Physical Therapy in Sport 14 (2013) 54–59

Contents lists available at SciVerse ScienceDirect

Physical Therapy in Sport

journal homepage: www.elsevier.com/ptsp

Clinical test for diagnosis of patellofemoral pain syndrome: Systematic review with meta-analysis

Guilherme S. Nunes [a], Eduardo Luiz Stapait [a,b,*], Michel Hors Kirsten [a], Marcos de Noronha [a,c], Gilmar Moraes Santos [a]

[a] Santa Catarina State University, Department of Physiotherapy, Rua Pascoal Simone, 358, 88080-350 Coqueiros, Florianópolis, Brazil
[b] West of Santa Catarina State University, Department of Physiotherapy, Rua Getúlio Vargas, 2125, Flor da Serra, Joaçaba 89600-000, Brazil
[c] La Trobe Rural Health School, Department of Allied Health, PO Box 199, Bendigo, Victoria, Australia

ARTICLE INFO

Article history:
Received 25 July 2012
Received in revised form
30 October 2012
Accepted 25 November 2012

Keywords:
Patellofemoral pain syndrome
Knee injuries
Sensitivity and specificity

ABSTRACT

The high incidence and diversity of factors attributed to the etiology of patellofemoral pain syndrome (PFPS) makes the diagnosis of this problem somewhat complex and susceptible to misinterpretation. Currently, there is not a defined set of procedures considered as ideal to diagnose PFPS. To investigate the diagnostic accuracy of clinical and functional tests used to diagnose PFPS through a systematic review. We searched relevant studies in the databases Medline, CINAHL, SPORTDiscus and Embase. The QUADAS score was used to assess the methodological quality of the eligible studies. We analyzed data that indicated the diagnostic properties of tests, such as sensibility, specificity, positive (LR+) and negative (LR−) likelihood ratio, and predictive values. The search identified 16,169 potential studies and five studies met the eligibility criteria. The 5 studies analyzed 25 tests intending to accurately diagnose PFPS. Two tests were analyzed in two studies and were possible to perform a meta-analysis. Within the five studies included, one study had high methodological quality, two studies had good methodological quality and two studies had low methodological quality. Two tests, the patellar tilt (LR+ = 5.4 and LR− = 0.6) and squatting (LR+ = 1.8 and LR− = 0.2), had values that show a trend for the diagnosis of PFPS (LR+ >5.0 and LR− <0.2), however their values do not represent clear evidence regarding diagnostic properties as suggested in the literature (LR+ >10 and LR− <0.1). Future diagnostic studies should focus on the sample homogeneity and standardization of tests analyzed so future systematic reviews can determine with more certainty the accuracy of the tests for diagnosis of PFPS.

1. Introduction

In the absence of other intra-articular disorders, there is currently consensus that anterior knee pain, which limits activities of daily living that demand knee flexion such as climbing and descending stairs, squatting or remaining seated, is defined as patellofemoral pain syndrome (PFPS) (Bohnsack, Hurschler, Demirtas, Rühmann, Stukenborg-Colsman, & Wirth, 2005; Bohnsack et al., 2009; Loudon, Wiesner, Goist-Foley, Asjes, & Loudon, 2002; Reid, 1993; Thomeé, Augustsson, & Karlsson, 1999). There is a high incidence of this condition among physically active populations; it affects 8.75% of the individuals involved in intense physical training and has a significant impact on their occupational activities (Wills, Ramasamy, Ewins, & Etherington, 2004). Besides the series of factors that negatively contributes to symptom reduction, there is a diversity of terms used to classify the types of PFPS. This diversity may lead to an erroneous evaluation and it could consequently influence treatment. Some terms have been used synonymously to classify individuals reporting symptoms in the knee region, such as chondromalacia patellae, patellar arthralgia, patellar pain and patellofemoral syndrome.

Even when analyzed within the general population, PFPS has an incidence between 15 and 25% (Boling, Padua, Marshall, Guskiewicz, Pyne, & Beutler, 2010; Tállay et al., 2004; Wood, Muller, & Peat, 2011) and is more predominant among female adolescents and physically active young adults (Boling, Padua, Marshall, Guskiewicz, Pyne, & Beutler, 2009; Ivković, Franić, Bojanić, & Pećina, 2007; Myer et al., 2010; Tenforde, Sayres, McCurdy, Collado, Sainani, & Fredericson, 2011). Several factors have been linked to PFPS, including a decrease in quadriceps strength (Lankhorst, 2011; Pattyn et al.,

* Corresponding author. Santa Catarina State University, Department of Physiotherapy, Rua Pascoal Simone, 358, 88080-350 Coqueiros, Florianópolis, Brazil. Tel.: +55 48 33218600.
E-mail address: eduardo.stapait@gmail.com (E.L. Stapait).

1466-853X/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ptsp.2012.11.003

2011), decreased flexibility (Miyamoto, Soriano, & Cabral, 2010; Piva et al., 2006), asynchrony in the electrical activity between the vastus medialis oblique and the vastus lateralis longus muscles (Santos, Ries, Sperandio, Say, Pulzatto, & Monteiro-Pedro, 2011), rotations between the femur and the tibia (Boling et al., 2009; Levinger, Gilleard, & Coleman, 2007), excessive subtalar pronation (Boling et al., 2009), and alterations in the position of the patella (Hunter et al., 2007; Ward & Powers, 2004).

