Send Orders for Reprints to reprints@benthamscience.net

The Reliability and Meaningfulness of the Anterior Knee Pain and Lower Extremity Functional Scales in Patellofemoralpain Syndrome

Konstantinos. D. Papadopoulos*, Dr Jeanette M Thom, Professor Jane Noyes, Jeremy G Jones and Dimitris Stasinopoulos

School of Sciences, Department of Health Sciences, Physiotherapy Programme; European University of Cyprus, School of Sports, Health and Exercise Sciences, Bangor University, UK

Abstract: Two very common scales used in the assessment of patellofemoral pain syndrome are the anterior knee pain scale and the lower extremity functional scale. There is only limited evidence regarding how specifically reliable and meaningful these scales are when assessing the syndrome. The purpose of this study was to assess which questions in both scales are suitable for patellofemoral pain syndrome patients. 20 patients with patellofemoral pain were recruited from the physiotherapy waiting list of the local hospital and asked to complete the anterior knee pain scale and the lower extremity functional scale on two occasions at least one week apart. A general test-retest reliability of the scales was measured in addition with test-retest and internal consistency of each single question. Finally, the questions marked as 'no problem' in both sessions were also measured. The total scores of the two scales were found to be highly reliable. However, the anterior knee pain scale revealed five questions with moderate test retest reliability, two questions with less internal consistency whilst it included three less meaningful questions. The lower extremity functional scale showed four questions with moderate test retest reliability, one question with less internal consistency and six meaningless questions. This study agrees with previous research stating that there are questions in both scales that can be considered meaningless and less reliable and should probably be excluded or replaced with other questions. The study provides useful information for the development of a more appropriate patellofemoral pain syndrome scale or a modified anterior knee pain scale and lower extremity functional scale for patellofemoral pain syndrome use only.

Keywords: Anterior Knee Pain Scale, Lower Extremity Functional Scale, Patellofemoral pain syndrome, Reliability, Assessment.

INTRODUCTION

Patellofemoral pain syndrome (PFPS) is also known as the runner's knee and is a challenge for both sport and therapeutic communities [1]. The anterior knee pain scale(AKPS) [2] and the lower extremity functional scale (LEFS) [3], are very common scales in the assessment of PFPS [4] and have been used as outcome measurements in studies with PFPS patients [5-8] testing disability, dysfunction and pain. The AKPS set out to be a specific scale for PFPS while, the LEFS is a generic scale for any lower limb pain. Cochrane database of systematic reviews [9] contains AKPS as one of the scales focusing on knee pain whilst, both scales are included in a recent review [4] as two out of ten outcome measurements with sufficient quality and evidence for knee assessment. The latter review presented outcome measurements for patients with musculoskeletal conditions. The authors reported that the AKPS shows evidence of content validity and responsiv-eness, however the LEFS revealed excellent reliability and better responsiveness than AKPS; however, it does not include questions such as locking, swelling and giving way [4].

*Address correspondence to this author at the School of Sports, Health and Exercise Sciences, Bangor University, George Building Bangor Gwynedd LL57 2PZ, UK; Tel:+44 (0)1248 388256; Fax: +44(0)1248 371053; E-mail:pepa01@bangor.ac.uk

Although these scales are used widely, there is currently not enough evidence regarding their specific reliability in PFPS patients. Crossley, *et al.* [5] reported that the AKPS was one of the most efficient measures for detecting a treatment effect, however, Bennell, *et al.* [6] demonstrated that although the AKPS was a reliable tool amongst others, the size of the measurement error should be considered. Other researchers [10, 11] tested the test re-test reliability of both scales finding them extremely reliable for PFPS patients. However, they commented that both scales include questions considered meaningless from many patients whilst, other questions should be included (e.g. about kneeling). In addition, other authors [12] challenged the specificity of those questions and whether they can differentiate PFPS from other knee condition patients. Finally, all the above studies call for further research to determine whether modification of these scales would produce a more sensitive and specific tool. The aim of the current study was to report the questions which keep the scales from being more specific and appropriate for use in PFPS cases. This would be identified by checking the test-retest reliability of the general scores of the AKPS and the LEFS, and reporting the questions that had less internal and test-retest reliability and the questions that were less meaningful.

