.2, 95% confidence interval (CI), 0.99–5.2. Clusters of test findings were substantially more accurate with any combination of two or three positive findings of pain encountered during resisted muscle contraction, pain during squatting, and pain during palpation yielding the following values: PPV = 89%, LR+ = 4.0, 95% CI: 1.8–10.3.

The diagnosis of PFPS is typically made based on the presence of anterior or retropatellar knee pain associated with prolonged sitting or with weight-bearing activities that load the patellofemoral joint such as squatting, kneeling, running, and ascending or descending stairs. Because no gold standard for a diagnosis of PFPS currently exists, inclusion of the step-down test may increase the likelihood of a diagnosis of PFPS from 40% to 65% (LR+ = 2.34).[10,11] Nijs et al[11] examined the validity of the following five clinical patellofemoral tests used in the diagnosis of PFPS: the vastus medialis coordination test, Patellar Apprehension test, Waldron's test, Clarke's test, and the Eccentric Step test. The LR+ was 2.26 for both the vastus medialis coordination test and the Patellar Apprehension test. For the Eccentric Step test, the LR+ was 2.34. A positive outcome on the vastus medialis coordination test, the Patellar Apprehension test, or the Eccentric Step test increases the probability of PFPS to a small, but somewhat significant degree.

Valid and reliable questionnaires are essential to evaluate the effectiveness of interventions and monitor the treatments in daily practice. Various measurement instruments for specific symptoms of the knee have been developed such as the Lysholm Knee Questionnaire[12], and the VISA-P questionnaire for patellar tendinopathy.[13] Neither of these instruments, however, specifically focuses on patellofemoral disorders and activities that are painful with PFPS. Only a few

instruments focus specifically on PFPS: the Patellofemoral Rating Scale (PRS) developed by Fulkerson et al[14], and the Kujala Patellofemoral Score (KPS).[15]

The PRS[14] and the KPS measure a similar concept but there are some differences. The major difference between the PRS and the KPS is the weight of the separate items of the score. The KPS gives a maximum of 5 points for six of its items and 10 points for seven of its items, which hypothetically can result in 100 points for pain. The PRS only gives 45 out of 100 points to pain. Furthermore, the PRS only measures a total of seven activities and symptoms, whereas the KPS measures a total of 13 items and places more emphasis on pain related to activities then does the PRS.

The KPS was developed in 1993 to evaluate subjective symptoms and functional limitations in PFPS. The objective of Kujala et al was to develop a disease-specific scoring questionnaire for anterior knee patients based on the following criteria: some questions should assess anterior knee symptoms specifically, the patient should complete the questionnaire independently to exclude investigator bias which also makes it possible to use the questionnaire in association with outpatient clinics, and the total scores should be easily and quickly calculated [15]. Kujala et al used the Oretrup modification of the Larson scale as a basis of the questionnaire an included new questions, most of which were specifically focused on anterior knee pain symptoms [15,16]. The KPS comprises 13 questions (Appendix 1). These questions inquire about activity related pain while walking up and down stairs, squatting, running, jumping, carrying out weight-bearing activity, and prolonged sitting with the knee in flexion. It also inquires about symptoms: limping, swelling, subluxation of the patella, atrophy in quadriceps muscle, flexion deficiency, and flexion pain. The total score ranges from 0 to 100, a higher score indicates fewer complaints.[8,15]

Several studies in the past have reported on the reliability and concurrent validity of the KPS. In 2004, Crossley et al[17] examined the test-retest reliability, concurrent validity, and responsiveness of the English version of the Anterior Knee Pain Scale (another name for the KPS); In 2005, Watson et al[18] found a good reliability for the KPS and Kuru et al[8] examined the test-retest reliability and the internal consistency of the Turkish version of KPS. All of these studies showed the KPS have good reliability and validity.

Evaluation of patient-reported symptoms and functional limitations would help Dutch physiotherapists better assess and monitor patellofemoral symptoms. There are no Dutch-language disease-specific questionnaires for patients with PFPS. The aim of this study was to translate the original disease-specific KPS from English into Dutch and evaluate its reliability in Dutch patients who have PFPS.