The high incidence and diversity of factors attributed to the etiology of PFPS make its diagnostics complex and susceptible to errors of interpretation. Currently, there is not a defined set of procedures considered as ideal to diagnose PFPS. However, Fredericson and Yoon (2006) recommend an association of tests and functional evaluations to aid in the diagnosis of PFPS. Even though a number of studies have proposed tests for diagnosing PFPS, the comparative validity of these tests is not clear. Thus, the objective of this study is to investigate, as a systematic review, the accuracy of these clinical and functional tests for diagnosing PFPS.

2. Methods

This systematic review was conducted according to PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) recommendations (Costa, Maher, Lopes, de Noronha, & Costa, 2011; Moher, Liberati, Tetzlaff, & Altman, 2010).

2.1. Eligibility criteria

This review included studies evaluating the accuracy of clinical and functional tests for diagnosing PFPS. No limits regarding date of publication or language were established, although the following types of studies were excluded: studies where the patients had undergone surgery in lower limbs affected by PFPS; studies evaluating the accuracy of diagnostic imaging tests; studies in which diagnosis was determined using questionnaires; studies in which the participants had other associated diseases (such as osteoarthritis and ligament injuries). We also excluded studies evaluating the accuracy of tests in individuals with chondromalacia patellae, because in this condition there is structural injury to the cartilage and it is thus not considered PFPS (Reid, 1993; Thomeé et al., 1999).

2.2. Search strategy

The electronic search was carried out in the following databases: *Medline* via *OVID*, *CINAHL* (*Cumulative Index to Nursing and Allied Health Literature*) via *EBSCO*, *SPORTDiscus* via *EBSCO* and *Embase*. The most recent search date was March 8, 2012. The search filters developed by the *Scottish Intercollegiate Guidelines Network* (SIGN, http://www.sign.ac.uk/methodology/filters.html) for diagnostic studies were combined with a specific search strategy for PFPS studies developed by the authors. The same approach was used for all searches, adapted as necessary according to the specifics of each database (Table 1).

2.3. Selection of the studies

After searching the databases, two independent evaluators selected articles first by title, then by abstract. Disagreement between the evaluators was solved by consensus. In the cases were no consensus was reached, a third evaluator was consulted to decide about the eligibility of the study. Only studies that were potentially appropriate in light of the inclusion and exclusion criteria were fully analyzed (Fig. 1).

Table 1
Medline search.[a]

1. exp "Sensitivity and Specificity"/
2. sensitivity.tw
3. specificity.tw
4. ((pre-test or pretest) adj probability).tw
5. post-test probability.tw
6. predictive value$.tw
7. likelihood ratio$.tw
8. or/1-7
9. Patellofemoral Pain Syndrome/
10. Patellofemoral Joint/
11. Patella/
12. Femur/
13. Knee/
14. Knee Injuries/
15. Joint Diseases/
16. Pain/
17. Arthralgia/
18. ((patell$ or femor$ or femoro-patell$ or retropatell$) adj3 (pain or syndrome or dysfunction)).mp
19. pain adj3 anterior knee.mp
20. patell$ or femor$ or femoro-patell$ or retropatell$.mp
21. Patellofemoral Pain Syndrome.mp
22. anterior knee pain.mp
23. ((chondromalac$ or chondropath$) adj3 (knee$ or patell$ or femor$ or femoro-patell$ or retropatell$)).mp
24. or/9-23
25. 8 and 24

[a] This strategy was used to search Medline. The strategy was modified for searches of other databases.

2.4. Evaluation of the methodological quality

The methodological quality of the diagnostic studies was evaluated using the *Quality Assessment of Diagnostic Accuracy Studies* (QUADAS) scale (Whiting, Rutjes, Reitsma, Bossuyt, & Kleijnen, 2003). This scale consists of 14 items that allow three types of answer: "yes", "no", and "not clear". The sum of "yes" answers is the final score, with a maximum of 14 points. Item 1 refers to the representativeness of the sample, and item 10 refers to the

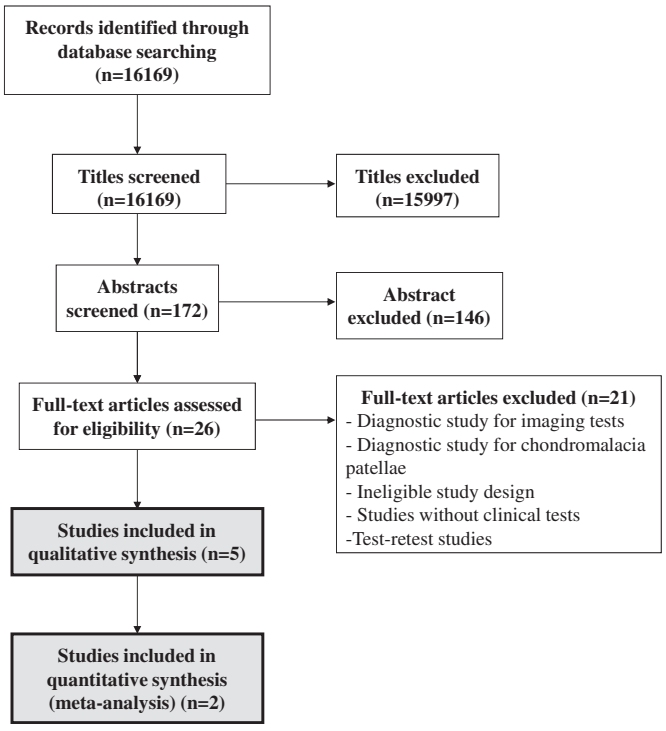

Fig. 1. Flow chart of the search process.

