



OPEN ACCESS

Comparative effectiveness of treatments for patellofemoral pain: a living systematic review with network meta-analysis

Marinus Winters ¹, Sinéad Holden ^{1,2}, Carolina Bryne Lura,¹ Nicky J Welton,³ Deborah M Caldwell,³ Bill T Vicenzino ⁴, Adam Weir,^{5,6,7} Michael Skovdal Rathleff^{1,2}

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/bjsports-2020-102819>).

¹Centre for General Practice at Aalborg University, Aalborg, Denmark

²SMI, Department of Health Science and Technology, Aalborg University, Aalborg, Denmark

³Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

⁴School of Health and Rehabilitation Sciences: Physiotherapy, University of Queensland, Brisbane, Queensland, Australia

⁵Sports Groin Pain Centre, Aspetar Orthopaedic and Sports Medicine Hospital, Doha, Qatar

⁶Sports Medicine and Exercise Clinic Haarlem (SBK), Haarlem, Netherlands

⁷Department of Orthopaedics, Erasmus MC University Medical Center for Groin Injuries, Rotterdam, Netherlands

Correspondence to

Dr Marinus Winters, Center for General Practice at Aalborg University, Aalborg, Denmark; marinuswinters@hotmail.com

Accepted 4 October 2020
Published Online First
26 October 2020



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Winters M, Holden S, Lura CB, et al. *Br J Sports Med* 2021;**55**:369–377.

ABSTRACT

Objective To investigate the comparative effectiveness of all treatments for patellofemoral pain (PFP).

Design Living systematic review with network meta-analysis (NMA).

Data sources Sensitive search in seven databases, three grey literature resources and four trial registers.

Eligibility criteria Randomised controlled trials evaluating any treatment for PFP with outcomes ‘any improvement’, and pain intensity.

Data extraction Two reviewers independently extracted data and assessed risk of bias with Risk of Bias Tool V.2. We used Grading of Recommendations, Assessment, Development and Evaluation to appraise the strength of the evidence.

Primary outcome measure ‘Any improvement’ measured with a Global Rating of Change Scale.

Results Twenty-two trials (with forty-eight treatment arms) were included, of which approximately 10 (45%) were at high risk of bias for the primary outcome. Most comparisons had a low to very low strength of the evidence. All treatments were better than wait and see for any improvement at 3 months (education (OR 9.6, 95% credible interval (CrI): 2.2 to 48.8); exercise (OR 13.0, 95% CrI: 2.4 to 83.5); education+orthosis (OR 16.5, 95% CrI: 4.9 to 65.8); education+exercise+patellar taping/mobilisations (OR 25.2, 95% CrI: 5.7 to 130.3) and education+exercise+patellar taping/mobilisations+orthosis (OR 38.8, 95% CrI: 7.3 to 236.9)). Education+exercise+patellar taping/mobilisations, with (OR 4.0, 95% CrI: 1.5 to 11.8) or without orthosis (OR 2.6, 95% CrI: 1.7 to 4.2), were superior to education alone. At 12 months, education or education+any combination yielded similar improvement rates.

Summary/conclusion Education combined with a physical treatment (exercise, orthoses or patellar taping/mobilisation) is most likely to be effective at 3 months. At 12 months, education appears comparable to education with a physical treatment. There was insufficient evidence to recommend a specific type of physical treatment over another. All treatments in our NMA were superior to wait and see at 3 months, and we recommend avoiding a wait-and-see approach.

PROSPERO registration number PROSPERO registration CRD42018079502.

INTRODUCTION

Musculoskeletal disorders are one of the largest global contributors to years lived with disability, and costly to society.¹ Patellofemoral Pain (PFP) is one

of the most common knee complaints in individuals between 10 and 50 years of age.² PFP impacts function, ability to participate in leisure time activities, work, sport and reduces quality of life.³ It is a clinical diagnosis made when patients present with pain around or behind the patella during daily activities such as stair walking, squatting or running. Similar to low back pain, PFP is characterised by a high degree of persistency and recurrence of symptoms. Nearly 40% of those with PFP continue to experience symptoms after 2 years, which is associated with frequent use of pain killers, lowering of physical activity levels and low quality of life.^{4,5}

Many different treatments are used in clinical practice to help patients with PFP.⁶ While there are several systematic reviews evaluating treatments for PFP,^{7–9} the comparative effectiveness of all available treatments has never been examined. This makes deciding on the most appropriate treatment challenging and may explain the variation in clinical practice.⁶ In this study, we use network meta-analysis (NMA), a technique which allows the simultaneous comparison of multiple interventions in a single coherent analysis. Traditional systematic review quickly become outdated.¹⁰ Living systematic reviews are continuously updated and incorporate new evidence when available.¹¹ A living systemic review with NMA would enable clinicians to consult a contemporary, comprehensive overview of the comparative effectiveness of treatments for a given condition. This living systematic review with NMA evaluates the comparative effectiveness of all available treatments for patients with PFP, providing a comprehensive and up-to-date overview of evidence-based treatments.

METHODS

Protocol registration

The living systematic review with NMA was prospectively registered on PROSPERO and a full protocol was published.¹² The findings are reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses checklist extension for NMA.¹³

Deviations from protocol

Deviations from the protocol relate to including a minimum number per treatment arm (>10) (to avoid resource intensive review work for very little gain), exclusion of adverse effects as an outcome (due to a lack of available data) and removal of threshold analyses (due to substantial overlap in

credible intervals (CrIs)). Deviations are outlined in more detail in online supplemental appendix 1.

Administration, dissemination and updating the living systematic review

This review will be administered at the Center for General Practice at Aalborg University, Denmark, and we plan to update the NMA annually for a minimum of 5 years. As described in our protocol, we will screen the literature annually to identify new data that may alter our conclusions and recommendations. When new data have become available, we will update the analysis and present the updated findings at the website of Aalborg University (<https://www.almenmedicin.aau.dk/forskning/forskningsomraade1/>). Here, we will also provide a plain-language summary for patients and clinicians dealing with PFP.

Patient and public involvement

Seven patients with PFP formed a patient reference group to select the hierarchy of outcomes (Global Rating of Change (GROC) Scale and pain scales). One researcher not involved in the study (acknowledgements) explained the various outcomes and participants indicated what they considered the most relevant instrument. Six of the seven indicated a preference for the GROC Scale. Consequently, the outcomes were prioritised as follows:

Primary outcome measure

Any improvement on a GROC Scale, a commonly used outcome with excellent reliability, intraclass correlation coefficients (ICCs) range from 0.90 to 0.99.^{14 15} Any improvement is defined as any¹⁵ degree of recovery or improvement. ‘Unchanged’ to ‘worse than ever’ are considered a treatment failure.