1875-399X/13

METHODS

National Health Service patients referred by their general practitioner or a consultant with a diagnosis of PFPS were recruited from the waiting list of the physiotherapy department at Ysbyty Gwynedd, a district general hospital. An extended scope physiotherapist identified the patients who should not have any other knee conditions (e.g. knee ligament conditions/menisci conditions, history of trauma, previous knee surgery, history of true locking, history of patellar dislocation, history of arthritis, knee joint effusion, patellar tendinopathy) [13] and sent a participant information sheet to the address the patients had provided. Once they verbally accepted to take part, they contacted the research team at the School of Sports, Health and Exercise Sciences of Bangor University to set an appointment. Patients agreed to visit university premises twice. The second time took place after no less than seven and no more than 10 days. There, theywere asked to sign a written consent form and then to complete the AKPS first and then the LEFS by following the instructions and without any assistance. At the first visit they were also asked to report whether the pain they experienced was permanent or appeared occasionally (on/off) and how long they had the pain for. None of the patients received any treatment while they participated in the study. The patient recruitment took place from June 2011 to October 2011. Ethical approval was granted by Betsi Cadwaladr University Health Board (09/WNo01/29).

STATISTICAL METHODS

Test-retest reliability of the scales was performed by Intra-Class Correlation (ICC). The overall scores of the scales that participants completed in the first session were correlated with the overall scores of those in the second session.

To identify how reliable each of the questions was, the internal consistency of each scale was reported along with how this would change if each of the questions was deleted. This was performed by reporting the Cronbach's alpha on SPSS (IBM New York, USA,v.20). Cronbach's alpha determines the internal consistency or average correlation of any items within a questionnaire or scale to gauge their reliability [14]. If the reliability of the scale becomes larger when a questionnaire is deleted, this means that that question lowers the overall reliability. In addition to this, test-retest reliability was also measured for each single question between the first and the second session. This would reveal whether each question can report the same value at two different times [14]. A value of 0.70 was set as cut-off point. In addition to the ICC, the standard error of measurement was also calculated for each single question to assess how confident we can be with the 'true' score of each question [15].

To find which of the questions were less meaningful, the questions within the two scales that were mostly answered as 'no problem' in both sessions were reported. Such questions cannot change the overall score of the scales; therefore, they have no clinical value and should be excluded from a scale which measures PFPS conditions. According to the answers we received it was decided that the cut-off point for a question to be considered as meaningful would be when at least 10 out of 20 participants reported a question as 'no problem in both sessions.

RESULTS

Twenty patients (10 males and 10 females) were included in this study. They were aged between 18 and 40 (29.0±6.6 years). Most of the patients (17/20) reported to have on/off pain (rather than permanent) which was aggravated with several activities (e.g. sports). In addition, the patients had been in pain for 62.2 ± 61.9 months. The total scores of the two scales were highly reliable [AKPS; ICC = 0.828, F (19, 20) = 5.821, sig<0.00; LEFS; ICC=0.816, F (19, 20) = 5.440, sig<0.00] and so was the internal consistency for each scale [AKPS; Cronbach's alpha = 0.791, LEFS; Cronbach's alpha = 0.921. However, within the reliability (test-retest) of each separate question there were some less reliable questions (<0.7=moderate reliability) as highlighted in Table **1**. The standard error of measurement was satisfactory as it was found very low in most cases. The AKPS revealed five questions with moderate test re-test reliability (questions one, two, five, nine, and 11) with four in the LEFS (questions nine, 11, 13 and 19). The AKPS revealed two questions (questionone & 12) that if they were 'deleted' from the scale the overall internal consistency would be increased. The LEFS showed only one (question 13). Additionally, both scales included questions that may be considered as meaningless as patients with PFPS answered as 'no problem' on both occasions; In the AKPS these were question number 10, 12 and 13 and in the LEFS the questions were three, four, five, seven, 10 and 20 (Table **1**).

DISCUSSION

The aim of the current study was to report the general test-retest reliability of the AKPS and the LEFS in addition with the test-retest, internal consistency and meaningfulness of each question. This study reports high test-retest reliability for the total scores of AKPS and LEFS in PFPS patients. This finding agrees with previous studies which used similar methodology to detect test-retest reliability and responsiveness of the two scales [10, 11]. In addition, both scales revealed high overall internal consistency. Although the AKPS is supposed to be more specific for PFPS patients [6], it revealed a lower Cronbach's alpha (0.791) than the more generic LEFS (0.921). This probably occurred because some of the questions in the AKPS focuses on pain (question nine), whilstotherson function (question three) and self-assessment (question 12). On the other hand the LEFS focuses only on function.