evaluator blinding procedures with respect to diagnosis. For the purposes of this review, studies in which items 1 and 10 received a "yes" and the total score was over 10 were considered to have high methodological quality. In the cases where the study obtained a score above 10 on the QUADAS scale and did not receive a "yes" for items 1 and 10, it was considered to have intermediate methodological quality. Studies with a score less than or equal to 10 receiving a "no" and/or a "not clear" for items 1 and 10 were considered to have low methodological quality. The classification into low, intermediate and high methodological quality, based on items 1 and 10, was performed because these items address issues considered to have higher impact in diagnostic studies regarding bias avoidance (Jaeschke, Guyatt, & Sackett, 1994; Leeflang, Deeks, Gatsonis, & Bossuyt, 2008). Two independent evaluators applied the QUADAS scale to the included studies; in case of disagreement, a third evaluator was consulted so that consensus could be reached (Table 2).

2.5. Data extraction and analysis

The studies were examined for quantitative data indicating the following diagnostic test properties:

- *Sensitivity*, which measures the test's capacity to identify individuals affected by the disease. It is calculated by dividing the true positive results by the true positive results plus the false negative results and is expressed as a percentage. The higher the value, the higher the test's capacity to identify affected individuals (Deeks, 2001; Jaeschke et al., 1994).
- *Specificity*, which indicates the test's capacity to identify individuals who are not affected by the disease in question. It is calculated by dividing the true negative results by the true negative results plus the false positive results and is expressed as a percentage. The higher the value, the higher the chance that the test will identify individuals who are not affected by the disease (Deeks, 2001; Jaeschke et al., 1994).
- *Positive* (*LR+*) *and Negative* (*LR−*) *Likelihood Ratio*, which refer to the discriminatory measurements of the tests, indicating how many times more (LR+) or less (LR−) likely the test results will be in affected than non-affected individuals (Deeks, 2001; Hayden & Brown, 1999; Jaeschke et al., 1994). LR+ is calculated as follows: sensitivity/1 − specificity. LR− is calculated as follows: 1 − sensitivity/specificity (Hayden & Brown, 1999). LR+ results greater than 10 and LR− less than 0.1 indicate convincing diagnostic evidence. However, an LR+ greater than 5 and an LR− less than 0.2 already indicate a strong diagnostic tendency (Jaeschke et al., 1994).
- *Predictive Value*, which refers to the percentage of times that the test will correctly diagnose the evaluated condition. Positive predictive value (PV+) refers to the proportion of affected individuals with positive results, and negative predictive value (PV−) refers to the proportion of non-affected individuals with negative results (Fritz & Wainner, 2001). Predictive values generally are expressed as a percentage. PV+ is the ratio of true positive results to all positive results obtained, whether true or false; PV− is the ratio of the true negative results to all negative results (Fritz & Wainner, 2001).

When a test was performed similarly in more than one study, a meta-analysis was run based on previous studies that have run meta-analysis for diagnostic tests (Deeks, 2001; Harbord et al., 2008; Hayden & Brown, 1999). For that, we used the true positive, true negative, false positive and false negative results for each test, from the included studies, to calculate a single final value for each property above mentioned that represent the test in question. Some studies did not present all of the above-mentioned analyses (Haim, Yaniv, Dekel, & Amir, 2006; Näslund, Näslund, Odenbring, & Lundeberg, 2006; Nijs, Van Geel, Van der Auwera, & Van de Velde, 2006), however they presented sufficient data that could be used to perform such analyses ourselves (Greenhalgh, 1997; Jaeschke et al., 1994; Leeflang et al., 2008).

3. Results

After searching a total of 16,169 titles; five articles conformed to the adopted eligibility criteria and were included in this review

Table 2
Quadas scores.

Study	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Total
Haim et al., 2006	N	Y	Y	NC	Y	Y	Y	Y	Y	NC	Y	Y	N	Y	10[c]
Näslund et al., 2006	Y	Y	Y	NC	Y	Y	Y	Y	Y	NC	Y	Y	N	Y	11[b]
Nijs et al., 2006	Y	Y	Y	NC	N	Y	Y	Y	Y	Y	Y	Y	N	Y	11[a]
Cook et al., 2010	Y	Y	Y	Y	Y	Y	Y	Y	NC	NC	Y	Y	N	Y	11[b]
Sweitzer et al., 2010	Y	Y	NC	Y	Y	Y	NC	Y	NC	Y	NC	Y	N	N	8[c]

Items QUADAS tool

1. Was the spectrum of patients representative of the patients who will receive the test in practice?
2. Were selection criteria clearly described?
3. Is the reference standard likely to correctly classify the target condition?
4. Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?
5. Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?
6. Did patients receive the same reference standard regardless of the index test result?
7. Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?
8. Was the execution of the index test described in sufficient detail to permit replication of the test?
9. Was the execution of the reference standard described in sufficient detail to permit its replication?
10. Were the index test results interpreted without knowledge of the results of the reference standard?
11. Were the reference standard results interpreted without knowledge of the results of the index test?
12. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?
13. Were uninterpretable/intermediate test results reported?
14. Were withdrawals from the study explained?

