A Systematic Review of Reviews in Patellofemoral Pain Syndrome. Exploring the Risk Factors, Diagnostic Tests, Outcome Measurements and Exercise Treatment

Konstantinos Papadopoulos*1,2, Demetris Stasinopoulos1 and Dimitar Ganchev2

1School of Sciences, Department of Health Sciences, Physiotherapy programme, European University of Cyprus 6, Diogenes Str. Engomi, Nicosia, Cyprus
2Department of Kinesiotherapy, National Sports Academy (NSA) Vassil Levski, zh.k. Studentski grad, 1700, Sofia, Bulgaria

Abstract: Background: Literature has shown a growing number of published studies on Patellofemoral Pain Syndrome every year. The increasing evidence base has revealed a significant number of reviews which makes it confusing for clinicians and researchers to choose from the best evidence. This study aimed to gather the reviews on Patellofemoral Pain Syndrome and provide information about the most common clinical tests, risk factors, exercise treatment and outcome measures. In addition, secondary questions aimed to report the study settings and patient characteristics of the primary included studies.

Methods: Studies eligible for this Review of Reviews were those published from 1993 to July 2013. Databases searched included PubMed, CINAHL, SPORTDiscus, Pedro and the Cochrane Library. Four key areas were searched using the words: PFPS, Anterior Knee Pain (AKP) or Chondromalacia patella (CP), plus a keyword. The keywords for the four research topics were a) risk factors; b) exercise treatment; c) diagnostic clinical tests and d) psychometric outcome measurements. Only reviews with clear methodological strategy were included. A two-stage evaluation was performed in each review assessment. At first, the level of evidence was graded and then the methodological quality of each review was assessed.

Results: Full text screening revealed that only 18 reviews were eligible for use. The selected reviews included 213 primary studies. After excluding duplicates, 144 primary studies were screened to answer the secondary questions. A gold standard clinical test for PFPS assessment cannot be reached and the use of functional tasks should be considered. The quadriceps strength deficits are still the only evidence based risk factors along with the dynamic malalignment of the lower limb. More research is still required on strength and flexibility deficits of other lower limb muscles. The quadriceps-based exercises are still the only ones to have strong evidence together with hamstrings, quadriceps, gastrocnemius and anterior hip muscles stretching. Finally, the usage of Activities of Daily Living Scale is recommended as the best outcome measure.

Conclusion: There is no evidence on whether the above treatment and assessment methods should be used in sedentary people or differently across population groups or gender. More RCTs with large populations, powered correctly, in clinical environments are called for in further research.

Keywords: Anterior knee pain, clinical tests, exercise treatment, outcome measures, patellofemoral pain syndrome, participant characteristics, risk factors.

BACKGROUND

The literature shows a growing number of published studies on healthcare interventions every year [1]. One category of healthcare intervention that contains nebulous pathophysiology is Patellofemoral Pain Syndrome (PFPS) [2]. A quick database search shows that in 2012 there are more than 70 research studies published on PFPS, while there were less than 50 in 2008. This rapidly growing evidence base makes it hard for people involved in providing care to choose from the best quality studies when making decisions [3]. Researchers have identified this problem and started gathering these studies in systematic reviews in order to appraise and summarise evidence [3]. Recently, a need for ‘rapid reviews’ to provide decision-makers with evidence has become apparent, however; these reviews can be sometimes problematic (in terms of their quality) compared to full systematic reviews [4]. As the number of reviews began to grow, so have the number of protocols for the conduct of systematic reviews. One such an example is the Cochrane Collaboration which in 2008 included only one review on PFPS. Since then, eight more reviews have been published. Lately, researchers and decision makers have
started to confront an overflow of reviews [5, 6]. From July 2012 to March 2013 the British Journal of Sport Medicine alone has published 4 systematic reviews on PFPS [7-10]. Because there are several reviews on similar topics and because it is likely for reviews to be of varied scope and quality, a growing interest of systematic review of reviews (RoR) has been called for [3]. For example, there are several reviews on PFPS conservative/non-operative treatment [11, 12], suggested exercises for PFPS treatment [13, 14], clinical tests for PFPS diagnosis [15, 16], and factors associated with PFPS [9, 17]. Only one previous RoR on PFPS has been published to date and that was about the quality of the systematic reviews on nonpharmacological conservative treatment for patellofemoral pain syndrome [18].

RATIONALE

This systematic RoR was conducted to provide background information as a basis for the design of future studies but also as evidence to compare and contrast recent results.

A previous mixed method study [19] showed that physiotherapists used a lot of different ways to assess, treat and measure outcomes in PFPS. This RoR was designed to report what literature suggests regarding PFPS assessment, treatment and outcome measures. In addition, literature showed inability to answer secondary questions such as whether subjects in the research studies had specific characteristics (for example gender, physical activity levels and attitude towards sports) and whether these characteristics are the same as those clinical physiotherapists treat in the clinic. Additionally, other questions which arose after searching the literature included whether researchers conduct their studies in clinical or research laboratory-environments and whether these settings vary across different countries. Finally, the RoR was designed to investigate how other researchers assessed strength and whether they used a dynamometer.