Secondary outcome measures

Pain intensity measured by ‘worst pain in the past week’ on a Visual Analogue Scale (VAS; 0–10/0–100) or Numerical Rating Pain Scale (NRS; 0–10/0–100). The reliability is excellent, ICC=0.76.^{15 16}

Patient-rated pain during specific activities of daily life and during sporting activities measured¹⁶ on a VAS or NRS as outlined above. Reliability for pain during activity is excellent, ICC=0.83.

Patient-reported outcome measures were not included as they have not been used for PFP until very recently.¹⁷ These may be included in future updates.

Research question

Based on the main purpose of the study and the input from our patient reference group, we formulated the following research question: Which treatment(s)/treatment category (class(es)) is most likely to be effective for PFP on any improvement and patient-rated pain?

Eligibility criteria

Type of studies

Published or unpublished randomised controlled trials (RCTs) (including randomisation through minimisation, or clustering) were eligible for inclusion providing a full-text report was available.

Type of population

All patients with a clinical diagnosis of PFP were included. Studies were included if they used synonyms for PFP, but as minimum criterion, described patients with retropatellar or peripatellar pain, of at least 6-week duration, and a non-traumatic onset. The diagnostic criteria used in the original studies were followed, given that pain was described as being retropatellar or peripatellar pain. Studies examining other conditions were excluded (eg, patellar dislocations, patellofemoral osteoarthritis, patellar tendinopathy, Osgood-Schlatter, iliotibial band syndrome, Sinding-Larsen-Johansson syndrome). Trials that included participants diagnosed with PFP, but with concomitant pain around the patella caused by other conditions (eg, patellar tendinopathy), were eligible for inclusion. No age restrictions were imposed.

Type of treatments and control treatments

Any conservative or invasive treatment, control treatment, placebo, wait-and-see or no treatment group studied were eligible for inclusion.

Type of outcomes

Studies assessing the treatment effect after a minimum of 6 weeks were included. Studies assessing the primary and secondary outcomes above were included (GROC Scale, worst pain intensity during the previous week and pain during activities).

Search strategy

We developed a sensitive search strategy which included a mix of indexed and free-text terms (see Winters *et al*¹²). No restrictions (eg, language or full-text availability) were applied. We searched conventional databases, grey literature databases and trial registers. The following sources were searched from their date of inception up till 4 June 2019: Embase, PubMed (including MEDLINE), Cochrane Central Register of Controlled Trials, Scopus, Web of Science, CINAHL, SPORTDiscus. OpenGrey.eu and WorldCat.org were searched for unpublished studies and conference proceedings were identified from all Patellofemoral Research Retreats (2009, 2011, 2013, 2015 and 2017). For unpublished or ongoing studies, we searched the WHO International Clinical Trials Registry Platform (<http://apps.who.int/trialsearch/>) Clinical Trials.gov, The European Union Clinical Trials Register and the ISRCTN registry. Finally, we screened reference lists of all Cochrane reviews (N=6) on PFP and the reports included in this review for possible relevant studies that were not identified by our search.

Study selection

Two researchers (MW and CBL) screened titles and abstracts independently after duplicate removal. Consensus was sought in cases of initial disagreement. If consensus could not be reached, the report was included for full-text evaluation. Both investigators independently applied our inclusion and exclusion criteria to the full-text reports. In case of disagreement, consensus was sought; however, if disagreement persisted a third author (MSR or AW) arbitrated the decision.

We used Covidence (Melbourne, Australia) for independent study selection, data extraction and risk of bias assessment.

Data extraction

Data were extracted by two researchers using standardised extraction forms adapted from the Cochrane Collaboration.¹⁸

Any disagreements were resolved by consensus. We extracted the following data:

- ▶ Publication and study details: For example, authors, year of publication, funding source, possible conflicts of interest, study aim, design and unit of allocation.
- ▶ Population: Number of patients included, population characteristics for age, sex, body mass index, activity level, setting where population was recruited, baseline scores for outcome measures (mean, SDs, standard errors extracted for continuous outcomes and number and percentage for categorical outcomes).
- ▶ Eligibility criteria and diagnostic criteria used for PFP.
- ▶ Treatments: For example, number randomised to group, detailed description of for example, application, dose, intensity, frequency, number of sessions, delivery, tailoring (individual/group), duration of treatment, providers, cotreatments, modification (change to treatment), adherence. We used items from the Template for Intervention Description and Replication checklist^{12,19} to assure comprehensive data extraction in this section of the extraction form.¹⁷
- ▶ Outcomes: Timepoints measured, and the timepoints reported on, outcome definition, person measuring, unit of measurement, scales (upper and lower limits), imputation of missing data, primary and secondary outcomes used in the original trials, unintentional outcomes (eg, adverse events, adverse effects, side effects).
- ▶ Data and analysis: Comparisons, outcomes, subgroups, timepoints, results (central estimates and measures of dispersion; eg, mean for both groups, mean difference (MD), SDs/95 CIs/standard errors), number of missing patients, statistical methods used and appropriateness of these.
- ▶ Other information: Key conclusions of study authors.

Risk of bias assessment

The Cochrane Risk of Bias Tool V.2 was used to assess the risk of bias for each outcome per study. We assessed bias following the 'intention-to-treat' principle (ie, assignment to intervention).²⁰ This tool has a fixed set of items to use for the risk of bias appraisal, that is, 'bias arising from the randomisation process', 'bias due to deviations from intended interventions', 'bias due to missing outcome data', 'bias in measurement of the outcome', 'bias in selection of the reported result' and overall risk of bias judgement for each outcome.

Pairs of two (MW and CBL; MW and AW) independently assessed all RCTs included. Each major domain of bias was appraised in light of each included study outcome. The tool's signalling questions and criteria were followed to inform a domain-based appraisal of the risk of bias. The risk of distortion of the outcome estimate was appraised as at 'low', 'some' or 'high' risk of bias. We made judgements regarding the direction of distortion 'favours experimental', 'favours comparator', 'towards null', 'away from null' or 'unpredictable'. Each outcome within a study received an overall risk of bias judgement based on the individual domains; 'low', 'some' or 'high' risk of bias, following the guidance from the tool. In case of disagreements between reviewers, consensus was reached through discussions in all cases.

Data synthesis and statistical methods

We constructed network plots using Stata software (StataCorp. V.2017. Stata Statistical Software:118 Release 15. College Station, Texas: StataCorp LLC) to visualise all head-to-head comparisons for all outcomes.²¹ Networks of treatment comparisons were constructed for the primary and secondary outcome separately. Three authors (MW, SH, MSR) appraised the clinical homogeneity before the start of the analysis, by tabulating study and population characteristics and inspecting them for differences in potential effect modifiers. This informed assessment of the assumption of exchangeability required for NMA. Treatments were also assigned to categories (ie, classes) (table 1).

