Journal of Orthopaedic & Sports Physical Therapy
Official Publication of the Orthopaedic and Sports Physical Therapy Sections of the American Physical Therapy Association

Systematic Review of the Quality of Randomized Controlled Trials for Patellofemoral Pain Syndrome

Mario Bizzini, PT, MS[1]
Capt. John D. Childs, PT, MS, MBA, OCS, CSCS, FAAOMPT[2]
Sara R. Piva, PT, MS, OCS, FAAOMPT[3]
Anthony Delitto, PT, PhD, FAPTA[4]

Study Design: Systematic review of the literature.
Objectives: To develop a grading scale to judge the quality of randomized clinical trials (RCTs) and conduct a systematic review of the published RCTs that assess nonoperative treatments for patellofemoral pain syndrome (PFPS).
Background: Systematic reviews of the quality and usefulness of clinical trials allow for efficient synthesis and dissemination of the literature, which should facilitate clinicians' efforts to incorporate principles of evidence-based practice in the clinical decision-making process.
Methods and Measures: Using a scale based on criteria in the Cochrane Collaboration Handbook, we sought to critically appraise the methodology used in RCTs related to the nonoperative management of PFPS, synthesize and interpret our results, and report our findings in a user-friendly fashion. A scale to assess the methodological quality of trials was designed and pilot tested for its content and reliability. Published RCTs identified during a literature search were then selected and rated by 6 raters. We used predefined cutoff scores to identify specific weaknesses in the clinical research process that need to be improved in future clinical trials.
Results: The quality scale we developed was demonstrated to be sufficiently reliable to warrant interpretation of the reviewers' findings. The percentage of trials that met a minimum level of quality for each specific criterion ranged from a low of 25% for the adequacy of the description of the randomization procedure to a high of 95% for the description and standardization of the intervention.
Conclusions: Based on the results of trials exhibiting a sufficient level of quality, treatments that were effective in decreasing pain and improving function in patients with PFPS were acupuncture, quadriceps strengthening, the use of a resistive brace, and the combination of exercises with patellar taping and biofeedback. The use of soft foot orthotics in patients with excessive foot pronation appeared useful in decreasing pain. In addition, at a short-term follow-up, patients who received exercise programs were discharged earlier from physical therapy. Unfortunately, most RCTs reviewed contained qualitative flaws that bring the validity of the results into question, thus diminishing the ability to generalize the results to clinical practice. These flaws were primarily in the areas of randomization procedures, duration of follow-up, control of cointerventions, assurance of blinding, accountability and proper analysis of dropouts, number of subjects, and the relevance of outcomes. Also, given the limited number of high-quality clinical trials, recommendations about supporting or refuting specific treatment approaches may be premature and can only be made with caution. *J Orthop Sports Phys Ther. 2003;33:4–20.*

Key Words: *bias, decision making, evidence, grading, methodology*

[1] *Physical therapist, Department of Physical Therapy, Schulthess Clinic, Zurich, Switzerland*
[2] *Doctoral candidate, U.S. Air Force, Department of Physical Therapy, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, PA and Physical Therapy Element, Wilford Hall Medical Center, Lackland AFB, TX.*
[3] *Doctoral candidate, Department of Physical Therapy, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, PA.*
[4] *Associate professor and chairman, Department of Physical Therapy, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, PA.*
Exempt from review by the University of Pittsburgh IRB based on the study being a literature review. The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the U.S. Air Force of Department of Defense.
Send correspondence to Capt. John D. Childs, 59th Medical Wing/Wilford Hall Medical Center, Physical Therapy Element/MMKPP, 2200 Bergquist Drive, Suite 1, Lackland AFB, TX 78236-5300. E-mail: childsjd@bigfoot.com

Patellofemoral pain syndrome (PFPS) is a common source of anterior knee pain in active individuals.[4,23,54,68] It accounts for 25% to 40% of all knee problems seen in sports medicine centers.[7,54] Although the etiology of PFPS is unclear, some have suggested that the pain and discomfort is likely to be the result of abnormal muscular and biomechanical factors that alter the distribution of shearing and compressive forces on the patellofemoral joint during normal activity.[61] Dye et al[14] recently suggested that the onset of PFPS may be due to a complex pathophysiologic process that may include peripatellar synovitis, increased intraosseous pressure, and increased intraosseous remodeling.

Because of the multifactorial nature of PFPS, many conservative treatment options have been proposed to treat this condition. However, no single intervention has been demonstrated to be the most effective. Consequently, several authors have conducted critical reviews of studies of the nonoperative treatment of PFPS.[3,10,48,68] Although the authors of these reviews offer many opinions of others' studies, their judgments are not based on a systematic approach to reviewing the literature. For example, Crossley et al[10] published a systematic re-

view of physical therapy treatments for PFPS. Sixteen clinical trials published as of October 2000 were selected for review. Although the effectiveness of physical therapy interventions such as strengthening, stretching, bracing, and taping was discussed, the specific criteria used to assess the methodological quality of these trials were not included.

A systematic review is a method that many believe minimizes bias and allows for better clinical decision making based on the evidence.[9] When several trials have been conducted that demonstrate efficacy for more than 1 treatment intervention, a systematic review of the literature has been suggested to provide the clinician with more information than the results of a single study.[56] Di Fabio[1,12] discussed the various ways to gather evidence from the literature using traditional, systematic, and masked reviews, and meta-analyses. Di Fabio also cautioned readers about the interpretation of systematic reviews, emphasizing the lack of definitions and the high degree of subjectivity when using these techniques. Jadad and his colleagues[25] reported that 80% of publications identified as systematic reviews had serious or extensive flaws, thus corroborating Di Fabio's view that such reviews are highly variable in their methodological rigor.

The purpose of this study was 2-fold: first, we sought to develop a grading scale to judge the quality of randomized clinical trials (RCTs); secondly, we wanted to use this scale to conduct a systematic review that critically appraises the methodological quality of the published RCTs that assess nonoperative treatments of PFPS, synthesize and interpret our results, and report our findings in a user-friendly fashion.

METHODS

To judge the quality of the RCTs, we developed a quality scale based on the Cochrane Collaboration Handbook.[9] The scale included factors that have been demonstrated to elicit bias and other factors that might affect the ability of clinicians to incorporate the results into their clinical practice. There is a general consensus that the Cochrane Collaboration guidelines[9] offer a relatively high degree of rigor to systematically examine the quality of clinical trials. While there is frequently little evidence to support the various weights given to components of quality scales, we believe that scales should ideally include empirically-based criteria that have been associated with bias in clinical trials. This is in accord with the view given in the Cochrane Collaboration Handbook, which states that the guidelines for systematic reviews are not to be used to dictate arbitrary standards, but rather to "help reviewers make good decisions about the methods they use."[9]

Design of Scale

One meeting was initially conducted among 6 experienced physical therapists with knowledge in factors that lead to bias in clinical trials and who were familiar with determining the best evidence upon which to base treatment strategies for their patients. During this meeting, the group discussed how to combine the relevant factors that might influence the methodological quality of a clinical trial into a meaningful scale. As an initial framework, the guidelines for systematic reviews proposed by the Cochrane Collaboration were followed.[70] Based on this discussion, the group agreed on the importance of 4 main criteria: (1) population, (2) interventions, (3) effect size, and (4) data presentation and analysis.

Scoring System of Methodological Quality Scale

Twenty-five points were assigned for each of the 4 main criteria for a total of 100 points. More specific criteria within each main criterion were then developed. Depending on our weighting of the criteria, a maximum of 5 or 10 points per specific criteria were given. All criteria ranged from 0 to 5 or 0 to 10 points, with 0 points for an inadequate description and the maximum number of points for an appropriately detailed description and adherence to the prospectively determined rules where applicable. If the rater believed that at least an attempt had been made to satisfy any criteria, even if inadequate, a partial score could be assigned. This latitude was provided to allow the rater freedom to score criteria that were partially complete and partially described. For example, description of exercise interventions that did not provide sufficient details to facilitate replication of the study (ie, information regarding technique, intensity, duration, frequency, etc) could receive partial credit. Each of the specific criteria were then further defined and weighted consistently with other published scales.

Considerations Within Each of the 4 Criteria

A brief description of the items that contributed to the assignment of points within each of the 4 main criteria and the relative weights given to the more specific criteria within each main criterion is outlined below.

Population (25 Points Total)

Inclusion Criteria (5 Points) The subjects had to be clearly defined in terms of the localization of the symptoms (anterior part of the knee) and type of symptoms (pain during squatting, kneeling, stairs, walking, and sitting with flexed knee) to substantiate a diagnosis of PFPS. Commonly used diagnostic sub-

stitutions that were also accepted included anterior knee pain, patella pain, chondromalacia patellae, and patellofemoral chondral lesions.

Exclusion Criteria (5 Points) The trial had to clearly define the exclusion criteria for potential subjects. Commonly utilized exclusion criteria were previous knee surgeries (tibiofemoral and patellofemoral) and knee instabilities (tibiofemoral and patellofemoral).

Adequate Number (10 Points) If the null hypothesis was rejected, discussion of power was considered not necessary and 5 points were given. However, if the null hypothesis was not rejected, the rater searched for the author to comment on the study's power. Up to an additional 5 points were given based on a scale considering the number of subjects per group: more than 25 subjects = 5 points; 21 to 25 = 4 points; 16 to 20 = 3 points; 11 to 15 = 2 points; 6 to 10 = 1 point; and less than 5 = 0 points.

Homogeneity (5 Points) The baseline characteristics of important prognostic indicators (sex, age, pain level, strength, and activity level) had to be similar among the groups.

Interventions (25 Points Total)

Standardized and Described (10 Points) The type of intervention used in the study had to be described explicitly, such that other researchers would be able to replicate the study. Important details of the intervention such as type, application technique, intensity, duration, and frequency of sessions had to be clearly stated to obtain a score of 10 points.

