PHYSIOTHERAPY THEORY AND PRACTICE
2024, VOL. 40, NO. 4, 843–855
https://doi.org/10.1080/09593985.2022.2158053

Taylor & Francis
Taylor & Francis Group

RESEARCH REPORT

Examining the diagnostic accuracy of common physical examination and functional tests in the diagnosis of patellofemoral pain syndrome among patients with anterior knee pain

Neda Mostafaee PhDPT[a], Melika Pashaei-Marandi MScPT[a], Hossein Negahban PhDPT[a,b], Nahid Pirayeh PhDPT[c,d], Amal Saki Malehi[e], and Mohammad Hosein Ebrahimzadeh MD[b]

[a]Department of Physical Therapy, School of Paramedical Sciences, Mashhad University of Medical Sciences, East Door of Ferdowsi University, Khorasan Razavi, Mashhad, Iran; [b]Orthopedic Research Center, Mashhad University of Medical Sciences, Ghaem Hospital, Ahmad Abad St., Khorasan Razavi, Mashhad, Iran; [c]Musculoskeletal Rehabilitation Research Center, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran; [d]Department of Physiotherapy, School of Rehabilitation Sciences, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran; [e]Department of Biostatistics and Epidemiology, Faculty of Public Health, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran

ABSTRACT

Purpose: The aims of this study were to evaluate the diagnostic accuracy of common physical examination and functional evaluation tests, and to determine a set of tests with the highest diagnostic accuracy for diagnosing patellofemoral pain syndrome (PFPS) in patients with anterior knee pain.

Methods: Based on careful evaluation of clinical findings and imaging methods by orthopedic physicians, 162 patients with anterior knee pain were classified into two groups of PFPS and non-PFPS. The physical examination and functional tests were performed by two physiotherapists. The accuracy of these measures was determined by calculating sensitivity, specificity, area under the receiver operating characteristic (ROC) curve (AUC), likelihood ratio (LR), and predictive value (PV).

Results: Our results showed the most sensitive tests in identifying patients with PFPS were as follows: eccentric step test [0.82 (95%CI: 0.72–0.89)]; palpation test [0.81(95%CI: 0.70–0.88)]; and prolonged sitting [0.73 (95%CI: 0.62–0.82)]. The palpation test, patellar tilt test, eccentric step test, navicular drop test, squatting, and stair descending tests had an acceptable accuracy (AUC ≥ 70). The strongest combination of the physical examination and functional tests included pain severity between 3 and 10 during stair descending test and pain severity between 6 and 10 during prolonged sitting test. This combination showed a positive LR of 19.47 (95% CI: 6.36–59.65) and a posttest probability of 95%.

Conclusion: Our findings provide evidence for the good accuracy of the palpation test, patellar tilt test, eccentric step test, navicular drop test, squatting, and stair descending and prolonged sitting tests for diagnosing PFPS. Also, the combination of stair descending test and prolonged sitting test could be very useful for ruling in PFPS patients.

ARTICLE HISTORY
Received 01 August 2022
Revised 02 December 2022
Accepted 02 Dec 2022

KEYWORDS
Patellofemoral pain syndrome; physical examination; functional tests; sensitivity; specificity

Introduction

Anterior knee pain is a common clinical symptom that is manifested in patellofemoral pain syndrome (PFPS). In the absence of other pathologies, anterior knee pain that is exacerbated by activities such as sitting, squatting, ascending and descending stairs, jumping, or running is identified as PFPS (Witvrouw et al., 2014). According to recent studies, this syndrome is not regarded as a simple self-limiting disease, and it has been reported that more than 50% of PFPS suffer from advanced chronic pain (Lankhorst et al., 2016; Rathleff et al., 2016; Wyndow et al., 2016). Concerns over the long-term consequences of anterior knee pain in adolescence and young adulthood include a predisposition to patellofemoral osteoarthritis in later life (Utting, Davies, and Newman, 2005). A study of patellofemoral osteoarthritis patients showed that 22% of them described preceding PFPS in their adolescence and early adult years (Utting, Davies, and Newman, 2005). Emerging evidence has suggested that these two conditions are in a continuum (Lack, Neal, De Oliveira Silva, and Barton, 2018). Therefore, an accurate diagnosis of PFPS is imperative to begin appropriate management early in order to prevent lingering issues (Crossley et al., 2016a, 2016b; Kasitinon, Li, Wang, and Fredericson, 2021).

Physical examination and functional evaluation are necessary to diagnose musculoskeletal disorders

CONTACT Melika Pashaei-Marandi melikapashaeimarandipt@gmail.com Department of Physical Therapy School of Paramedical Sciences, Mashhad University of Medical Sciences, East Door of Ferdowsi University, Khorasan Razavi, Mashhad, Iran

including PFPS (Malanga, Andrus, Nadler, and McLean, 2003). To draw appropriate conclusions from the findings of physical examination and functional evaluation, it is very important that the clinicians know about the significance of the test results (i.e. the diagnostic value of each test) (Nijs, Van Geel, Van der Auwera, and VanVanVan de Velde, 2006). Investigating the diagnostic accuracy of assessment tests plays an important role in evidence-based medicine and leads to favorable patient outcomes by enabling clinicians to choose an appropriate treatment strategy (Fritz and Wainner, 2001; Knottnerus and Muris, 2003).

In a systematic review, the following factors were reported as important factors causing PFPS: quadriceps weakness; increased quadriceps angle; incorrect mechanics and improper alignment of the lower extremities, overactivity and stiffness of the external retinaculum, excess patellar tilt angle, reduced strength of abductors, and external rotators of the hip (Lankhorst, Bierma-Zeinstra, and van Middelkoop, 2012). Further, a recent systematic review and meta-analysis identified that quadriceps weakness in military recruits and higher hip strength in adolescents were risk factors for PFPS (Neal et al., 2019). The high prevalence and wide variety of factors contributing to the etiology of PFPS make its diagnosis complex and miss-interpretable. No definitive agreement has so far been observed in the literature on the uniform use of physical examination and functional tests aimed at an ideal diagnosis of PFPS (Nunes et al., 2013).

Moreover, a few studies specifically investigating the diagnostic accuracy of physical examination tests for diagnosis of PFPS have been published since 2000, indicating that research in this area is still in its infancy (Cook et al., 2010; Décary et al., 2017; Haim, Yaniv, Dekel, and Amir, 2006; Nijs, Van Geel, Van der Auwera, and VanVanVan de Velde, 2006; Nunes et al., 2013). In fact, the current literature offers limited evidence about how valid different physical examination and functional tests are for the accurate diagnosis of PFPS (Décary et al., 2017). Additionally, there are a number of limitations in the existing studies including case–control design, studying patients from only one medical center, no assessment of the reliability of physical examination tests, and small sample size (Cook et al., 2010; Haim, Yaniv, Dekel, and Amir, 2006; Nijs, Van Geel, Van der Auwera, and VanVanVan de Velde, 2006; Nunes et al., 2013). In other words, the methodological quality of most of these studies ranges from low to moderate, and due to the presence of different types of bias in these studies, the diagnostic validity of the studied tests might have been overestimated (Cook et al., 2010; Décary et al., 2017; Nunes et al., 2013). Moreover, according to previous studies, when used individually physical examination tests may not provide an accurate diagnosis of PFPS (Cook et al., 2010; Décary et al., 2017; Nunes et al., 2013). Recently, Décary et al. (2018) evaluated the diagnostic validity of a set of patient history information and physical examination tests to diagnose PFPS. They indicated that a limitation of their study was that the functional lower limb assessments were not included. To the best of our knowledge, not enough consistent findings have been made for the diagnosis of PFPS (Décary et al., 2017). Thus, the objectives of this study were to evaluate the diagnostic accuracy of common physical examination and functional evaluation tests and to determine a set of tests with the highest diagnostic accuracy in diagnosing PFPS in a sample of patients with anterior knee pain.