OBJECTIVES

The overarching aim was to conduct a systematic RoR to identify all published systematic reviews, reporting evidence on PFPS risk factors, diagnostic clinical tests, the clinometric properties of outcome measures and treatment (primary questions). The secondary aim was to determine the context and characteristics of participants in included studies.

Following the Smith et al. [3] approach for conducting a systematic review of systematic reviews in healthcare interventions, the PICOS (participants, interventions, comparisons, outcomes and study design) structure was used:

REVIEW OBJECTIVES

1. To determine the risk factors for patients with PFPS reported in Randomised Control Trials (RCTs), case control, cohort, cases studies, case series, expert’s opinion and formal consensus studies.

2. To identify the clinical tests used for the diagnosis of PFPS reported in RCTs, case control, cohort, cases studies, case series, expert’s opinion and formal consensus studies.

3. To identify the outcome measurements used in PFPS, and their clinometric properties reported in RCTs, case control, cohort, cases studies, case series, expert’s opinion and formal consensus.

4. To determine the effectiveness of exercise-based interventions for PFPS reported in RCTs, case control, cohort, cases studies, case series, expert’s opinion and formal consensus studies.

5. To determine the characteristics of included study participants (adults) and the individual included study context.

REVIEW QUESTIONS

1. ‘What are the risk factors for adult patients with PFPS across study types and patient groups’?

2. ‘What diagnostic clinical tests are used for adult patients with PFPS’?

3. ‘What outcome measures are used in adult patients with PFPS’ across study types and patient groups’?

4. ‘What exercises are effective for adult patients with PFPS’?

5. ‘What are the characteristics/demographics of patients that studies recruit’?

6. ‘In what settings and in which countries were studies conducted’?

7. ‘What was the sample size in each included study’?

8. ‘Was a dynamometer used to measure strength, and if so what type and what setting was it used’?

METHODS

The systematic RoR approach of Smith et al. [3] was used, with additional data mining to extract specific evidence of interest from the primary studies included in the systematic reviews.

The Smith approach [3] uses standard systematic review processes to identify and appraise reviews, describe the quality of the evidence base, summarise and compare the review’s conclusions and discuss the strength of these conclusions. The Preferred Reporting Items of Systematic reviews Meta-Analyses (PRISMA) guidelines formerly known as QUOROM (Quality Of Reporting Of Meta-analysis) was adopted when conducting this review [6]. PRISMA consists of a 27-item checklist and a four-phase flow diagram about identification, screening, eligibility and inclusion. PRISMA items which cannot be used in this systematic RoR (e.g. question 21 about meta-analysis results) are not presented. All other questions included in the PRISMA checklist were used.
Protocol and Registration

The protocol and the questions of the current study were designed beforehand; however, the protocol was not published or uploaded on the web.

Eligibility Criteria

Inclusion criteria

- Only reviews with a clear search strategy and detailed references [20] which attempted to collate all empirical evidence that fitted pre-specified eligibility criteria in order to answer a specific questions, were selected [21].
- Only reviews in English language were obtained.
- No restriction regarding the origin of the systematic reviews (country) was imposed.
- Only reviews with available full text were included.
- Reviews which include the following study design papers: systematic reviews, meta-analyses, randomized controlled trials (RCTs), case-control studies, cohort studies, case series, formal consensus and expert’s opinion were included.
- Study population: adults with PFPS. Participant can be either patients of National Health Service or patients visiting a private clinic or private practice physiotherapists.
- Case definition: PFPS, AKP (Anterior Knee Pain), CP (Chondromalacia Patellae), if the authors intended AKP or CP to be a description for PFPS.

Exclusion Criteria

- Study design: Studies with no clear search strategy and detailed reference. Studies that did not report clear methodology.
- Study population: non-humans, or people under the age of 18.
- Case definition: Studies focusing on other named knee pathologies (such as Osgood Schlatter disease, Sinding Larsen-Johansson’s disease, tendinitis or bursitis, intra-articular pathologies, plica syndromes and rarely occurring pathologies) were excluded.

Studies eligible for this RoR were those published from 1993 to July 2013. The reason for the 1993 year selection was that according to Lichtenstein et al. [22] the search needs to cover at least two decades. In addition, the earliest PFPS reviews identified by quick search were published in 1993. In earlier years (1955 to 1980) surgical interventions were more common than non-operative treatment and most of the articles were written from an orthopaedic standpoint [23, 24].

Information Sources/Search

Databases searched included PubMed, CINAHL, SPORTDiscus, Pedro and the Cochrane Library. Reference lists of highlighted recent reviews were manually searched to identify additional reviews [25]. Four key areas were searched using the words: PFPS, Anterior Knee Pain (AKP) or Chondromalacia patella (CP), plus a keyword. The keywords for the four research topics were a) risk factors; b) exercise treatment; c) diagnostic clinical tests and d) psychometric outcome measurements. Because this study aimed to identify reviews only, the last keyword in every search was the word ‘review’.

Study Selection

For the first level of screening, one reviewer KP read the titles of all the available citations obtained from the electronic database search and reference lists and removed all the citations which were not related to PFPS. The second level of screening involved the screening of abstracts, and was conducted by two reviewers KP and DS. Along with the abstracts, full-text articles were also obtained for the reviews which did not clearly meet the eligibility criteria. When, even after analysing the full text, the eligibility of an article remained uncertain, it was planned to ask a third reviewer to undertake a full analysis; However, this contingency was not required [3].