Control and Placebo Adequate (10 Points) Five points were assigned each for the presence of a control group and a placebo group.

Cointerventions Avoided (5 Points) No other interventions parallel to the main intervention were allowed, unless the rater determined that the cointerventions were applied equally to both the experimental and control groups.

Effect Size (25 Points Total)

Relevant Outcome (10 Points) Outcome measurements should be explicitly described, symptom and region specific, and relevant to the research question and clinical practice. Raters searched for the instruments' reliability, validity, and responsiveness. Raters looked for outcome measures of physical performance and self-reported measures of pain, function, and disability.

Blinded Outcome Assessment (10 Points) Outcome assessment done by a person masked to group assignment must be explicit and sufficiently described.

Follow-up Period Adequate (5 Points) The following scale was used: more than 24 months = 5 points; 13 to 23 months = 4 points; 7 to 12 months = 3 points; 4 to 6 months = 2 points; 1 to 3 months = 1 point; and less than 1 month = 0 points.

Data presentation and Analysis (25 Points Total)

Randomization Described (5 Points) The randomization procedure had to be explicitly described. To avoid excessively penalizing the authors who stated the word "randomize" but did not provide details of the randomization procedure used, we allowed partial credit.

Dropouts (5 Points) The authors had to clearly describe the rate of withdrawal in each group. In addition, the reason for subjects initially enrolled in the study who did not complete the trial or were lost to follow-up (eg, change of residence, job change, treatment intolerance, or adverse effect of the intervention) had to be clearly stated to allow determination of risks associated with the treatments.

Intention to Treat (5 Points) The data analysis had to include those subjects who were noncompliant to treatment or were lost to follow-up.

Proper Statistical Procedures Described (10 Points) The research hypothesis had to be analyzed with the appropriate statistical procedures and provide a clear description, including presentation of point estimates and measures of variability.

Pilot Testing of the Scale

To pilot test the scale, the same group of 6 experienced physical therapists involved in the development of the scale reviewed and scored 5 published trials and then met to discuss the results. The goal of this meeting was to assess the agreement on the criteria list and the weights assigned to each item. Additionally, this process was used to clarify any confusion regarding the interpretation of each item and to determine if any significant items were missing. The same 5 trials were then rescored together during the following meeting. Disagreements between reviewers were resolved by consensus under the guidance of a third party "referee," author AD, who was the senior researcher of the group. The idea was not to force the reviewers to assign the same score, but rather to agree on a common interpretation of each criterion. A range of 2 to 3 points of deviation for each main criterion was considered acceptable. Table 1 provides a copy of the final scale that was used for scoring the trials.

Literature Search

Because of the known bias associated with designs other than the RCT,[56] we chose to limit the review to trials that at least were purported to allocate subjects randomly to treatment groups. The focus was on published RCTs in the English peer-reviewed literature. Moher et al[39] found no differences between the quality of reporting in RCTs published in English, French, German, Italian, and Spanish. Thus there is a consensus that to minimize potential bias, all trials

TABLE 1. Scoring scale used to grade trials.

Criteria	Points	Comments
Population (25)		
Inclusion criteria (5)		
Exclusion criteria (5)		
Adequate number (10)		
Homogeneity (5)		
Intervention (25)		
Standardized and described (10)		
Control and placebo adequate (10)		
Cointerventions avoided (5)		
Effect size (25)		
Relevant outcome (10)		
Blinded outcome assessment (10)		
Follow-up period adequate (5)		
Data presentation and analysis (25)		
Randomization described (5)		
Dropouts accounted for (5)		
Intention to treat (5)		
Proper statistical procedures (10)		
Total score		

should be included in systematic review regardless of the language in which they were published.[19] However, the translation of trials published in languages other than English would have been time- and cost-prohibitive for the purposes of this review.

We conducted a search on the MEDLINE, CINAHL, and Web of Science databases, and the Cochrane Database of Systematic Reviews for the period between 1966 and October, 2000. The following key words were utilized singularly and in all possible combinations: patellofemoral pain syndrome, patella pain, peripatellar pain, retropatellar pain, patella, anterior knee pain, knee pain, chondromalacia patellae, chondropathy, physical therapy, physiotherapy, rehabilitation, controlled clinical trial, and randomized controlled trial. The reference list provided in 4 previous reviews,[3,10,48,68] and the reference lists in the identified trials were also screened. Redundant trials were then removed and the remaining 20 trials listed as RCTs in the databases were retrieved for review. Despite the fact that many studies lack an explicit randomization procedure, we kept all 20 trials, and did not apply a methodological filter to the RCTs based on randomization, as proposed by Sackett et al[55] for 2 reasons: first, a small number of studies were available, and second, the randomization criteria from the scale would account for the appropriateness of the randomization procedure used.

To obtain trials related to PFPS, the inclusion criteria for the trial had to provide some framework for determining that the subject had PFPS. However, because the inclusion criteria in some trials were not specific enough to rule out the possibility of competing diagnoses related to anterior knee pain, some of the patients in these trials may have had other sources of anterior knee pain. We only included trials that used nonoperative interventions such as therapeutic exercises, taping, braces, orthotics, injection, and medication.

Scoring of Trials

The trials were then reviewed by a second group of physical therapists, consisting of 6 graduate students and 2 seasoned faculty members. During the first meeting, the articles were distributed and training was conducted by one of the authors to ensure that all reviewers understood the components of the scale and how to score the various trials. A second meeting was conducted 1 week later to clarify any questions about the scoring criteria. This meeting served to minimize disagreements or misunderstandings, thus to ensure common interpretation of the scoring criteria. We planned to resolve any disagreement of more than 3 points for a main criterion between 2 reviewers in a consensus fashion. If the disagreement persisted, a third party "referee" would be consulted to assist in the resolution. However, because the ranges of the scores were within limits, referee consultation was not necessary. Finally, the results of the 20 scored trials were forwarded to one of the authors, and the results were entered into a computerized database (SPSS, Version 10.1, SPSS, Inc., Chicago, IL) for analysis.

Assessors were not masked to the authors, institution, and journal of publication for the trial under review. Several studies have assessed the importance of masking articles for a quality assessment of clinical trials.[25,37,40–42,71] Although there is some evidence that failure to mask these factors results in higher and less consistent scoring,[25] a recent study found no association between masking and unmasking with respect to factors related to quality assessment.[71] Both the logistical difficulty of masking and the inconsistent magnitude and direction of the effects of masking contributed to our decision not to mask the trials.

Interrater Reliability of the Scale

The intraclass correlation coefficient ($ICC_{1,1}$) for each component of the scale and for the total score of the trial was calculated according to the method described by Shrout and Fleiss.[60] Formula 1,1 was used because each trial reviewed was rated once by a different set of raters.[60] In this case, 6 out of 8 reviewers judged each trial and the combination of raters varied for each trial.

Identification of Areas of Improvement

To identify weaknesses in the methodology used in the RCTs in the treatment of PFPS, it was necessary to establish a minimum level of quality for each of

the 14 criteria in the quality scale. Minimum quality was operationally defined as a trial scoring at least half of the maximum possible score for the criterion under review. For example, if a specific criterion was assigned a maximum score of 10 points, trials that achieved a score of at least 5 points for this item were considered to meet a minimum level of quality. Table 2 depicts whether the minimum level of quality was satisfied for each criterion across all of the trials reviewed. Then, for each of the 14 criteria, we calculated the percentage of the trials (number of trials out of the 20 trials) that met the minimum level of quality. A calculation of the percentage of trials that met this cutoff was used to identify specific areas of the clinical research process that needed improvement. We defined that if 70% or less of the trials did not meet the minimum level of quality in a specific criterion, that criterion reflected a problematic area that needed to be more adequately addressed in future trials.

The cutoff value selected to define a minimum level of quality and the definition of areas that needed improvement based on the percentage of trials that met the minimum level of quality are admittedly somewhat arbitrary. However, we chose a minimum level of quality of half of the maximum points possible to protect against the possibility of being unreasonably critical in the review. Moreover, based on our judgment, this minimum level seemed to differentiate those trials with sufficient quality to provide valid results with adequate generalizability. It is entirely possible that a different set of reviewers may have selected different cutoff scores.

RESULTS

The ICCs for each criterion in the scale and the total score of the methodological quality of the trials are reported in Table 3. The ICCs for the items within each of the 4 main criteria ranged from 0.64 for the intervention's standardization and description to 0.99 for the item related to the blinding of outcome assessors. The ICC for the total quality score was 0.97.

Descriptive statistics (mean and standard deviation) of each trial with respect to the individual criterion score and total methodological quality score can be seen in Table 4. The total methodological quality scores ranged from a low score of 19 to a high score of 82.

Table 5 illustrates the relative scores of each trial and the description of interventions assessed in each of the RCTs. The display of the scores was divided into quartiles to facilitate visualization of the distribution of the scores. Two trials scored in a range of 0 to 24 points, 5 scored between 25 to 49, 8 scored between 50 to 74, and 5 scored between 75 to 100. Whether the trial detected a significant difference between interventions and what intervention was shown to be superior is also noted in Table 5. To facilitate the discussion of the results of the review, we classified the interventions into 7 groups, based on the primary goal of the intervention. We classified separately the studies in which more than 1 intervention was tested (eg, an exercise program versus education, and versus education with taping). The groups are as follows: (1) orthotics (including foot

TABLE 2. Descriptive statistics for each trial and percentage of trials meeting the minimum level of quality for each area. Yes indicates the the last column.