Our primary outcome measure, any improvement, was summarised using an OR for improvement versus non-improvement, with a 95% CrI. This was in line with our protocol, and done because various GROC Scales, with different number of response options and descriptors (eg, improvement vs recovery), did not allow for data to be analysed in a proportional odds regression model. We used a conservative intention-to-treat approach to the analyses where missing data in original studies was handled as a treatment failure.

For our secondary outcome measures, worst pain and pain descending stairs intervention effects were expressed as MDs, with 95% CrIs, when outcomes were measured with the same instrument. Pain measured on a 0–100 scale was converted to 0–10 where necessary. Mean and median ranks and 95% credibility limits were used to estimate the likelihood of individual

Table 1 Treatments and treatment categories (ie, classes) for the 'any improvement'

Categories (ie, classes)	Category definition	Treatments	Studies
Education	Education consisted of information/advice given by a healthcare practitioner on patellofemoral pain, (aggravating) activity/exercise and management of pain/symptoms. Delivered either face to face or in the form of written materials	Education	Rathleff <i>et al.</i> ²⁷ Linschoten <i>et al.</i> ²⁵ Collins <i>et al.</i> ²⁶
Education+exercise therapy+patellar taping/mobilisations	Education (as above), combined with resistance exercise (as described below) and application of taping/movement to the patella	Education+exercise therapy+patellar taping/mobilisations	Rathleff <i>et al.</i> ²⁷ Linschoten <i>et al.</i> ²⁵ Collins <i>et al.</i> ²⁶
Education+orthosis	Education (as above) in combination with prefabricated orthotics to be placed under the foot in the shoe to support the arch	Education+orthosis	Collins <i>et al.</i> ²⁶ Matthews <i>et al.</i> ⁴³ Mills <i>et al.</i> ⁴⁴
Education+exercise therapy+patellar treatments+orthosis	A combination of these treatments outlined above/below	Education+exercise therapy+patellar treatments+orthosis	Collins <i>et al.</i> ²⁶
Wait and see	No treatment or continuing with current/planned management	Wait and see	Mills <i>et al.</i> ⁴⁴ Drew <i>et al.</i> ³²
Exercise therapy	Resistance exercise, that is, muscles contracting against resistance provided in the form of weights, bands or body/limb weight with the goal of improving muscle strength, endurance or mass. Programmes targeted the lower limb and/or trunk muscles (with/without general aerobic conditioning as warm-up/cooldown). This could be delivered face to face or as a home exercise programme	Exercise hip/knee Exercise hip/knee trunk Exercise hip/knee+blood flow restriction Exercise hip/knee with feedback	Matthews <i>et al.</i> ⁴³ Baldon <i>et al.</i> ³⁰ Giles <i>et al.</i> ³⁹ Riel <i>et al.</i> ⁴⁸ Drew <i>et al.</i> ³² Baldon <i>et al.</i> ³⁰ Giles <i>et al.</i> ³⁹ Riel <i>et al.</i> ⁴⁸

treatments being superior to the other treatments for the individual with PFP.

NMA models were fitted for both treatment categories (ie, classes'; overarching treatment categories outlined in [table 1](#)), and individual treatments. For all analyses, we fitted fixed and random effects NMA models²² at the treatment level, and compared model fit using the deviance information criterion and posterior mean residual deviance (lower measures of both statistics are preferred, with differences of 3 or more considered meaningful). In accordance with our protocol, we grouped outcome follow-ups based on the available data, that is, 3 and 12 months. If there were multiple timepoints available for an outcome, and these were equally close to the timepoint to be synthesised across studies, the last follow-up in this timeframe was used. For the class-level models, we fitted hierarchical (ie, random) and fixed effect between treatment within-class models and compared model fit as described above. For the secondary outcome measures worst pain and pain descending stairs, we also attempted to fit bivariate models to capture the correlation in treatment effects on these related outcomes.

If >10 studies were available per comparison, we assessed statistical heterogeneity by inspecting the between-study SD, and by comparing fit of the fixed and random effect models. We assessed the consistency assumption for each network by comparing model fit between the NMA model and an unrelated mean effects model that relaxes the consistency assumption.²³ We assessed small study bias using comparison-adjusted funnel plots if 10 or more trials were available for 1 comparison.²⁴ We ran a sensitivity analysis to test if our findings were robust for our decision to pool the study arms (education vs education+exercise) in van Linschoten *et al*²⁵ with the study arms in Collins *et al*²⁶ and Rathleff *et al*²⁷ (both education vs education+exercise+patellar treatments).

NMA models were fitted in a Bayesian framework using Markov chain Monte Carlo simulations in WinBUGS (V.1.4, Medical Research Council, UK, and Imperial College of Science, Technology and Medicine, University of Cambridge, UK). A Bayesian analysis estimates a posterior distribution which we summarise with the mean or median, and 95% CrIs. 95% CrIs are interpreted as a range of values within which a parameter lies with a 95% chance. CrIs are very similar to CIs in frequentist analyses but may be wider for random effects models due to Bayesian analyses incorporating uncertainty around the between-studies SD.

Certainty of the evidence (Grading of Recommendations, Assessment, Development and Evaluation (GRADE) approach)

After the analysis, we used the GRADE approach to assess the certainty of the evidence from the NMA.^{28 29} Two researchers graded (MW and AW, see acknowledgements) the evidence on the basis of the risk of bias assessments, inconsistency, indirectness, imprecision and publication bias. We assigned an overall judgement about the certainty of the evidence for all comparisons, and for the evidence from the NMA as a whole.

RESULTS

Selection process

We included 22 RCTs. [figure 1](#) details the search and selection process. Online supplemental appendix 2 shows the studies excluded+the reasons for exclusion. Thirty-three trials were found through register searches; sixteen were completed, nine were ongoing and eight trials' status was unknown (online supplemental appendices 3 and 4).

Characteristics of the studies included

Twenty-one treatments were investigated in twenty-two trials.^{25-27 30-48} Forty-eight treatment arms were investigated, with sample sizes from 10 to 109 per treatment arm. Thirty-six (75%) of arms included exercise. Other common intervention categories included patient education, orthotics and wait and see (or a combination). For full details see [table 1](#) and online supplemental appendix 5. Eleven RCTs used a GROSC Scale to measure patient-reported outcomes, and nineteen trials used a worst pain scale or measured pain during a specific activity. Cumulatively, 1472 patients with PFP were included. Online supplemental appendix 5 details all study and patient characteristics of the studies included.

Risk of bias and certainty of the evidence

The majority of outcomes were at high risk of bias (79%), with some concerns in 21% (online supplemental appendix 6, [tables 1 and 2a,b](#)). Outcome measurement (45%), deviations from the intended intervention (45%), missing outcome data (32%) and reporting (23%) were the most common sources of bias.

The certainty of evidence for all comparisons was low to very low, except for hip/knee exercises versus education+orthosis, for which there was moderate evidence (online supplemental appendix 7). Evidence for comparisons was rated down for risk of bias (n=100, 100%; ie, 'serious' to 'very serious') and imprecision (n=85, 85%).