Abbreviation: N = No; NC = Unclear; Y = Yes.
[a] High methodological quality.
[b] Intermediate methodological quality.
[c] Low methodological quality.

(Cook, Hegedus, Hawkins, Scovell, & Wyland, 2010; Haim et al., 2006; Näslund et al., 2006; Nijs et al., 2006; Sweitzer, Cook, Steadman, Hawkins, & Wyland, 2010) (Fig. 1). Of the included articles, one presented high, two presented intermediate and two presented low methodological quality, according to the QUADAS evaluation scale, previously described (Table 2).

Table 3 presents the accuracy analyses for 24 tests evaluated by the studies included in this review. The best diagnostic measures are highlighted in Table 3, showing that squatting was the most sensitive test (91%), with the lowest LR− (0.2) and highest PV− (74%). The vastus medialis coordination test had the best specificity among all tests (93%); the patellar tilt had the highest LR+ (5.4) and the active instability test had the highest PV+ (100%). A meta-analysis was carried out for the patellar apprehension test, and its accuracy measurements are presented in bold in Table 3. Of the five articles included, two presented accuracy analyses for the combination of tests (Table 4).

4. Discussion

The reviewed studies did not offer consistent evidence regarding the accuracy of the PFPS diagnostic tests. However, the results from the patellar tilt test (LR+ >5) (Haim et al., 2006) and pain during squatting (LR− <0.2) (Cook et al., 2010), suggest a strong tendency toward PFPS diagnosis.

In the present study, methodological quality was evaluated using the QUADAS scale, which is commonly used in systematic reviews of diagnostic tests (Alqarni, Schneiders, & Hendrick, 2011; Hegedus, Cook, Hasselblad, Goode, & McCrory, 2007; Reneker, Paz, Petrosino, & Cook, 2011). However, certain items scored on the QUADAS scale are crucial for quality assurance. It was considered that items 1 and 10 are the most relevant since they address important precautions regarding the methodological validity of diagnostic studies. Therefore, in order to avoid error, a separate criterion, i.e., positive responses to questions 1 and 10, was adopted for determining high methodological quality besides the QUADAS scale score. Evaluated in this manner, only Nijs et al. (2006) was considered to have high methodological quality, while Cook et al. (2010) and Näslund et al. (2006) were considered to have intermediate methodological quality. These studies were not considered to have high quality because they did not clarify whether the evaluator was blinded regarding the subjects diagnosis while applying the tests, i.e., they failed to comply with the 10th item of the QUADAS scale. These two articles (Cook et al., 2010; Näslund et al., 2006) would have being considered of high quality had they been qualified without special consideration to item 10, however blinding of evaluators is an issue commonly raised as great source of bias (Herbert, Jamtvedt, Mead, & Hagen, 2005; Manske & Lehecka, 2012; Page, 2012).

Even though only one study presented high methodological quality, some tests did obtain results suggestive of PFPS diagnosis, although none presented consistent diagnostic evidence. The squatting test presented the highest sensitivity among tests (Cook et al., 2010). This test is used in clinical practice because it considerably increases the load on the knee joint and consequently exacerbates symptoms in this area (Richards, Thewlis, Selfe, Cunningham, & Hayes, 2008). The active instability test detected

Table 3
Diagnostic accuracy of the test.