Criteria	Antich[2]	Fulkerson[18]	Finestone[17]	Miller[38]	Kowall[32]	Eburne[15]	Rowlands[53]	Roush[52]	Rogvi-Hansen[51]
Population (25)									
Inclusion criteria (5)	no	no	yes	yes	yes	yes	yes	yes	yes
Exclusion criteria (5)	no	no	no	yes	no	yes	yes	yes	yes
Adequate number (10)	no	yes	no	no	no	yes	no	no	no
Homogeneity (5)	no	no	no	no	yes	yes	no	no	yes
Intervention (25)									
Standardized and described (10)	yes	yes	yes	no	yes	yes	yes	yes	yes
Control and placebo adequate (10)	yes	no	yes	yes	yes	no	yes	yes	yes
Cointerventions avoided (5)	no	no	yes	no	no	no	yes	no	no
Effect size (25)									
Relevant outcome (10)	no	no	no	no	yes	no	yes	no	no
Blinded outcome assessment (10)	no	no	no	no	no	yes	no	yes	yes
Follow-up period adequate (5)	no	no	no	no	no	no	no	no	no
Data presentation and analysis (25)									
Randomization described (5)	no	no	no	no	yes	no	no	no	no
Dropouts accounted for (5)	no	no	no	no	yes	yes	no	yes	no
Intention to treat (5)	no	no	no	yes	yes	no	no	no	no
Proper statistical procedures (10)	no	no	no	no	yes	yes	yes	no	yes

orthoses, patellar braces, elastic sleeves, and patellar taping), (2) manual therapy, (3) modalities, (4) medications, (5) acupuncture, (6) strength training methods, and (7) combined interventions. The theoretical rationale and the results of each study will be discussed with respect to the treatment category to which the trial was assigned.

The percentage of trials that met the minimum level of quality for each specific criterion based on this definition ranged from a low of 25% for the adequacy of the description of the randomization procedure to a high of 95% for the description and standardization of the intervention (Table 2).

Several areas included in our quality scale were consistently scored high by most of the trials, suggesting a good understanding of these criteria by the researchers. Five out of the 14 areas scored well, defined as greater than 70% of the trials meeting the minimum level of quality. These 5 criteria include: (1) standardization and description of intervention, (2) adequate control and placebo, (3) description of inclusion criteria, (4) description of exclusion criteria, and (5) proper statistical procedures. Table 2 includes the specific percentage for each of these items. Because these items were consistently scored high by the reviewers, these components of the clinical research process will not be discussed further.

Nine out of 14 specific areas of the scale did not score well, suggesting a poor understanding of these criteria by the researchers. They include the following: (1) adequate description of randomization, (2) adequate follow-up, (3) sufficient control of cointerventions, (4) assurance of blinding, (5) consideration of an intention-to-treat analysis (ITT), (6) adequate number of subjects, (7) inclusion of relevant outcomes, (8) proper accounting for withdrawals and dropouts, and (9) homogeneity of groups at baseline.

DISCUSSION

Reliability of the Scale

The ICC ranges from 0 to 1, with higher values representing better reliability. Portney and Watkins[47] suggest that ICCs above 0.75 represent good reliability, ICCs from 0.75 to 0.5 represent moderate reliability, and those below 0.5 indicate poor reliability. In this study, the consistency of rating each criterion and rating the total score using the qualitative scale was considered acceptable. With exception of 3 criteria that demonstrated moderate reliability (inclusion criteria, homogeneity of groups, and description of interventions), all other criteria demonstrated good reliability. The total score of quality of the trials seems to represent consistency of rating among all criteria. Although we recognize that the content of any grading scale is determined by the consensus of individuals who developed it, we believe that the total scores offer some valuable information as to the methodological quality of the trials, based on the good reliability demonstrated across the ratings.

Summary of the Trials

The remaining discussion provides a summary of the results of the systematic review, with each trial

minimum level of quality was met and no indicates that it was not; percentages of trials meeting minimum level of quality are denoted in

Eng[16]	Thomee[67]	Suter[64]	Harrison[20]	Timm[69]	Witvrouw[74]	Raatikainen[50]	Kannus[27]	Clark[8]	Kannus[28]	Jensen[26]	% Meeting Minimum Quality
yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	90%
yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	80%
yes	no	yes	yes	yes	yes	yes	yes	yes	yes	yes	60%
yes	yes	yes	yes	yes	yes	yes	yes	yes	no	yes	65%
yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	95%
yes	no	yes	yes	yes	yes	yes	yes	no	yes	yes	80%
no	no	yes	yes	yes	yes	yes	yes	no	no	yes	45%
no	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	60%
no	no	no	no	no	no	yes	no	yes	yes	yes	35%
no	yes	no	no	no	no	yes	yes	yes	yes	yes	30%
no	yes	no	no	no	yes	no	no	yes	yes	no	25%
yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	70%
yes	yes	yes	yes	yes	yes	no	no	yes	no	yes	50%
yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	75%

TABLE 3: Summary of intraclass correlation coefficients (ICCs) for scale criteria and total score.

Criteria	ICC (Formula 1,1)
Population	
Inclusion criteria	0.66
Exclusion criteria	0.89
Adequate number	0.91
Homogeneity	0.74
Intervention	
Standardized and described	0.64
Control and placebo adequate	0.95
Cointerventions avoided	0.87
Effect size	
Relevant outcome	0.87
Blinded outcome assessment	0.99
Follow-up period adequate	0.92
Data presentation and analysis	
Randomization described	0.93
Dropouts accounted for	0.96
Intention to treat	0.95
Proper statistical procedures	0.91
Total score	0.97

being discussed within the treatment group to which the trial was assigned (based on the primary goal of the intervention). The theoretical rationale for each intervention is discussed to provide a context that might explain why the investigators in each clinical trial selected the particular treatment or combination of treatments in their study. In addition, the items in the scale in which the minimum level of quality was met or was not met for each trial is reported.

Orthotic Interventions

Orthotic interventions such as foot orthoses, patellar braces, elastic sleeves, and patellar taping were included in the same intervention group because an orthosis is defined as any external appliance worn to restrict or to enhance motion.[44] The use of foot orthotics for PFPS is based on the assumption that excessive foot pronation during stance phase would cause abnormal tibial internal rotation that consequently would disrupt the normal tracking of the patella inside the trochlear groove.[13,16,30] Therefore, the foot orthotic is believed to modify foot and leg position during stance phase of gait, potentially restoring normal patellar tracking.

Eng et al[16] conducted a RCT assessing the efficacy of soft foot orthotics in a group of adolescent females exhibiting excessive foot pronation, which was defined as calcaneal varus or forefoot varus greater than 6°. Patients who completed 16 sessions of a physical therapy program consisting of exercises and the wearing of soft foot orthotics were demonstrated to have significant reductions in pain compared to a control group who completed the same physical therapy program, except they did not receive the soft foot orthotics. Therefore, it seems that there is some evidence for using this intervention in female adolescents who exhibit excessive foot pronation. This trial was assigned a score of 61 points (Table 5). This trial did not meet the minimum level of quality for the following items in the scale: cointerventions avoided, relevant outcome, blinded outcome assessment, follow-up period adequate, and randomization described (Table 2).

Patellar braces were initially suggested as an intervention purported to apply a sustained medially displacing force to the lateral border of the patella.[45] The theoretical rationale of patellar bracing is to im-

TABLE 4. Descriptive statistics (mean ± standard deviation) for the reviewers' scores on each criterion in the quality scale for all of the

Criteria	Antich[2]	Fulkerson[18]	Finestone[17]	Miller[38]	Kowall[32]	Eburne[15]	Rowlands[53]	Roush[52]
Population (25)								
Inclusion Criteria (5)	0.3 ± 0.5	2.0 ± 0.9	2.5 ± 0.8	4.0 ± 1.0	4.5 ± 0.5	4.0 ± 0.9	4.4 ± 0.5	3.1 ± 1.1
Exclusion Criteria (5)	0.0 ± 0.0	0.4 ± 0.7	1.0 ± 0.8	4.7 ± 0.5	0.3 ± 0.5	2.7 ± 0.5	4.9 ± 0.4	4.1 ± 0.9
Adequate Number (10)	2.7 ± 0.5	6.0 ± 0.9	3.5 ± 0.5	2.9 ± 0.7	2.6 ± 0.5	6.7 ± 0.5	2.9 ± 0.7	4.6 ± 0.8
Homogeneity (5)	0.0 ± 0.0	1.8 ± 1.0	1.5 ± 0.5	2.1 ± 0.9	4.5 ± 0.5	3.3 ± 0.5	0.7 ± 1.0	2.3 ± 0.5
Intervention (25)								
Standardized and Described (10)	6.0 ± 1.5	8.5 ± 1.1	7.0 ± 0.9	4.9 ± 0.9	5.4 ± 0.7	5.3 ± 1.4	8.7 ± 1.1	9.1 ± 0.9
Control and Placebo Adequate (10)	5.3 ± 0.5	0.0 ± 0.0	5.0 ± 0.0	5.3 ± 0.5	5.0 ± 0.0	4.3 ± 1.0	5.1 ± 0.4	5.0 ± 0.0
Cointerventions Avoided (5)	1.0 ± 0.9	0.4 ± 0.5	4.6 ± 0.5	1.4 ± 0.8	0.3 ± 0.5	1.0 ± 0.9	5.0 ± 0.0	1.1 ± 1.1
Effect Size (25)								
Relevant Outcome (10)	1.7 ± 1.0	3.0 ± 1.3	3.0 ± 1.1	3.3 ± 1.0	5.4 ± 0.7	0.3 ± 0.5	5.9 ± 0.9	4.1 ± 0.9
Blinded Outcome Assessment (10)	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	10.0 ± 0.0	0.0 ± 0.0	10.0 ± 0.0
Follow-up Period Adequate (5)	0.7 ± 1.0	0.3 ± 0.5	1.9 ± 0.4	1.3 ± 0.5	0.5 ± 0.5	0.3 ± 0.5	0.6 ± 0.5	1.6 ± 0.5
Data Presentation and Analysis (25)								
Randomization Described (5)	1.0 ± 0.9	0.6 ± 0.5	0.1 ± 0.4	0.3 ± 0.5	4.4 ± 0.7	1.3 ± 0.5	1.4 ± 0.5	0.6 ± 0.5
Dropouts Accounted For (5)	0.3 ± 0.5	0.0 ± 0.0	0.3 ± 0.7	0.0 ± 0.0	4.5 ± 0.5	2.7 ± 0.5	0.0 ± 0.0	4.7 ± 0.5
Intention To Treat (5)	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	4.0 ± 1.0	4.4 ± 0.9	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Proper Statistical Procedures (10)	0.0 ± 0.0	0.9 ± 1.0	1.0 ± 1.1	4.0 ± 1.0	5.9 ± 1.0	5.7 ± 1.0	8.6 ± 1.0	3.9 ± 1.3
Total Score (100)	19.0 ± 1.8	23.8 ± 3.5	31.4 ± 3.2	38.1 ± 3.9	47.5 ± 2.9	47.7 ± 1.0	48.1 ± 2.0	54.3 ± 4.3

