

BMJ Open Which treatment is most effective for patients with patellofemoral pain? A protocol for a living systematic review including network meta-analysis

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

To cite: 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. doi:10.1136/bmjopen-2018-022920

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2018-022920>).

Received 14 March 2018
Revised 23 August 2018
Accepted 8 October 2018



© Author(s) (or their employer(s)) 2018. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Marinus Winters;
marinuswinters@hotmail.com

ABSTRACT

Introduction Patellofemoral pain (PFP) affects 1 in every 14 adults. Many treatments for PFP have been evaluated, but the comparative effectiveness of all available treatments has never been examined. Network meta-analysis is the only design to study the comparative effectiveness of all available treatments in one synthesis. This protocol describes the methods for a systematic review including network meta-analysis to assess which treatment is most likely to be effective for patients with PFP.

Methods and analysis The primary outcome measures of this network meta-analysis are the global rating of change scale at 6–12 weeks, 13–52 weeks and >52 weeks. The secondary outcome measures are patient-rated pain scales at 6–12 weeks, 13–52 weeks and >52 weeks. Completed published and unpublished randomised controlled trials with full-text reports are eligible for inclusion. We will search Embase, PubMed (including MEDLINE), CENTRAL, Scopus, Web of Science, and CINAHL, SPORTDiscus, OpenGrey, WorldCat, conference Proceedings and multiple trial registers for relevant reports. Two researchers will appraise the study eligibility and perform data extraction. Risk of bias will be assessed with the Cochrane Risk of Bias Tool V.2.0. Bayesian network meta-analyses will be constructed for global rating of change scale and patient-rated pain. Consistency between direct and indirect comparisons will be assessed. Between study variability will be explored, and a threshold analysis for the credibility of the network meta-analyses' conclusions will be performed.

Ethics and dissemination Ethical approval is not required, as this study will be based on published data. The study commenced at 1 February 2018, and its expected completion date is 15 January 2019. Full publication of the work will be sought in an international peer-reviewed journal, as well as translational articles to disseminate the work to clinical practitioners.

PROSPERO registration number CRD42018079502.

INTRODUCTION

Patellofemoral pain (PFP) affects 1 in every 14 adolescents and 1 in every 8 adults.¹ PFP

Strengths and limitations of this study

- This living systematic review will include thorough search methods, searching conventional databases, grey literature resources and trial registers.
- Risk of bias in randomised controlled trials will be appraised using the new Cochrane Risk of Bias tool V.2.0 for intervention studies.
- This living systematic review and network meta-analysis enables clinicians to consult a contemporary, comprehensive overview of the comparative effectiveness of treatments for patellofemoral pain.
- The feasibility of this study is depending on the availability and the homogeneity of the trials and the consistency between direct and indirect evidence.

is characterised by diffuse pain around or behind the knee cap, provoked during activities that load the knee-joint, such as stair climbing, running and jumping.² One in every two patients with PFP continue to suffer from knee pain, which can impact their quality of life and physical activity.^{3,4}

Similar to other chronic musculoskeletal pain conditions, there are many different treatments. Recent recommendations from an expert panel based on the available evidence are for the use of exercise of the hip and knee, foot orthoses and combinations that include patellar taping or manual therapy.⁵ Patient education and gait retraining have been recently promoted as well but with little research support.^{5,6} While there are several systematic reviews that focus on different treatments for PFP,^{7–12} the comparative effectiveness of all available treatments has never been examined. This is challenging for clinicians and patients who are faced with uncertainty when presented with so many potentially beneficial treatment options.