Methods

Participants

From March 2021 to April 2022, a total of 162 patients with anterior knee pain were recruited from multicenter physiotherapy and orthopedics clinics in Mashhad, Iran. Patients aged between 18 and 50 with anterior knee pain of more than 3 months' duration were included in the study (Décary et al., 2018; Lankhorst, Bierma-Zeinstra, and van Middelkoop, 2012; Loudon et al., 2002; Piva et al., 2009). Patients were excluded if they had a history of lower limb surgery or knee arthroplasty in the past 6 months, malignancy, systemic inflammatory disorders associated with knee complaints, musculoskeletal or neurological involvement of the lower limb that interferes with independent physical activity, back pain, sacroiliac pain, referral pain due to trigger point in rectus femoris muscle, and pregnancy (Lankhorst, Bierma-Zeinstra, and van Middelkoop, 2012; Loudon et al., 2002; Piva et al., 2009). The patients' activity levels ranged between 2 and 5 based on the Tegner activity scale (Negahban et al., 2011). The present study conforms to the Standards for Reporting Diagnostic Accuracy Studies 2015 (Bossuyt et al., 2015; Simel, Rennie, and Bossuyt, 2008). The study was approved by the ethics committees of Mashhad University of Medical Sciences (Ref. ID: IR.MUMS.REC.1399.662), and an informed consent form was signed by patients prior to commencement of the study.

Procedure

Reference standard definition

In this study, the clinical diagnosis of PFPS as a reference standard was performed by three orthopedic physicians. The orthopedic physicians had more than 10 years of experience. To standardize the collection of data, the

three orthopedic physicians attended a 30-minute educational session offered by the investigators in advance.

The criteria used to diagnose PFPS were based on those used in other PFPS studies and included patients who: 1) were diagnosed by an orthopedic physician as a case of PFPS; 2) had anterior or retropatellar knee pain; 3) reported exacerbation of symptoms during at least two of the following activities: prolonged sitting, ascending or descending stairs, squatting, and kneeling; and 4) had no clinical evidence of current knee problem other than PFPS (Cook et al., 2010; Décary et al., 2018; Nijs, Van Geel, Van der Auwera, and VanVanVan de Velde, 2006). To diagnose problems related to anterior knee pain other than PFPS (i.e. osteoarthritis, bursitis, tendonitis, meniscus tears, ligament or joint capsule tear, patellar dislocations, and Baker's cysts) standard clinical examinations were performed for all subjects in the study. These clinical examinations involved history taking, inspection, and a set of standard orthopedic tests (i.e. assessment of active and passive range of motion of the knee joint, palpation of the knee joint, varus/valgus and pivot stress tests, anterior and posterior drawer tests, Lachman test, McMurray test, and Apley test for internal and external meniscus) (Cleland, Koppenhaver, and Su, 2015; Cook and Hegedus, 2014; Magee, 2014). After clinical examination, the physicians were presented with reviewed imaging results and radiology reports and performed their own analysis of the relevant imaging diagnoses. All participants were required to have a radiograph of their knee that included the following three views: 1) anterior-posterior; 2) lateral; and 3) skyline views. Magnetic resonance imaging was required when the physicians suspected any ligament injury or a meniscal injury or when the physicians needed to exclude another pathology (Décary et al., 2018). The orthopedic physicians made their final clinical diagnosis only after they carefully evaluated clinical findings and imaging results. All patients underwent the imaging procedure in the same week when they were subject to clinical examination (Cook et al., 2010; Hinman, Lentzos, Vicenzino, and Crossley, 2014; van der Heijden et al., 2016). Imaging was not used to confirm PFPS, but rather to rule out competing diagnoses (Cook et al., 2010; Décary et al., 2018). Participants who did not meet at least two criteria listed above (i.e. criteria 1 and 4) were placed in the non-PFPS group. Thus, based on the final diagnosis of the orthopedist physicians, the patients were categorized into two groups, PFPS and non-PFPS (Cook et al., 2010; Décary et al., 2018; Nijs, Van Geel, Van der Auwera, and VanVanVan de Velde, 2006). Subsequently, two evaluators (physiotherapists) who were masked for group allocation carried out a battery of physical examination tests and asked the participants about pain intensity while doing functional tests. These physical examination and functional evaluation tests were regarded as index tests. The battery of physical examination and functional evaluation tests were performed as follows.

Physical examination tests

These tests were selected as common tests for physical examination of patients based on published evidence, and textbooks (Cook et al., 2010; Décary et al., 2018, 2017; Nunes et al., 2013). Two evaluators (physiotherapists) with more than 5 years of experience independently performed the physical examination. Physical examination tests related to PFPS included patellar tilt test (Haim, Yaniv, Dekel, and Amir, 2006), tenderness at palpation of the postero-medial and postero-lateral borders of the patella (Cook et al., 2010; Décary et al., 2018; Manske and Davies, 2016), pain with resisted isometric extension at 60° knee flexion (Cook et al., 2010; Décary et al., 2018), and eccentric step test (Manske and Davies, 2016; Nijs, Van Geel, Van der Auwera, and VanVanVan de Velde, 2006). The results of these tests were reported as positive or negative. We also measured quadriceps angle (Piva et al., 2006), foot pronation by navicular drop test (Piva et al., 2006), amount of tibia torsion measuring the angle formed between the knee axis, imaginary line from the medial to lateral epicondyle, imaginary line through the malleolus using a standard goniometer (Piva et al., 2006), and the amount of anti-version femoral by Craig's test (Choi and Kang, 2015; Piva et al., 2006) (Figure 1). The results of these tests were reported numerically.

Functional tests

After the patients performed these tests, the evaluators asked them about the severity of pain during a particular functional activity. The results of these tests are reported numerically based on pain intensity using the 10-point Visual Analogue Scale (VAS) in which 0 indicates no pain and 10 indicates the worst pain imaginable (Boonstra, Schiphorst Preuper, Balk, and Stewart, 2014). Functional tests related to PFPS included double leg squatting at 60° knee flexion angle (Cook et al., 2010; Manske and Davies, 2016), stair climbing and descending from 9 steps with a height of 20 cm and a width of 22 cm (Dobson et al., 2013; Tolk et al., 2019), double stance kneeling at 90° knee flexion angle (Cook et al., 2010), and prolonged sitting with knees 90° flexed (Cook et al., 2010) (Figure 1). In all of these tests, the patients reported maximum pain intensity based on VAS while performing the test.

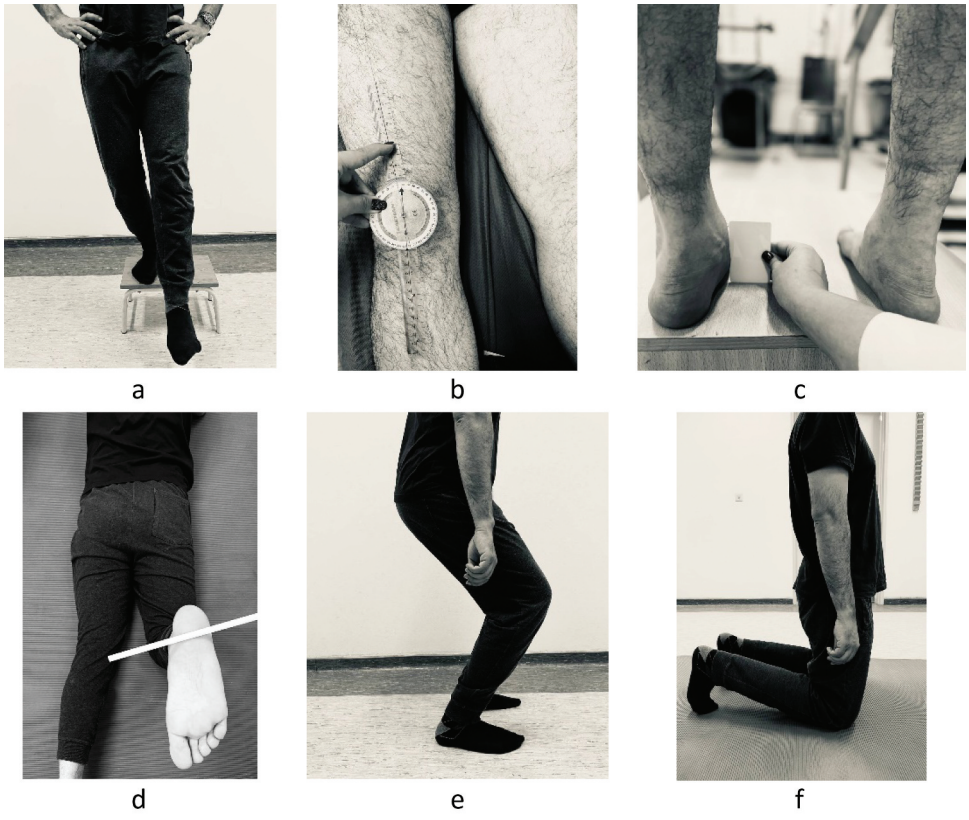

Figure 1. Illustration of some physical examination and functional tests. (A) Eccentric step test; (B) Q angle; (C) Navicular drop test; (D): Tibial torsion test; (E): Squatting; (F): Kneeling.