prove tracking of the patella in the trochlear groove, thus reducing the pain believed to be associated with patellar malalignment. Two trials[17,38] in our review incorporated a patellar brace for the purpose of improving pain. Finestone et al[17] used a sample of individuals diagnosed as having overuse patellofemoral pain to compare the efficacy of an elastic knee sleeve that has a silicone patellar ring with a simple elastic sleeve with a control group that received no treatment. They reported no difference in pain reduction between the groups and reported that wearing the sleeve with a silicone ring resulted in significantly more skin abrasion. Miller et al[38] compared what they referred to as a "dynamic patellar brace" versus a knee strap versus a no-brace condition. They found no significant difference in pain between the groups throughout the study. The studies by Finestone et al[17] and Miller et al[38] were assigned scores of 31 and 38 points, respectively (Table 5). Both of these trials did not meet the minimum level of quality for adequate number of subjects, homogeneity of groups, use of relevant outcome, blinded outcome, blinded outcome assessment, follow-up period adequate, randomization described, accounting for dropouts, and proper statistical procedures. Furthermore, the Finestone et al[17] trial did not meet the minimum level of quality for the exclusion criteria and intention to treat, whereas the Miller et al[38] study did not meet the minimum level of quality for standardization and description of intervention and avoidance of cointervention (Table 2). Based on the relatively low quality scores of these studies, there does not appear to be adequate evidence to support or refute the use of patellar bracing to improve patellar pain in the management of PFPS.

Timm[69] conducted a study in which he compared a group that used the Protonics brace (Inverse Technology Corporation, Lincoln, NE) during daily activities against a control group that received no treatment. The brace was designed to provide progressive resistance to knee motion in the sagittal plane, with the rationale that a high volume of submaximal concentric contractions of the quadriceps and hamstring muscles may perhaps facilitate patellar alignment and reduce abnormal patellar congruence and pain. They reported that patients in the experimental group had significant reduction in pain and improvement in function compared to the control group. Based on these results, there seems to be some evidence that this resistive brace may reduce pain and improve function. This study was assigned a score of 72 points (Table 5). This trial did not meet the minimum level of quality for the following items in the scale: blinded outcome assessment, follow-up period adequate, and randomization described (Table 2).

Patellar taping is suggested to correct the position of the patella to promote alignment, reduce pain, and reduce joint reaction forces.[11,36] A recent review of nonrandomized studies of patellar taping concluded that, although there is some evidence to support the application of patellar taping to produce immediate reduction in pain during provocative tasks, there is insufficient evidence to confirm any carryover in pain reduction or any other proposed effect.[11]

We found only 1 RCT that isolated the effect of patellar taping as the treatment being studied. Kowall et al[32] found no significant differences at a 1-month follow-up between a physical therapy program over 8 sessions that incorporated stretching

trials included in the review.

Rogvi-Hansen[51]	Eng[16]	Thomee[67]	Suter[64]	Harrison[20]	Timm[69]	Witvrouw[74]	Raatikainen[50]	Kannus[27]	Clark[8]	Kannus[28]	Jensen[26]
4.0 ± 0.0	4.8 ± 0.5	4.9 ± 0.4	5.0 ± 0.0	5.0 ± 0.0	5.0 ± 0.0	4.9 ± 0.4	4.2 ± 1.0	4.9 ± 0.4	4.5 ± 0.8	5.0 ± 0.0	4.5 ± 0.8
3.3 ± 0.5	4.3 ± 0.9	4.9 ± 0.4	5.0 ± 0.0	5.0 ± 0.0	5.0 ± 0.0	4.4 ± 0.8	4.0 ± 0.9	4.7 ± 0.5	4.8 ± 0.5	4.9 ± 0.4	3.3 ± 0.9
3.0 ± 0.0	6.8 ± 0.9	4.7 ± 0.5	10.0 ± 0.0	10.0 ± 0.0	10.0 ± 0.0	10.0 ± 0.0	7.2 ± 1.0	8.1 ± 1.5	8.3 ± 1.2	7.7 ± 1.1	10.0 ± 0.0
5.0 ± 0.0	4.6 ± 0.7	4.3 ± 1.0	4.4 ± 0.8	4.4 ± 0.8	4.4 ± 0.8	4.0 ± 0.8	4.3 ± 0.5	4.0 ± 1.0	4.8 ± 0.5	1.1 ± 0.9	4.5 ± 0.8
9.3 ± 1.0	9.3 ± 1.0	9.9 ± 0.4	10.0 ± 0.0	10.0 ± 0.0	10.0 ± 0.0	8.6 ± 1.4	7.5 ± 0.8	9.1 ± 1.2	7.8 ± 1.2	9.1 ± 0.9	8.9 ± 1.2
10.0 ± 0.0	5.0 ± 0.0	0.3 ± 0.5	5.0 ± 0.0	5.0 ± 0.0	5.0 ± 0.0	5.0 ± 0.0	9.0 ± 0.9	9.0 ± 0.8	4.8 ± 0.5	9.3 ± 0.5	5.6 ± 1.2
0.3 ± 0.5	0.5 ± 0.5	1.0 ± 1.0	3.6 ± 0.5	3.6 ± 0.5	3.6 ± 0.5	4.0 ± 0.8	2.5 ± 0.5	4.0 ± 0.8	1.3 ± 0.9	0.6 ± 0.5	4.3 ± 0.9
4.0 ± 0.9	4.4 ± 0.9	9.7 ± 0.5	7.3 ± 1.0	7.3 ± 1.0	7.3 ± 1.0	8.6 ± 0.8	7.3 ± 1.0	9.0 ± 1.0	8.4 ± 0.9	8.7 ± 1.0	7.8 ± 0.5
10.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.1 ± 0.4	10.0 ± 0.0	4.6 ± 0.5	9.4 ± 0.9	10.0 ± 0.0	9.8 ± 0.7
2.0 ± 0.0	1.5 ± 0.5	4.4 ± 0.5	1.4 ± 0.5	1.4 ± 0.5	1.4 ± 0.5	1.9 ± 0.4	4.3 ± 0.5	3.7 ± 0.5	2.9 ± 0.8	5.0 ± 0.0	4.0 ± 0.0
0.7 ± 0.5	0.6 ± 0.5	4.6 ± 0.5	1.0 ± 0.8	1.0 ± 0.8	1.0 ± 0.8	5.0 ± 0.0	1.0 ± 0.6	0.4 ± 0.5	5.0 ± 0.0	4.9 ± 0.4	0.5 ± 0.5
0.0 ± 0.0	4.6 ± 0.5	5.0 ± 0.0	5.0 ± 0.0	5.0 ± 0.0	5.0 ± 0.0	4.7 ± 0.8	4.7 ± 0.5	4.9 ± 0.4	4.1 ± 1.0	4.9 ± 0.4	5.0 ± 0.0
0.0 ± 0.0	4.8 ± 0.5	5.0 ± 0.0	5.0 ± 0.0	5.0 ± 0.0	5.0 ± 0.0	4.4 ± 0.8	0.0 ± 0.0	1.0 ± 0.8	4.8 ± 0.5	0.0 ± 0.0	4.6 ± 0.5
6.0 ± 0.9	9.6 ± 0.7	10.0 ± 0.0	9.3 ± 1.0	9.3 ± 1.0	9.3 ± 1.0	9.0 ± 1.0	9.3 ± 0.8	8.9 ± 1.1	9.5 ± 0.9	9.3 ± 1.0	10.0 ± 0.0
57.7 ± 1.4	60.6 ± 4.5	68.6 ± 2.2	72.0 ± 2.2	72.0 ± 2.2	72.0 ± 2.2	74.6 ± 2.5	75.3 ± 4.1	76.3 ± 3.5	80.0 ± 2.7	80.4 ± 3.0	82.6 ± 3.6

TABLE 5. Summary of the mean total scores between the reviewers for each trial in the study and the results displayed relative to the intervention. When a significant difference was detected between treatment groups, the horizontal bar appears as a solid bar and is placed next to the intervention shown to be more effective. When no differences were found, the horizontal bar appears as a striped bar and is placed in the center of the cell. Score (divided in quartiles), with the total possible score from 0 to 100 points (0 meaning poor quality and 100 meaning high quality).