Traditional systematic reviews present fragmented pairwise ‘head to head’ comparisons, for example, treatment A versus B, and treatment B versus C. The limitation with this approach is that multiple treatments cannot be compared simultaneously (ie, treatment A vs B vs C). The traditional approach may lead to invalid interpretations regarding the comparative effectiveness of treatments.¹³ Clinicians are left to speculate on which treatment is most effective, based on multiple, independent ‘head to head’ comparisons. Network meta-analyses offer the opportunity to combine both direct and indirect treatment comparisons in a single analysis, which overcomes main limitations of pairwise systematic reviews. They do this by allowing for:

- ▶ A coherent comparison of effectiveness of multiple treatments in one statistical model while maintaining the randomised nature of the evidence.
- ▶ Comparison of treatments even if the treatments have not been investigated directly in a randomised controlled trial (RCT).^{14–16}

Based on the network meta-analysis (NMA), a ranking from ‘most likely to be effective’ to ‘least likely to be effective’ treatment (for a given outcome) can be estimated. In this way, the results from the NMA can directly feed into shared decision making in clinical practice.

A common critique on systematic reviews is that they are soon out of date.¹⁷ Living NMAs are particularly suitable to control for this issue as they are regularly updated, preferably as open access content. This enables clinicians to consult a comprehensive overview of the comparative effectiveness of treatments for a given condition, while ensuring a contemporary evidence synthesis for clinical practice (table 1).^{18 19}

The comparative effectiveness of all studied treatments for patients with PFP has never been examined. The aim of this living systematic review with NMA is to evaluate 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 protocol for this living systematic review with NMA is registered on PROSPERO (CRD42018079502). This protocol follows the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols and Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for NMA checklist for reporting systematic review protocols and NMA.^{20–22}

Patient involvement and prioritising outcomes

Patients with PFP (n=7) from a patient reference group have been involved in setting a hierarchy of outcomes (global rating of change scale (GROC) and pain scales) for this NMA. One researcher, otherwise not involved in the study (see acknowledgements), contacted the patient panel members by phone. He explained the various outcomes. All participants were subsequently sent a list and asked to indicate the most relevant instrument to judge their knee pain. Six out of seven (86%) indicated a preference for the GROC over pain outcomes. Consequently, the outcomes selected are as follows:

Primary outcome measure

- ▶ GROC: this scale usually has seven descriptors for perceived change: completely recovered, strongly recovered, slightly recovered, unchanged, slightly worse, strongly worse and worse than ever. The reliability of the GROC is excellent with intraclass correlation coefficient (ICC) from 0.90 to 0.99.^{23 24}

Secondary outcome measures

- ▶ Pain intensity, measured by ‘worst pain in the previous week’ on a visual analogue scale (VAS) (0–10/0–100) or numerical rating pain scale (NRPS) (0–10/0–100). The reliability is excellent, ICC=0.76.^{24 25}
- ▶ Patient-rated pain during specific activities of daily life (ADLs) and during sporting activities. We will synthesise one pain outcome for ADL and one for sporting activities. The choice for these outcomes will be made based on availability, an outcome that allows for inclusion of the highest number of comparisons. Pain will

Table 1 The advantages of a living network meta-analysis (NMA) compared with traditional systematic reviews

	Traditional systematic review+meta-analysis	Systematic review+NMA	Living systematic review+NMA
Direct comparison between treatments	X	X	X
Indirect comparisons between treatments that have never been compared in an RCT		X	X
Research question	X	X	X
▶ Which treatment is most effective: A or B?		X	X
▶ Which of the many available treatments that have been tested in randomised trials are most effective?			
Always up-to-date best evidence synthesis to inform clinical practice			X

RCT, randomised controlled trial.

be expressed a VAS (0–10/0–100) or NRPS (0–10/0–100). Reliability for pain during activity is excellent, ICC=0.83.^{24 25}

Research questions

1. Which treatment(s) is most likely to be effective for patients with PFP on global rating of change and patient-rated pain?
2. Which treatment class(es) is most likely to be effective for patients with PFP on global rating of change and patient-rated pain? The study of treatment classes is relevant when more than one subtype for a treatment is available, for example, multiple types of exercise regimes, which can be grouped together to answer this question.