METHODS

After obtaining permission from Kujala et al,[15] the KPS was translated into Dutch (Appendix 2). The translation process was conducted according to the guidelines for the process of cultural adaption of self-report measures from Beaton et al.[19] In this clinimetric study, the reliability was determined in a sample of PFPS subjects with an interval of one week.

Subjects

The study included 65 patients with patellofemoral complaints. From January 2012 through May 2012, the data were collected at 16 physical therapy clinics in the community in the Netherlands and from students with patellofemoral complaints at the faculty physiotherapy and occupational therapy of the Zuyd University of Applied Science in Heerlen, the Netherlands. All clinics and

physiotherapists were familiar with patellofemoral complaints and the use of questionnaires. All subjects completed a signed informed consent form prior to participation in the study. This study was conducted according to the regulations of the Medical Ethical Committee at Atrium Medical Centre, Orbis Medical Centre and Zuyd University in Heerlen, the Netherlands.

The patient inclusion criteria were: age between 14 and 60 years, sufficient knowledge of the Dutch language, grinding of the patella, a positive clinical patella test such as Clarke's test or the patellar femoral grinding test, and the presence of at least three of the following symptoms: pain when ascending or descending stairs, pain when squatting, pain when running, pain when cycling, pain when sitting with knees flexed for a prolonged period. Subjects were excluded if they had patellar tendinopathy, Osgood-Schlatter disease, or other defined pathological conditions of the knee, or if they had previous knee injuries or surgery.[20,21]

Procedures and measurements

Patients who met the selection criteria, and who were willing to participate in the study, completed the questionnaire individually without assistance from the therapist after the first intake or (current) treatment session. To assess the reliability, the same questionnaire was re-administered to the patients by the same therapists after a one-week interval during which patients received therapy as usual.

To evaluate the reliability it was necessary to include a stable patient population. To ensure this, the patients independently scored the Global Perceived Effect (GPE) on a 7-point scale during the second measurement.[22,23] The responses on the 7-point scale are: completely recovered, moderate improved, slightly improved, unchanged, slightly worsened, moderate worsened, and worse than ever. Only patients who reported slightly improved, unchanged and slightly worsened were included in the study.

All participating therapists (n = 21) received a written manual with instructions about inclusion and exclusion criteria, the one-week interval, informed consent and the questionnaires.

Data analysis

Reliability is defined as "the degree to which the measurement is free from measurement error."[24] In addition to the general definition, there is an extended definition: "the extent to which scores for patients who have not changed are the same for repeated measurement under several conditions: for example, using different sets of items from the same multi-item measurement instrument (internal consistency); over time (test-retest); by different persons on the same occasions (inter-rater); or by the same persons (ie raters or responders on different occasions (intra-rater)."[24] The internal consistency, along with test-retest reliability and measurement error, is also an aspect of reliability.[25]

The reliability of the translated KPS was tested by means of internal consistency, test-retest reliability (how well patients can be distinguished from each other), measurement error and Limits of Agreement (LoA). The intraclass correlation coefficient ($ICC_{agreement}$) tests and evaluates the agreement between baseline and retest scores (in the case of the present study, after an interval of one week). The measurement error (the standard deviation of errors of measurement that is associated with the test scores for a specified group of test takers)[26] expressed by SEM and LoA which were calculated by means of a Bland and Altman plot.

Internal consistency is defined by the COSMIN panel as the degree of interrelatedness among items.[24] In a unidimensional scale or subscale of a multi-item instrument, internal consistency is a measure of the extent to which items assess the same construct. Cronbach's alpha is the best

known parameter for assessing the internal consistency of a scale. A well-accepted guideline for good internal consistency of a scale is a Cronbach's alpha value between 0.70 and 0.90.[25]

The ICC is defined as the ratio of the variance between patients to the total variance and is sample-dependent. ICC values can theoretically range from 0 to 1, with a higher value indicating that less variance is due to other factors such as differences between observations or measurements. An intraclass correlation coefficient of at least 0.70 is considered to be satisfactory for group comparisons, and a value between 0.90 and 0.95 is satisfactory for individual comparisons.[25] The ICC was calculated from a two-way random effects model, for absolute agreement.