To evaluate the inter-rater reliability of these tests, the two evaluators independently performed physical examination tests and queried a random sample of 15 participants about the pain severity during each functional test. The 15 participants were randomly selected using a table of random numbers. In addition, to evaluate the intra-rater reliability of physical examination and functional evaluation tests, one of the evaluators performed the physical examination tests and queried a sample of 15 participants about the pain severity during each functional test at the first session and again 7–10 days after the first visit. Thus, the 15 participants with no change in their knee functions within the time interval were included in the retest day. In order to calculate inter-rater and intra-rater reliability, the Cohen's Kappa coefficient was used for qualitative physical examination tests, while intra-class correlation coefficient (ICC) was used for quantitative physical examination and functional evaluation tests (Landis and Koch, 1977; Weir, 2005).

Statistical analysis

Kolmogorov–Smirnov (K-S) test was used to test the normal distribution of the quantitative data in the primary analysis. The results showed normal distribution of variables. In order to check homogeneity of demographic and clinical characteristics of the PFPS and non-PFPS patients, and to compare the two groups in terms of their physical examination and functional tests, independent-samples *t* test and Chi-square test were used for quantitative and qualitative data, respectively.

To determine the diagnostic accuracy of the diagnostic tests, including physical examination and functional evaluation tests, indicators of accuracy are defined as follows: Sensitivity or true positive rate is the proportion of subjects with the condition (i.e. PFPS) who have a positive test result. Specificity or true negative rate is the proportion of subjects without the condition (i.e. non-PFPS) who have a negative test result (Florkowski, 2008; Fritz and Wainner, 2001). ROC analysis was used to determine the accuracy of evaluative measurements. Based on the reference standard and the results of the quantitative physical examination and functional evaluation tests, the diagnostic accuracy of Q angle, navicular drop test, tibia torsion test, Craig test, squatting, stair climbing, stair descending, kneeling, and prolonged sitting were defined as follows. At first, a pair of sensitivity and specificity were calculated for a range of possible scores of these tests. Then, the ROC curve was

obtained by plotting the sensitivity and 1-specificity for all possible cutoff scores on the vertical axis (Y) and horizontal axis (X), respectively. The area under the curve (AUC) obtained from the ROC curve was used to determine the discrimination ability of the two groups. A traditional academic point scale was used to classify the accuracy of the AUC: 0.9 to 1: excellent; 0.8 to 0.89: good; 0.70 to 0.79: acceptable; 0.60 to 0.69: poor; and 0.00 to 0.59: failure (Greiner, Pfeiffer, and Smith, 2000; Wikstrom et al., 2012). To estimate the optimal cutoff point, the ROC curve analysis was also performed. The optimal cutoff point was determined on the ROC curve by using Youden Index. The Youden Index provides a criterion for choosing the optimal cutoff value, the threshold for which sensitivity + specificity – 1 is maximized (Fluss, Faraggi, and Reiser, 2005; Perkins and Schisterman, 2006). Thus, the optimal cutoff point was found by identifying the point with the highest sensitivity and specificity in the ROC analysis that usually takes a point with the highest Youden Index (Fluss, Faraggi, and Reiser, 2005; Perkins and Schisterman, 2006). Therefore, the Youden Index is considered as a good quantitative estimate of a cutoff point that maximizes correct classification and minimizes misclassification (Fluss, Faraggi, and Reiser, 2005; Perkins and Schisterman, 2006).

In addition, the LR for a positive test (PLR), the LR for a negative test (NLR), the positive predictive value (PPV), the negative predictive value (NPV) and their 95% confidence intervals were calculated. The usefulness of the diagnostic test was summarized in the statistics of LR (Fritz and Wainner, 2001). PLR is the ratio of the proportion of patients who have both PFPS and positive test result to the proportion of patients having no PFPS who also have positive test result. NLR is the ratio of the proportion of patients who have PFPS and negative test result to the proportion of patients having no PFPS who also have negative test result. A PLR > 10 and an NLR < 0.1 were reported to provide strong evidence for ruling in and ruling out the diagnoses, respectively (Eusebi, 2013; Florkowski, 2008; Šimundić, 2009). PPV is the percentage of patients with a positive test who have a positive reference standard (PFPS). NPV is the percentage of patients with a negative test who have a negative reference standard (non-PFPS). If a test has high PPV and NPV, it will have a higher chance of correct identification of patients with PFPS and non-PFPS, respectively (Eusebi, 2013; Florkowski, 2008; Šimundić, 2009).

Finally, diagnostic clusters for physical examination and functional tests were also developed to determine combinations that could improve diagnostic accuracy. Posttest probability was calculated using LR+ and the prevalence of PFPS in the study sample was calculated as the number of participants who met the physicians' criteria for diagnosis of PFPS. Classification and regression tree (CART) was used to identify the diagnostic clusters. CART is one of the tree-based methods and a nonparametric technique for exploratory modeling. This algorithm first develops the maximum tree and then prunes it to avoid overfitting. This pruning rule finds a tree that has a smaller size with minimum estimation of the misclassification error. This method determines the clusters based on contrast effects of different diagnostic tests and clinical factors by recursive partitioning of the subjects. All statistical analyses were performed using SPSS (version 21; SPSS Inc., Chicago, IL) and STATA 13 (STATA Corporation, College Station, Texas). The level of statistical significance was set at $P < .05$.

Results

A total of 162 participating patients with anterior knee pain determined based on reference standard were divided into two groups of PFPS (n = 80) and non-PFPS (n = 82). Patients in the Non-PFPS group were diagnosed with knee osteoarthritis (n = 27), meniscal tear (n = 14), anterior cruciate ligament tear (n = 21), or other knee disorders (n = 20). There were no significant differences between the two groups in terms of their demographic and clinical characteristics (Table 1). The results of descriptive statistics for the physical examination and functional tests are reported in Table 2. As shown in this table, all of the physical examination tests, except for measures of tibia torsion and Q angle, and all of the functional tests except for the stair climbing test were statistically different in the two groups ($p < .05$).

The inter- and intra-rater reliabilities of physical examination and functional tests are depicted in Table 3. The ICC and Cohen's Kappa coefficient for inter- and intra-rater reliability ranged between 0.72 and 0.99 for these physical examination and functional tests, respectively.

The sensitivity, specificity, AUC value, optimal cutoff point, LR, PV, and posttest probability for each measure are shown in Table 4. The results of diagnostic accuracy of these measures show that among physical examination and functional tests, eccentric step test, palpation test and prolonged sitting had the highest sensitivity in identifying patients with PFPS ([0.82 (95%CI 0.72–0.89)], [0.81 (95%CI:0.70–0.88)] and [0.73 (95%CI 0.62–0.82)], respectively). Based on ROC analysis, physical examination tests including palpation test, patellar tilt test, eccentric step test, and navicular drop test had

Table 1. Demographic and clinical characteristics of patients in the PFPS and non-PFPS groups.

	PFPS group	Non-PFPS group	P-value
Age (y) Mean ± SD	41.99 ± 9.29	39.93 ± 12.14	0.40
Weight (kg) Mean ± SD	73.11 ± 13.25	75.35 ± 10.67	0.43
Height (cm) Mean ± SD	164.61 ± 9.38	165.82 ± 10.16	0.40
Pain (VAS) Mean ± SD	6.45 ± 1.57	6.06 ± 2.0	0.25
Gender, n (%)	20 (25%)	24 (29.3%)	0.16
Male	60 (75%)	58 (70.7%)	
Female			
Symptomatic leg, n (%)	43 (53.8%)	38 (46.3%)	0.11
Right	37 (46.3%)	44 (53.7%)	
Left			
Dominant leg, n (%)	76 (95%)	75 (91.5%)	0.37
Right	4 (5%)	7 (8.5%)	
Left			
Education level, n (%)	1 (1.3%)	4 (4.9%)	0.42
Illiterate	24 (30%)	27 (32.9%)	
Middle school	24 (30%)	25 (30.5%)	
Diploma	30 (37.5%)	23 (28%)	
Bachelor's degree	1 (1.3%)	3 (3.7%)	
Master's degree			

PFPS = patellofemoral pain syndrome; VAS = visual analog scale.