Eligibility criteria

Type of studies

Published or unpublished RCTs (including randomisation through minimisation or clustering), for which a full-text report or full-text protocol of a completed trial is available, are eligible for inclusion.

Type of population

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

Type of treatments and control treatments

Any treatment, control treatment, placebo, wait-and-see or no treatment group studied in an RCT is eligible for inclusion. Examples of treatment classes are exercise therapy, orthoses, braces, patient education, pain medicine or surgery.

Type of outcomes

Studies assessing the treatment effect after a minimum of 6 weeks will be included. Studies assessing the following outcomes will be included:

- ▶ GROC.
- ▶ Worst pain in the previous week, measured with a VAS (0–10) or NRPS (0–100).
- ▶ Patient-rated pain during activities of daily living and sporting activities, measured with a VAS (0–10) or NRPS (0–100).

Search strategy

A sensitive search strategy has been developed for each of the data sources by a research librarian and one investigator (MW). We used the Cochrane sensitive search strategy for RCTs and modified this for the purpose of our study.²⁶ The search strategy includes a mix of indexed and free text terms, where applicable (online supplementary appendix 1). No restrictions (eg, language or full-text availability) were applied to the search.

One investigator (MW) will search conventional databases, grey literature databases and trial registers from their date of inception. Online supplementary appendix 1 provides a detailed explanation of how the search is built and with source-specific search strategies for each database, grey literature sources and trial registers.

Conventional databases

Conventional electronic databases Embase, PubMed (including MEDLINE), Cochrane Central Register of Controlled Trials (CENTRAL), Scopus, Web of Science, and CINAHL and SPORTDiscus (both via EBSCO) will be searched for relevant reports.

Identifying grey literature and ongoing studies

Databases

OpenGrey.eu and WorldCat.org will be searched for studies that have remained unpublished.

Conference proceedings

We will search the conference proceedings from all Patellofemoral Research Retreats (2009, 2011, 2013, 2015 and 2017) for relevant reports and request authors to make available their full reports or protocols for unpublished studies.

Trial registers

We will search the WHO International Clinical Trials Registry Platform (<http://apps.who.int/trialsearch/>) ClinicalTrials.gov, The European Union Clinical Trials Register and the ISRCTN registry for unpublished or ongoing studies.

Hand searching

We will screen reference lists of all Cochrane reviews (n=6) on PFP for possible relevant studies that were not identified by our search. We will also screen reference lists of all the reports included in our systematic review.

Study selection

Two researchers will screen titles and abstracts independently, after duplicate removal by one of the investigators. Consensus will be sought in case of initial disagreement. If consensus cannot be reached, the report will be included for full-text evaluation.

Both investigators will independently apply inclusion and exclusion criteria to the full-text reports. In case of disagreement, consensus will be sought; however, if disagreement persists, a third author (AW) will take the decision.

Data extraction

Data will be independently extracted by two researchers using standardised extraction forms adopted from the Cochrane Collaboration (see online supplementary appendix 2).²⁷ Disagreements will be resolved by seeking consensus and by a third reviewer (AW) in case of persistent disagreement. The following data will be extracted:

- ▶ Publication and study details: for example, authors, year of publication, funding source, possible conflicts of interest, aim study, design and unit of allocation.
- ▶ Population: number of included patients, population characteristics for age, sex, body mass index, activity level, setting where population was recruited, baseline scores for outcome measures (mean, SDs, SEs 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) and adherence. We used items from the Template for Intervention Description and Replication checklist to assure comprehensive data extraction in this section of the extraction form.²⁸
- ▶ Outcomes: time points measured, and the time points 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 and side effects).
- ▶ Data and analysis: comparisons, outcomes, subgroups, time points, results (central estimates and measures of dispersion; for example, mean for both groups, mean difference (MD), SDs/95% CIs/SEs), 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 2.0 will be used to assess the risk of bias for each outcome per study and for outcomes across a (direct) comparison. In this tool risk of bias can be assessed following the 'intention-to-treat' principle (ie, assignment to intervention) or 'per protocol' (ie, adherence to intervention). We will assess risk of bias on the basis of 'assignment to intervention'. This new tool has a fixed set of items to use for the risk of bias appraisal, that is, 'bias arising from the randomization 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.^{29 30}