Study	QUADAS	PFPS group (*n*)	Control group (*n*)	Test	SEN % (95%IC)	SPE % (95%IC)	LR+ (95%IC)	LR− (95%IC)	PV+ % (95%IC)	PV− % (95%IC)
Haim et al., 2006	10	61	25	Patellar tilt	43 (31–55)	92 (75–98)	5.4 (1.4–20.8)	0.6 (0.5–0.8)	93 (75–99)	40 (27–53)
				Active instability	25 (17–37)	100 (87–100)	[a]	0.8 (0.6–0.9)	100 (75–100)	35 (25–48)
				Patella alta test	49 (37–61)	72 (52–86)	1.8 (0.9–3.5)	0.7 (0.5–1.0)	81 (64–91)	37 (24–52)
				Patellar apprehension test	7 (3–16)	92 (75–98)	0.9 (0.2–4.2)	1.0 (0.9–1.2)	67 (24–94)	29 (20–40)
Näslund et al., 2006	11	29	17	Compression test	83 (66–92)	18 (6–41)	1.0 (0.8–1.3)	1.0 (0.3–3.6)	63 (46–77)	38 (10–74)
				Tenderness medial	48 (31–66)	71 (47–87)	1.6 (0.7–3.8)	0.7 (0.5–1.2)	74 (49–90)	44 (26–64)
				Tenderness lateral	41 (26–59)	71 (47–87)	1.4 (0.6–3.3)	0.8 (0.5–1.3)	71 (44–89)	41 (24–61)
				Passive gliding patellar	48 (31–66)	47 (26–69)	0.9 (0.5–1.6)	1.1 (0.6–2.0)	61 (39–80)	35 (17–57)
Nijs et al., 2006	11	31	28	Vastus medialis coordination test	16 (6–35)	93 (75–99)	2.3 (1.9–2.9)	0.9 (0.6–0.9)	71 (30–95)	50 (36–64)
				Patellar apprehension test	32 (17–52)	86 (66–95)	2.3 (2.1–2.5)	0.8 (0.8–1.0)	71 (42–90)	53 (38–68)
				Waldron's test phase 1	45 (28–64)	68 (48–83)	1.4 (0.6–3.2)	0.8 (0.4–1.8)	61 (39–80)	53 (36–69)
				Waldron's test phase 2	23 (10–42)	79 (59–91)	1.1 (1.0–1.1)	1.0 (0.9–1.1)	54 (26–80)	48 (33–63)
				Clarke's test	48 (31–67)	75 (55–89)	1.9 (1.1–3.6)	0.7 (0.4–1.3)	68 (45–85)	57 (40–73)
				Eccentric step test	42 (25–61)	82 (62–93)	2.3 (1.9–2.9)	0.7 (0.6–0.9)	72 (46–89)	56 (40–71)
Cook et al., 2010	11	54	24	Manual compression	68 (54–79)	54 (35–72)	1.5 (0.9–2.3)	0.6 (0.3–1.0)	75 (67–82)	46 (33–57)
				Palpation	47 (33–60)	68 (47–82)	1.5 (0.9–2.8)	0.8 (0.6–1.1)	75 (63–85)	39 (30–46)
				Resisted isometric quadriceps muscle contraction	39 (27–52)	82 (64–93)	2.2 (1.0–5.2)	0.8 (0.6–1.1)	82 (67–91)	40 (32–44)
				Squatting	91 (79–96)	50 (31–69)	1.8 (1.3–2.3)	0.2 (0.1–0.4)	79 (73–82)	74 (54–87)
				Stair climbing	75 (62–85)	43 (25–61)	1.3 (1.0–1.9)	0.6 (0.0–1.1)	73 (66–79)	46 (30–60)
				Kneeling	84 (73–92)	50 (31–69)	1.7 (1.2–2.4)	0.3 (0.2–0.6)	79 (71–83)	61 (44–75)
				Prolonged sitting	72 (58–82)	57 (39–76)	1.7 (1.1–2.7)	0.5 (0.3–0.8)	77 (60–84)	50 (37–60)
Sweitzer et al., 2010	8	59	23	Patellar translation superior-inferiorly	63 (56–69)	56 (39–72)	1.4 (0.9–2.5)	0.7 (0.4–1.1)	79 (70–86)	37 (26–47)
				Patellar translation medial-laterally	54 (47–59)	69 (52–83)	1.8 (0.9–3.6)	0.7 (0.5–1.0)	82 (72–90)	37 (27–45)
				Patellar inferior pole tilt	19 (13–22)	83 (68–93)	1.1 (0.4–3.0)	0.9 (0.8–1.3)	73 (51–89)	28 (23–32)
				Patellar tendon mobility	49 (43–53)	83 (66–93)	2.8 (1.3–7.3)	0.6 (0.5–0.9)	88 (76–95)	39 (31–44)
Meta-analyses	N/A	92	53	Patellar apprehension test (Haim et al., 2006; Nijs et al., 2006)	15 (9–24)	89 (77–95)	1.3 (0.6–3.3)	1.0 (0.8–1.1)	70 (46–87)	38 (29–47)

Numbers in bold represent the best diagnostic measure. Abbreviation: SEN = Sensitivity; SPE = Specificity; LR+ = Positive likelihood ratio; LR− = Negative likelihood ratio; PV+ = Positive predictive value; PV− = Negative predictive value.

[a] Because specificity was 100%, it was not possible to calculate LR+.

Table 4
Diagnostic accuracy of the combination tests.

Study	Tests	Combination tests	SEN % (95%IC)	SPE % (95%IC)	LR+ (95%IC)	LR− (95%IC)	PV+ % (95%IC)	PV− % (95%IC)
Cook et al., 2010	Manual compression Palpation	Pain in the resisted isometric and squatting (2 of 2)	35 (23–48)	89 (69–96)	3.3 (1.2–9.2)	0.7 (0.6–0.9)	87 (71–95)	40 (34–43)
	Resisted isometric quadriceps muscle contraction Squatting Stair climbing Kneeling Prolonged sitting	Pain in two test: resisted isometric. squatting and/or kneeling (2 of 3)	60 (46–72)	85 (64–93)	4.0 (1.8–10.3)	0.5 (0.4–0.7)	89 (79–95)	50 (40–46)
		Pain in the resisted isometric. squatting and palpation (3 of 3)	33 (22–46)	89 (69–96)	3.1 (1.1–9.5)	0.7 (0.6–0.9)	86 (69–95)	39 (34–43)
Sweitzer et al., 2010	Patellar translation superior-inferiorly	1 of 4 positive findings	75 (68–81)	48 (32–64)	1.4 (1.0–2.2)	0.5 (0.3–1.0)	79 (71–85)	42 (30–56)
	Patellar translation medial-laterally	2 of 4 positive findings	53 (46–58)	69 (52–83)	1.7 (0.9–3.5)	0.7 (0.5–1.0)	82 (71–89)	36 (27–44)
	Patellar inferior pole tilt	3 of 4 positive findings	41 (34–44)	83 (67–93)	2.4 (1.0–6.2)	0.7 (0.6–1.0)	86 (72–94)	35 (28–39)
	Patellar tendon mobility	4 of 4 positive findings	17 (12–19)	91 (78–97)	1.9 (0.5–7.7)	0.9 (0.8–1.1)	83 (58–95)	30 (25–32)