Critical Appraisal of Reviews

In order to assess the quality of a paper, there is a need to consider not only the type of evidence (i.e. randomised trials, pilots, cross-sectional or others) but also how good the methodology of each review is. Therefore, a two-stage evaluation was performed in each review assessment. At first, the level of evidence was graded [3] and then the methodological quality of each review was assessed [26]. In order to assess the level of evidence, systematic reviews with randomized controlled trials were established as first in the hierarchy (gold standard) and then an evidence grade was given to each review based on the Scottish Intercollegiate Guidelines Network (SIGN) [27]. This framework is a widely used critical appraisal and evidence hierarchy which has the advantage of being simple and clear to use. The level of evidence can be graded from 1++ to 2- depended on the study design. Only the first 4 categories assess review studies; therefore, the reviews of this study were graded from 1++ to 2++.

According to Clarke [28], the methodological conduct of a review plays an important role in the successful interpretation of results from systematic reviews. Therefore, in order to present high methodological evidence this review used a MeaSurement Tool to Assess systematic Reviews (AMSTAR). The AMSTAR tool [26] was selected as the most acceptable to critically appraise the methodology of systematic reviews. AMSTAR is used by a number of groups such as the Canadian Agency for Drugs and Technologies in Health and The Cochrane Effective Practice and Organization of Care Group. The tool consists of 11 items and was created to assess the methodological quality of systematic reviews and found to have good inter-rater reliability (Kappa scores >0.8) and good face, content and construct validity [26]. Each item is given a score of 1 if the specific criterion is met, or a score of 0 if the criterion is not met, is unclear, or is not applicable. According to the total score that each systematic review received the methodological quality of included reviews was assessed and rated as low, medium or high quality. AMSTAR characterises quality at three levels. Levels 0-3 can be considered as low quality, 4-7 as medium quality and 8-11 as high quality [29]. In this RoR high, moderate and low quality levels were utilised.
Although PRISMA is not a quality assessment instrument for systematic reviews it can be useful for critical appraisal purposes. Consequently, the study also set out to identify which of the systematic reviews used PRISMA to report their data and how many of the 27 items have been used. This information was considered supplementary and was not used as a criterion to include or exclude any reviews. In case of any disagreement regarding grading of evidence, quality appraisal of reviews or effectiveness of the intervention, consensus was reached by discussion between KP and DS.

Data Collection Process

Titles and abstracts were screened for eligibility according to the aforementioned inclusion and exclusion criteria. When the appropriateness of some reviews was not clear, the full text was obtained.

Data Items

The variables for which data were sought included reviews of RCT’s, case-control or cohort studies referred to the risk factors, outcome measures, exercise treatment and clinical tests of patients with PFPS. It was assumed that the AKP and CP patients used by some reviews had the same characteristics with PFPS patients. Studies could be funded by an external source or not.

Risk of Bias in Individual Studies

A specific number of studies in a review were not an inclusion criterion. Therefore no risk of bias or publication bias was assessed.

Summary Measures

Data with and without significant differences were presented in this study. The level of significance was reported according to the statistical methods each of the review used i.e. for strength difference p value was used or percentage of deficiency, for reliability and validity intra-class correlation coefficient or Pearson correlation was used.

Additional Analyses

The additional analyses included the search for evidence regarding the secondary questions of this review (5-8 from the review questions). All secondary questions that this review was designed to answer were not reported according to statistical significance but according to the categories each of the question could be separated into (e.g. research place=clinic or research centre). Only systematic reviews with clear study report were used in this section.

RESULTS

Study Selection

The comprehensive search strategy identified 246 reviews; 86 on exercise treatment, 95 on outcome measures, 48 on risk factors and 17 on clinical tests. Five more reviews were identified through references of other reviews. After the titles of the reviews were screened only 59 of them met the inclusion criteria (Supplementary Material 1).

After screening titles and abstracts 31 reviews remained. Full text screening revealed that only 18 reviews were eligible for use; 2 regarding outcome measures, 3 about clinical tests, 7 about exercise treatment, 4 about risk factors, 1 about clinical tests and risk factors and 1 regarding all 4 components of this study. The flow diagram in Fig. (1) shows the procedure from identification to inclusion.

Study Characteristics

Examination of the 13 reviews that were excluded after full texts were obtained revealed 11 with no clear methodology about how the included studies were gathered, one review with combined exercise treatment and other treatment not suitable for this review and one review with combined exercise and drug treatment. From the 18 included studies, 7 were entitled as systematic reviews, another 7 as systematic review with meta-analysis, one as a critical review, one as an invited review and the other two as narrative reviews. The four studies which were not entitled as systematic [30-33] were included because they reported systematic methodology in the way the studies were gathered. However, their purpose was not to answer a single question, therefore in some cases the subject matter related to more than one of the research questions of the present review. Supplementary Material 2 shows all 31 full-text assessed reviews, their topic, design and which reviews were included and which not.