Two experienced reviewers will independently assess the risk of bias for each outcome within the study for each follow-up. They will trial the approach by assessing 20 RCTs in other musculoskeletal conditions, before the study starts. Each major domain of bias will be appraised in light of each outcome. The tool's signalling questions and criteria will be followed to inform a domain-based appraisal of the risk of bias.^{29 30} The risk of distortion of the outcome estimate by the methodology will be appraised as at 'low', 'some' or 'high' risk of bias. Judgements will be made regarding the direction of distortion 'favours experimental', 'favours comparator', 'towards null', 'away from null' or 'unpredictable'. Each outcome within a study will receive an overall risk of bias judgement based on the individual domains: 'low', 'some' or 'high' risk of bias.^{29 30}

In case of disagreements between reviewers, consensus will be sought through discussions. If consensus is not met, a third reviewer (AW) will take the decision.

Data synthesis and statistical methods

We plan a NMA to assess which treatment for PFP is most effective. Networks of treatment comparisons will be constructed for the primary and (each) secondary outcome separately.

Three authors (MW, SH and MSR) will appraise the clinical homogeneity before any analysis is commenced by tabulating study and population characteristics and inspecting them for differences in potential effect modifiers. This is to assess the assumption of exchangeability required for NMA. In addition, treatments will be assigned to a class, for example, exercise therapy, surgery and drug therapy.

Bayesian NMA

We will model networks following the Bayesian approach, 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). Direct, pair-wise comparisons will be estimated first. For treatments that are connected in a network of comparisons from our included studies, we will estimate relative treatment effects using NMA, and hierarchical NMA using classes if possible.^{31 32}

Our primary outcome measure, the GROC, will be synthesised using a proportional odds model and expressed with an OR and their 95% credible interval, if GROC across studies are similar. Otherwise, GROC will be dichotomised at a common cut-off point where all scales coincide, for example, improved/not improved, recovered/not recovered. In the latter case, a logistic regression model will be run.

For our secondary outcome measures, continuous outcomes will be presented as MD, with their 95% credible intervals when outcomes are measured with the same instrument. We will present standardised MDs if different continuous measures are used to evaluate the same construct.

For all analyses, we will fit both fixed and random effects models and compare model fit using the deviance information criterion and posterior mean residual deviance. A lower deviance depicts a better model fit. We will group outcome follow-ups based on the available data, seeking the following approximate timeframes: 6–12 weeks, 13–52 weeks and >52 weeks. If there are multiple time points available for an outcome, and these are equally close to the time point to be synthesised across studies, the last follow-up in this timeframe will be used. For >52 weeks, a slightly different approach will be followed, where multiple time points will be synthesised following available data. We will make attempts to model a time course function for pain scales instead of analysis for multiple timeframes, if possible.

Surface under the cumulative ranking curves and probability ranks will be used to estimate the likelihood of individual treatments being superior than the other treatments for the individual with PFP.

Assessing statistical heterogeneity and exploring it with individual patient data

Statistical heterogeneity will be assessed by inspecting the between study SD, comparing fit of the fixed and random effect models. Depending on resources and data availability, individual patient data from a previous RCT by our group will be used together with study level data to explore statistical heterogeneity.³³ Otherwise, only study level data will be used. The following factors are considered for exploration when sufficient data are available (>10 studies/events per variable), in the following order: diagnostic approach used (clinical vs imaging), pain intensity, symptom duration, active or sedentary population, age, sex (male/female), quality of life, unilateral versus bilateral pain and publication status (published/unpublished).^{34 35}

Exploring inconsistency in the network

The consistency assumption will be tested for each network. We will compare results from a model that assumes consistency with a model that relaxes the consistency assumption to assess whether there is evidence of inconsistency. For this purpose, we will compare the models' residual deviance and deviance information criterion to examine model fit. If we identify evidence of inconsistency, we will use the node-split method to identify where in the network the inconsistency is.³⁶ We will use a Bonferroni correction for interpreting multiple p values.