All questions included in the scales did not report the same internal consistency. The two questions about limping and atrophy of thigh muscles (questions one& 12)in the AKPS, would increase the overall internal consistency if they were excluded ('deleted') from the scale, In the LEFS, this question was about going up and down 10 steps (question 13). This may suggests that the above three questions were found to correlate poorly with the other questions within the scales and to diverge from the consistency of results across questions [16]. The reason that these questions lowered the overall internal consistency is probably because of the great divergence in the answers of

Table 1. Intraclass Correlation Coefficient, Cronbach's alpha if Each Question was 'deleted' and Meaningfulness of each Question for Patients with PFPS

Questions in AKPS	Mean ± SD Occasion 1	Mean ± SD Occasion 2	ICC	Standard Error of Measurement	Cronbach's Alpha if Item Deleted	N/A Questions
Question 1 (Limp)	3.9±1.0	3.3±1.1	0.69*	0.07	0.794+	4
Question 2 (Support)	3.4±1.2	3.6±0.9	0.45*	1.92	0.781	1
Question 3 (Walking)	3.3±1.3	3.2±1.3	0.83	0.50	0.758	5
Question 4 (Stairs)	4.1±2.3	3.1±2.4	0.85	0.34	0.762	0
Question 5 (Squatting)	3.5±1.0	3.1±1.3	0.46*	1.14	0.767	1
Question 6 (Running)	4.0±2.2	5.4±2.8	0.79	1.21	0.778	1
Question 7 (Jumping)	5.9±2.7	5.3±2.9	0.81	1.22	0.719	1
Question 8 (Prolonged sitting with knees flexed)	6.4±2.3	6.5±2.4	0.95	1.30	0.733	3
Question 9 (Pain)	5.4±2.3	5.1±2.4	0.66*	1.66	0.780	0
Question 10 (Swelling)	9.0±2.4	8.9±2.4	0.95	0.31	0.748	13#
Question 11 (Instability giving way in the knees)	6.8±2.6	6.6±2.4	0.63*	1.89	0.816	4
Question12 (Atrophy of thighs)	4.5±1.3	4.6±1.2	0.97	0.12	0.900+	17#
Question13 (Flexion deficiency)	4.3±1.3	4.4±1.3	0.90	0.33	0.775	13#
Questions in LEFS						
Question1 (Usual work activities)	3.0±1.1	2.7±1.1	0.83	0.45	0.916	2
Question 2 (Hobbies)	1.5±1.2	1.5±1.3	0.77	0.63	0.920	1
Question 3 (Into out of the bath)	3.6±0.9	3.6±0.8	0.92	0.17	0.917	12#
Question 4 (Walking between rooms)	3.8±1.1	3.8±1.0	0.86	0.37	0.918	15#
Question 5 (Putting on/off socks)	3.3±0.9	3.3±0.9	0.93	0.17	0.921	10#
Question 6 (Squatting)	2.1±1.4	1.8±1.3	0.78	0.68	0.918	1
Question 7 (Lifting an object)	3.7±1.0	3.7±0.9	0.97	0.02	0.919	13#
Question 8 (Light activities at home)	3.3±1.0	3.3±0.8	0.70	0.57	0.918	7
Question 9 (Heavy activities at home)	2.5±1.1	2.6±0.9	0.55*	0.82	0.918	2
Question 10 (Getting into car)	3.3±0.9	3.2±1.0	0.88	0.30	0.919	10#
Question 11 (Walking 2 blocks)	2.9±1.2	3.0±1.1	0.59*	0.92	0.915	6
Question 12 (Walking a mile)	2.3±1.1	2.6±1.1	0.80	0.52	0.917	3
Question 13 (Up or down 10 steps)	2.4±0.8	2.5±1.1	0.64*	0.71	0.923×	1
Question 14 (Standing 1 hour)	2.8±1.1	2.5±1.2	0.91	0.26	0.916	6
Question 15 (Sitting 1 hour)	2.8±1.3	2.9±1.0	0.75	0.61	0.916	4
Question 16 (Running on even ground)	1.9±1.2	2.0±1.1	0.88	0.38	0.915	1
Question 17 (Running on uneven ground)	1.6±1.1	1.6±1.1	0.89	0.32	0.916	1
Question 18 (Making sharp turns while running)	1.6±1.4	1.8±1.4	0.91	0.31	0.914	3
Question 19 (Hopping)	2.1±1.4	2.2±1.5	0.61*	1.12	0.914	2
Question 20 (Rolling over in bed)	3.7±1.0	3.8±0.9	0.93	0.19	0.918	15#