NETWORK META-ANALYSES

[Figure 2](#) shows direct treatment comparisons in the field of PFP for any improvement. Nine studies could be included in the NMA at 3 months, and three studies in the NMA at 12 months. Six classes were included in network analyses for the primary outcome ([table 1](#)). Model fit statistics are reported in online supplemental appendix 8. On the basis of model fit, we report results for the fixed effect model with random between treatment within class effects for any improvement at 3 months. We report results for a fixed effect treatment model without class for any improvement at 12 months, as it was not possible to conduct class-level analyses for any outcomes at 12 months, or for pain while descending stairs at 3 months, as only single treatments were available. Estimated individual treatment effects for 'any improvement', 'worst pain' and 'pain while descending stairs' are presented in online supplemental appendix 9. No other measures for pain during an activity could be synthesised in an NMA. None of the networks showed evidence of inconsistency between direct and indirect comparisons, where data from direct and indirect evidence were available. Too few trials were available to investigate heterogeneity, or small study bias. Six studies that could not be included in an NMA are presented in a descriptive synthesis in online supplemental appendix 10; among the six there were two studies that could have been connected to the network if data were available.

Online supplemental appendix 11 provides the findings for the sensitivity analysis. It shows that our main findings are robust for our decision to pool the education+exercise arm in van Linschoten *et al*²⁵ with the education+exercise+patellar treatments arms in Collins *et al*²⁶ and Rathleff *et al*.²⁷

Comparative treatment (class) effectiveness on the primary outcome ('any improvement')

Any improvement at 3 months

The fixed effect model with hierarchical (ie, random) effect between treatments within the class showed that all treatment

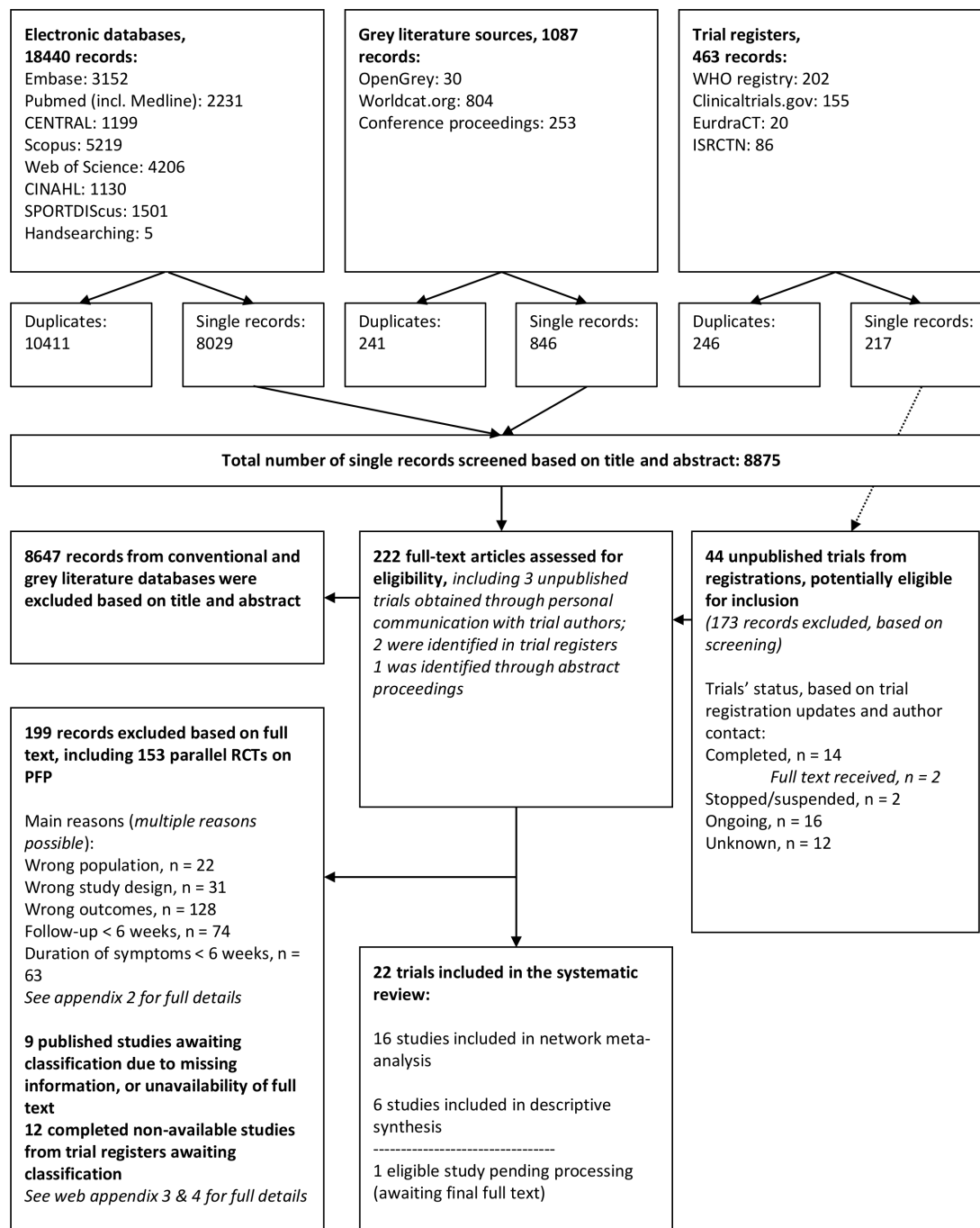


Figure 1 Flow diagram, search on 4 June 2019. PFP, patellofemoral pain; RCTs, randomised controlled trials.

classes were superior to wait and see: education (OR 9.6, 95% CrI 2.2 to 48.8), exercise (OR 13.0, 95% CrI: 2.4 to 83.5), education+orthosis (16.5, 95% CrI: 4.9 to 65.8), education+exercise+patellar taping/mobilisations (25.2, 95% CrI: 5.7 to 130.3) and education+exercise+patellar taping/mobilisations+orthosis (38.8, 95% CrI: 7.3 to 236.9) (table 2a).

Education+exercise+patellar taping/mobilisations, with or without orthosis, was superior to education alone (2.6, 95% CrI: 1.7 to 4.2, and, 4.0, 95% CrI: 1.5 to 11.8, respectively). Neither exercise (prescribed on its own) nor orthosis+education was superior to education alone (1.3, 95% CrI: 0.3 to 6.3, and, 1.7, 95% CrI: 0.8 to 4.0, respectively). No specific type of exercise was superior to another type of exercise (online supplemental appendix 9).