Assessing small study bias

Where possible, we will use comparison-adjusted funnel plots to examine small study bias. In this case, we assume that small study bias is consistent across comparisons, and experimental treatments are more likely to be favoured in small studies compared with control treatments/groups. The funnel plot will be evaluated for its distribution, where missing small studies are expected favouring

the control treatment in the presence of small study bias. Funnel plots will be generated for each outcome, but only when ≥ 10 studies are available.³⁷ Conventional funnel plots for pairwise comparisons are constructed if comparison-adjusted funnel plots cannot be constructed.²⁶

Threshold analysis for credibility of the NMA's conclusions

Risk of bias in the pair-wise estimates may distort the reliability of the network's estimate and can, therefore, affect the credibility of the NMA's conclusion. We will investigate if bias in the estimate for global rating of change and pain would change the posterior mean treatment effect and hence the recommended treatment based on the probability ranks.³⁸ We will perform a threshold analysis where the variance around the bias estimate is assumed to be 0. We assume bias for both measures to overestimate or underestimate treatment effects by maximally 20%, following empirical estimations of bias by Page *et al*, Wood *et al* and Armijo-Olivo *et al*.^{39–41} The threshold analysis will be run with steps of 5% to detect the level at which bias may attenuate rankings.

Potential limitations of the planned work

NMA allows multiple interventions to be compared simultaneously and can form a coherent basis for intervention recommendations. Notwithstanding this, with any evidence synthesis, the quality of the planned work is dependent on the availability of study data and the comparisons investigated to allow the construction of a network. NMA relies on connected networks of evidence; it is not possible to make comparisons between interventions that are unconnected. The method assumes that the evidence is consistent, so that the intervention effects observed directly in head-to-head studies are in agreement with those obtained indirectly via the network of comparisons. It is therefore essential to check the consistency assumption when possible (ie, both direct and indirect evidence are available). As with all evidence syntheses, the NMA estimates reflect the evidence available including the limitations in that evidence. Assessment of risk of bias of the included studies is therefore essential. Exploration of heterogeneity through subgroup analysis is limited by the evidence available with limited power to detect effects and may suffer from aggregation bias. There are also limitations to the living nature of the proposed research. Living reviews are labour intensive and rely on regular updates. Moreover, the chance of type 1 errors, that is, incorrectly concluding there is a significant effect in the meta-analysis, increases with the growing number of updates.

Administration, dissemination and updating the living systematic review

The living systematic review will be administered at the Research Unit for General Practice in Aalborg, and we plan to update the NMA for at least 5 years. The study started on 1 February 2018, and the expected completion date for its first version is 15 January 2019. The search

and review process will be updated every 12 months, if needed. When new data have become available, we will update the analysis and present the updated findings at the website of Aalborg University. Here, we will also provide a plain language summary for patients and clinicians dealing with PFP. If there is a change in conclusions, republication will be sought in an international peer-reviewed journal. We will seek presentation of the study results on national and international conferences, and we will submit the full-text report for 'open access' publication in an international peer-reviewed journal.

PERSPECTIVES

Systematic reviews should inform clinical practice and treatment decisions. When multiple treatments exist, traditional systematic reviews come shorthanded. NMA is the only design that can study the comparative effectiveness of *all* available treatments for a condition. Patients and clinicians dealing with PFP are in urgent need of evidence rather than expert opinion-based guidance for the treatment of this often long-living condition. NMA will rank treatments according to their probability of being the most effective treatment. In this way, it directly informs the clinician and patient when making a shared decision-making on how to treat PFP. The 'living' nature of this NMA facilitates to make an informed shared decision in clinical practice based on the latest level 1 evidence.