The measurement error can be adequately expressed by the standard error measurement (SEM) and smallest detectable change (SDC). The SEM is a measure of how far apart the outcomes of repeated measurements are; it is the standard deviation (SD) around a single measurement.[25,27] Several methods can be used to obtain the SEM value. The SEM was calculated from the SD of the differences between two measurements ($SD_{difference}/\sqrt{N}$).

Another additional parameter of measurement error can be found in the LoA, proposed by Bland and Altman[28] . Bland and Altman designed a plot in which not only systematic errors can be seen easily but also, relating the LoA to the range of the scale provides an impression of the magnitude of the measurement error. Measures of agreement refer to the absolute measurement error (presented in the units of measurement of the instrument) that is associated with one measurement taken from an individual patient.[29,30] The mean difference between the two measurements and the SD of this difference were calculated. The magnitude of the SD expresses the extent to which the same value was achieved within two measurements.[31] Subsequently, the 95% LoA were calculated, defined as the mean difference between two measurements ± 1.96 SD of this mean difference ($\mu \pm 1.96 \times SD_{difference}$).[28] By definition, 95% of the differences between repeated measurements fall between the LoA. If a patient's KPS changes outside the borders of the LoA, it is improbable that this is due to measurement error, and it possibly indicates a real change. Therefore, LoA provide information about the SDC.[25] To indicate a real change, the clinical implications for an individual patient is calculated with the SDC, and for groups by the SEM.

All analyses in this study were performed using SPSS program, version 19, (Armonk, New York, United States), except for computing the SEM and SDC, which were performed in Microsoft Excel 2007. Missing values in the KPS were imputed using the mean method per item. When the number of missing values in a questionnaire exceeded 20% the patient was excluded on a list-wise basis.[32]

RESULTS

In total, 65 patients with PFPS were recruited by the participating physiotherapists. Of these individuals, fifteen patients were excluded because their symptoms changed as measured with the GPE 7-point scale. Therefore, data of these 50 patients were used for calculation of reliability parameters. **(TABLE 1)**. There were no significant differences for gender, age, right or left knee duration of symptoms in years or the KPS baseline between the included and excluded patients. The KPS after one week showed significant differences for the excluded group between the KPS at the baseline and after one week.

TABLE 1. Baseline characteristics of the study population

	N = 50 (included)	N = 15 (excluded)	P value
Gender (% male)	16 (32%)	7 (47%)	P = .297 [1]
Age (mean (SD))	27.8 (13.1)	29.1 (17.8)	P = .759 [2]
Right/Left PFPS (% right)	40%	53%	P = .360 [1]
Duration of symptoms in months (median)	23,6	15,8	
IQR (25-50-75 percentile)	7,8 - 23,6 - 63,0	3,9 - 15,8 - 47,2	P = .602 [3]
Kujala Patellofemoral Score baseline (mean (SD)) (min./max. score)	68.5 (15.2) 19/88	65.1 (8.5) 45/74	P = .414 [1]
Kujala Patellofemoral Score after one week (mean (SD)) (min./max. score)	69.1 (13.8) 22/89	80.1 (9.8) 50/91	P = .005 [1*]
GPE (number (%))			
Slightly improved	14 (28%)		
Unchanged	32 (64%)		
Slightly worsened	4 (8%)		

1.Pearson Chi-Square, 2. T-Test, 3. Mann- Whitney U Test, *significant difference (P < .05) difference between excluded and included group Kujala Patellofemoral Score after one week.

Internal consistency was calculated using Cronbach's alpha, the Cronbach's alpha was 0.78 for the first assessment and 0.80 for the second assessment. The $ICC_{agreement}$ between the first and second assessments was 0.98 (95% Confidence interval (CI) 0.96-0.98). The mean difference between the first and second measurements is 0.64 and SD is 5.51. The Bland and Altman plot **(FIGURE 1)** shows that the LoA are -10.37 and 11.65. The SEM is 0.78 and SDC is ±11.01. To indicate a real change, the clinical implications for an individual patient is that the change must be more than ±11.01 points, and for groups it must be more than 0.78 points (on a scale of 100 points).