Any improvement at 12 months

At 12 months, four treatment arms were compared in a fixed effects treatment model. No differences were found for education+exercise+patellar taping/mobilisations (OR 1.5, 95% CrI: 0.9 to 2.4), education+orthosis (2.3, 95% CrI: 1.0 to 6.2) or education+exercise+patellar taping/mobilisations+orthosis (1.9, 95% CrI: 0.8 to 4.9) when compared with education alone (table 2b).

Treatment (class) rankings 'any improvement'

Education+exercise+patellar taping/mobilisations, either with or without orthosis, was the best combination of treatments for PFP at 3 months (median ranking 1, median's 95% CrI: 1 to 3;

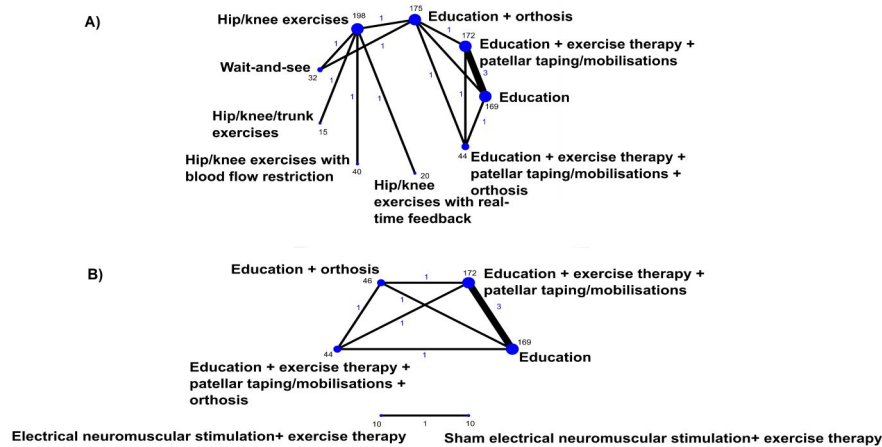


Figure 2 Network graphs for direct treatment comparisons for any improvement at 3 months (A) and 12 months (B). Blue text represents the number of treatment comparisons, and the text in black represents the number of participant that received the respective treatment. The thickness of the lines and the size of the dots are proportional to the number of trial comparisons and the number of participants in the treatment arms, respectively.

and, median ranking 2, 95% CrI: 1 to 4, respectively). Wait and see was least likely to be effective (median ranking 6, 95% CrI: 6 to 6) (table 3).

At 12 months, education+exercise+patellar taping/mobilisations, either with (median ranking 2, 95% CrI: 1 to 4) or without (median ranking 3, 95% CrI: 1 to 4) orthosis, and education+orthosis (median ranking 1, 95% CrI: 1 to 4) showed similar rankings. Education alone seemed least likely to be effective (median ranking 4, 95% CrI: 3 to 4).

Comparative treatment effectiveness on the secondary outcomes

Comparative treatment effects for worst pain at 3 months, as estimated by random effects model with random between treatments within class effect, are reported in online supplemental appendix. We report on the results from a fixed effects model without class effect for worst pain at 12 months. The pooled findings for ‘worst pain’ at 3 months show that none of the treatments was found to be superior to any other treatment or to wait

Table 2a Treatment effects for the primary outcome: Comparative treatment class effects for any improvement at 3 months (fixed effects model with random between treatment within class effect)

Wait-and-see					
9.6 (2.2 to 48.8)	Education				
13.0 (2.4 to 83.5)	1.3 (0.3 to 6.3)	Exercise			
16.5 (4.9 to 65.8)	1.7 (0.8 to 4.0)	1.3 (0.4 to 4.5)	Education+orthosis		
25.2 (5.7 to 130.3)	2.6 (1.7 to 4.2)	2.0 (0.4 to 8.5)	1.5 (0.7 to 3.6)	Education+exercise+patellar taping/mobilisations	
38.8 (7.3 to 236.9)	4.0 (1.5 to 11.8)	3.0 (0.6 to 16.0)	2.3 (0.8 to 7.5)	1.5 (0.6 to 4.6)	Education+exercise+patellar taping/mobilisations+orthosis

ORs with their 95% credible intervals from the network meta-analysis are shown.

For any cell, an OR<1 favours the upper-left treatment, and an OR>1 favours the lower-right treatment. Comparative treatment effect differences are shown in bold.

Table 2b Treatment effects for the primary outcome: Comparative treatment effects for any improvement at 12 months (fixed effects model without class effect)

Education				
1.5 (0.9 to 2.4)	Education+exercise+patellar taping/mobilisations			
2.3 (1.0 to 6.2)	1.5 (0.6 to 4.2)	Education+orthosis		
1.9 (0.8 to 4.9)	1.3 (0.5 to 3.3)	0.8 (0.3 to 2.4)	Education+exercise+patellar taping/mobilisations+orthosis	

ORs with their 95% credible intervals (CrIs) from the network meta-analysis are shown in the lower left triangle, and ORs with their 95% CrIs from the pairwise meta-analyses (ie, direct evidence from randomised controlled trials) in the upper-right triangle. For any cell, an OR<1 favours the upper-left treatment, and an OR>1 favours the lower-right treatment.

Table 3 Treatment rankings from the network meta-analyses for any improvement.

Treatment (class)	3 months*		12 months†	
	Mean rank	Median rank (95% CrI)	Mean rank	Median rank (95% CrI)
Education+exercise+patellar taping/mobilisations+orthosis	1.36	1 (1 to 3)	2.03	2 (1 to 4)
Education+exercise+patellar taping/mobilisations	2.12	2 (1 to 4)	2.56	3 (1 to 4)
Education+orthosis	3.17	3 (2 to 5)	1.57	1 (1 to 4)
Exercise	3.77	4 (1 to 5)	NA	NA
Education	4.58	5 (3 to 5)	3.84	4 (3 to 4)
Wait and see	5.99	6 (6 to 6)	NA	NA

*Results from a fixed effects model with random between treatment within class effect.

†Results from a fixed effects model without class effect.

95% CrI, 95% credible interval; NA, not applicable.

and see. At 12 months, education+exercise+patellar taping/mobilisations appears superior to education alone. Education+exercise+patellar taping/mobilisations+orthosis appears better than education alone but was not found to be superior to education+exercise+patellar taping/mobilisations.

For pain while descending stairs, three treatments could be compared. All analyses were performed using fixed effects models without class effects. At 3 months, an exercise programme including hip, knee and trunk exercises was superior to hip and knee exercises alone and to a programme including ‘minimal’ hip/knee exercises. No difference was found between minimal hip/knee exercises and usual hip/knee exercises. At 12 months, hip, knee and trunk exercises were superior to a combination of hip/knee exercises and arthroscopy, and also superior to hip/knee exercises alone. No difference was found between hip/knee exercises+arthroscopy or hip/knee exercises alone.