Trial	Interventions	Methodologic Quality Scores 0–24	25–49	50–74	75–100
Antich[2]	• Ice • Phonophoresis • Iontophoresis • Ultrasound/ice contrast				
Fulkerson[18]	• Nonsteroidal anti-inflammatory A (diflunisal) • Nonsteroidal anti-inflammatory B (naproxen)				
Finestone[17]	• Elastic sleeve (silicon ring) • Simple elastic sleeve • No sleeve				
Miller[38]	• Palumbo brace • Cho strap • No brace				
Kowall[32]	• Physical therapy (PT) program and patellar taping • PT program				
Eburne[15]	• Isometric quadriceps exercise and functional re-education • McConnell's regimen				
Rowlands[53]	• Patella mobilization • Detuned ultrasound				
Roush[52]	• Home program • PT program • Home program ("vastus medialis oblique special")				
Rogvi-Hansen[51]	• Low-level laser • Sham laser				
Eng[16]	• Physical therapy program/foot orthotics • Physical therapy program				
Thomee[67]	• Isometric quadriceps exercise • Eccentric quadriceps exercise				
Suter[64]	• Sacroiliac joint manipulation • Sacroiliac/lower back "assessment"				
Harrison[20]	• Home strength/stretching program • Program monitored by PT • Comprehensive PT program				
Timm[69]	• Exercise with "resisted brace" • No treatment				
Witvrouw[74]	• Open kinetic chain exercise • Closed kinetic chain exercise				
Raatikainen[50]	• Injections glycoaminoglycan polysulfate • Placebo injection				
Kannus[27]	• Injection glycoaminoglycan polysulfate/quadriceps exercise • Placebo injection/quadriceps exercise • Quadriceps exercise				
Clark[8]	• Exercise/taping/education • Exercise/education • Taping/education • Education				
Kannus[28]	• Injection glycoaminoglycan polysulfate/quadriceps exercise • Placebo injection/quadriceps exercise • Quadriceps exercise				
Jensen[26]	• Acupuncture • No treatment				

and isometric, isotonic, and isokinetic quadriceps strengthening exercises versus the same program plus patellar taping applied during the exercises. This trial was assigned a score of 48 points (Table 5). The study did not meet the minimum level of quality for the following items in the scale: description of exclusion criteria, adequacy of number of participants, avoidance of cointerventions, blinded outcome assessment, and follow-up period adequate (Table 2). Based on the results of this study and in light of its quality score, the usefulness of patellar taping in the management of PFPS is not clear.

Manual Therapy

Manual therapy for the treatment of PFPS has been used in 1 RCT in the form of manual stretching of the lateral patellar structures and in another RCT as manipulation directed to the sacroiliac joint (SIJ). Studies have suggested that patients with PFPS may exhibit tightness of structures that attach to the lateral border of the patella, such as the lateral retinaculum.[31,57] Tightness of the lateral retinacular structures, perhaps as a result of increased tension in the iliotibial band, may adversely alter tracking of the patella in the trochlear groove. The theoretical rationale for manual stretching of the lateral patellar structures is that the application of a low-load, prolonged stretch may increase the length of these tissues and decrease the excessive lateral pull of these structures over the patella, allowing a normal tracking of the patella inside the trochlear groove. The biological plausibility of lengthening soft tissue to facilitate a permanent elongation has been previously reported in the literature.[65]

Rowlands et al[53] compared a group of patients who received detuned ultrasound with a group that received a patellar mobilization procedure consisting of manual sustained glide followed by high-velocity, low-amplitude manipulation. They reported that although no difference was found in functional outcome between the groups, the patellar mobilization group demonstrated significantly lower levels of pain than the control group at a 1-month follow-up. This study was assigned a score of 48 points (Table 5). The study did not meet the minimum level of quality for the following items in the scale: adequate number of subjects, homogeneity of groups, blinded outcome assessment, follow-up period adequate, randomization described, dropouts accounted for, and use of intention-to-treat analysis (Table 2). It seems there may be some usefulness for manual therapy in the treatment of PFPS, however, the results must be interpreted in light of the study's quality score.

Suter et al[63] documented the presence of quadriceps activation failure (QAF) in patients with anterior knee pain. QAF is the inability to fully activate a muscle during a maximum voluntary contraction.[22] Although there is no evidence documenting the role of the SIJ in the maintenance of normal patellofemoral joint mechanics, the authors speculate that SIJ dysfunction may adversely alter patellofemoral biomechanics. They reported that the patients who received a manipulation theoretically directed at the SIJ demonstrated short-term results in decreasing QAF compared with patients who did not receive manipulation.[64] This study was assigned a score of 71 points (Table 5). This trial did not meet the minimum level of quality for the following items in the scale: blinded outcome assessment, follow-up period adequate, and randomization described (Table 2). Therefore, there seems to be some evidence that the proposed manipulation approach decreases QAF. However, the direct effect of this finding on pain and function has not been established.

Modalities

Antich et al[2] published the first RCT to investigate the effect of different modalities on strength and subjective improvement for patients with chondromalacia patella, infrapatellar tendonitis, and peripatellar pain. Ice, phonophoresis, iontophoresis, and ultrasound-ice contrast were compared. This study was assigned a score of 19 points (Table 5). The results suggested that the use of ultrasound-ice was the most effective modality for treatment of pain in these patients. However, we suggest caution the use of these results as evidence because of the study's low-quality score. The only items in which this trial met the minimum level of quality were the standardization and description of intervention and use of adequate control and placebo group (Table 2). Rogvi-Hansen et al[51] found no difference in symptoms between patients with arthroscopically diagnosed chondromalacia patellae who received real or sham low-level laser. This trial was assigned a score of 58 points (Table 5). This trial did not meet the minimum level of quality for the following items of the scale: adequate number of subjects, cointerventions avoided, use of relevant outcome, follow-up period adequate, description of randomization, dropouts accounted for in the analysis, and use of an intention-to-treat analysis (Table 2). The results of this study seem to indicate that low-level laser treatment is not effective in the management of PFPS, but the low quality of the study does not allow a definitive conclusion.

Medications

Nonsteroidal anti-inflammatory drugs (NSAIDs) are commonly used as an adjunct to physical therapy in the management of PFPS.[43] Fulkerson et al[18] compared 2 different NSAIDs, diflunisal and naproxen, in patients described as having anterior knee pain

and found significant levels of pain relief for both drugs. This study was assigned a score of 24 points (Table 5). The only items in which this trial met the minimum level of quality were the adequacy of the number of subjects and standardization and description of interventions (Table 2). It appears that some of the patients in this study may have been diagnosed with anterior knee pain other than PFPS. Therefore, the subjects could have had a variety of conditions that are primarily inflammatory in nature which might respond well to NSAIDs.

Kannus et al[27,28] assessed the effect of intra-articular and intramuscular injections of glycoaminoglycan polysulfate (GAGPS) in the management of PFPS. Experimental research has shown that GAGPS can inhibit degradative enzyme reactions, thus inhibiting the inflammatory cascade of events associated with inflammation, and can stimulate the metabolism of chondrocytes and synovial cells.[27] They compared the use of intra-articular injections of GAGPS with placebo injections and no injections. The results were published in 2 studies, 1 reporting the results at follow-ups of 6 weeks and 6 months,[27] and the other reporting the results at a follow-up of 7 years.[28] The 3 groups of patients also received an exercise program (isometric quadriceps and stretching) and oral doses of the NSAID piroxicam during the initial 6-week period of rehabilitation. These studies reported that at 6-week,[27] 6-month,[27] and 7-year follow-ups,[28] the use of intra-articular injections of GAGPS does not influence the outcome of rehabilitation in patients with PFPS participating in a quadriceps exercise program. These trials were assigned scores of 76[27] and 80 points,[28] respectively (Table 5). Both trials did not meet the minimum level of quality for the criteria related to use of an intention-to-treat analysis. Additionally, the first trial[27] did not meet the minimum level of quality for blinded outcome assessment, whereas the second trial[28] did not meet the minimum level of quality for homogeneity of groups and cointerventions avoided (Table 2). Another study[50] that investigated a similar hypothesis reported results that conflicted with the one above. In a trial that was assigned a score of 75 points (Table 5), Raatikainen et al[50] reported positive results with intramuscular injections of GAGPS in patients with arthroscopically verified damage of patellar cartilage. However, the type of injection (intramuscular versus intra-articular) and patient selection criteria were different between the 2 studies. This trial did not meet the minimum level of quality for the description of randomization and use of an intention-to-treat analysis (Table 2). Although the above studies seem to possess a sufficient level of quality, the role of injection of GAGPS remains unclear based on the contradictory findings of the studies.

Acupuncture

Although the mechanism by which acupuncture reduces pain is unclear, it is believed to be related to the gate and endorphin theories of pain reduction.[26] Jensen et al[26] assessed the effect of acupuncture in the treatment of PFPS and found significant improvements in pain and function in those patients who received acupuncture. This trial was assigned a score of 83 points (Table 5). The only item in the scale for which this trial did not meet the minimum level of quality was the item related to the description of randomization (Table 2). Based on the results of this study and in light of its quality score, acupuncture appears to be effective in the treatment of PFPS.

Strength Training Methods

Werner et al[72] demonstrated that patients with PFPS exhibited decreased strength and reduced electromyographic activity during a seated knee extension exercise compared to subjects who did not have PFPS. The restoration of quadriceps muscle strength has been found to correlate with the long-term outcome in patients with PFPS.[43]

Two RCTs[67,74] assessed quadriceps training methods and found no measurable difference of any relevant outcome between the different approaches investigated. Witvrouw et al[74] evaluated the efficacy of non–weight-bearing exercises versus weight-bearing exercises in patients with patellofemoral pain. Although they reported decreased pain and increased function in both groups, no difference in pain, muscle performance, and functional outcomes between groups was observed. Thomee et al[67] compared isometric and eccentric training programs and observed no overall differences in physical activity, pain, or muscle function between groups. The studies by Witvrouw et al[74] and Thomee et al[67] were assigned scores of 75 and 69 points, respectively (Table 5). Both of these trials did not meet the minimum level of quality for the item related to blinded outcome assessment. Additionally, the Witvrouw et al[74] trial did not meet the criterion for adequacy of the follow-up period, whereas the Thomee et al[67] trial did not meet the minimum level of quality for adequate number of subjects, adequacy of control and placebo group, and cointerventions avoided (Table 2). The fact that all the groups in both trials demonstrated some improvement in function and decrease in pain, suggests that there is moderate evidence that non–weight-bearing, weight-bearing, isometric, and eccentric exercises may be useful in the management of PFPS; however, no single approach has been demonstrated to be superior to another.