Critical Appraisal

The level of evidence for the selected reviews is presented in Table 1. No level of evidence was reported for the four reviews which were not systematic or meta-analysis. The results showed two clinical test reviews of high quality systematic reviews of case control or cohort studies [34, 35], one high quality meta-analysis of RCT’s [36] and three high quality systematic reviews of case control or cohort studies on risk factors [9, 17, 37], three high quality meta-analyses or systematic reviews of RCT’s [11, 38, 39] and three high quality systematic reviews of case control or cohort studies on exercise treatment [12, 13, 40] and two high quality systematic reviews of case control or cohort studies on outcome measures [41, 42].

The methodological quality of the selected reviews showed that there was no review that answered all 11 questions of the AMSTAR tool or reported all 27 items of the PRISMA checklist. The only review which stood out from the rest was the one published by Collins et al. [11]. The reason was that this study reported to have a priori design (AMSTAR question-1, PRISMA checklist item-5). On the opposite side, the two studies from Lankhorst et al. [9, 36] were the only studies which assessed the likelihood of publication bias (AMSTAR Question-10, PRISMA checklist item-15). All studies performed a comprehensive literature search (question 3 AMSTAR) which was anticipated because this criterion was one of the inclusion criteria for the current review. In addition, most of the studies only reported the included studies and not those which were excluded (AMSTAR question-5). Overall, the
AMSTAR tool revealed five studies with low (0-3) four studies with moderate (4-7) and 9 with high (8-11) methodological quality (Table 1).

**Clinical Tests**

Three non-systematic reviews of low level of evidence and two high quality systematic reviews of case reports or
cohort studies were found (Supplementary Material 3). None of the non-systematic reviews could be appraised by the SIGN tool, their methodological quality was found to be low (less than 3/11 in the AMSTAR tool) whilst, all three reviews only reported a few items from the PRISMA checklist (less than 5). One of the two systematic reviews also reported meta-analysis on one clinical test [34]. Both systematic reviews were also graded with high methodological quality (more than 8) and with 15 out of 27 items of the PRISMA checklist. The non-systematic reviews, reported different clinical tests for PFPS. The most common were the q-angle, tilting and patellar compression. The q-angle and tilting tests were reported by Fredericson and Yoon [31] and Selfe [33] and were found to have low reliability. Patellar compression was reported by Fredericson and Yoon [31] and Malanga et al. [30]. The first authors reported low sensitivity and specificity. The latter suggested this test without evidence provided. It is important to mention that the Malanga et al. [30] study aimed to identify several clinical test of the knee; and some of them were for PFPS. Only Fredericson and Yoon [31] reported functional tests which were found to be highly reliable. Generally, the non-systematic reviews concluded that there is no evidence to support which is the best clinical test for PFPS and the reliability or sensitivity of the tests was low or untested.

The two recent systematic reviews examined a series of clinical tests (24 and 22) some of which were similar, i.e. squatting, active instability. Nunes et al. [34] reported than none of the tests were good enough for diagnostic purposes because of the lack of homogeneity and test standardisation. Cook et al. [35] suggested the active instability test, pain during stair climbing, Clarke’s test, pain during prolonged sitting, patellar inferior pole tilt and pain during squatting as the best tests. However, they also found important disadvantages across the studies i.e. blinding and different reference standards, therefore, they proposed that PFPS should be a diagnosis of exclusion.

Risk Factors

Six reviews were identified in this section (Supplementary Material 4). Two of them [31, 33] could not be identified as systematic, therefore, their level of evidence was low and methodological quality low as well (AMSTAR 3 out of 11). Larger q-angle, muscle strength deficits, muscle tightness and joint and patella laxity were the components that the two studies had in common. Most of the risk factors showed contradictory results and totally different methodology across the primary studies. This explains why in these two studies comparison across the included studies was difficult if not impossible. Selfe [33] also reported anthropometric risk factors such as body weight, age and sex, however the evidence was limited and in some cases absent.

One review was entitled as systematic with no meta-analysis [37]. Although this systematic review was identified as high quality systematic review of case control and cohort studies, the methodological quality was low (AMSTAR 3 out of 11) while only 6 out of 27 PRISMA items were reported.

Waryasz and McDermott [37] reported contradictory results in all reported risk factors except for quadriceps tightness, Electromyography (EMG) neuro-motor dysfunction and functional testing. Fredericson and Yoon [31] reported significant results for Iliotibial Band (ITB) tightness, mediolateral patellar mobility, tight quadriceps, hip abductor weakness and functional performance. Therefore, the only risk factors that both studies agreed on were quadriceps tightness and reduced functional performance of patients with PFPS.