Table 2. Descriptive statistics of physical examination and functional tests in PFPS and non-PFPS groups.

	PFPS group	Non-PFPS group	P-value
Patellar tilt test, n (%)	31 (38.8%)	64 (78.0%)	<0.001*
N	49 (61.3%)	18 (22.0%)	
P			
Palpation, n (%)	15 (18.8%)	54 (65.9%)	<0.001*
N	65 (18.3%)	28 (34.1%)	
P			
Resisted isometric, n (%)	31 (38.8%)	52 (63.4%)	<0.001*
N	49 (61.3%)	30 (36.6%)	
P			
Eccentric step test, n (%)	14 (17.5%)	53 (64.6%)	<0.001*
N	66 (82.5%)	29 (35.4%)	
P			
Q angle (Degree) Mean ± SD	12.61 ± 3.83	12.06 ± 3.42	0.51
Navicular drop test (mm) Mean ± SD	11.40 ± 5.27	7.63 ± 2.66	<0.001*
Tibial torsion test (Degree) Mean ± SD	21.45 ± 5.46	21.62 ± 6.43	0.99
Craig test (Degree) Mean ± SD	12.23 ± 6.42	9.47 ± 4.35	0.01*
Squatting (VAS) Mean ± SD	6.03 ± 2.25	3.34 ± 2.03	<0.001*
Stair climbing (VAS) Mean ± SD	3.92 ± 2.65	3.91 ± 2.49	0.86
Stair descending (VAS) Mean ± SD	6.15 ± 2.25	2.93 ± 2.16	<0.001*
Kneeling (VAS) Mean ± SD	4.23 ± 3.11	2.85 ± 2.53	<0.001*
Prolonged sitting (VAS) Mean ± SD	7.12 ± 2.31	3.15 ± 2.02	<0.001*

PFPS = patellofemoral pain syndrome; VAS = visual analog scale; N = negative; P = positive; P-value<0.05; *indicate significant difference.

Table 3. Intraclass correlation coefficient (ICC) and Cohen's Kappa coefficient for intra- and inter-rater reliability of physical examination and functional tests (n = 15).

Variable	Intra-rater reliability ICC unless indicated	Inter-rater reliability ICC unless indicated	Effect size
Patellar tilt test, Cohen's Kappa coefficient	0.99	0.98	0.63
Palpation, Cohen's Kappa coefficient	0.85	0.91	0.54
Resisted isometric contraction, Cohen's Kappa coefficient	0.93	0.96	0.59
Eccentric step test, Cohen's Kappa coefficient	0.93	0.94	0.58
Q angle (Degree)	0.82	0.88	0.42
Navicular drop test (mm)	0.93	0.90	0.53
Tibial torsion test (Degree)	0.93	0.89	0.45
Craig test (Degree)	0.72	0.80	0.38
Squatting (VAS)	0.89	0.93	0.59
Stair climbing (VAS)	0.88	0.92	0.57
Stair descending (VAS)	0.94	0.97	0.62
Kneeling (VAS)	0.96	0.96	0.60
Prolonged sitting (VAS)	0.95	0.98	0.64

Table 4. Diagnostic accuracy values of the physical examination and functional tests.

	Sensitivity (95% CI)	Specificity (95% CI)	AUC (95% CI)	Best cutoff	LR+ (95% CI)	LR- (95% CI)	PPV (95% CI)	NPV (95% CI)	Posttest probability
Patellar tilt test	0.61 (0.49–0.71)	0.78 (0.67–0.86)	0.73 (0.65–0.81)		2.79 (1.79–4.34)	0.49 (0.37–0.65)	0.73 (0.60–0.82)	0.67 (0.56–0.76)	73%
Palpation	0.81 (0.70–0.88)	0.65 (0.54–0.75)	0.79 (0.72–0.87)		2.37 (1.73–3.27)	0.28 (0.17–0.45)	0.69 (0.59–0.78)	0.78 (0.66–0.86)	69.5%
Resisted isometric	0.61 (0.49–0.71)	0.78 (0.67–0.86)	0.64 (0.56–0.74)		2.79 (1.79–4.34)	0.67 (0.56–0.76)	0.67 (0.56–0.76)	0.66 (0.51–0.73)	73%
Eccentric step test	0.82 (0.72–0.89)	0.64 (0.53–0.74)	0.78 (0.71–0.87		2.33 (1.71–3.17)	0.27 (0.16–0.44)	0.69 (0.59–0.78)	0.79 (0.67–0.87)	69%
Q angle	0.45 (0.33–0.56)	0.60 (0.48–0.70)	0.53 (0.44–0.62)	12.5	1.11 (0.78–1.59)	0.92 (0.74–1.14)	0.52 (0.39–0.64	0.52 (0.42–0.63)	52%
Navicular drop test	0.58 (0.47–0.69)	0.78 (0.67–0.86)	0.75 (0.67–0.82)	9.5	2.67 (1.71–4.18)	0.32 (0.40–0.69)	0.72 (0.59–0.89)	0.65 (0.55–0.75)	72%
Tibial torsion test	0.47 (0.36–0.58)	0.57 (0.45–0.68)	0.47 (0.38–0.56)	21.5	1.11 (0.79–1.56)	0.91 (0.73–1.14)	0.52 (0.40–0.63)	0.53 (0.41–0.63)	52%
Craig test	0.42 (0.31–0.54)	0.76 (0.65–0.85)	0.59 (0.50–0.68)	11.5	1.83 (1.14–2.93)	0.78 (0.61–0.91)	0.64 (0.49–0.76)	0.57 (0.47–0.67)	64%
Squatting	0.60 (0.48–0.70)	0.87 (0.78–0.93)	0.82 (0.75–0.89)	5.5	4.92 (2.67–9.03)	0.45 (0.34–0.59)	0.82 (0.70–0.90)	0.69 (0.59–0.77)	82%
Stair climbing	0.45 (0.33–0.56)	0.51 (0.40–0.62)	0.53 (0.43–0.62)	4.5	0.92 (0.66–1.28)	1.07 (0.86–1.33)	0.47 (0.35–0.59)	0.48 (0.37–0.59)	47%
Stair descending	0.62 (0.50–0.72)	0.87 (0.76–0.92)	0.86 (0.80–0.92)	5.5	4.65 (2.62–8.28)	0.43 (0.32–0.57)	0.81 (0.69–0.90)	0.70 (0.60–0.78)	82%
Kneeling	0.32 (0.22–0.44)	0.81 (0.71–0.89)	0.63 (0.53–0.72)	5.5	1.77 (1.01–3.09)	0.82 (0.70–0.96)	0.63 (0.46–0.77)	0.55 (0.46–0.64)	63%
Prolonged sitting	0.73 (0.62–0.82)	0.92 (0.84–0.96)	0.90 (0.84–0.95)	5.5	10.07 (4.61–22.01)	0.28 (0.19–0.40)	0.90 (0.80–0.96)	0.78 (0.68–0.85)	91%

CI = Confidence interval; AUC = area under the curve; LR+ = positive likelihood ratio; LR- = negative likelihood ratio; PPV = positive predictive value; NPV = negative predictive value.