Combined treatments

The rationale for trials to combine different treatment approaches likely involves an effort to simultaneously ameliorate as many impairments associated with PFPS as possible. Many studies have investigated the impairments associated with PFPS. Some studies[62,73] demonstrated that decreased quadriceps strength and decreased hamstrings and quadriceps flexibility are risk factors for the development of anterior knee pain.[13] It has been theorized that decreased flexibility of the iliotibial band/tensor fascia lata complex, and tightness of the lateral retinacular tissues may also contribute to PFPS.[4,6,31,57,73] In addition, some authors suggest that muscle imbalance between the vastus medialis oblique (VMO) and vastus lateralis contribute to a tendency for the patella to laterally sublux,[35] but there is little evidence to substantiate this hypothesis, and there is increasing evidence that this is in fact not the case.[5,29,49] The anatomical differentiation between the fibers of the VMO and fibers of the vastus medialis lateralis has also been recently questioned.[21] We found 4 RCTs that used a combination of different exercises, patellar taping, orthotics, biofeedback, and patient education.

Harrison et al[20] performed a study comparing 3 groups that received different treatments: (1) lower extremity strengthening and stretching exercises performed as a home program; (2) lower extremity strengthening and stretching exercises plus stretching of the lateral retinacular structures supervised by a physical therapist; (3) same as group 2 plus patellar taping and biofeedback technique directed to the VMO muscle. They found significant improvement in pain and function for group 3 at the 1-month follow-up. However, no differences were detected between the rehabilitation programs at a 1-year follow-up. This study was assigned a score of 72 points (Table 5). This trial did not meet the minimum level of quality for the following items in the scale: blinded outcome assessment, follow-up period adequate, and description of randomization (Table 2). Although no differences were detected at the 1-year follow-up, the improvements reported at 1-month may still justify intervention based on the potential to resolve the patient's complaints sooner.

Clark et al[8] found no significant differences at a 1-year follow-up between 4 groups that incorporated the following interventions: (1) exercise, taping, and education; (2) exercise and education; (3) taping and education; and (4) education alone. However, patients who received the exercise program were significantly more likely to be discharged after 3 months compared to patients who did not receive exercise. Patient satisfaction was used as the criterion for discharge. There were no significant differences in pain, anxiety and depression, quadriceps strength, and function at the 3-month and 1-year follow-ups. This study was assigned a score of 80 points (Table 5). This trial did not meet the minimum level of quality for the items related to adequate control and placebo group and cointerventions avoided (Table 2). Similar to Harrison et al,[20] based on the ability to quickly resolve a patient's complaints, even short-term improvements that are not necessarily maintained over a longer period of time may still be sufficient to justify intervention.

Roush et al[52] examined 3 treatments: (1) a home program that used traditional straight-leg-raising and pillow-squeezing exercises; (2) a physical therapy program conducted over 18 physical therapy sessions consisting of stretching, strengthening, patellar mobilization, taping, orthotics, and aerobic exercise; and (3) a special home program that incorporated a modified straight leg raise purported to be specific for isolating the VMO. Recruitment of the VMO was not investigated during the trial. They reported improved pain and function at 12 weeks for subjects in group 3. This study was assigned a score of 54 points (Table 5). This trial did not meet the minimum level of quality for the following items in the scale: adequate number of subjects, homogeneity of groups, cointerventions avoided, use of relevant outcome, follow-up period adequate, description of randomization, intention-to-treat analysis, and use of proper statistical procedures (Table 2).

Eburne et al[15] compared a group that received isometric quadriceps exercises and functional re-education with a group that received McConnell's regimen (ie, patellar taping, VMO training, eccentric muscle action in weight bearing, and functional activities). The only significant difference was that fewer patients who received McConnell's regimen exhibited a positive symptom provocation test upon completion of the treatment, however, there were no differences in the primary pain outcome measure. This trial was assigned a score of 48 points (Table 5). The study did not meet the minimum level of quality for the following items in the scale: control and placebo adequate, cointerventions avoided, use of relevant outcome, follow-up period adequate, description of randomization, and use of an intention-to-treat analysis (Table 2).

Several trials[8,15,20,52] have investigated the use of a combined treatment approach in the management of PFPS. Except for 1 trial,[15] each demonstrated a relatively short-term treatment effect in improving pain and function[20,52] or the rate of discharge from physical therapy based on patient satisfaction.[8] Two of the 4 trials[8,20] also had quality scores that exceeded 70 points (Table 5). Based on the evidence from these trials and in light of their quality scores, there appears to be some evidence for the use of a combined treatment approach in the management of PFPS.

Areas of Improvement and Practical Applications

A secondary purpose of this review was to objectively identify and discuss specific areas in the clinical research process that need to be improved in future clinical trials. If improvements in these areas could be incorporated into the design of future clinical trials, the end result would be an improvement in the validity and generalizability of the results from these studies. The areas of the clinical research process that may need to be improved based on the predefined cutoff scores (ie, those areas in which 70% or less of the trials satisfied the minimum level of quality for that criterion) are briefly discussed in the order of the lowest to highest percentage of trials for each criterion.

Failure to appropriately randomize subjects to treatment groups was the criteria that scored the lowest using our scale. Although all studies in this review were purported to be an RCT, only 25% of all trials met the minimum requirement for appropriate randomization (Table 2). Failure to randomly allocate patients to treatment groups, which may include using an inadequate randomization procedure, is believed to be the most important factor contributing to bias in clinical trials.[58] Random allocation among a sufficiently large sample tends to distribute important variables, both known and unknown, equally across treatment groups. This helps to ensure that at the beginning of treatment groups are similar with respect to characteristics other than treatment assignment that might influence the outcome. In this manner, one can be confident that the observed treatment effect can be attributed to the treatment of interest and that causal relationships can be identified.

The criterion that addressed the homogeneity of the treatment groups at baseline was met in 65% of the trials (Table 2). The failure to use an appropriate randomization procedure may partially account for the apparent lack of similarity between the groups at baseline. The reason it is necessary to test for homogeneity of the treatment groups under investigation is to assure the reader that the randomization process achieved its goal. In the absence of proper randomization, even when the groups are demonstrated to be homogenous for the chosen tested variables, it is possible that another important variable overlooked by the researcher (eg, level of physical activity, strength) could still affect the outcome and bring the validity of the results into question.

Only 30% of trials met the minimum requirement to follow subjects for a sufficiently long period after treatment (Table 2). In general, the follow-up period must at least be sufficiently long for the outcome of interest to appear. Treatment effects tend to "wash out" over time, meaning that as treatment groups get farther from the time in which they were exposed to treatment, they become more equal with respect to the outcome of interest. This may or may not influence one's interpretation of the clinical meaningfulness of the results of a study. For example, even short-term improvements in pain and function potentially may have implications for decreasing costs associated with treatment and prevention of future injury. In most trials assessing physical therapy interventions, a follow-up period of at least 6 months to 1 year is ideal, although follow-up for even longer periods is even better. However, if the outcome of interest is expected to appear within a few days, such as an improvement in pain, or the nature of the research question justifies a short-term outcome, a short-term follow-up period may be appropriate.

Only 35% of the trials met the minimum level of quality for adequate blinding of both clinicians and outcomes assessors (Table 2). Increased effect sizes have been associated with a failure to double-blind.[59] When at all possible, subjects, outcomes assessors (technicians, examining clinicians, etc), and even the investigators not participating in the data collection process, should be blind to treatment assignment and outcomes. However, because subjects must be made aware of potential treatment group assignments when they provide informed consent, blinding of subjects in trials assessing physical therapy interventions can be quite problematic as designing a genuine placebo can be difficult, if not impossible. However, investigators can still be extremely careful to avoid implying to subjects which treatment is hypothesized to be most efficacious. Investigators can also assure that outcomes assessors remain blinded to treatment group assignment. The use of self-report measures of outcome helps to minimize bias associated with the failure to blind outcomes assessors, as the patient's self-report is not readily subject to bias from the investigator or clinician. However, for measures subject to rater bias, such as range of motion or muscle strength, it is imperative to utilize separate "treating" and "examining" clinicians. This helps to ensure that the treating clinician is blinded from the outcome and the examining clinician is blinded from treatment group assignment. Subjects must also be given specific instructions and frequently reminded to avoid divulging blinded information to either the treating or examining clinician.

Adequate control of cointerventions was reported in only 45% of the reviewed trials (Table 2). Extraneous factors such as use of pain medication and participation in sports may influence the effect of the intervention under study on the outcome of interest in PFPS, resulting in the inability to attribute the outcome solely to the experimental treatment. If any cointervention cannot be eliminated or manipulated

to be equally applied to both the experimental and control group, and it is associated with the outcome of interest, then the cointervention should be accounted for in the data analysis (eg, used as covariate).

The criterion for intention-to-treat (ITT) was met by only 50% of the trials in this review (Table 2). The failure to analyze the data based on the ITT principle has been frequently demonstrated to lead to bias in clinical trials.[34,46,58] To conduct an ITT analysis, the researcher analyzes all subjects within the group to which they were originally assigned, regardless of whether they completed the original treatment, dropped out of the study completely, or crossed over to a different treatment group. By doing an ITT, the researcher accounts for any systematic reasons as to why subjects dropped out of the study.