The other three studies were systematic reviews with meta-analysis. Two of them [9, 17] were high quality systematic reviews of case control and cohort studies whilst Lankhorst et al. [36] also published a systematic review with meta-analysis of RCTs. Therefore, the level of evidence of the last review was the highest and should be taken into more consideration. In addition, all three studies were graded as studies with high methodology (8-11 in AMSTAR tool). The systematic review of RCT’s had one point less in the AMSTAR tool than the non RCT review as there was no declaration any conflict of interest. However, only the two studies from Lankhorst et al. [9, 36] reported most of the PRISMA items (24/27 for both reviews) whilst the study from Pappas et al. [17] only reported 17 out of 27. Pappas and Wong-Tom [17] concluded that because PFPS is multifactorial, clinicians should evaluate strength flexibility and dynamic alignment of the lower limb. Limited flexibility of quadriceps and gastrocnemius and knee extension weakness could detect a PFPS case, however, these components could not work for non-athletic population as these tests have not been tested in civilians but only in military people who are generally supposed to be more athletic than non-active populations. The other review of case control and cohort studies [9] reported that a larger Q-angle, larger sulcus angle, larger patellar tilt angle, less hip abduction and knee extension strength are associated as risk factors in PFPS. However no flexibility tests were reported as risk factors and they called for more research in high-risk groups such as athletes and military populations. Finally, the review of RCTs [36] reported that only knee extension deficits can be considered as risk factors whilst there is not enough evidence for flexibility deficits components. They also reported that clinicians should focus on dynamic lower limb malalignment because all studies reported in the review reported biomechanical and neuromuscular risk factors and not structural (static) risk factors.

Exercise Treatment

Eight reviews met the inclusion criteria of exercise treatment in PFPS (Supplementary Material 5). Two of the reviews were not systematic; the one was narrative [32] and the other was a critical review [33], respectfully. The other 5 reviews were identified as systematic reviews and three of them included meta-analysis [11, 12, 38]. Additionally, three reviews were identified as high quality meta-analysis or systematic reviews of RCTs [11, 38, 39] whilst the other two [12, 13] as high quality systematic reviews of case control or cohort studies. Therefore, the level of the first three reviews was found to have high methodological level (8-11) in the AMSTAR tool whilst the two reviews of case control or cohort studies had moderate methodological level (4-7 in the
AMSTAR tool). Only the reviews from Heintjes et al. [38] and Collins et al. [11] reported more than 20 items from the PRISMA checklist (24 and 26 out of 27 correspondingly) whilst the other three reported 14 or less. Both reviews from Fagan and Delahunt [40] and Bolgla and Malone [32] were found to have moderate methodological evidence (6/11 AMSTAR tool for both reviews). However the Bolgla and Malone [32], review reported more PRISMA items (13/27) than the systematic review from Fagan and Delahunt [40]. As mentioned before the review from Selfe [33] was identified to have low level of evidence and had only one out of 27 items from the PRISMA checklist.

The main finding from these three reviews was that open kinetic chain (OKC) exercises are as beneficial as the closed kinetic chain (CKC) exercises; however, Collins et al. [11] suggested that CKC exercises may be preferable for short-term periods. Selfe [33] finally concluded that since both OKC and CKC exercises are good for PFPS then functional exercises which combines both OKC and CKC is probably important in the rehabilitation of PFPS.

Additionally, all eight reviews showed the importance of quadriceps strengthening. The inclusion of hip abductor strengthening in a rehabilitation program was supported by Frye et al. [13] Bolgla and Boling [12] and Harvie et al. [14] but not from Collins et al. [11] Fagan and Delahunt [40] and Bolgla and Malone [32], who reported that adding hip strengthening on quadriceps-based programme did not change the outcomes measures. However, all reviews suggested hip strengthening as a part of a rehabilitation program and not as the main treatment. Bolgla and Malone [32] agreed with Selfe [33] on the evidence that isometric exercises of the quadriceps [Straight leg raises (SLRs)] along with eccentric and isokinetic exercises can be beneficial for PFPS patients.

The review from Heintjes et al. [38] was the only one which concluded that there is not enough evidence to support that exercise is better in terms of pain and function than no exercise. In contrast, Collins et al. [11] concluded that the RCTs included in their review support the use of exercise whilst Frye et al. [13] found one study which showed no improvement after an exercise programme. Contrasting results were reported between Frye et al. [13] and Collins et al. [11] regarding whether exercise should be supervised or not. The former reported significant results between supervised and not supervised exercise prescription whilst the latter concluded that there was no significant difference. Only two of the reviews [13, 14] included information about whether stretching is beneficial and which structures clinicians should aim for greater flexibility. They both agree that the iliotibial band is one of these components. Additionally, Harvie et al. [14] included even more structures to stretch (hamstrings, quadriceps, gastrocnemius and anterior hip).

### Outcome Measures

Three studies were identified in this section (Supplementary Material 6). The earliest one was the review from Selfe [33] which did not focus on outcome measures only. As mentioned previously the study was a critical review with low methodological evidence (1/11). The author reported his surprise that he only found 3 articles investigating outcome measures. Among several outcome measures [Anterior Knee Pain Scale (AKPS), Visual Analogue Scale (VAS), Function Index Questionnaire (FIQ), Modified FIQ (MFIQ), Flandry questionnaire and Pierrynowski] the FIQ was reported as the easiest to complete, the Flandry questionnaire as the most accurate for depicting symptoms and the MFIQ was recommended for clinical use.