an acceptable level of accuracy (AUC ≥ 70), while resistance isometric contraction test had poor level of differentiation [AUC = 0.64 (95%CI 0.56–0.74)] between the two groups of PFPS and non-PFPS patients (Figures 2 and 3). Among all functional tests, squatting and stair descending tests had a good level of accuracy (AUC ≥ 70), and prolonged sitting test had excellent accuracy [AUC = 0.90 (95% CI 0.84–0.95)] (Figure 4). The results of LRs and PV also show that eccentric step test and palpation test had the least NLR [0.27 (95% CI 0.16–0.44)] [0.28 (95% CI 0.17–0.45)], respectively, and a high NPV [0.79 (95% CI 0.67–0.87)] and [0.78 (95% CI 0.66–0.86)]. Additionally, among the physical examination tests, resisted isometric test, patella tilt test and navicular drop test had high specificity [0.78 (95% CI 0.67–0.86)] and high PLR [2.79 (95%CI 1.79–4.34), 2.79 (95%CI 1.79–4.34) and 2.67 (95% CI 1.71–4.18)], respectively. As far as functional tests were concerned, the prolonged sitting test had the highest rate in specificity [0.92 (95% CI 0.84–0.96)], PLR [10.07 (95%CI 4.61–22.01)] and PPV [0.90 (95%CI 0.80–0.96)]. The squatting test and stair descending test had high specificity [0.87 (95%CI 0.78–0.93)], high PLR [4.92 (95%CI 2.67–9.03), 4.65 (95% CI 2.62–8.28)], and high PPV [0.82 (95%CI 0.70–0.90), 0.81 (95%CI 0.69–0.90)], respectively. Optimal cutoff points for the measures were also determined using ROC curve analysis (Table 4).

In addition, the combination of the physical examination and functional tests demonstrated improved diagnostic accuracy. Table 5 shows the diagnostic clusters. Three clusters were determined. As seen in Table 5, the strongest combination of the physical examination and functional tests included pain severity between 3 and 10 during stair descending test and pain severity between 6 and 10 during prolonged sitting test. This combination showed a PLR of 19.47 (95% CI: 6.36–59.65) and a posttest probability of 95%.

Discussion

Our findings indicated that among physical examination tests, only palpation test, patellar tilt test, eccentric step test, and navicular drop test can discriminate between PFPS and non-PFPS (AUC > 0.70). Furthermore, the eccentric step test and palpation test have the highest sensitivity and the lowest NLR among physical examination tests, and thus have the best screening ability to rule out patients with PFPS. On the other hand, because of their highest specificity and the highest PLR, the resisted isometric test, patellar tilt test, and navicular drop test can be used to confirm the diagnosis of PFPS patients. In addition, these tests can be useful for the correct identification of PFPS patients because of their high PPV.

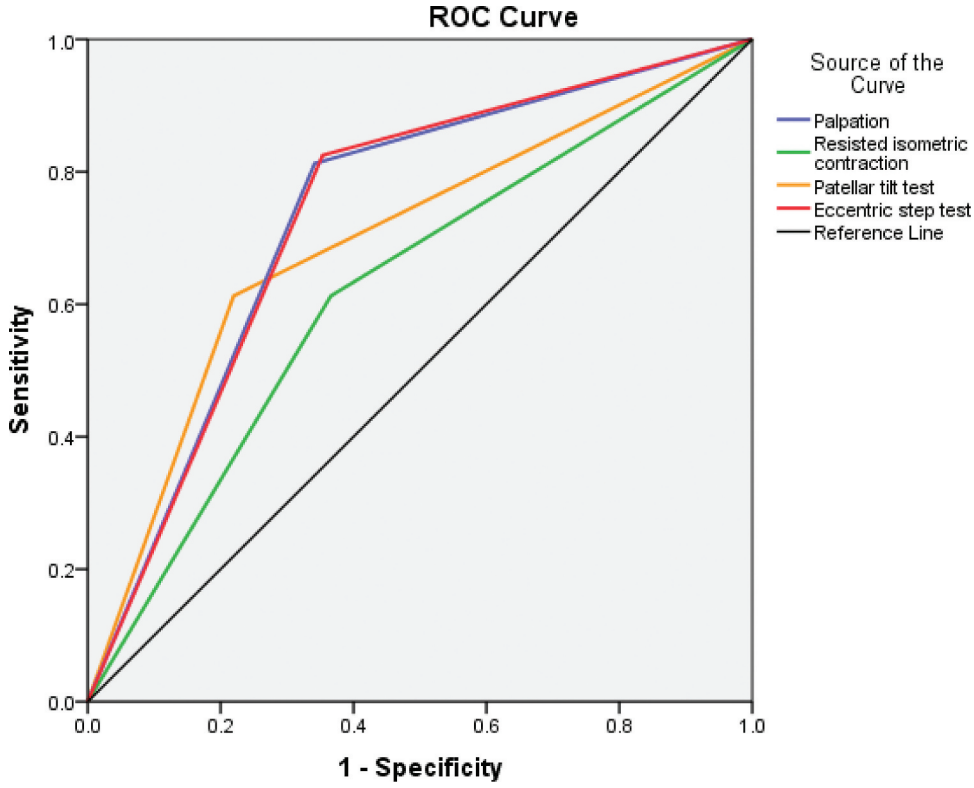

Figure 2. The receiver operating characteristic (ROC) curve for palpation, resisted isometric contraction, patellar tilt test, and eccentric step test.

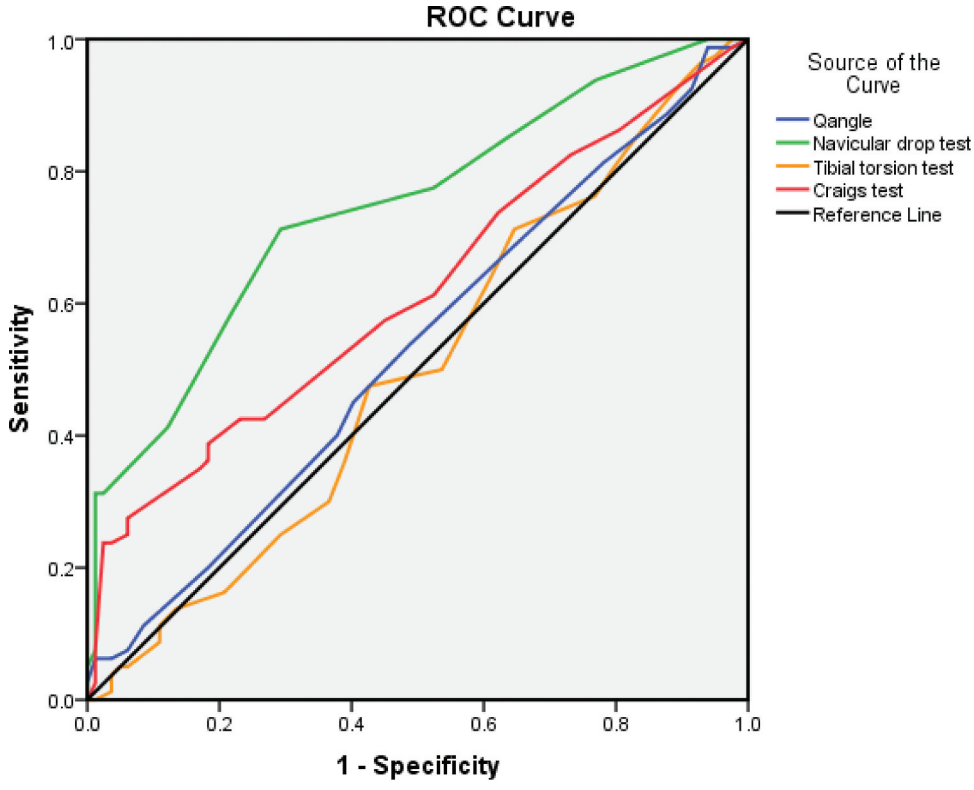

Figure 3. The receiver operating characteristic (ROC) curve for Q angle, navicular drop test, tibial torsion test, and Craig test.