The criteria that addressed the accountability for withdrawals and dropouts was met in 70% of the trials (Table 2). For clinicians to analyze the feasibility of a studied treatment, researchers must include details of the number of and specific reasons for subjects who drop out of each treatment group. It is imperative to have a clear description of the reasons for subjects dropping out and to acknowledge any adverse event to allow determination of the potential side effects and risks associated with the treatments. A dramatically different dropout rate between groups with reasons related to the side effects or lack of feasibility of the treatment under investigation provides the clinician a clue that the particular treatment, despite any efficacy that has been demonstrated, may not be practical or safe in clinical practice.

The size of the sample was considered appropriate in 60% of the trials (Table 2). The number of subjects in a study is important when interpreting the results. Sample size is directly proportional to the statistical power of the study. Failure to demonstrate statistically significant differences should not be assumed to mean that no difference truly exists without a determination as to whether the study had sufficient power. When the sample size is inadequate, the possibility of not finding a difference in the effect of the treatments under investigation, when in reality there is a difference, must be considered (ie, Type II error).

Of the appraised trials, only 60% of the trials met the minimal requirement for the use of relevant outcome (Table 2). To achieve meaningful results, it is essential to incorporate reliable, valid, and responsive outcome measures that rate symptoms, impairments, and functional limitations specific to patients with PFPS. Moreover, it is important that the outcome measure properly addresses the research question. In a recent publication, Marx et al[33] compared the reliability, validity, and responsiveness of 4 self-reported, knee-specific, health-related outcome scales in athletic patients with a variety of knee disorders, including PFPS. Although the 4 scales demonstrated adequate psychometric properties, they recommended the use of the Activities of Daily Living Scale of the Knee Outcomes Survey to assess physical function in this population of patients.[24,66]

In a busy clinical practice, it may be both unreasonable and unnecessary for clinicians to actually use a quality scale and assign an individual score for each consideration to obtain an overall quality score. However, clinicians can relatively quickly judge the overall quality of a treatment study with respect to these problematic areas that may be associated with bias in a study's results simply by looking for key words in the methodology and results section such as, "randomization," "blind," "cointerventions," "follow-up," "intention-to-treat," etc. This brief glance can give clinicians an initial impression as to whether the study is of sufficient quality to warrant applying the study's results to their patients with similar characteristics as those used in the study without having to assign a definitive quality score. However, clinicians should always use caution in the decision to accept or reject treatment approaches based on a limited number of trials, especially those that are not high-quality studies. On the other hand, one very high-quality RCT may in fact be sufficient to base a treatment decision on, so long as the clinician's patient is similar to the subjects included in the study.

Although several of the trials in this review demonstrated a treatment effect for an intervention, others did not. Perhaps the inability to demonstrate a therapeutic effect for some of the physical therapy interventions for the treatment of PFPS may be because the researchers did not account for the existence of different subgroups of patients who require a specific type of treatment approach. In such a classification, each subgroup of patients would be characterized by a set of key signs and symptoms that clinicians could use to match the patient's findings with a well-defined treatment approach tailored to that subgroup. If clinicians were able to stage patients with PFPS appropriately and then match each presentation with a targeted treatment based on those key signs and symptoms, significant improvements in the outcome of trials that assess nonoperative treatments for PFPS would be expected. Ideally, the results of the clinical examination would guide the clinician down one of a finite number of paths of an easy-to-use treatment-based algorithm demonstrated to be reliable and valid in patients with PFPS. Treatment progression would then be based on the patient's achievement of certain clinical milestones and the clinician's ability to detect clinically important improvements in impairments and measures of function and disability. In this manner, an evidence-based approach could be used for the management of the

patient across the entire spectrum of the patient's course of care, from the initial examination to the patient's return to full function. Developing effective, clinical methods for classifying patients with PFPS should improve clinical decision making and treatment outcomes by matching treatments to the patients most likely to benefit from them. Classification methods will also enhance the power of clinical research in patients with PFPS by permitting researchers to study more homogenous groups of patients.

CONCLUSION

Based on the results of trials that seemed to exhibit a sufficient level of quality, treatments that seem effective in decreasing pain and improving function in patients with PFPS are acupuncture, quadriceps strengthening, the use of a resistive brace, and the combination of exercises with patellar taping and biofeedback. The use of soft foot orthotics in patients with excessive foot pronation appears useful to decrease pain. In addition, at a short-term follow-up, patients who received exercise programs were more likely to be discharged earlier from physical therapy.

Most RCTs we reviewed contained qualitative flaws that may bring the validity of the results into question, thus diminishing the ability to generalize the results to clinical practice. These flaws were primarily in the areas of randomization procedures, duration of follow-up, control of cointerventions, assurance of blinding, accountability and proper analysis of dropouts, number of subjects, and relevance of the outcomes. More high-quality clinical trials are needed to elucidate the most effective treatment strategies for patients with PFPS. Trials of high quality will support the clinical decision-making process in the care of our patients and help to provide strong evidence that can contribute to the recognition of the value of physical therapy. Given the limited number of high-quality clinical trials, recommendations to support or refute specific treatment approaches may be premature and can only be made with caution.

ACKNOWLEDGEMENTS

The authors would like to thank Gogilavaani Pillai, Sherry Heiser, Erica Baum, and Mark Schmeler for assisting with the literature review and article retrieval. We would also like to thank Gregory Hicks, Steve George, Kelley Fitzgerald, and Jay Irrgang for assisting with scoring the trials that were included in the review.

REFERENCES

1. Altman DG, Bland JM. Absence of evidence is not evidence of absence. *BMJ.* 1995;311(7003):485.
2. Antich TJ, Randall CC, Westbrook RA, Morrissey MC, Brewster CE. Physical therapy treatment of knee extensor mechanism disorders: comparison of four treatment modalities. *J Orthop Sports Phys Ther.* 1986;8(5):255–259.
3. Arroll B, Ellis-Pegler E, Edwards A, Sutcliffe G. Patellofemoral pain syndrome. A critical review of the clinical trials on nonoperative therapy. *Am J Sports Med.* 1997;25(2):207–212.
4. Brody LT, Thein JM. Nonoperative treatment for patellofemoral pain. *J Orthop Sports Phys Ther.* 1998;28(5):336–344.
5. Cerny K. Vastus medialis oblique/vastus lateralis muscle activity ratios for selected exercises in persons with and without patellofemoral pain syndrome. *Phys Ther.* 1995;75(8):672–683.
6. Cesarelli M, Bifulco P, Bracale M. Study of the control strategy of the quadriceps muscles in anterior knee pain. *IEEE Trans Rehabil Eng.* 2000;8(3):330–341.
7. Chesworth BM, Culham EG, Tata GE, Peat M. Validation of outcome measures in patients with patellofemoral syndrome. *J Orthop Sports Phys Ther.* 1989;10(8):302–308.
8. Clark DI, Downing N, Mitchell J, Coulson L, Syzpryt EP, Doherty M. Physiotherapy for anterior knee pain: a randomised controlled trial. *Ann Rheum Dis.* 2000;59(9):700–704.
9. Clarke M, Oxman AD. *Cochrane Reviewers' Handbook 4.1.4.* The Cochrane Library; 2001.
10. Crossley K, Cowan SM, Bennell KL, McConnell J. Patellar taping: is clinical success supported by scientific evidence? *Man Ther.* 2000;5(3):142–150.
11. Crossley K, Bennell K, Green S, McConnell J. A systematic review of physical interventions for patellofemoral pain syndrome. *Clin J Sport Med.* 2001;11(2):103–110.
12. Di Fabio RP. What is "evidence"? *J Orthop Sports Phys Ther.* 2000;30(2):52–55.
13. Duffey MJ, Martin DF, Cannon DW, Craven T, Messier SP. Etiologic factors associated with anterior knee pain in distance runners. *Med Sci Sports Exerc.* 2000;32(11):1825–1832.
14. Dye SF, Saubli HU, Biedert RM, Vaupel GL. The mosaic of pathophysiology causing patellofemoral pain: therapeutic implications. *Oper Tech Sports Med.* 1999;7(2):46–54.
15. Eburne J, Bannister G. The McConnell regimen versus isometric quadriceps exercises in the management of anterior knee pain. A randomised prospective controlled trial. *Knee.* 1996;3(3):151–153.
16. Eng JJ, Pierrynowski MR. Evaluation of soft foot orthotics in the treatment of patellofemoral pain syndrome. *Phys Ther.* 1993;73(2):62–68; discussion 68–70.
17. Finestone A, Radin EL, Lev B, Shlamkovitch N, Wiener M, Milgrom C. Treatment of overuse patellofemoral pain. Prospective randomized controlled clinical trial in a military setting. *Clin Orthop.* 1993;293:208–210.
18. Fulkerson JP, Folcik MA. Comparison of diflunisal and naproxen for relief of anterior knee pain. *Clin Ther.* 1986;9(Suppl C):59–61.
19. Gregoire G, Derderian F, Le Lorier J. Selecting the language of the publications included in a meta-analysis: is there a Tower of Babel bias? *J Clin Epidemiol.* 1995;48(1):159–163.
20. Harrison EL, Sheppard MS, McQuarrie AM. A randomized controlled trial of physical therapy treatment programs in patellofemoral pain syndrome. *Physiother Can.* 1999;51(2):93–100.