The other two studies were identified as high systematic reviews of case controls or cohort studies. Their methodological evidence was moderate for Howe et al. [41] with 5 out of 11 and high for Esculier et al. [42] with 8 out of 11 in the AMSTAR tool, respectfully. In addition the reviews included 10 and 17 items of the PRISMA checklist. The review from Howe et al. [41] evaluated several outcome measures for different knee conditions and not just for PFPS. Among different outcome measures such as the Lower Extremity Functional Scale (LEFS), Patellofemoral Severity Scale (PSS), Patient Specific Functional Scale (PSFS), Visual Analogue Scale (VAS), Lysholm, FIQ and Activities of Day Living Scale (ADLS) only the AKPS was found to be designed for PFPS. Esculier et al. [42] identified the five most used outcome measures in PFPS and concluded that the ADLS was the most appropriate for PFPS patients because of its reliability, validity and responsiveness. AKPS and FIQ could be also recommended but they still need to be tested in larger populations. The scope of the two reviews was slightly different. Howe et al. [41] investigated which outcome measure was best for which knee symptom, whilst Esculier et al. [42] aimed to identify which of the PFPS was best for clinical use in PFPS patients.

### Synthesis of Results

Statistical pooling was not possible because the trials did not often present sufficient data and more importantly, there was not a common set of outcomes across the reviews.

### Additional Analysis

To answer the secondary questions of this review only the systematic reviews with clear study report were included. Therefore, the four narrative reviews of general scope which were excluded from this section were the following: Malanga et al. [30]; Fredericson & Yoon [31]; Bolgla & Malone [32]; Selfe [33]. The 14 systematic reviews included 213 studies; 69 duplicates were identified and 144 primary studies were screened to answer the secondary questions of this RoR. This backtracking search revealed 43 studies which included dynamometers for muscle strength. Most of them (29) used non-portable isokinetic dynamometers rather than hand-held ones. Only 24 studies were identified as randomised controlled trials, whilst there was a wide distribution in the countries were the studies took place. Most of them were conducted in Europe, USA, Canada, Australia and Brazil. In terms of the research setting, there were 21 studies which did not report where the research was conducted; however, most of the studies (67) took place at university laboratories and not in clinical environments. The rest were conducted either in military bases, research centres and university hospitals.
The population that most of the studies used was not specified. Most of the studies also did not mention the patients’ general activity levels or sport at all. Therefore in most cases the patient characteristics was unclear. However, there were 40 studies which reported that their patients were athletic or participated in sports such as running and dancing. In addition to those 40 studies, 14 studies included military populations which researchers assume to be athletic as well. Only 8 studies reported that they used outpatient participants or general populations. Therefore, there was a clear tendency from researchers to use athletic rather than sedentary patients. The 144 studies included 4141 PFPS patients (28 patient per study approximately). The average study participant number becomes even smaller if it is considered that most of the patients were from military studies where large numbers of participants were recruited. Additionally, the gender of participants was not identified in 30 studies; however, 1888 women were recruited in the other 114 studies with a smaller number of men (1507).

**DISCUSSION**

**Summary of Evidence**

The purpose of this systematic RoR was to identify the literature evidence on clinical tests, risk factors, exercise treatment and outcome measures for PFPS. Additionally, the search of the primary studies included in the systematic reviews brought answers to the secondary questions into light. These questions were about the number of participants, the systematic reviews used, the gender and the participants’ attitude towards sports (athletes or sedentary people), whilst the country, the clinical or laboratory setting and the use of dynamometer was also searched. Generally, this RoR showed that the recent reviews stronger evidence compared to the earlier ones, because the level of evidence and methodological quality of the reviews have been positively improved. An analytical discussion of all components of this systematic RoR is provided below.

**Clinical Tests**

Earlier studies mostly reported the use of specific clinical tests whilst more recent revealed that the idea of finding the best clinical test seems to be abandoned. There is a clear tendency to functional assessment using specific tasks, such as the squat, or measuring pain during climbing and sitting with the legs flexed. More clinical tests are also reported but because of the non-consistent definition of PFPS, the nebulous pathophysiology and the different methodology among the studies, clinical tests cannot be compared between studies [35]. This is probably the reason that functional tests are preferred for the assessment rather than specific tests of the patella. For the same reason it is not surprising that newer reviews suggest that PFPS may be a diagnosis of exclusion [34, 35].

**Risk Factors**

Some of the tests used for the assessment (large q-angle, pain on patellar compression) were also suggested as risk factors in the older studies. However, there were a large number of studies which assessed strength and flexibility differences between PFPS populations and healthy controls. The result revealed many contradictions on specific muscle groups (e.g. hip muscle strength, gastrocnemius, ITB and hamstring flexibility) and the only risk factor that they agreed on was the quadriceps strength deficit and the low functional ability in PFPS. Recent systematic reviews [9, 17] also revealed contradictions on strength and flexibility risk factors, however, the strong evidence from the review of RCTs [36] only reported that less quadriceps strength should be considered as risk factor whilst there was no evidence on flexibility deficits in PFPS population. Additionally, clinicians should focus on dynamic, rather than static, malalignment of lower limb because research showed that PFPS populations have neuromuscular risk factors. It is worth mentioning that these risk factors are for athletes with PFPS and they have not been tested in the general population. Moreover, researchers call for more research in athletic or military population. This is surprising as there is evidence to suggest that sedentary people are also predisposed to PFPS [43].