DISCUSSION

This living systematic review included 22 RCTs with a total of 1472 patients with PFP for more than 6 weeks. All treatments were superior to a wait-and-see approach for the primary outcome, any improvement, at 3 months. Patient education combined with a physical treatment appears most effective in the short term. After 12 months, education or education+additional exercises, orthosis or patellar treatment yielded similar improvement rates. However, no studies included wait and see for any improvement at 12 months precluding comparisons to determine if this was also inferior in the longer term. For the secondary outcome, pain intensity during previous week, no treatment was superior to wait and see.

Guidelines for musculoskeletal pain often recommend a wait-and-see approach in general practice.⁴⁹ Based on this NMA, wait and see is the least effective treatment available. Clinicians should, therefore, avoid a wait-and-see approach. Clinicians can consider a minimum of patient education at the first consultation and potentially add exercise and patellar taping/mobilisations if the patient and clinician agree on the time requirements and benefit. Costs may be another aspect to consider when offering add-on treatments. If additional treatments such as exercise and patellar taping/mobilisations have limited persisting benefit in the long term, these treatments may not represent value for money. Trials should include cost evaluations in the future.

One of the studies not included in NMA demonstrated that among runners, patient education on load management may be equally as effective as patient education and exercise

therapy or patient education and gait retraining. These findings are in contrast to the findings from the NMA that suggest education may be inferior to education+physical treatments at 3 months. Future trials may need to explore which types of education may be most effective, and how it can be improved even further as a low cost and scalable intervention. Education varied in terms of how it was delivered. Education in all trials included information about PFP, information on pain and guidance on how to manage activity in the context of pain, without stopping all exercise. This also has important implications for research, as many studies include some educational material as a control, with few available studies using a true control or placebo control. The ‘wait-and-see’ arms often also included patients continuing with their planned treatments and therefore may not reflect a true no treatment control. Despite all treatments being superior to wait and see for any improvement, there was no superior effect of active treatments compared with wait and see on pain intensity.

If a resistance exercise programme is chosen, the specific type of exercises prescribed may not be important. They mainly comprise strengthening exercises for the hip, the knee or both the hip and knee. This gives clinicians the opportunity to devise an individual prescription together with their patient. The additional benefit of adding patellar taping/mobilisations to an exercise programme is presently unknown. Shared decision making between the patient and their healthcare practitioner should guide administering the addition of a physical treatment to education. It appears that foot orthoses do not have an additive effect to education and exercise so clinicians should consider this when selecting adjuncts to education and exercise. On the other hand, orthosis and education combined had a similar effect as the combination of exercise, education and patellar treatments. This highlights two different treatment approaches with different time requirement, which appear equally effective.

Strengths and weaknesses in relation to previous studies

Previous systematic reviews and RCTs on the management of PFP were restricted to traditional comparisons of one treatment versus another, for example, patient education versus exercise. Similar to a Cochrane review,⁷ we found exercise to be superior to wait and see on any improvement, but not on pain. However, our NMA further suggests that combining exercise, (with or without foot orthoses) to education may be most effective for 3-month improvements. This is in keeping with a recent consensus statement from experts in the field of PFP that recommends exercise, orthosis, manual therapy and combined treatments in the management of PFP.⁵⁰ Our quantitative NMA supports the view that combined treatment is associated with the best outcomes, but that orthoses may not improve outcomes when added to education, exercises and patellar taping/mobilisations. Our living NMA will be updated when new evidence becomes available, ensuring a contemporary overview of the evidence for the best treatment of PFP for patients and clinicians dealing with the condition.

Strengths and weaknesses of this study

There was an overall lack of high-quality, large studies, with only one study including >100 subjects per trial arm.⁴³ Small trials with varying quality predominate in the field of musculoskeletal health. To ensure, we included credible research to inform the NMA we only included research where (1) patients had the condition for ≥6 weeks, and (2) studies included relevant

patient-rated outcomes and follow-ups (≥ 6 weeks). As a result, we excluded 128 RCTs. While the comparisons in those studies could add value to the body of knowledge, these were not evaluated by the patient—which is considered the gold standard of outcome measurement by WHO, and ‘essential’ by Cochrane.^{51 52} Consensus on which outcomes to use in RCTs through the development of a core outcome set has the potential to move the field of PFP forward. In our study, patients were involved in prioritising our outcomes making our results relevant to patients and clinical practice. We performed a comprehensive search covering published and unpublished studies, state-of-the-art risk of bias and GRADE assessments. We synthesised evidence in Bayesian hierarchical NMAs to optimally use all evidence to determine which treatment is best for PFP.

Limitations in the conclusions drawn by the NMA are primarily caused by the original data. Overall evidence was graded as very low to low. There was no evidence of heterogeneity between studies with fixed effect models preferred over random effects models based on model fit; however, this may be due to the limited number of studies available per comparison. Consistency between direct and indirect evidence could not be checked for all comparisons in the NMAs. However, where this was possible, by comparing consistency models to models that relaxed the consistency assumption, we did not find evidence of inconsistency. Treatment outcomes may be different on the basis of some characteristics (ie, effect modification). Lankhorst *et al*⁵³ suggested that sex and symptom duration may be effect modifiers in the relationship between exercise and function at 3 months, but no significant association was found. As there is no good evidence for any potential effect modifier, we did not include potential effect modifiers when planning the synthesis. This is a limitation that can be overcome in the future when new evidence for effect modifiers becomes available. Studies should break down their results on the basis of potential effect modifiers, such as sex and duration of symptoms, which would allow exploring modification effects in the future.

Treatments provided for PFP, as with any musculoskeletal condition, are diverse. For example, the mode of delivery, frequency and intensity may be different for various exercise regimes that we pooled together. This is a limitation; when more studies are done the effect of delivery, frequency and intensity of exercise could be formerly tested. We did subgroup exercise regimes per exercise region (eg, hip/knee or hip/knee/trunk exercises). The NMA suggests that it may not matter in which region exercises are performed. We ran a sensitivity analysis to test if our findings were robust for our decision to pool the study arms (education vs education+exercise) in van Linschoten *et al*²⁵ with the study arms in Collins *et al*²⁶ and Rathleff *et al*²⁷ (both education vs education+exercise+patellar treatments). The sensitivity analysis showed that this decision did not affect the NMA's comparative estimates, and second, that patellar treatments in addition to education+exercise appear not to be effective to education+exercises alone.

Publication bias could not be investigated due to a lack of studies.^{18 24} We found a number of trials in registries that may have remained unpublished but it is unclear if this was due to publication or small study bias. Collectively, bias and low study quality decrease the certainty of our findings. We chose a conservative statistical approach with the NMA, handling missing data as a treatment failure. This may balance out inflated effects in the original studies, or could even underestimate the comparative effects.

CONCLUSION

Education combined with a physical treatment (exercise, orthoses or patellar taping/mobilisation) is most likely to be effective. There was insufficient evidence to recommend a specific type of physical treatment, or a combination of physical treatments, over another. All treatments analysed in our NMA were superior to wait and see at 3 months, and we recommend avoiding a wait-and-see approach.