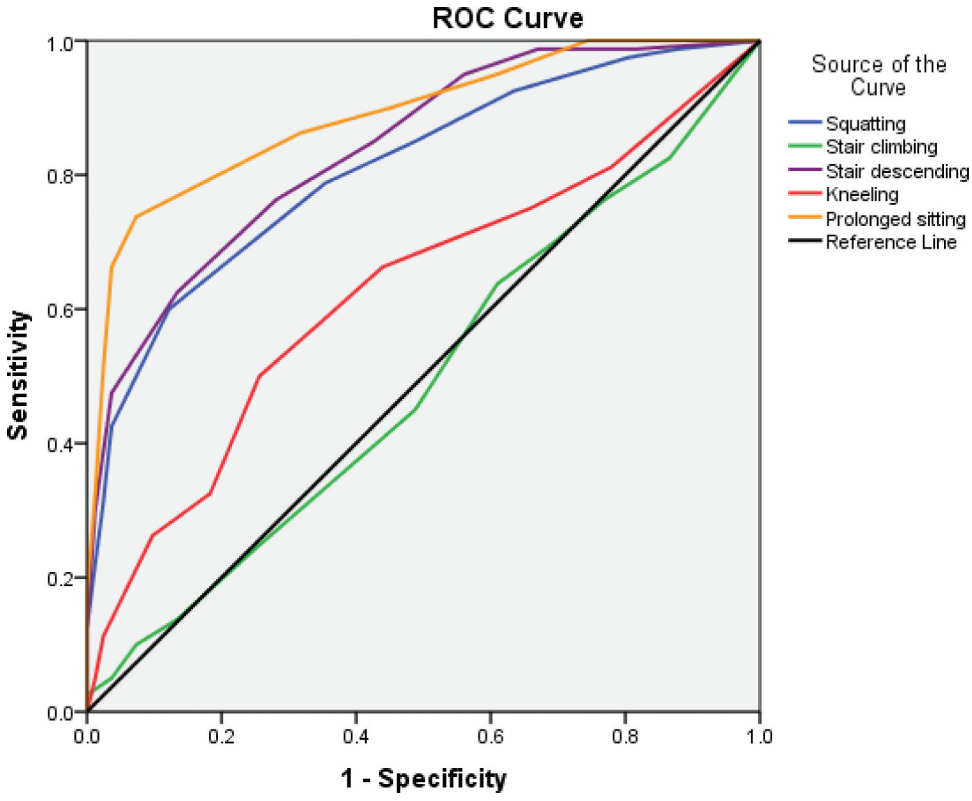

Figure 4. The receiver operating characteristic (ROC) curve for squatting, stair climbing, stair descending, kneeling, and prolonged sitting.

The results of our study were in agreement with those of previous studies examining the diagnostic accuracy of evaluation tests in patients with PFPS. Haim, Yaniv, Dekel, and Amir (2006) found that the patellar tilt test has the highest specificity, the highest PLR, and good accuracy in diagnosing PFPS patients. Cook et al. (2010) also showed that pain during the resisted isometric test has high PLR and PPV. However, the results of our study on the eccentric step test were not consistent with those of Nijs, Van Geel, Van der Auwera, and VanVanVan de Velde (2006) who showed that this test has high PLR and can help identify patients with PFPS. In our study, the eccentric step test was found to have low NLR, and since it has a relatively low false-negative result, its true negative result can be useful for ruling out patients with PFPS.

Also, Selfe et al. (2001) reported that the eccentric step test triggers pain in 74% of patients with PFPS which was confirmed by our results indicating provoked pain in 66 out of 80 PFPS knees after the eccentric step test (82%). According to Piva et al. (2006) who examined the reliability of the measures of impairments related to

Table 5. Diagnostic accuracy of best combinations (clusters) of the physical examination and functional tests.

		Sensitivity	Specificity	PPV (95%CI)	NPV (95%CI)	LR+ (95% CI)	LR- (95%CI)	Posttest Probability
Cluster 1	• 7≤ pain during stair descending≤10 • 0≤ pain during prolonged sitting≤5	12.5	97.56	83.33 (53.07–95.67)	53.33 (51.10–55.56)	5.12 (1.16–22.66)	0.9 (0.82–0.98)	83%
Cluster 2	• 3≤ pain during stair descending≤10 • 6≤ pain during prolonged sitting≤10	71.25	96.34	95 (86.12–98.31)	77.45 (70.81–82.94)	19.47 (6.36–59.65)	0.3 (0.21–0.42)	95%
Cluster 3	• Patellar tilt test = 1(positive) • Eccentric step test = 1(positive) • 0≤ pain during stair descending≤6 • 0≤ pain during prolonged sitting≤5	8.75	93.9	58.33 (31.67–80.88)	51.33 (49.15–53.51)	1.43 (0.48–4.33)	0.97 (0.89–1.06)	58%

PPV = positive predictive value; NPV = negative predictive value; LR+ = positive likelihood ratio; LR- = negative likelihood ratio.

PFPS, among the tests they examined, the navicular drop test had a very high reliability.

The reason for the difference between our results and others may be that our study utilized a cohort of participants from a relevant clinical population (i.e. with PFPS and other knee disorders causing anterior knee pain) while subjects included in other studies consisted of individuals with PFPS and without any knee disorders. When subjects without any symptoms, impairments, or disabilities are tested, this does not reflect the way most tests are applied clinically where distinctions between individuals with similar symptoms are required (Fritz and Wainner, 2001). Spectrum or selection bias may occur when study subjects are not representative of the population on whom the test is typically applied in practice. Spectrum bias can profoundly affect the results of a study. The best method of ensuring a representative sample and avoiding spectrum bias is to utilize a cohort design with a consecutive group of subjects from a relevant clinical population (Fritz and Wainner, 2001). Further, our study included larger sample size and more general population compared to other ones. Another reason may be that we recruited a sample with a broader age range than did previous PFPS diagnostic studies.

The results of our study also demonstrated that all functional tests except the stair climbing test and the kneeling test have a good or excellent level of differentiation between the two groups of PFPS and non-PFPS patients. In addition, all tests except the stair climbing test had a high specificity and a high PLR. Thus, they have good ability in ruling in patients with PFPS. Although the kneeling test had high specificity, the PLR of this test was 1.77. As a result, a positive test result of this test alters, to a small and rarely important degree, the probability that a patient with anterior knee pain has PFPS.

In addition, based on the results of the present study, it can be stated that among the functional tests examined, the prolonged sitting test has the highest specificity (0.92) and PLR (10.07). This means that the true positive result of this test (i.e. pain greater than 5.5 while doing this test) is useful to confirm the diagnosis of PFPS. The results of our study also showed that the presence of pain, while prolonged sitting has the highest PPV (0.90). This means that if during this test a patient experiences a pain level of greater than 5.5 (the optimal cutoff point) or, in other words, the result of this test is positive, there is a 90% chance that this patient has PFPS. Our results also indicated that the squatting test and stair descending test have a high specificity, a high PLR, and a high PPV.

The results of our study are contrary to those of Cook et al. (2010) who evaluated the diagnostic accuracy of common functional activities used in the clinical diagnosis of PFPS. In their study, they found that the presence of pain during squats and kneeling tests had high sensitivity and PPV, but due to their low PLR (less than 2), these tests alone did not have a high diagnostic value in diagnosing PFPS. The reason for this difference may be due to the case–control design of Cook et al. (2010) study.

Given the fact that LRs are the best indicators for determining the clinical meaningfulness of diagnostic tests (Bossuyt et al., 2004; Cook, Cleland, and Huijbregts, 2007; Denegar and Fraser, 2006; Eusebi, 2013; Fritz and Wainner, 2001) and since PFPS patients are correctly diagnosed after screening for other knee disorders, the diagnosis of PFPS requires tests with a high PLR for ruling in the presence of this disorder. The values of PLR for functional tests, including squatting, stairs descending, and prolonged sitting, ranged from 4.65 to 10.07, were greater than the ones for physical examination tests, including palpation test, patellar tilt test, eccentric step test, and navicular drop test, ranged from 2.33 to 2.79. Thus, it can be argued that the physical examination tests show only a slight increase in the probability of identifying the PFPS group, while the functional tests show a moderate to large increase in the probability of the identification of the PFPS group. Therefore, the presence of pain during functional tests has a higher value in the diagnosis of patients with PFPS. The results of the present study can be justified by the fact that the presence of pain while descending stairs, squatting, and prolonged sitting test is a diagnostic indicator of patients with PFPS, and that during these tests the load on the knee joint considerably increases and consequently exacerbates the symptoms of PFPS patients (Richards et al., 2008). We also found that the combination of stair descending test and prolonged sitting test can be very useful for ruling in PFPS patients. In other words, we achieved the highest posttest probability with regard to combined findings compared to the results obtained from using functional tests and physical examination tests independently.