ICC: Intra class Correlation Coefficient. SEm: Standard Error of measurement. N/A questions: the number of patients who answered the question as 'no problem' in both occasions.*: questions with moderate test-retest reliability. #: considered meaningless questions for PFPS. +: AKPS questions with internal consistency less than 0.791.×: LEFS question with internal consistency less than 0.921. The confidence interval was 95% in all cases.

these questions from different patients. Some patients had significant problem with limping and walking up/down the stairs. However, this would depend on whether they were on pain on the day they completed the scales. On the other hand, although the stair question is mentioned in both scales, the question was found to lower the internal consistency only in the LEFS. The reason was probably that the question in the LEFS mentions walking up and down the stairs whilst patients with PFPS usually have problem only when they go down the stairs. Therefore, if the question was reworded the results of this question would probably be different. Finally, regarding the muscle atrophy question some people could

detect their atrophy whilst some other although the atrophy was noticeable they could not.

The current study has also revealed several questions within the two scales with less test-retest reliability and could be considered less reliable for PFPS. These are questions that PFPS patients did not probably know how to answer because they were not clear to them (e.g instability giving way in the knees) or because they are not specific and every time patients completed the questions they may had a different activity in their mind (e.g. heavy activities at home).On the other hand, cultural adaption of the questions could be another reason for moderate test-retest reliability. Such an example could be the question about walking. This question revealed high test-retest reliability in the AKPS (0.83) but moderate in LEFS (0.59). The latter scale asks about the problem that patients have when they have to walk two blocks. The word 'block' is mostly used in American English not used by British people and is probably not comprehended in a rural area where there are no 'blocks'. Consequently, when patients were asked to complete these questions at two different occasions the answers they gave were different.

As these scales were not designed specifically for PFPS alone, they include questions that could be considered as meaningless, i.e. where the patients completed them as 'no problem' on both occasions. This probably reduces the ability to discriminate improvements. Such an example would be question number 12 (atrophy of thigh muscles) of the AKPS which revealed too high reliability (0.97) probably because 17 out of the 20 patients reported it as 'no problem'. This study verifies Kujala, *et al.* [2] who also found extremely small or no differences for questions 10, 12 and 13 in a PFPS group compared with healthy controls in the AKPS. This study agrees with suggestions that a modified version of the above scales might be needed excluding the less reliable and meaningful questions [9] and replacing them with questions which can discriminate PFPS from other knee pain conditions [12]. Therefore, it has provided information regarding the identification of the questions which could be replaced or reworded (e.g. the question about squatting was more reliable in LEFS than in AKPS). Also, as all people do not perform the same activities (e.g. running and jumping) and cannot assess themselves (e.g. atrophy of thighs) it is suggested that all questions in a modified scale need to include a 'not applicable answer' choice. In addition, questions should focus more on function in the activities (as the LEFS does) and not on pain (AKPS focuses more on pain) as the PFPS patients of this study did not report consistent pain (17/20 reported on/off pain rather than permanent). Thus, the questions that had lower reliability may have resulted from not having a time scale for pain specified in the scales.

STRENGTH AND LIMITATIONS

Previous literature review has not shown relative evidence regarding this way of analysis. Previous researchers analysed the reliability of the final scores of the scales and not of each question separately. On the other hand, a limitation of the study could be the small number of patients included in this study because the district general hospital could not provide us with more patients during that period of recruitment. We are planning to conduct a study with more participants in the future. In addition, the patients of this study were asked to complete the scales only twice. If they participated more times the analysis would probably provide us with stronger results. The reason for this decision was that all patients were recruited from the waiting list and if they were asked to participate longer in this study they would probably have to start their physiotherapy treatment. If this happened, the above analysis of the scales would not be possible as the parameters would not be the same between sessions. Finally, the analysis would be complemented if we included a control group. If participants with no PFPS reported a question as a 'problem' this question should also be ruled out from a PFPS scale.