ETHICS AND DISSEMINATION

Ethical approval is not required, as this study will be based on published data. The study commenced at 1 February 2018, and its expected completion date is 15 January 2019. Full publication of the work will be sought in an international peer-reviewed journal, as well as translational articles to disseminate the work to clinical practitioners.

Author affiliations

¹Research Unit for General Practice in Aalborg, Department of Clinical Medicine, Aalborg University, Aalborg, Denmark

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

³School of Health and Rehabilitation Sciences: Physiotherapy, The University of Queensland, Queensland, Queensland, Australia

⁴Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

⁵Department of Orthopaedics, Erasmus MC University Medical Center for Groin Injuries, Rotterdam, The Netherlands

⁶Sports Groin Pain Centre, Aspetar Orthopaedic and Sports Medicine Hospital, Doha, Qatar

Acknowledgements We would like to thank Christian Lund Straszek, MSc, and the Research Unit for General Practice's PFP patient reference group for their help with developing our research questions. We also would like to extend our gratitude to Mette Buje Grundsøe, research librarian, for her help with developing a sensitive search strategy. We would like to thank Dr Roy Elbers, from the Department for Population Health Sciences, University of Bristol, for his highly valuable feedback on the risk of bias approach.

Contributors MW, AW and MSR came up with the study idea. MW, SH, BV, AW and MSR designed the study. MW and CBL designed the risk of bias approach, and MW, SH, NJW, DMC and MSR designed the statistical analysis plan. MW, SH and

MSR drafted the manuscript. All authors provided feedback and gave important intellectual input. All authors read and consented to the content of the article.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests MW receives funding outside this project from Trygfonden, a non-profit organisation in Denmark.

Patient consent Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

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/>.

REFERENCES

- 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.
- Crossley KM, van Middelkoop M, Callaghan MJ, *et al*. 2016 Patellofemoral pain consensus statement from the 4th International Patellofemoral Pain Research Retreat, Manchester. Part 2: recommended physical interventions (exercise, taping, bracing, foot orthoses and combined interventions). *Br J Sports Med* 2016;50:844–52.
- 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.
- 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.
- Barton CJ, Crossley KM. Sharing decision-making between patient and clinician: the next step in evidence-based practice for patellofemoral pain? *Br J Sports Med* 2016;50:833–4.
- van der Heijden RA, Lankhorst NE, van Linschoten R, *et al*. Exercise for treating patellofemoral pain syndrome. *Cochrane Database Syst Rev* 2015;1:CD010387.
- Barton CJ, Lack S, Hemmings S, *et al*. The 'best practice guide to conservative management of patellofemoral pain': incorporating level 1 evidence with expert clinical reasoning. *Br J Sports Med* 2015;49:923–34.
- Heintjes E, Berger MY, Bierma-Zeinstra SM, *et al*. Pharmacotherapy for patellofemoral pain syndrome. *Cochrane Database Syst Rev* 2004(3):CD003470.
- Smith TO, Drew BT, Meek TH, *et al*. Knee orthoses for treating patellofemoral pain syndrome. *Cochrane Database Syst Rev* 2015;8:CD010513.
- Hossain M, Alexander P, Burls A, *et al*. Foot orthoses for patellofemoral pain in adults. *Cochrane Database Syst Rev* 2011;1:CD008402.
- Callaghan MJ, Selve J. Patellar taping for patellofemoral pain syndrome in adults. *Cochrane Database Syst Rev* 2012;4:CD006717.
- Edwards SJ, Clarke MJ, Wordsworth S, *et al*. Indirect comparisons of treatments based on systematic reviews of randomised controlled trials. *Int J Clin Pract* 2009;63:841–54.
- Jansen JP, Fleurence R, Devine B, *et al*. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR task force on indirect treatment comparisons good research practices: part 1. *Value Health* 2011;14:417–28.
- Hoaglin DC, Hawkins N, Jansen JP, *et al*. Conducting indirect-treatment-comparison and network-meta-analysis studies: report of the ISPOR task force on indirect treatment comparisons good research practices: part 2. *Value Health* 2011;14:429–37.
- Ioannidis J. Next-generation systematic reviews: prospective meta-analysis, individual-level data, networks and umbrella reviews. *Br J Sports Med* 2017;51:1456–8.