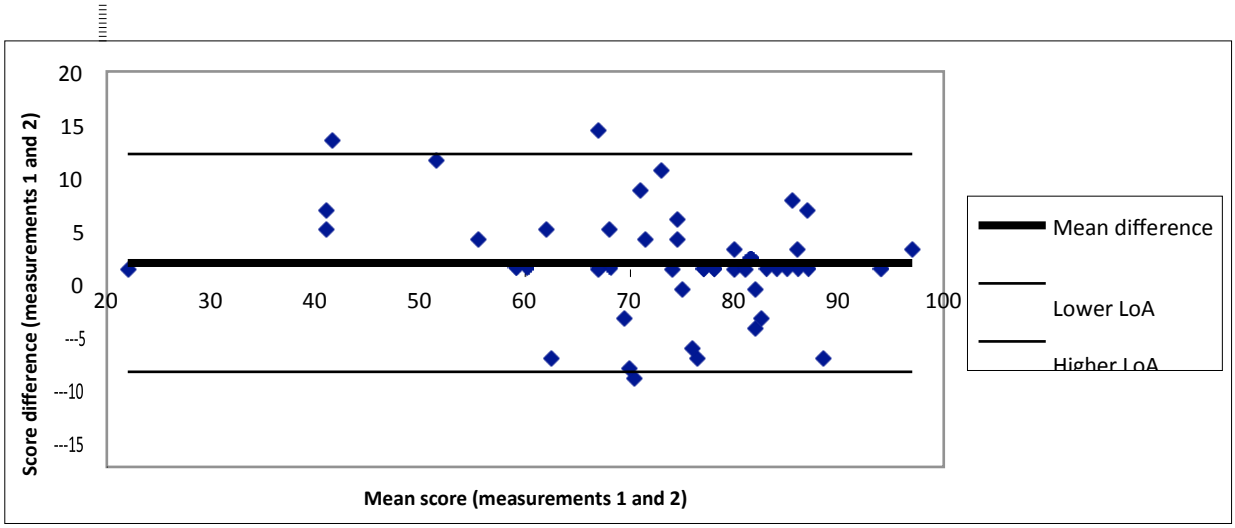

FIGURE 1. Bland and Altman plot. Differences between two measurements, plotted against the mean values for both measurements for each patient. The figure shows the mean difference between measurements (solid line at center) and the LoA (dashed outer lines) corresponding to ±1.96 SDs of the mean difference between the first and second measurements.

DISCUSSION

The aim of this study was to translate the original disease-specific KPS into Dutch and evaluate its reliability in patients in the Netherlands who have PFPS. The test-retest reliability (ICC = 0.98 (95% CI 0.96-0.98)) and the internal consistency (Cronbach's alpha = 0.78 for the first measurement, and 0.80 for the second measurement) show a good reliability of the KPS. The standard error of measurement was calculated using a Bland and Altman plot. The SEM gave a score of 0.78 and the SDC was ±11.01.

The included patients for gender, age and duration of symptoms and the results for the translated version of the KPS are similar to those of international studies on the clinimetric properties of the English and Turkish version of KPS. Crossley et al[17] examined the test-retest reliability, concurrent validity, and responsiveness of the Anterior Knee Pain Scale and found an ICC of 0.81. The minimum clinical difference in this study was calculated using by two methods: median change score and ROC curve, the minimal clinical difference was 10 points. Crossely et al selected the Visual Analog Scale for usual pain (VAS-U) outcome measure as reference measure for determining the relative efficiency of each out- come measure. Watson et al[18] found an ICC of 0.95 for the KPS. In 2010, Kuru et al[8] examined the test-retest reliability and the internal consistency of the KPS in Turkey. The test-retest reliability of the items in that study ranged from r_s 0.61 to 1.000, with a mean correlation coefficient of 0.94. They found an internal consistency of 0.84, calculated with Cronbach's alpha.