The results of the present study are in line with those of Cook et al. (2010) who showed that the combination of pain during isometric contraction of the quadriceps muscle and pain during squatting leads to increased specificity and increased ability in identifying PFPS patients. Also, the presence of pain in at least two of the three tests (i.e. pain during isometric quadriceps contraction and pain during squatting and palpation test) leads to an increase in the LR and an increase in the ability to identify PFPS patients (Cook et al., 2010).

Unlike Cook et al. (2010), our study included different tests which might explain the differences observed in the obtained results and the magnitude of LR+.

Our results were also inconsistent with those of Sweitzer et al. (2010) who found that the combination of tests did not improve the diagnostic accuracy of indicators and was not a better alternative for clinical evaluation. Of course, Fredericson and Yoon (2006) claimed that combining several tests is the best form of diagnosis for people with PFPS, but they reported no measurement data regarding the accuracy of the tests, nor did they propose any specific combination (Fredericson and Yoon, 2006). The reason for the difference may be that, in the Sweitzer et al. (2010) study, the patellar mobility scales were used alone for investigating the diagnosis of PFPS.

In summary, this study provided new insights into the diagnostic accuracy of physical examination and functional tests for the diagnosis of PFPS in anterior knee pain patients. Previous studies investigating the diagnostic accuracy of selected clinical tests for PFPS have concluded that individual physical examination tests do not help diagnose PFPS (Cook et al., 2010; Décary et al., 2017; Nijs, Van Geel, Van der Auwera, and VanVanVan de Velde, 2006). However, Haim, Yaniv, Dekel, and Amir (2006) reported LR+ values of 5.4 for the tilt test. Of course, it should be noted that the case–control design of their study and its sample which included only male soldiers may have biased the measures of diagnostic accuracy. On the other hand, our study demonstrated that when used alone, some physical examinations or functional tests can contribute to the better diagnosis of PFPS (because of their high LR+). This discrepancy in results can be explained by the different diagnostic criteria of PFPS and the included tests for diagnostic accuracy in the two studies.

Strengths and limitations

We recruited our patients from multicenter physiotherapy and orthopedics clinics, allowing for maximum variation in patients with different diagnoses. Our data collection procedure ensured blinding of the evaluators between the index tests and the reference standard. The composite reference standard used by the expert orthopedic physicians included history elements, physical examination tests, and imaging procedures. In addition to physical examination tests, we also included functional lower limb evaluation tests as index tests, which had not already been used in other studies. However, our study was limited in that we included the most common, but not all, physical examination tests used by clinicians to assess the knee joint, which might have affected the results. Another limitation was that the number of participants was not sufficient for evaluating inter-rater reliability, so we suggest investigating the inter-rater reliability of these tests in a larger sample size in future studies.

Conclusion

Measures of palpation test, patellar tilt test, eccentric step test, and navicular drop test, as well as functional tests including squatting, stair descending, and prolonged sitting test have the best ability to discriminate between PFPS and non-PFPS (AUC > 0.70). Furthermore, these measures are clinically valuable in ruling out PFPS. Our results also showed that the combination of the stair descending test and the prolonged sitting test, as opposed to single tests, could be very useful in ruling out PFPS patients.

Acknowledgments

Special thanks to Mashhad University of Medical Sciences for financial support (Grant number: 991003).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Mashhad University of Medical Sciences (Grant number: 991003).

ORCID

Hossein Negahban PhDPT http://orcid.org/0000-0001-8387-1781

References

Boonstra AM, Schiphorst Preuper HR, Balk GA, Stewart RE 2014 Cut-off points for mild, moderate, and severe pain on the visual analogue scale for pain in patients with chronic musculoskeletal pain. Pain 155: 2545–2550 10.1016/j.pain.2014.09.014.

Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Lijmer JG, Moher D, Rennie D, de Vet HC 2004 Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. Family Practice 21: 4–10 10.1093/fampra/cmh103.

Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, Lijmer JG, Moher D, Rennie D, de Vet HC, et al. 2015 STARD 2015: An updated list of essential items for

reporting diagnostic accuracy studies. BMJ 351: h5527 10.1136/bmj.h5527.

Choi BR, Kang SY 2015 Intra and inter-examiner reliability of goniometer and inclinometer use in Craig's test. Journal of Physical Therapy Science 27: 1141–1144 10.1589/jpts.27.1141.

Cleland J, Koppenhaver S, Su J 2015 Netter's orthopaedic clinical examination. Philadelphia: Elsevier Canada.

Cook C, Cleland J, Huijbregts P 2007 Creation and critique of studies of diagnostic accuracy: Use of the STARD and QUADAS methodological quality assessment tools. Journal of Manual and Manipulative Therapy 15: 93–102 10.1179/106698107790819945.

Cook C, Hegedus E 2014 Orthopedic physical examination tests: An evidence-based approach 2nd. Harlow. Essex: Pearson Education.

Cook C, Hegedus E, Hawkins R, Scovell F, Wyland D 2010 Diagnostic accuracy and association to disability of clinical test findings associated with patellofemoral pain syndrome. Physiotherapy Canada 62: 17–24 10.3138/physio.62.1.17.

Crossley KM, Stefanik JJ, Selfe J, Collins NJ, Davis IS, Powers CM, McConnell J, Vicenzino B, Bazett-Jones DM, Esculier JF, et al. 2016a Patellofemoral pain consensus statement from the 4th international patellofemoral pain research retreat, manchester. Part 1: Terminology, definitions, clinical examination, natural history, patellofemoral osteoarthritis and patient-reported outcome measures. British Journal of Sports Medicine 50: 839–843 10.1136/bjsports-2016-096384.

Crossley KM, van Middelkoop M, Callaghan MJ, Collins NJ, Rathleff MS, Barton CJ 2016b Patellofemoral pain consensus statement from the 4th international patellofemoral pain research retreat, manchester. Part 2: recommended physical interventions (exercise, taping, bracing, foot orthoses and combined interventions). British Journal of Sports Medicine 50: 844–852 10.1136/bjsports-2016-096268.

Décary S, Frémont P, Pelletier B, Fallaha M, Belzile S, Martel-Pelletier J, Pelletier J-P, Feldman D, Sylvestre M-P, Vendittoli P-A, et al. 2018 Validity of combining history elements and physical examination tests to diagnose patellofemoral pain. Archives of Physical Medicine and Rehabilitation 99: 607–614 10.1016/j.apmr.2017.10.014.

Décary S, Ouellet P, Vendittoli PA, Roy JS, Desmeules F 2017 Diagnostic validity of physical examination tests for common knee disorders: An overview of systematic reviews and meta-analysis. Physical Therapy in Sport 23: 143–155 10.1016/j.ptsp.2016.08.002.

Denegar CR, Fraser M 2006 How useful are physical examination procedures? Understanding and applying likelihood ratios. Journal of Athletic Training 41: 201–206.

Dobson F, Hinman RS, Roos EM, Abbott J, Stratford P, Davis AM, Buchbinder R, Snyder-Mackler L, Henrotin Y, Thumboo J, et al. 2013 OARSI recommended performance-based tests to assess physical function in people diagnosed with Hip or knee osteoarthritis. Osteoarthritis and Cartilage / OARS, Osteoarthritis Research Society 21: 1042–1052 10.1016/j.joca.2013.05.002.

Eusebi P 2013 Diagnostic accuracy measures. Cerebrovascular Diseases 36: 267–272 10.1159/000353863.

Florkowski CM 2008 Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: Communicating the performance of diagnostic tests. Clinical Biochemist Reviews 29 Suppl 1: S83–87.

Fluss R, Faraggi D, Reiser B 2005 Estimation of the youden index and its associated cutoff point. Biometrical Journal 47: 458–472 10.1002/bimj.200410135.

Fredericson M, Yoon K 2006 Physical examination and patellofemoral pain syndrome. American Journal of Physical Medicine & Rehabilitation 85: 234–243 10.1097/01.phm.0000200390.67408.f0.

Fritz JM, Wainner RS 2001 Examining diagnostic tests: An evidence-based perspective. Physical Therapy 81: 1546–1564 10.1093/ptj/81.9.1546.

Greiner M, Pfeiffer D, Smith RD 2000 Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. Preventive Veterinary Medicine 45: 23–41 10.1016/S0167-5877(00)00115-X.