Abbreviations: SEN = Sensitivity; SPE = Specificity; LR+ = Positive likelihood ratio; LR− = Negative likelihood ratio; PV+ = Positive predictive value; PV− = Negative predictive value.

all of those who did not suffer from PFPS, i.e., it was 100% specific (Haim et al., 2006). Nevertheless, its specificity should be considered with caution since it was based on results with a control group that had no alterations in the evaluated knee. Therefore, if Haim et al. (2006) are not considered, the test with the highest specificity for diagnosing PFPS was the vastus medialis coordination test (93%) (Nijs et al., 2006). Taking into consideration the positive and negative likelihood ratio analyses, that actually determine the accuracy of tests (Jaeschke et al., 1994), the squatting test (Cook et al., 2010) continued to present good results. It demonstrated a strong diagnostic tendency (LR− = 0.2) that was not observed in the vastus medialis coordination test. Another test with a strong diagnostic tendency was the patellar tilt test (LR+ = 5.4) (Haim et al., 2006), although this result should also be considered cautiously since, as previously cited, Haim et al. (2006) control group was of healthy individuals.

Due to the tests' lack of accuracy, an association of tests has been proposed to evaluate PFPS (Fredericson & Yoon, 2006), and two of the studies included in this review considered the joint results of the tests (Cook et al., 2010; Sweitzer et al., 2010). These studies presented higher accuracy values, but still no diagnostic evidence (Jaeschke et al., 1994). In Cook et al. (2010), three possible combinations were presented: 1) pain during an isometric contraction of the quadriceps and while squatting (2 out of 2); 2) pain in at least two of the following movements: isometric contraction of the quadriceps, while squatting and during kneeling (2 out of 3); 3) pain in isometric contraction of the quadriceps, while squatting and during palpation (3 out of 3). The following results stood out: a specificity of 89% in the 2 out of 2 and 3 out of 3 combinations, an LR+ of 4.0 in the 2 out of 3 combination and the ability to identify affected individuals (PV+ = 87% in 2 out of 2; PV+ = 89% in 2 out of 3). In Sweitzer et al. (2010) there was great specificity in a combination of the positive results of all four evaluated tests (91%), and each combination presented high PV+ values (≥79%).

Nevertheless, combinations of tests were not able to diagnose PFPS. The accuracy of the combinations proposed by Cook et al. (2010) and Sweitzer et al. (2010) was no better than the squatting test. This suggests that the combinations of tests presented by Cook et al. (2010) and Sweitzer et al. (2010) are not the best form of clinical evaluation, in contrast to the recommendations of Fredericson and Yoon (2006). In their narrative review, these authors searched for the main alterations presented by individuals affected by PFPS and suggested that the best form of diagnosis would be an association of several tests. However, no measurement data regarding the accuracy of the tests or their proposed combinations was reported. Thus, two possibilities can be inferred: 1) the combinations of movements and tests are no better than the individual tests and, therefore, their joint analysis is not necessary; 2) other possible test combinations should be measured with respect to diagnostic properties.

Among the tests in the included studies, one was analyzed in two different studies. Haim et al. (2006) and Nijs et al. (2006) applied the patellar apprehension test in an identical fashion, which allows their results to be combined. The meta-analysis of this test revealed no measurement indicative of diagnostic accuracy. A relevant result was observed for specificity (89%), but this as well as the other data from the meta-analysis should be interpreted cautiously due to the low methodological quality of the study by Haim et al. (2006). Within the five included studies, 24 different tests were analyzed and only the patellar apprehension test was performed in similar fashion in two occasions (Haim et al., 2006; Nijs et al., 2006), in which case a meta-analysis was viable. Meta-analyses are used to strengthen a final conclusion in regards to what is being analyzed, thus, the unviability of meta-analyses in the current study raises the concern that tests for PFPS lack standard procedures and that could help to explain consistent results within studies.

Based on the accuracy values observed in the studies included in this review, the presence of pain during the squatting test comes closest to diagnosing PFPS, although it cannot be considered sufficient as a basis for PFPS diagnosis. Therefore, according to the results, diagnosing PFPS solely from the results of these clinical tests is not possible; the addition of complementary exams, such as imaging tests, as well as interpretation of the information obtained during evaluation are required. The lack of diagnostic evidence for the tests involved in this review may stem from searching in only four databases. It could be that this approach overlooked other eligible studies indexed in different databases. However, the selected databases are the most representative for the studied conditions since they have the most health-related articles indexed. Another possible reason for the lack of tests conclusively diagnosing PFPS could be the heterogeneity in sample selection in the included studies. Such heterogeneity is likely to come from the difficulty in defining PFPS because currently a patient is most likely to be diagnosed with PFPS only when nothing else is found.