As mentioned in the data analysis, an essential requirement of all measurements in clinical practice and research is that they must be reliable. But repeated measurements can display variation arising from several sources: measurement instruments, patients filling out the questionnaire, or circumstances under which the measurements are taken.[25] To ensure these requirements were met in this study we used clear inclusion criteria, a well-defined protocol and the COSMIN checklist (Box A, internal consistency, Box B, reliability and Box C, measurement error).[33]

In the absence of a golden standard, we used the GPE as an external criterion to assess the stability of patients. The question is whether the GPE 7-point scale is a suitable questionnaire to assess stable patients with patellofemoral complaints. Two patients scored significantly better (higher) on KPS (subscales running, pain and swelling), while the score at the GPE 7-point scale indicated only slightly improved. One patient scored "significantly worsened" (lower) on the KPS (subscale running), and the GPE 7-point scale indicated unchanged. The subjects were asked to fill in the GPE 7-point scale independently after one week. There have been many studies published in which various questionnaires are validated with the GPE 7-point scale as construct validation. In that sense, we are able to speak of concurrent validity.[23,34,35,36,37,38] Kamper et al[39] found in a meta-analysis in several population good test-retest reliability scores (ICC = 0.90–0.99) of the GPE. The patients' current status is likely to be strongly influential when patients rate GPE and this tendency increases as the period over which the transition spans increases. These findings, however, do not preclude the use of the GPE. Regarding the use of GPE for clinimetric assessment of other instruments, particular care must be taken if the GPE is to be used as an external criterion of change with which to measure the responsiveness or minimally important change of other assessment instruments.

De Vet et al recommend an interval of two weeks to minimize the effect of recognition.[25] However, to keep the subjects as stable as possible in this study an one-week interval was chosen because patellofemoral symptoms vary strongly and to decrease the risks of excessive compressive stress to the patellar facets.

All subjects were in different treatment phases. They varied from patients who came for the intake to patients who had been treated several times already. All patients received care as usual. The written manual did not contain instructions on which information (for instance about progression of the complaints) the physiotherapist should give during the treatment. Because the KPS also inquires about pain, swelling for example, it is possible that the patients were biased when filling out the questionnaire for the second time. In addition, the subjects showed a wide range of duration of complaints (median 23,6 months; IQR (25,50,75 Percentile) 7,8 - 23,6 - 63,0). Thus, both subjects with acute and chronic patellofemoral complaints were included, resulting in a heterogeneous patient population included in this study.

The implication for clinical practice is that the questionnaire is manageable; the sum-score of the KPS is easy to calculate and subjects need about 5 minutes to complete the 13 questions. Based on the LoA, a change of ±11.01 points (±1.96 × 5.51 (SD)) between the KPS of the first and second measurements is needed to be able to speak of a minimal detectable change on individual level. To investigate whether the score is sufficiently responsive, KPS must be further examined. In addition to the responsiveness, it is also recommended that the diagnostic validity of the Dutch version of the KPS be examined to ensure that all clinimetric properties of the questionnaire have been examined.

CONCLUSION

Considering the absence of disease-specific questionnaires for Dutch patients with PFPS, we investigated the internal consistency, test-retest reliability and the measurement error of the KPS. The results of the present study indicated that the translated Dutch version of the KPS questionnaire is equivalent to its original version, has good test-retest reliability and good internal consistency.