17. Ioannidis JP. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q* 2016;94:485–514.
18. Vandvik PO, Brignardello-Petersen R, Guyatt GH. Living cumulative network meta-analysis to reduce waste in research: a paradigmatic shift for systematic reviews? *BMC Med* 2016;14:59.
19. 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.
20. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.
21. Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 2015;349:g7647.
22. 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;162:777–84.
23. Kamper SJ, Ostelo RW, 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.
24. 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.
25. 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.
26. In: Higgins JPT, Green S, eds. *Cochrane handbook for systematic reviews of interventions, Version 5.1.0*: The Cochrane Collaboration, 2011. (updated March 2011).
27. The Cochrane Collaboration. Data Collections Forms for intervention reviews. <http://training.cochrane.org/resource/data-collection-forms-intervention-reviews>
28. 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.
29. Higgins JPT, Sterne JAC, Savović J, et al. A revised tool for assessing risk of bias in randomized trials. In: Chandler J, McKenzie J, Boutron I, Welch V, eds. *Cochrane methods. Cochrane database of systematic reviews*, 2016.
30. Higgins JPT, Savović J, Page MJ, et al. on behalf of the development group for ROB 2.0. Revised Cochrane risk of bias tool for randomised trial (RoB 2.0). 2016;23 <https://sites.google.com/site/riskofbiastool/welcome/rob-2-0-tool>
31. Ades AE, Caldwell DM, Reken S, et al. *Evidence Synthesis of Treatment Efficacy in Decision Making: A Reviewer's Checklist [Internet]* London: National Institute for Health and Care Excellence (NICE); 2012 Jan. NICE DSU Technical Support Document No. 7. London: NICE Decision Support Unit Technical Support Documents.
32. Owen RK, Tincello DG, Keith RA. Network meta-analysis: development of a three-level hierarchical modeling approach incorporating dose-related constraints. *Value Health* 2015;18:116–26.
33. 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.
34. Matthews M, Rathleff MS, Claus A, et al. Can we predict the outcome for people with patellofemoral pain? A systematic review on prognostic factors and treatment effect modifiers. *Br J Sports Med* 2017;51.
35. Winters M, Weir A. Grey matters; on the importance of publication bias in systematic reviews. *Br J Sports Med* 2017;51:488–9.
36. Dias S, Welton NJ, Caldwell DM, et al. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010;29(7-8):932–44.
37. Chaimani A, Higgins JP, Mavridis D, et al. Graphical tools for network meta-analysis in STATA. *PLoS One* 2013;8:e76654.
38. Caldwell DM, Ades AE, Dias S, et al. A threshold analysis assessed the credibility of conclusions from network meta-analysis. *J Clin Epidemiol* 2016;80:68–76.
39. Page MJ, Higgins JP, Clayton G, et al. Empirical evidence of study design biases in randomized trials: systematic review of meta-epidemiological studies. *PLoS One* 2016;11:e0159267.
40. Wood L, Egger M, Gluud LL, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ* 2008;336:601–5.
41. Armijo-Olivo S, Saltaji H, da Costa BR, et al. What is the influence of randomisation sequence generation and allocation concealment on treatment effects of physical therapy trials? A meta-epidemiological study. *BMJ Open* 2015;5:e008562.