**Exercise Therapy**

With regards to the exercise therapy, there are no many changes over the last 20 years. Earlier studies agree that there was strong evidence that PFPS physiotherapeutic interventions should focus on quadriceps strengthening. Additionally, isometric exercises of knee extensions appeared to be beneficial for patients with PFPS, whilst, there was no evidence that strengthening the hip muscles could be beneficial. Although recent studies of high-level systematic reviews of RCTs with meta-analysis [11, 38, 39] were published, only limited further evidence was added to literature as many methodological contradictions between primary studies were observed. In addition, fundamental principles such as whether exercise is better than no exercise were still debated. Heinjes et al. [38] reported that there is still not enough evidence to report that patients would benefit from exercise than just from rest alone.

In terms of the use of closed or open kinetic chain exercises, OKC exercises were found to be as good as CKC exercises [11, 33, 38]. However, Collins et al. [11] revealed that CKC exercises are better for short-time period treatment. Additionally, recent reviews still suggested the use of quadriceps strengthening, whilst the use of hip strengthening is still under debate. Although four of the systematic reviews (two of which were systematic reviews of RCTs with meta-analysis and high level of evidence; [38, 39]) reported positive use of hip strengthening in the treatment of PFPS patients, the third high level systematic review of RCTs reported that recruiting the hip muscles would not make any difference to the PFPS patients. Stretching was mentioned in two systematic reviews only [13, 39]. These two reviews agreed that ITB stretching was useful for PFPS patients. However, taking the higher level evidence and methodology from Harvie et al. [14] would conclude that the use of hamstrings, quadriceps, gastrocnemius and anterior hip muscles stretching is also necessary. Finally, since the study of Collins et al. [11] was of higher level of evidence compared to Frye et al. [13] it could also be concluded that supervised exercising in PFPS is no better that unsupervised exercise.
A Systematic Review of Reviews in Patellofemoral Pain Syndrome

The Open Sports Medicine Journal, 2015, Volume 9

15

Outcome Measures

Only recently authors reported outcome measures on PFPS. The earlier review from Selfe [33], recommending the MFIQ as the most appropriate outcome measure cannot be considered as strong evidence because no other studies had ever validated the applicability of that scale on PFPS patients. The two recent high-level of evidence systematic reviews that were identified cannot be compared because of their different scope and suitability for PFPS. Howe et al. [41] investigated several outcome measures for different knee conditions some of which were for PFPS whilst, Esculier et al. [42] reported which of the 5 most used PFPS outcome measures were best for clinical use according to their reliability, validity, responsiveness and cultural adaption. Howe et al. [41] reported that only the AKPS was designed for PFPS whilst the LEFS was not, although the latter had better responsiveness. The AKPS was also problematic because it included questions not suitable for PFPS when other questions, such kneeling, should be included. The review from Esculier et al. [42] had better methodological evidence and suggested the AKPS could be used as one of the most appropriate for PFPS if further tested in more participants. The FIQ which was suggested from Selfe [33], and Lysholm scale should be excluded from PFPS use. The only scale which was the most appropriate for PFPS patients because of its reliability, validity and responsiveness was deemed the ADLS.

Secondary Questions

The systematic reviews revealed a large number of studies on PFPS which mostly took place in university settings. Most of the studies recruited participants from physiotherapy clinics or outpatients hospital departments. However the actual research took place in university laboratories. Only a few studies were conducted in a clinical environment, therefore, the question that could be raised is whether the evidence base regarding tests and techniques can be translated from the laboratories to the routine physiotherapy clinic.

In addition, the fact that only 24 studies out of 144 were found to be RCTs shows a strong need for more strong evidence research studies. The total number of participants divided by the number of studies revealed that approximately 28 PFPS participants per study. This is a small number especially if it is considered that this number would be much smaller if the large military studies with the hundreds of participants were excluded.

As women are more likely to get PFPS it was not surprising that female population was the larger. However, there was no evidence regarding whether women should be assessed or treated differently to men. Additionally, one of the most important observations was that researchers tend to use athletic rather than non-athletic populations, whilst Lankhorst et al. [36] reported that more research should be conducted in athletic populations. However, there is no evidence about whether sedentary patients with PFPS have the same risk factors, should be assessed with the same outcome measures and be treated in the same way as athletic patients. A stronger case could be made about the patient characteristics if it is considered that most of the primary studies did not report whether their patients with PFPS were athletic or not.

As mentioned previously, most of the studies were conducted in universities where the required for testing equipment and the research environments are usually the same and do not differ from country to country. Therefore, the research evidence from different countries (as Supplementary Material 7 showed) can easily be adopted from different research centres for research use. However, when it comes to research in clinical environment the different conditions of different national health systems may differentiate the applicability of the findings across countries.

Finally, strength assessment was mostly done by non-portable isokinetic dynamometers, probably because most of the research took place in university laboratories where dynamometers did not have to be portable. These dynamometers may not be available in clinics or military bases. Therefore, other studies reported the use of portable hand-held dynamometers. However, the reliability and validity and therefore the applicability of those tools were not mentioned in the studies.