References

1. Taunton, J.E., Ryan, M.B., Clement, D.B., McKenzie, D.C., Lloyd-Smith, D.R., Zumbo, B.D., (2002) A retrospective case-control analysis of 2002 running injuries, Br J Sports Med;36:95-101
2. van Middelkoop, M., van Linschoten, R., Berger, MY., Koes BW., Bierma-Zienstra, SM. (2008) Knee complaints seen in general practice: active sport participants versus non-sport participants. BMC MusculoskeletDisord: 9:36
3. Myer, G.D., Ford, K.R., Barber Foss, K.D., Goodman, A., Ceasar, A., Rauh, M.J., Divine, J.G., Hewett, T.E., (2010) The incidence and potential pathomechanics of patellofemoral pain in female athletes, Clin Biomech (Bristol, Avon); 25(7): 700-707. Doi:10.1016/j.clinbiomech.2010.04.001
4. Linschoten, van, R. Middelkoop van M, Berger, M.Y., et al. (2009) Supervised exercise therapy versus usual care for patellofemoral pain syndrome: an open label randomised controlled trail, BMJ; 339:b4074
5. Hewett, T.E., Myer, G.D., Ford, K.R., Heidt, R.S., Colosimo, A.J., McLean, S.G., van den Bogert, A.J., Paterno, M.V., Succop, P., (2005) Biomechanical measures of neuromusculair control and valgus loading of the knee predict anterior cruciate ligament injury risk in female athletes: a prospective study. Am J Sports Med; 33(4): 492-501
6. Bolgla, L., Boling, C.B., (June 2011) An update for the conservative management of patellofemoral pain syndrome: A systematic review of the literature from 2000 to 2010.The international Journal of sports physical therapy. Volume 6, number 2.
7. Prins, M.R., van der Wurff, P. (2009) Females with patellofemoral pain syndrome have weak hip muscles: a systematic review. Aust J Physiother. 55(1):9-15
8. Tugba Kuru, Elif Elcin Dereli, et.al. (2010) Validity of the Turkisch version of the Kujala patellofemoraal score in patellofemoraal pain syndrome, Acta Orthop Traumatol Turc, 44(2):152-156.
9. Cook, C., Hegedus, E., Hawkins, R., Scovell, F., Wyland, D., (2010) Diagnostic accuracy and association to disability of clinical test findings associated with patellofemoral painsyndroma. Physiother Can; 62:17-24
10. Lowry, D.C., Cleland, J.A., Kelly, D. (2008) Management of Patients with patellofemoral pain syndrome using a multimodal approach: A case Series. Journal of orthopaedic & sports physical therapy, volume 38, number 11.
11. Nijs J., Van Geel C., Van der Auwera C., Van de Velde B., (2006) Diagnostic value of five clinical tests in patellofemoral pain syndrome. Man Ther. ;11:69-77.
12. Lysholm, J., Qillquist, J. (1982) Evaluation of knee ligament surgery results with special emphasis on use of a scoring scale. Am. J Sports Med: 10:150-4.
13. Zwerver, J., Kramer, T., van den Akker-Scheek, I., (2009) Validity and relaibilty of the Dutch translation of the VISA-P questionnaire for pattellar tendinopathy. BMC Musculoskeletal Disorders; 10:102
14. Fulkerson, J.P., Becker, G.J., Meaney, J.A., Miranda, M., Folcik, MA., (1990) Anteromedial tibial tubercle transfer without bone graft. Am J sports Med: 18:490-6.
15. Urho M. Kujala, Laura H. Jaakkola et. al., (1993) Scoring of Patellofemoral Disorders, Arthoscopy: the Journal of Arthoscopy and Related Surgery (2):159-163
16. Oretorp, N., Gillquist, J., Lilijedahl, S-O., (1979) Long term results of surgery for non-acute anteromedial rotatory instability of the knee. Acta Orthop Scand; 50; 329-36
17. Crossley, K.M., Bennell,K.,L., Cowan,S.,M., Green, S. (2004) Analysis of Outcome Measures for Persons With Patellofemoral Pain: Which Are Reliable and Valid? Arch Phys Med Rehabil Vol85.
18. Watson, C. J., Propps, M., Ratner, J., Zeigler, D.L., Horton, P., Smith, S.S. (2005) Reliability and Responsiveness of the Lower Extremity Functional Scale and the Anterior Knee Pain Scale in Patients With Anterior Knee Pain. J Orthop Sports Phys Ther Vo lume 35 Number 3.
19. Beaton, D., Bombardier, C., Guiilemin, F., Ferraz, M. (2000) Guidelines for the Process of Cultural Adaptation of Self-Report Measures, Spine volume 24, Number 24, pp 3186-3191.
20. Linschoten, van R. Middelkoop van M, Berger, M.Y., et al. (2009) Supervised exercise therapy versus usual care for patellofemoral pain syndrome: an open label randomised controlled trail, BMJ; 339:b4074.
21. Lenssen, T.& Wijnen, A., (2003) Fysiotherapeutische behandeling bij anterieure kniepijn, Jaarboek Fysiotherapie, Kinesitherapie, Bohn Stafleu Van Loghem; Houten.
22. Hudak P.L, Wright J.G. (2000) The characteristics of patient satisfaction measures. Spine; 25:3167-3177.