What is already known

- ▶ Patellofemoral is a persisting condition, four in every ten patients continue to have symptoms after 2 years. It is disabling and it impacts on quality of life.
- ▶ Many patients receive multiple treatments that impacts healthcare consumption.
- ▶ The comparative effectiveness of all available treatments is currently unknown.

What are the new findings

- ▶ Our living network meta-analysis (NMA) of randomised controlled trials suggests that education in combination with a physical treatment (exercise, orthoses or patellar taping/mobilisation) is most likely to be effective at 3 months.
- ▶ At 12 months education alone is comparable to education combined with a physical treatment (exercises, orthoses or patellar taping/mobilisation).
- ▶ All treatments analysed in our NMA were superior to wait and see, a common first-line approach currently administered by general practitioners.

Twitter Marinus Winters @marinuswinters, Sinéad Holden @Sinead_Holden and Bill T Vicenzino @Bill_Vicenzino

Acknowledgements We would like to thank Christian Lund Straszek, MSc, and the Research Unit for General Practice's patellofemoral pain patient reference group for their help with developing our research questions. We thank Rebecca Mellor, PhD, for her help with extracting data, and Arco van der Vlist, MD, for his help with the Grading of Recommendations, Assessment, Development and Evaluation assessments. We thank Negar Pourbordbari, MD, Asiah Rahi Sherzaman and Hedaïat Saei for translating Persian articles to English, and Xu Wang for translating Chinese articles to English.

Contributors MW, AW and MSR came up with the study idea. MW, SH, BT, V, AW and MSR designed the study. MW, NJW, DMC, SH and MSR designed the statistical analyses plan. MW constructed the search with help of a research librarian. MW and CBL performed the search and selection process. MW, CBL, MSR, AW, SH, and Rebecca Mellor performed data extraction. MW, CBL and AW performed risk of bias assessments. MW and NJW performed all data analyses. MW and Arco van der Vlist performed Grading of Recommendations, Assessment, Development and Evaluation assessments. MW, SH and MSR drafted the first version of the manuscript. MW and MSR are the study guarantors. All authors provided feedback and gave important intellectual input. All authors read and consented to the content of the article.

Funding The Tryg Foundation is acknowledged for provided support for this project (Grant ID: 118547). The foundation had no role in the planning, conduct or reporting of this work.

Competing interests NJW lead a research project in collaboration with Pfizer plc. Pfizer part-funded a junior researcher. The projects is purely methodological, using historical data on treatments for pain relief. NJW has no other conflicts. All other authors report to have no conflicts of interest.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content

includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Marinus Winters <http://orcid.org/0000-0001-5742-7441>