21. Hubbard JK, Sampson HW, Elledge JR. Prevalence and morphology of the vastus medialis oblique muscle in human cadavers. *Anat Rec.* 1997;249(1):135–142.
22. Hurley MV, Jones BW, Wilson D, Newham DJ. Rehabilitation of quadriceps inhibited due to isolated rupture to the anterior cruciate ligament. *J Orthop Rheumatol.* 1992;5:145–154.
23. Insall J. Current Concepts Review: patellar pain. *J Bone Joint Surg Am.* 1982;64(1):147–152.
24. Irrgang JJ, Snyder-Mackler L, Wainner RS, Fu FH, Harner CD. Development of a patient-reported measure of function of the knee. *J Bone Joint Surg Am.* 1998;80(8):1132–1145.
25. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials.* 1996;17(1):1–12.
26. Jensen R, Gothesen O, Liseth K, Baerheim A. Acupuncture treatment of patellofemoral pain syndrome. *J Altern Complement Med.* 1999;5(6):521–527.
27. Kannus P, Natri A, Niittymaki S, Jarvinen M. Effect of intraarticular glycosaminoglycan polysulfate treatment on patellofemoral pain syndrome. A prospective, randomized double-blind trial comparing glycosaminoglycan polysulfate with placebo and quadriceps muscle exercises. *Arthritis Rheum.* 1992;35(9):1053–1061.
28. Kannus P, Natri A, Paakkala T, Jarvinen M. An outcome study of chronic patellofemoral pain syndrome. Seven-year follow-up of patients in a randomized, controlled trial. *J Bone Joint Surg Am.* 1999;81(3):355–363.
29. Karst GM, Willett GM. Onset timing of electromyographic activity in the vastus medialis oblique and vastus lateralis muscles in subjects with and without patellofemoral pain syndrome. *Phys Ther.* 1995;75(9):813–823.
30. Kendall FP, McCreary EK, Provance PG. *Muscles: Testing and Function.* 4th ed. Baltimore, MD: Williams & Wilkins; 1993.
31. Kolowich PA, Paulos LE, Rosenberg TD, Farnsworth S. Lateral release of the patella: indications and contraindications. *Am J Sports Med.* 1990;18(4):359–365.
32. Kowall MG, Kolk G, Nuber GW, Cassisi JE, Stern SH. Patellar taping in the treatment of patellofemoral pain. A prospective randomized study. *Am J Sports Med.* 1996;24(1):61–66.
33. Marx RG, Jones EC, Allen AA, et al. Reliability, validity, and responsiveness of four knee outcome scales for athletic patients. *J Bone Joint Surg Am.* 2001;83-A(10):1459–1469.
34. May GS, DeMets DL, Friedman LM, Furberg C, Passamani E. The randomized clinical trial: bias in analysis. *Circulation.* 1981;64(4):669–673.
35. McConnell JS. The management of chondromalacia patellae: a long term solution. *Austr J Physiother.* 1986;32(4):215–233.
36. McConnell JS, Fulkerson JP. The knee: patellofemoral and soft tissue injuries. In: Zachazewski JE, Magee DJ and Quillen WS, eds. *Athletic Injuries and Rehabilitation.* Philadelphia, PA: W.B. Saunders Co.; 1996:693–728.
37. McNutt RA, Evans AT, Fletcher RH, Fletcher SW. The effects of blinding on the quality of peer review. A randomized trial. *JAMA.* 1990;263(10):1371–1376.
38. Miller MD, Hinkin DT, Wisnowski JW. The efficacy of orthotics for anterior knee pain in military trainees. A preliminary report. *Am J Knee Surg.* 1997;10(1):10–13.
39. Moher D, Fortin P, Jadad AR, et al. Completeness of reporting of trials published in languages other than English: implications for conduct and reporting of systematic reviews. *Lancet.* 1996;347(8998):363–366.
40. Moher D, Jadad AR, Nichol G, Penman M, Tugwell P, Walsh S. Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. *Control Clin Trials.* 1995;16(1):62–73.
41. Moher D, Jadad AR, Tugwell P. Assessing the quality of randomized controlled trials. Current issues and future directions. *Int J Technol Assess Health Care.* 1996;12(2):195–208.
42. Moher D, Pham B, Jones A, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet.* 1998;352(9128):609–613.
43. Natri A, Kannus P, Jarvinen M. Which factors predict the long-term outcome in chronic patellofemoral pain syndrome? A 7-yr prospective follow-up study. *Med Sci Sports Exerc.* 1998;30(11):1572–1577.
44. O'Sullivan SB, Schmitz TJ. *Physical Rehabilitation: Assessment and Treatment.* 2nd ed. Philadelphia, PA: F.A. Davis Co.; 1998.
45. Palumbo PM, Jr. Dynamic patellar brace: a new orthosis in the management of patellofemoral disorders. A preliminary report. *Am J Sports Med.* 1981;9(1):45–49.
46. Peduzzi P, Wittes J, Detre K, Holford T. Analysis as randomized and the problem of nonadherence: an example from the Veterans Affairs Randomized Trial of Coronary Artery Bypass Surgery. *Stat Med.* 1993;12(13):1185–1195.
47. Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice.* Norwalk, CT: Appleton & Lange; 1993.
48. Powers CM. Rehabilitation of patellofemoral joint disorders: a critical review. *J Orthop Sports Phys Ther.* 1998;28(5):345–354.
49. Powers CM, Landel R, Perry J. Timing and intensity of vastus muscle activity during functional activities in subjects with and without patellofemoral pain. *Phys Ther.* 1996;76(9):946–955; discussion 956–967.
50. Raatikainen T, Vaananen K, Tamelander G. Effect of glycosaminoglycan polysulfate on chondromalacia patellae. A placebo-controlled 1-year study. *Acta Orthop Scand.* 1990;61(5):443–448.
51. Rogvi-Hansen B, Ellitsgaard N, Funch M, Dall-Jensen M, Prieske J. Low level laser treatment of chondromalacia patellae. *Int Orthop.* 1991;15(4):359–361.
52. Roush MB, Sevier TL, Wilson JK, et al. Anterior knee pain: a clinical comparison of rehabilitation methods. *Clin J Sport Med.* 2000;10(1):22–28.
53. Rowlands BW, Brantingham JW. The efficacy of patella mobilization in patients suffering from patellofemoral pain syndrome. *J Orthop Rheumatol.* 1999;7(4):142–149.
54. Rubin B, Collins R. Runner's knee. *Physician Sportsmed.* 1980;8(6):49–58.
55. Sackett DL, Haynes RB, Guyatt GH, Tugwell PD. *Clinical Epidemiology: A Basic Science for Clinical Medicine.* Boston, MA: Little, Brown, & Co.; 1991.
56. Sackett DL, Strause SE, Richardson WS, Rosenberg W, Haynes RB. *Evidence-Based Medicine: How to Practice and Teach EBM.* 2nd ed. New York, NY: Churchill Livingstone, Inc.; 2000.
57. Sanchis-Alfonso V, Rosello-Sastre E, Martinez-Sanjuan V. Pathogenesis of anterior knee pain syndrome and functional patellofemoral instability in the active young. *Am J Knee Surg.* 1999;12(1):29–40.
58. Schulz KF. Subverting randomization in controlled trials. *JAMA.* 1995;274(18):1456–1458.

59. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA*. 1995;273(5):408–412.
60. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86(2):420–428.
61. Sikorski JM, Peters J, Watt I. The importance of femoral rotation in chondromalacia patellae as shown by serial radiography. *J Bone Joint Surg Br*. 1979;61–B(4):435–442.
62. Smith AD, Stroud L, McQueen C. Flexibility and anterior knee pain in adolescent elite figure skaters. *J Pediatr Orthop*. 1991;11(1):77–82.
63. Suter E, Herzog W, DeSouzak K, Bray R. Inhibition of the quadriceps muscles in patients with anterior knee pain. *J Appl Biomech*. 1998;14:360–373.
64. Suter E, McMorland G, Herzog W, Bray R. Conservative lower back treatment reduces inhibition in knee-extensor muscles: a randomized controlled trial. *J Manipulative Physiol Ther*. 2000;23(2):76–80.
65. Taylor DC, Dalton JD, Jr., Seaber AV, Garrett WE, Jr. Viscoelastic properties of muscle-tendon units. The biomechanical effects of stretching. *Am J Sports Med*. 1990;18(3):300–309.
66. Thachuk C, Irrgang JJ, Starz TW. Responsiveness of a specific measure of physical function of the knee as compared to a general measure of health status. *American College of Rheumatology, 62nd National Meeting*. San Diego, CA; 1998.
67. Thomee R. A comprehensive treatment approach for patellofemoral pain syndrome in young women. *Phys Ther*. 1997;77(12):1690–1703.
68. Thomee R, Augustsson J, Karlsson J. Patellofemoral pain syndrome: a review of current issues. *Sports Med*. 1999;28(4):245–262.
69. Timm KE. Randomized controlled trial of Protonics on patellar pain, position, and function. *Med Sci Sports Exerc*. 1998;30(5):665–670.
70. van Tulder MW, Assendelft WJ, Koes BW, Bouter LM. Method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for Spinal Disorders. *Spine*. 1997;22(20):2323–2330.
71. Verhagen AP, de Vet HC, de Bie RA, Kessels AG, Boers M, Knipschild PG. Balneotherapy and quality assessment: interobserver reliability of the Maastricht criteria list and the need for blinded quality assessment. *J Clin Epidemiol*. 1998;51(4):335–341.
72. Werner S. An evaluation of knee extensor and knee flexor torques and EMGs in patients with patellofemoral pain syndrome in comparison with matched controls. *Knee Surg Sports Traumatol Arthrosc*. 1995;3(2):89–94.
73. Witvrouw E, Lysens R, Bellemans J, Cambier D, Vanderstraeten G. Intrinsic risk factors for the development of anterior knee pain in an athletic population. A two-year prospective study. *Am J Sports Med*. 2000;28(4):480–489.
74. Witvrouw E, Lysens R, Bellemans J, Peers K, Vanderstraeten G. Open versus closed kinetic chain exercises for patellofemoral pain. A prospective, randomized study. *Am J Sports Med*. 2000;28(5):687–694.