23. Swinkels, R., & Engelen van, E., Global perceived effect. Geraadpleegd op 27 december 2011, via http://www.meetinstrumentenzorg.nl/Portals/0/bestanden/212_1_N.pdf
24. Mokkink, L. B., Terwee, C. B., Patrick, D, L., et.al (2010a). International consensus on taxonomy, terminology, and definitions of measurements properties for health-related patient-reported outcomes: results of the COSMIN study. Journal of clinical epidemiology, 63, 737-45.
25. Vet de, Henrica, C.W., Terwee, Caroline, B., Mokkink, Lidwine, B., Knol, Dirk, L. (2011) Measurement in Medicine, Cambridge university press, Cambrigde, United kingdom.
26. Harvill, M.L., (1991) Standard Error of measurement. Educational Measurement: Issues and Practice
27. Terwee, C.B., Bot, S.D.M., de Boer, M.R., van der Windt, D.A.W.M., Knol, D.L., Dekker, J., Bouter, L.M., de Vet, H.C.W. (2007) Quality criteria are proposed for clinimetric studies of health status questionnaires. Journal of clinical epidemiology; 60: 34-42
28. Bland, J.M., Altman D.G., (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986, 1(8476): 307-310.
29. de Winter A.F, Heemskerk M.A, Terwee C.B, Jans M.P, Deville W, van Schaardenburg D.J, Scholten R.J, Bouter L.M. (2004) Inter-observer reproducibility of measurements of range of *Disord* 2004, 5:18.
30. de Vet H.C., (1998) Observer reliability and agreement. In *Encyclopedia of biostatistics Volume 4*. Edited by: Armitage P, Colton T. Boston , John Wiley & sons; 1998: 3123-3128.
31. Terwee, C.B., de Winter, A.F., Scholten, R.J., Jans, M.P., Deville, W., van estimation of range of motion of the shoulder. *Arch Phys Med Rehabil*, 86(7): 1356-1361.
32. Nunnally, J.C. and Bernstein, I.H., (1994) Psychometric Theory. 3rd edition. New York: McGraw-Hill
33. Terwee, C. B, Cosmin checklist. Geraadpleegd op 18 augustus 2012 via http://www.cosmin.nl/images/upload/File/checklist%20final%20feb%202010.pdf
34. Norman G.R., Stratford P., Regher G., (1997) Methodological problems in the retrospective computation of responsiveness to change: The lesson of Cronbach. J Clin Epidemiol Vol 50(8):869879.
35. Ostelo R.W.J.G., de Vet H.C.W., (2005) Clinically important outcomes in low back pain. Best Pract & Res Clin Rheum; 19(4); 593-607.
36. Hagg O., Fritzell P., Nordwall A., (2003) The clinical importance of changes in outcome scores after treatment for chronic low back pain. Eur Spine J; 12:12-20.
37. Ostelo R.W.J.G., de Vet H.C.W., Knol D.L., van den Brandt P.A., (2004) 24-item Roland-Morris Disability Questionnaire was preferred out of six functional status questionnaires for post-lumbar disc surgery. J Clin Epidemiol 2004;57:268-276.
38. Davidson M., Keating J.L., (2002) A comparison of five low back pain disability questionnaires: reliability and responsiveness. Phys Ther; 82:8-24.
39. Kamper, J.S., Ostelo, R.W.J.G., Knol, D.L., Maher, C.G., Vet de, H.C.W., Hancock, M.J. (2010) Global Perceived Effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. Journal of Clinical Epidemiology 63; 76.0766

Appendix: 1

Kujala Patellofemoral Score

Name: Date:

Age:

Knee: L/R

Duration of symptoms: Years Months

For each question, circle the latest choice (letter), which corresponds to your knee symptoms.