Haim A, Yaniv M, Dekel S, Amir H 2006 Patellofemoral pain syndrome: Validity of clinical and radiological features. Clinical Orthopaedics and Related Research 451: 223–228 10.1097/01.blo.0000229284.45485.6c.

Hinman RS, Lentzos J, Vicenzino B, Crossley KM 2014 Is patellofemoral osteoarthritis common in middle-aged people with chronic patellofemoral pain? Arthritis Care and Research 66: 1252–1257 10.1002/acr.22274.

Kasitinon D, Li WX, Wang EX, Fredericson M 2021 Physical examination and patellofemoral pain syndrome: An updated review. Current Reviews in Musculoskeletal Medicine 14: 406–412 10.1007/s12178-021-09730-7.

Knottnerus JA, Muris JW 2003 Assessment of the accuracy of diagnostic tests: The cross-sectional study. Journal of Clinical Epidemiology 56: 1118–1128 10.1016/S0895-4356(03)00206-3.

Lack S, Neal B, De Oliveira Silva D, Barton C 2018 How to manage patellofemoral pain – Under-standing the multifactorial nature and treatment options. Physical Therapy in Sport 32: 155–166 10.1016/j.ptsp.2018.04.010.

Landis JR, Koch GG 1977 The measurement of observer agreement for categorical data. Biometrics 33: 159–174 10.2307/2529310.

Lankhorst NE, Bierma-Zeinstra SM, van Middelkoop M 2012 Risk factors for patellofemoral pain syndrome: A systematic review. Journal of Orthopaedic and Sports Physical Therapy 42: 81–94 10.2519/jospt.2012.3803.

Lankhorst NE, van Middelkoop M, Crossley KM, Bierma-Zeinstra SM, Oei EH, Vicenzino B, Collins NJ 2016 Factors that predict a poor outcome 5-8 years after the diagnosis of patellofemoral pain: A multicentre observational analysis. British Journal of Sports Medicine 50: 881–886 10.1136/bjsports-2015-094664.

Loudon JK, Wiesner D, Goist-Foley HL, Asjes C, Loudon KL 2002 Intrarater reliability of functional performance tests for subjects with patellofemoral pain syndrome. Journal of Athletic Training 37: 256–261.

Magee DJ 2014 Orthopedic physical assessment. st. Louis, MO: Saunders.

Malanga GA, Andrus S, Nadler SF, McLean J 2003 Physical examination of the knee: A review of the original test description and scientific validity of common orthopedic tests. Archives of Physical Medicine and Rehabilitation 84: 592–603 10.1053/apmr.2003.50026.

Manske RC, Davies GJ 2016 Examination of the patellofemoral joint. International Journal of Sports Physical Therapy 11: 831–853.

Neal BS, Lack SD, Lankhorst NE, Raye A, Morrissey D, van Middelkoop M 2019 Risk factors for patellofemoral pain: A systematic review and meta-analysis. British Journal of Sports Medicine 53: 270–281 10.1136/bjsports-2017-098890.

Negahban H, Mostafaee N, Sohani SM, Mazaheri M, Goharpey S, Salavati M, Zahednejad S, Meshkati Z, Montazeri A 2011 Reliability and validity of the tegner and marx activity rating scales in Iranian patients with anterior cruciate ligament injury. Disability and Rehabilitation 33: 2305–2310 10.3109/09638288.2011.570409.

Nijs J, Van Geel C, Van der Auwera C, VanVanVan de Velde B 2006 Diagnostic value of five clinical tests in patellofemoral pain syndrome. Manual Therapy 11: 69–77 10.1016/j.math.2005.04.002.

Nunes GS, Stapait EL, Kirsten MH, de Noronha M, Santos GM 2013 Clinical test for diagnosis of patellofemoral pain syndrome: Systematic review with meta-analysis. Physical Therapy in Sport 14: 54–59 10.1016/j.ptsp.2012.11.003.

Perkins NJ, Schisterman EF 2006 The inconsistency of "optimal" cutpoints obtained using two criteria based on the receiver operating characteristic curve. American Journal of Epidemiology 163: 670–675 10.1093/aje/kwj063.

Piva SR, Fitzgerald GK, Irrgang JJ, Fritz JM, Wisniewski S, McGinty GT, Childs JD, Domenech MA, Jones S, Delitto A 2009 Associates of physical function and pain in patients with patellofemoral pain syndrome. Archives of Physical Medicine and Rehabilitation 90: 285–295 10.1016/j.apmr.2008.08.214.

Piva SR, Fitzgerald K, Irrgang JJ, Jones S, Hando BR, Browder DA, Childs JD 2006 Reliability of measures of impairments associated with patellofemoral pain syndrome. BMC Musculoskeletal Disorders 7: 33 10.1186/1471-2474-7-33.

Rathleff MS, Rathleff CR, Olesen JL, Rasmussen S, Roos EM 2016 Is knee pain during adolescence a self-limiting condition? Prognosis of patellofemoral pain and other types of knee pain. American Journal of Sports Medicine 44: 1165–1171 10.1177/0363546515622456.

Richards J, Thewlis D, Selfe J, Cunningham A, Hayes C 2008 A biomechanical investigation of a single-limb squat: Implications for lower extremity rehabilitation exercise. Journal of Athletic Training 43: 477–482 10.4085/1062-6050-43.5.477.

Selfe J, Harper L, Pedersen I, Breen-Turner J, Waring J 2001 Four outcome measures for patellofemoral joint problems. Physiotherapy 87: 507–515 10.1016/S0031-9406(05)65448-X.

Simel DL, Rennie D, Bossuyt PM 2008 The STARD statement for reporting diagnostic accuracy studies: Application to the history and physical examination. Journal of General Internal Medicine 23: 768–774 10.1007/s11606-008-0583-3.

Šimundić AM 2009 Measures of diagnostic accuracy: Basic definitions. Electronic Journal of the International Federation of Clinical Chemistry and Laboratory Medicine 19: 203–211.

Sweitzer BA, Cook C, Steadman JR, Hawkins RJ, Wyland DJ 2010 The inter-rater reliability and diagnostic accuracy of patellar mobility tests in patients with anterior knee pain. Physician and Sportsmedicine 38: 90–96 10.3810/psm.2010.10.1813.

Tolk JJ, Janssen RP, Prinsen CA, Latijnhouwers D, van der Steen MC, Bierma-Zeinstra SM, Reijman M 2019 The OARSI core set of performance-based measures for knee osteoarthritis is reliable but not valid and responsive. Knee Surgery, Sports Traumatology, Arthroscopy 27: 2898–2909 10.1007/s00167-017-4789-y.

Utting MR, Davies G, Newman JH 2005 Is anterior knee pain a predisposing factor to patellofemoral osteoarthritis? Knee 12: 362–365 10.1016/j.knee.2004.12.006.

van der Heijden RA, Oei EH, Bron EE, van Tiel J, van Veldhoven PL, Klein S, Verhaar JA, Krestin GP, Bierma-Zeinstra SM, van Middelkoop M 2016 No difference on quantitative magnetic resonance imaging in patellofemoral cartilage composition between patients with patellofemoral pain and healthy controls. American Journal of Sports Medicine 44: 1172–1178 10.1177/0363546516632507.

Weir JP 2005 Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. Journal of Strength and Conditioning Research 19: 231–240 10.1519/15184.1.

Wikstrom EA, Tillman MD, Chmielewski T, Cauraugh JH, Naugle KE, Borsa PA 2012 Discriminating between copers and people with chronic ankle instability. Journal of Athletic Training 47: 136–142 10.4085/1062-6050-47.2.136.

Witvrouw E, Crossley K, Davis I, McConnell J, Powers CM 2014 The 3rd International Patellofemoral Research Retreat: An international expert consensus meeting to improve the scientific understanding and clinical management of patellofemoral pain. British Journal of Sports Medicine 48: 408 10.1136/bjsports-2014-093437.

Wyndow N, Collins N, Vicenzino B, Tucker K, Crossley K 2016 Is there a biomechanical link between patellofemoral pain and osteoarthritis? A narrative review. Sports Medicine 46: 1797–1808 10.1007/s40279-016-0545-6.