LIMITATIONS

A major limitation of this RoR was that meta-analysis was not possible because of the different methodologies that both primary and review studies used. Additionally, some of the primary studies were reported in more than one review and this duplication of evidence is shown in Supplementary Material 7. Although duplicates were reported, the identification of the duplicate evidence in the reviews was in most cases impossible. In addition, some of the primary studies were difficult to identify because they were either very old or not available (e.g. the journals had closed down). In some other cases the primary papers were in languages that the author of this RoR is not familiar with. These factors may help explain why some of the questions were unresolved.

The current RoR focused on the diagnostic clinical tests for PFPS but not on imaging which can also be used for diagnosis. Future reviews may also search the sensitivity and specificity of x-ray and MRI findings in patients with PFPS.

CONCLUSION

To conclude, early studies on PFPS were trying to find a gold standard clinical test for PFPS assessment while, nowadays the use of functional tasks such as the squat or the measurement of pain during functional task are suggested for clinical assessment instead. With regards to the risk factors, there are still a lot of contradictions in terms of muscle strength deficits. Today, the quadriceps strength deficits are still the only evidence based risk factors along with the dynamic malalignment of lower limb. More research is still required on the strength of other muscle and flexibility deficits. Isometric exercises of quadriceps along with OKC and CKC exercises of the lower limb were suggested were suggested in the early studies, whilst today the quadriceps-based exercises are still the only ones to have strong evidence together with hamstrings, quadriceps, gastrocnemius
and anterior hip muscles stretching. There is a need of more research on hip muscle exercises to establish their benefits in PFPS treatment whilst exercise dosage focusing on endurance and high repetitions of hip exercises should also be considered. Evidence on outcome measures was absent in the earlier studies whilst the usage of ADLS is recommended today for clinical use. Finally, there is no evidence on whether the above treatment and assessment methods should be used in sedentary people or differently across population groups or gender. More RCTs with large populations, powered correctly, in clinical environments are called for in further research, whilst the country where the evidence comes from and the use of portable dynamometers for strength assessment should also be justified and considered.

AUTHORS’ CONTRIBUTIONS

KP has been involved in conducting the main research study and drafting the manuscript.

DS ensured that questions related to the accuracy or integrity of any part of the work were appropriately investigated and resolved. Additionally, he was the second author who critically appraised the review articles.

DG reviewed the final draft of the manuscript and made sure that all research questions were resolved.

All three authors have given approval of the version to be published.

AUTHORS’ INFORMATION

KP is a research active physiotherapist, Lab/ Clinical Trainer Administrator, Physiotherapy Program, EUC, Cyprus

DS is an Assistant Professor, Physiotherapy Program, EUC, Cyprus.

DG is an Assistant Professor, National Sports Academy, Bulgaria.

LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADLS</td>
<td>Activities of Day Living Scale</td>
</tr>
<tr>
<td>AKP</td>
<td>Anterior Knee Pain</td>
</tr>
<tr>
<td>AMSTAR</td>
<td>A Measurment Tool to Assess systematic Reviews</td>
</tr>
<tr>
<td>CKC</td>
<td>Closed Kinetic Chain</td>
</tr>
<tr>
<td>CP</td>
<td>Chondromalacia Patellae</td>
</tr>
<tr>
<td>EMG</td>
<td>Electromyography</td>
</tr>
<tr>
<td>FIQ</td>
<td>Function Index Questionnaire</td>
</tr>
<tr>
<td>ITB</td>
<td>Iliotibial Band</td>
</tr>
<tr>
<td>LEFS</td>
<td>Lower Extremity Functional Scale</td>
</tr>
<tr>
<td>MFIQ</td>
<td>Modified Function Index Questionnaire</td>
</tr>
<tr>
<td>OKC</td>
<td>Open Kinetic Chain</td>
</tr>
<tr>
<td>PFPS</td>
<td>Patellofemoral Pain Syndrome</td>
</tr>
<tr>
<td>PICOS</td>
<td>Participants, Interventions, Comparisons, Outcomes and Study design</td>
</tr>
<tr>
<td>PRISMA</td>
<td>Preferred Reporting Items of Systematic reviews Meta-Analyses</td>
</tr>
<tr>
<td>PSS</td>
<td>Patellofemoral Severity Scale</td>
</tr>
<tr>
<td>QUOROM</td>
<td>Quality Of Reporting Of Meta-analysis</td>
</tr>
<tr>
<td>RCTs</td>
<td>Randomized Control Trials</td>
</tr>
<tr>
<td>RoR</td>
<td>Review of Reviews</td>
</tr>
<tr>
<td>SIGN</td>
<td>Scottish Intercollegiate Guidelines Network</td>
</tr>
<tr>
<td>SLRs</td>
<td>Straight Leg Raises</td>
</tr>
<tr>
<td>VAS</td>
<td>Visual Analogue Scale</td>
</tr>
</tbody>
</table>

CONFLICT OF INTEREST

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

ACKNOWLEDGEMENTS

The authors would like to thank the librarians of the School of Healthcare Sciences, Bangor University for their help in gathering the reviews and the primary research studies.

SUPPLEMENTARY MATERIAL

Supplementary material is available on the publisher’s web site along with the published article.

REFERENCES

A Systematic Review of Reviews in Patellofemoral Pain Syndrome

The Open Sports Medicine Journal, 2015, Volume 9