Limp None (5) Slight or periodical (3) Constant (0)	**Prolonged sitting with the knees flexed** No difficulty (10) Pain after exercise (8) Constant pain (6) Pain forces to extend knees temporarily (4) Unable (0)
Support Full support without pain (5) Painful (3) Weight bearing impossible (0)	**Pain** None (10) Slight and occasional (8) Interferes with sleep (6) Occasionally severe (3) Constant and severe (0)
Walking Unlimited (5) More than 2 km (3) 1-2 km (2) Unable (0)	**. Swelling** None (10) After severe exertion (8) After daily activities (6) Every evening (4) Constant (0)
Stairs No difficulty (10) Slight pain when descending (8) Pain both when descending and ascending (5) Unable (0)	**. Abnormal painful kneecap (patellar) movements (subluxations)** None (10) Occasionally in sports activities (6) Occasionally in daily activities (4) At least one documented dislocation (2) More than two dislocations (0)
Squatting No difficulty (5) Repeated squatting painful (4) Painful each time (3) Possible with partial weight bearing (2) Unable (0)	**. Atrophy of thigh** None (5) Slight (3) Severe (0)
Running No difficulty (10) Pain after more than 2 km (8) Slight pain from start (6) Severe pain (3) Unable (0)	**. Flexion deficiency** None (5) Slight (3) Severe (0)
Jumping No difficulty (10) Slight difficulty (7) Constant pain (2) Unable (0)	

Referentie: Urho M. Kujala, Laura H. Jaakkola et. al., (1993), Scoring of Patellofemoral Disorders, Arthoscopy: the Journal of Arthoscopy and Related Surgery (2):159-163

Appendix: 2

Kujala Patellofemoral Score – Dutch translated VersionVertaling: P.E.J. Ummels

Naam: Datum:

Geboortedatum: Geslacht: Man/Vrouw

Knie: L/R

Duur van de klachten: Jaar Maanden Weken

Omcirkel bij elke vraag de keuze (letter) die het beste past bij uw knieklachten.

Mank lopen: Niet (5) Af en toe of een beetje (3) Altijd (0)	**Langdurig zitten met gebogen knieën:** Geen probleem (10) Pijn na langdurig zitten met gebogen knieën (8) Continu pijn (6) Ik moet nu en dan mijn knie strekken vanwege de pijn (4) Niet mogelijk (0)
Belastbaarheid: Staan op één been is niet pijnlijk (5) Staan op één been is pijnlijk (3) Staan op één been is niet mogelijk (0)	**Pijn:** Geen (10) Af en toe een beetje (8) Het hindert bij het slapen (6) Soms hevig (3) Altijd hevig aanwezig (0)

Wandelen:
Onbeperkt (5)
Meer dan 2 km (3)
1-2 km (2)
Niet mogelijk (0)

Traplopen:
Geen probleem (10)
Lichte pijn bij trap aflopen (8)
Zowel trap op als trap aflopen is pijnlijk (5)
Niet mogelijk (0)

Hurken
Geen probleem (5)
Herhaald hurken is pijnlijk (4)
Hurken is iedere keer pijnlijk (3)
Alleen mogelijk indien niet volledig belast (2)
Niet mogelijk (0)

Hardlopen:
Geen probleem (10)
Pijn na meer dan 2 km hardlopen (8)
Lichte pijn vanaf begin hardlopen(6)
Zeer Pijnlijk (3)
Niet mogelijk (0)

Springen:
Geen probleem (10)
Enige moeite (7)
Altijd pijnlijk (2)
Niet mogelijk (0)

Zwelling:
Geen (10)
Na forse in spanning (8)
Na dagelijkse activiteiten (6)
Iedere avond (4)
Altijd (0)

Voelt u uw knieschijf (patella) wel eens pijnlijk wegschieten (dislocatie)?
Nooit (10)
Soms tijdens het sporten (6)
Soms bij dagelijkse activiteiten (4)
1 of 2 vastgestelde dislocaties (2)
Meer dan twee vastgestelde dislocaties (0)

Zijn uw bovenbeenspieren dunner geworden?
Nee (5)
Ja, een beetje (3)
Ja, veel (0)

Kunt u de knie volledig buigen?:
Ja (5)
Een beetje beperkt (3)
Heel erg beperkt (0)

Referentie: Urho M. Kujala, Laura H. Jaakkola et. al., (1993), Scoring of Patellofemoral Disorders, Arthoscopy: the Journal of Arthoscopy and Related Surgery (2):159-163