5. Conclusion

Due to the multifactorial etiology of PFPS, a number of tests have been developed for its diagnosis. This review found no PFPS test

with diagnostic consistency, which thus prohibits inferences about the best test to use. Future studies should focus on or address sample homogeneity and test standardization so that new systematic reviews with meta-analysis can more clearly determine the tests' accuracy in diagnosing PFPS.

Conflict of interest

The authors report no conflict of interest.

Funding

None declared.

References

Alqarni, A. M., Schneiders, A. G., & Hendrick, P. A. (2011). Clinical tests to diagnose lumbar segmental instability: a systematic review. *The Journal of Orthopaedic and Sports Physical Therapy, 41*(3), 130–140.

Bohnsack, M., Hurschler, C., Demirtas, T., Rühmann, O., Stukenborg-Colsman, C., & Wirth, C. J. (2005). Infrapatellar fat pad pressure and volume changes of the anterior compartment during knee motion: possible clinical consequences to the anterior knee pain syndrome. *Knee Surgery, Sports Traumatology, Arthroscopy: Official Journal of the ESSKA, 13*(2), 135–141.

Bohnsack, M., Klages, P., Hurschler, C., Halcour, A., Wilharm, A., Ostermeier, S., et al. (2009). Influence of an infrapatellar fat pad edema on patellofemoral biomechanics and knee kinematics: a possible relation to the anterior knee pain syndrome. *Archives of Orthopaedic and Trauma Surgery, 129*(8), 1025–1030.

Boling, M. C., Padua, D. A., Marshall, S. W., Guskiewicz, K., Pyne, S., & Beutler, A. (2009). A prospective investigation of biomechanical risk factors for patellofemoral pain syndrome: the Joint Undertaking to Monitor and Prevent ACL Injury (JUMP-ACL) cohort. *The American Journal of Sports Medicine, 37*(11), 2108–2116.

Boling, M., Padua, D., Marshall, S., Guskiewicz, K., Pyne, S., & Beutler, A. (2010). Gender differences in the incidence and prevalence of patellofemoral pain syndrome. *Scandinavian Journal of Medicine & Science in Sports, 20*(5), 725–730.

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*(1), 17–24.

Costa, L. O. P., Maher, C. G., Lopes, A. D., de Noronha, M. A., & Costa, L. C. M. (2011). Transparent reporting of studies relevant to physical therapy practice. *Revista Brasileira de Fisioterapia, 15*(4), 267–271.

Deeks, J. J. (2001). Systematic reviews in health care: systematic reviews of evaluations of diagnostic and screening tests. *British Medical Journal (Clinical Research Ed.), 323*(7305), 157–162.

Fredericson, M., & Yoon, K. (2006). Physical examination and patellofemoral pain syndrome. *American Journal of Physical Medicine & Rehabilitation/Association of Academic Physiatrists, 85*(3), 234–243.

Fritz, J. M., & Wainner, R. S. (2001). Examining diagnostic tests: an evidence-based perspective. *Physical Therapy, 81*(9), 1546–1564.

Greenhalgh, T. (1997). How to read a paper. Papers that report diagnostic or screening tests. *British Medical Journal (Clinical Research Ed.), 315*(7107), 540–543.

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.

Harbord, R. M., Whiting, P., Sterne, J. A. C., Egger, M., Deeks, J. J., Shang, A., et al. (2008). An empirical comparison of methods for meta-analysis of diagnostic accuracy showed hierarchical models are necessary. *Journal of Clinical Epidemiology, 61*(11), 1095–1103.

Hayden, S. R., & Brown, M. D. (1999). Likelihood ratio: a powerful tool for incorporating the results of a diagnostic test into clinical decisionmaking. *Annals of Emergency Medicine, 33*(5), 575–580.

Hegedus, E. J., Cook, C., Hasselblad, V., Goode, A., & McCrory, D. C. (2007). Physical examination tests for assessing a torn meniscus in the knee: a systematic review with meta-analysis. *Journal of Orthopaedic and Sports Physical Therapy, 37*(9), 541–550.

Herbert, R., Jamtvedt, G., Mead, J., & Hagen, K. B. (2005). *Practical evidence-based physiotherapy* (1st ed.). Oxford: Butterworth-Heinemann.

Hunter, D., Zhang, Y., Niu, J., Felson, D., Kwoh, K., Newman, A., et al. (2007). Patella malalignment, pain and patellofemoral progression: the Health ABC Study. *Osteoarthritis and Cartilage/OARS, Osteoarthritis Research Society, 15*(10), 1120–1127.

Ivković, A., Franić, M., Bojanić, I., & Pećina, M. (2007). Overuse injuries in female athletes. *Croatian Medical Journal, 48*(6), 767–778.

Jaeschke, R., Guyatt, G. H., & Sackett, D. L. (1994). Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA: The Journal of the American Medical Association, 271*(9), 703–707.

Lankhorst, N. E. (2011). Risk factors for patellofemoral pain syndrome: a systematic review. *Journal of Orthopaedic and Sports Physical Therapy, 42*(2), 81–94.

Leeflang, M. M. G., Deeks, J. J., Gatsonis, C., & Bossuyt, P. M. M. (2008). Systematic reviews of diagnostic test accuracy. *Annals of Internal Medicine, 149*(12), 889–897.