CONCLUSION

To conclude, although the two scales were found to be generally reliable, both scales have been shown to include non-specific PFPS questions. The AKPS has revealed 10 questions with less reliability or meaningfulness for PFPS whilst the LEFS revealed 11. These questions could be reworded or replaced with other questions more appropriate for PFPS use. This study provides valuable information for the development of a modified or anew PFPS scale which will assess when patients get better and if they are ready to return to sports. Further research with more patients is called to support this evidence.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflicts of interest.

ACKNOWLEDGEMENTS

The authors would like to thank the Ysbyty Gwynedd physiotherapist Moyra Barnes who identified the patients of this study, the European University of Cyprus statistician Nassios Orinos for helping with the analysis of this study and Loutsia Nardi for her valuable advice.

REFERENCES

[1] Dixit S, Difiori JP. Management of patellofemoral pain syndrome. Am Fam Physician 2007; l75: 194-202.

[2] Kujala UR, Jaakkola LH, Koskinen SK, Taimela S, Hurme M, Nelimarkka O. Scoring of patellofemoral disorders. Arthroscopy 1993; 9: 159-63.

[3] Binkley JA, Stratford PW, Lott SA, Riddle DL. The lower extremity functional scale (LEFS): Scale development, measurement properties, and clinical application. Phys Ther 1999; 79: 371-83.

[4] Howe TE, Dawson LJ, Syme G, Duncan L, Reid J. Evaluation of outcome measures for clinical practice for adults with musculoskeletal conditions of the knee: A systematic review. Man Ther 2012; 17: 100-18.

[5] Crossley KM, Bennell KL, Cowan SM, Green S. Analysis of outcome measures for persons with patellofemoral pain syndrome: which are reliable and valid? Arch Phys Med Rehabil 2004; 85: 815-22.

[6] Bennell K, Bartam S, Crossley K, Green S. Outcome measures in patellofemoral pain syndrome: test retest reliability and inter-relationships. Phys Ther Sport 2000; 1: 22-41.

[7] Callaghan M, Selfe J. Patellar taping for patellofemoral pain syndrome in adults. Cochrane Database Syst Rev 2012; 4: CD006717.

[8] Long-Rossi F, Salsich SB. Pain and hip lateral rotator muscle strength contribute to functional status in females with patellofemoral pain. Physiother Res Int 2010; 15: 57-60.

[9] Heintjes EM, Berger M, Bierma–Zeinstra SMA, Bernsen RMD, Verhaar JAN, Koes BW. Exercise therapy for patellofemoral pain syndrome. Cochrane Database Sys Rev 2003; 4: CD003472.

[10] Watson CJ, Propps M, Ratner J, Ziegler DL, Horton P, Smith SS. 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 2005; 35: 136-46.

[11] Paxton EW, Fithian DC, Stone ML, Silva P. The reliability and validity of knee specific and general health instruments in assessing acute patellar dislocation outcomes. Am J Sports Med 2003; 31: 487-92.

[12] Callaghan M, Selfe J, Dey MP. Activity-associated pain in patellofemoral pain syndrome: How does it inform research and practice? Physiotherapy 2009; 95: 321-2.

[13] Crossley KM, Bennell, KL, Green S, Cowan SM, McConnell J. Physical therapy for patellofemoral pain: a randomized, double blinded, placebo-controlled trial. Am J Sports Med 2002; 30: 857-65.

[14] Reynaldo J, Santos A. Cronabach's alpha: A tool for assessing the reliability of scales. J Extension 1999; 37. Available from: http://www.joe.org/joe/1999april/tt3.html [Cited: 8th May 2013].

[15] Mollenkopf GW. Variation of the standard error of measurement. Psychometrika 1949; 14: 189-229.

[16] Rattray J, Jones CM. Essential elements of questionnaire design and development. J Clin Nurs 2005; 16: 234-43.

Received: February 18, 2013 Revised: June 12, 2013 Accepted: June 13, 2013