Sinéad Holden <http://orcid.org/0000-0002-7314-2152>

Bill T Vicenzino <http://orcid.org/0000-0003-0253-5933>

REFERENCES

- Fatoye F. The economic impact of musculoskeletal pain. *Pain and Rehabilitation* 2018;3–4.
- Smith BE, Selve J, Thacker D, et al. Incidence and prevalence of patellofemoral pain: a systematic review and meta-analysis. *PLoS One* 2018;13:e0190892.
- Coburn SL, Barton CJ, Filbay SR, et al. Quality of life in individuals with patellofemoral pain: a systematic review including meta-analysis. *Phys Ther Sport* 2018;33:96–108.
- Rathleff MS, Rathleff CR, Olesen JL, et al. Is knee pain during adolescence a self-limiting condition? prognosis of Patellofemoral pain and other types of knee pain. *Am J Sports Med* 2016;44:1165–71.
- Lankhorst NE, van Middelkoop M, Crossley KM, et al. Factors that predict a poor outcome 5-8 years after the diagnosis of patellofemoral pain: a multicentre observational analysis. *Br J Sports Med* 2016;50:881–6.
- Smith BE, Hendrick P, Bateman M, et al. Current management strategies for patellofemoral pain: an online survey of 99 practising UK physiotherapists. *BMC Musculoskelet Disord* 2017;18:181.
- van der Heijden RA, Lankhorst NE, van Linschoten R, et al. Exercise for treating patellofemoral pain syndrome. *Cochrane Database Syst Rev* 2015;1:Cd010387.
- Callaghan MJ, Selve J. Patellar taping for patellofemoral pain syndrome in adults. *Cochrane Database Syst Rev* 2012;4:Cd006717.
- Hossain M, Alexander P, Burls A, et al. Foot orthoses for patellofemoral pain in adults. In: *Cochrane database of systematic reviews*. John Wiley & Sons, Ltd, 2010.
- Shojania KG, Sampson M, Ansari MT, et al. How quickly do systematic reviews go out of date? A survival analysis. *Ann Intern Med* 2007;147:224–33.
- Elliott JH, Turner T, Clavisi O, et al. Living systematic reviews: an emerging opportunity to narrow the evidence-practice gap. *PLoS Med* 2014;11:e1001603.
- Winters M, Holden S, Vicenzino B, et al. Which treatment is most effective for patients with patellofemoral pain? A protocol for a living systematic review including network meta-analysis. *BMJ Open* 2018;8:e022920.
- Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med* 2015;162:777.
- Kamper SJ, Ostelo RWJG, Knol DL, et al. Global perceived effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. *J Clin Epidemiol* 2010;63:760–6.
- Crossley KM, Bennell KL, Cowan SM, et al. Analysis of outcome measures for persons with patellofemoral pain: which are reliable and valid? *Arch Phys Med Rehabil* 2004;85:815–22.
- Cicchetti DV, Sparrow SA. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *Am J Ment Defic* 1981;86:127–37.
- Crossley KM, Macri EM, Cowan SM, et al. The patellofemoral pain and osteoarthritis subscale of the KOOS (KOOS-PF): development and validation using the COSMIN checklist. *Br J Sports Med* 2018;52:1130–6.
- Green S, Higgins JPT. *Preparing a cochrane review*. *Cochrane Handbook for systematic reviews of interventions*. John Wiley & Sons, Ltd: 11–30.
- Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014;348:g1687.
- Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:l4898.
- Chaimani A, Higgins JPT, Mavridis D, et al. Graphical tools for network meta-analysis in STATA. *PLoS One* 2013;8:e76654.
- Dias S, Sutton AJ, Ades AE, et al. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Med Decis Making* 2013;33:607–17.
- Dias S, Welton NJ, Sutton AJ, et al. Evidence synthesis for decision making 4: inconsistency in networks of evidence based on randomized controlled trials. *Med Decis Making* 2013;33:641–56.
- Chaimani A, Higgins JPT, Mavridis D, et al. Graphical tools for network meta-analysis in STATA. *PLoS One* 2013;8:e76654.
- van Linschoten R, van Middelkoop M, Berger MY, et al. Supervised exercise therapy versus usual care for patellofemoral pain syndrome: an open label randomised controlled trial. *BMJ* 2009;339:b4074.
- Collins N, Crossley K, Beller E, et al. Foot orthoses and physiotherapy in the treatment of patellofemoral pain syndrome: randomised clinical trial. *BMJ* 2008;337:a1735.
- Rathleff MS, Roos EM, Olesen JL, et al. Exercise during school hours when added to patient education improves outcome for 2 years in adolescent patellofemoral pain: a cluster randomised trial. *Br J Sports Med* 2015;49:406–12.
- Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924–6.
- The GRADE Working Group. *Grade Handbook for grading quality of evidence and strength of recommendations*, 2013. Available: <https://gdt.gradepro.org/app/handbook/handbook.html>
- Baldon RdeM, Serrão FV, Scattone Silva R, et al. Effects of functional stabilization training on pain, function, and lower extremity biomechanics in women with patellofemoral pain: a randomized clinical trial. *J Orthop Sports Phys Ther* 2014;44:240–8.
- Demirci S, Kinikli GI, Callaghan MJ, et al. Comparison of short-term effects of mobilization with movement and Kinesiotaping on pain, function and balance in patellofemoral pain. *Acta Orthop Traumatol Turc* 2017;51:442–7.
- Drew BT, Conaghan PG, Smith TO, et al. The effect of targeted treatment on people with patellofemoral pain: a pragmatic, randomised controlled feasibility study. *BMC Musculoskelet Disord* 2017;18:338.
- Emamvirdi M, Letafatkar A, Khaleghi Tazji M. The effect of valgus control instruction exercises on pain, strength, and functionality in active females with Patellofemoral pain syndrome. *Sports Health* 2019;11:223–37.
- Eng JJ, Pierrynowski MR. Evaluation of soft foot orthotics in the treatment of patellofemoral pain syndrome. *Phys Ther* 1993;73:62–8.
- Esculier J-F, Bouyer LJ, Dubois B, et al. Is combining gait retraining or an exercise programme with education better than education alone in treating runners with patellofemoral pain? A randomised clinical trial. *Br J Sports Med* 2018;52:659–66.
- Foroughi F, Sobhani S, Yoosefinejad AK, et al. Added Value of Isolated Core Postural Control Training on Knee Pain and Function in Women With Patellofemoral Pain Syndrome: A Randomized Controlled Trial. *Arch Phys Med Rehabil* 2019;100:220–9.
- Fukuda TY, Melo WP, Zaffalon BM, et al. Hip posterolateral musculature strengthening in sedentary women with patellofemoral pain syndrome: a randomized controlled clinical trial with 1-year follow-up. *J Orthop Sports Phys Ther* 2012;42:823–30.
- Glaviano NR, Marshall AN, Mangum LC, et al. Impairment-Based rehabilitation with patterned electrical neuromuscular stimulation and lower extremity function in individuals with Patellofemoral pain: a preliminary study. *J Athl Train* 2019;54:255–69.
- Giles L, Webster KE, McClelland J, et al. Quadriceps strengthening with and without blood flow restriction in the treatment of patellofemoral pain: a double-blind randomised trial. *Br J Sports Med* 2017;51:1688–94.
- Hart JM, Kuenze C, Norte G, et al. Prospective, randomized, double-blind evaluation of the efficacy of a single-dose hyaluronic acid for the treatment of Patellofemoral chondromalacia. *Orthop J Sports Med* 2019;7:232596711985419.
- Hott A, Brox JI, Pripp AH, et al. Effectiveness of isolated hip exercise, knee exercise, or free physical activity for Patellofemoral pain: a randomized controlled trial. *Am J Sports Med* 2019;47:1312–22.
- Kettunen JA, Harilainen A, Sandelin J, et al. Knee arthroscopy and exercise versus exercise only for chronic patellofemoral pain syndrome: a randomized controlled trial. *BMC Med* 2007;5:38.
- Matthews M, Rathleff MS, Claus A, et al. Does foot mobility affect the outcome in the management of patellofemoral pain with foot orthoses versus hip exercises? a randomised clinical trial. *Br J Sports Med* 2020:bjsports-2019-100935 (published Online First: 2020/03/29).
- Mills K, Blanch P, Dev P, et al. A randomised control trial of short term efficacy of in-shoe foot orthoses compared with a wait and see policy for anterior knee pain and the role of foot mobility. *Br J Sports Med* 2012;46:247–52.
- Petersen W, Ellermann A, Rembitzki IV, et al. Evaluating the potential synergistic benefit of a realignment brace on patients receiving exercise therapy for patellofemoral pain syndrome: a randomized clinical trial. *Arch Orthop Trauma Surg* 2016;136:975–82.
- Witvrouw E, Lysens R, Bellemans J, et al. Open versus closed kinetic chain exercises for patellofemoral pain. A prospective, randomized study. *Am J Sports Med* 2000;28:687–94.
- Yilmaz GD. The effect of postural stabilization exercises on pain and function in females with patellofemoral pain syndrome. *Acta Orthop Traumatol Turc* 2015.
- Riel H, Matthews M, Vicenzino B, et al. Efficacy of live feedback to improve objectively monitored compliance to prescribed, home-based, exercise therapy-dosage in 15 to 19 year old adolescents with patellofemoral pain- a study protocol of a randomized controlled superiority trial (The XRCISE-AS-INSTRUCted-1 trial). *BMC Musculoskelet Disord* 2016;17:242.
- Belo J, Bierma-Zeinstra S, Kuijpers T, et al. Niet-traumatische knieklachten; NHG-standaard. *Huisarts Wet* 2016;59:62–6.

- 50 Collins NJ, Barton CJ, van Middelkoop M, *et al.* 2018 consensus statement on exercise therapy and physical interventions (orthoses, taping and manual therapy) to treat patellofemoral pain: recommendations from the 5th International Patellofemoral pain research retreat, gold Coast, Australia, 2017. *Br J Sports Med* 2018;52:1170–8.
- 51 World Health Organization, Regional Office for Europe. Patient outcome measures in mental health: report on a WHO consensus meeting, Stockholm, 23–24 November 1995. Copenhagen WHO Regional Office for Europe; 1996.
- 52 Johnston BC, Patrick DL, Devji T, *et al.* Chapter 18: Patient-reported outcomes. In: Higgins JPT, Thomas J, Chandler J, *et al.*, eds. *Cochrane Handbook for systematic reviews of interventions version 6.0*, 2019. www.training.cochrane.org/handbook
- 53 Lankhorst NE, van Middelkoop M, van Trier YDM, *et al.* Can we predict which patients with patellofemoral pain are more likely to benefit from exercise therapy? A secondary exploratory analysis of a randomized controlled trial. *J Orthop Sports Phys Ther* 2015;45:183–9.