Levinger, P., Gilleard, W., & Coleman, C. (2007). Femoral medial deviation angle during a one-leg squat test in individuals with patellofemoral pain syndrome. *Physical Therapy in Sport, 8*(4), 163–168.

Loudon, J. K., Wiesner, D., Goist-Foley, H. L., Asjes, C., & Loudon, K. L. (2002). Intrarater reliability of functional performance tests for subjects with patellofemoral pain syndrome. *Journal of Athletic Training, 37*(3), 256–261.

Manske, R. C., & Lehecka, B. J. (2012). Evidence – based medicine/practice in sports physical therapy. *International Journal of Sports Physical Therapy, 7*(5), 461–473.

Miyamoto, G. C., Soriano, F. R., & Cabral, C. M. N. (2010). Alongamento muscular segmentar melhora função e alinhamento do joelho de indivíduos com síndrome femoropatelar: estudo preliminar. *Revista Brasileira de Medicina do Esporte, 16*(4), 269–272.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2010). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *International Journal of Surgery (London, England), 8*(5), 336–341.

Myer, G. D., Ford, K. R., Barber Foss, K. D., Goodman, A., Ceasar, A., Rauh, M. J., et al. (2010). The incidence and potential pathomechanics of patellofemoral pain in female athletes. *Clinical Biomechanics (Bristol, Avon), 25*(7), 700–707.

Näslund, J., Näslund, U.-B., Odenbring, S., & Lundeberg, T. (2006). Comparison of symptoms and clinical findings in subgroups of individuals with patellofemoral pain. *Physiotherapy Theory and Practice, 22*(3), 105–118.

Nijs, J., Van Geel, C., Van der Auwera, C., & Van de Velde, B. (2006). Diagnostic value of five clinical tests in patellofemoral pain syndrome. *Manual Therapy, 11*(1), 69–77.

Page, P. (2012). Research designs in sports physical therapy. *International Journal of Sports Physical Therapy, 7*(5), 482–492.

Pattyn, E., Verdonk, P., Steyaert, A., Van den Bossche, L., Van den Broecke, W., Thijs, Y., et al. (2011). Vastus medialis obliquus atrophy: does it exist in patellofemoral pain syndrome? *The American Journal of Sports Medicine, 39*(7), 1450–1455.

Piva, S. R., Fitzgerald, K., Irrgang, J. J., Jones, S., Hando, B. R., Browder, D. A., et al. (2006). Reliability of measures of impairments associated with patellofemoral pain syndrome. *BMC Musculoskeletal Disorders, 7*(33).

Reid, D. C. (1993). The myth, mystic, and frustration of anterior knee pain. *Clinical Journal of Sport Medicine, 3*(3), 139.

Reneker, J., Paz, J., Petrosino, C., & Cook, C. (2011). Diagnostic accuracy of clinical tests and signs of temporomandibular joint disorders: a systematic review of the literature. *The Journal of Orthopaedic and Sports Physical Therapy, 41*(6), 408.

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*(5), 477–482.

Santos, G. M., Ries, L. G. K., Sperandio, F. F., Say, K. G., Pulzatto, F., & Monteiro-Pedro, V. (2011). Onset of electromyographic activity of patellar stabilizers during gait in subjects with and without patellofemoral pain syndrome. *Fisioterapia em Movimento, 24*(1), 125–132.

Sweitzer, B. A., Cook, C., Steadman, J. R., Hawkins, R. J., & Wyland, D. J. (2010). The inter-rater reliability and diagnostic accuracy of patellar mobility tests in patients with anterior knee pain. *The Physician and Sportsmedicine, 38*(3), 90–96.

Tállay, A., Kynsburg, A., Tóth, S., Szendi, P., Pavlik, A., Balogh, E., et al. (2004). Prevalence of patellofemoral pain syndrome. Evaluation of the role of biomechanical malalignments and the role of sport activity. *Orvosi Hetilap, 145*(41), 2093–2101.

Tenforde, A. S., Sayres, L. C., McCurdy, M. L., Collado, H., Sainani, K. L., & Fredericson, M. (2011). Overuse injuries in high school runners: lifetime prevalence and prevention strategies. *PM&R: The Journal of Injury, Function, and Rehabilitation, 3*(2), 125–131.

Thomeé, R., Augustsson, J., & Karlsson, J. (1999). Patellofemoral pain syndrome: a review of current issues. *Sports Medicine (Auckland, N.Z.), 28*(4), 245–262.

Ward, S. R., & Powers, C. M. (2004). The influence of patella alta on patellofemoral joint stress during normal and fast walking. *Clinical Biomechanics, 19*(10), 1040–1047.

Whiting, P., Rutjes, A. W. S., Reitsma, J. B., Bossuyt, P. M. M., & Kleijnen, J. (2003). The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Medical Research Methodology, 3*(25).

Wills, A. K., Ramasamy, A., Ewins, D. J., & Etherington, J. (2004). The incidence and occupational outcome of overuse anterior knee pain during army recruit training. *Journal of the Royal Army Medical Corps, 150*(4), 264–269.

Wood, L., Muller, S., & Peat, G. (2011). The epidemiology of patellofemoral disorders in adulthood: a review of routine general practice morbidity recording. *Primary Health Care Research & Development, 12*(2), 157–164.