Risk factors for patellofemoral pain: a systematic review and meta-analysis

Bradley S Neal,1,2 Simon D Lack,1,2 Nienke E Lankhorst,3 Andrew Raye,1 Dylan Morrissey,1,4 Marienke van Middelkoop3

ABSTRACT
Background Patellofemoral pain (PFP) is a prevalent condition commencing at various points throughout life. We aimed to provide an evidence synthesis concerning predictive variables for PFP, to aid development of preventative interventions.

Methods We searched Medline, Web of Science and SCOPUS until February 2017 for prospective studies investigating at least one potential risk factor for future PFP. Two independent reviewers appraised methodological quality using the Newcastle–Ottawa Scale. We conducted meta-analysis where appropriate, with standardised mean differences (SMD) and risk ratios calculated for continuous and nominal scaled data.

Results This review included 18 studies involving 4818 participants, of whom 483 developed PFP (heterogeneous incidence 10%). Three distinct subgroups (military recruits, adolescents and recreational runners) were identified. Strong to moderate evidence indicated that age, height, weight, body mass index (BMI), body fat and Q angle were not risk factors for future PFP. Moderate evidence indicated that quadriceps weakness was a risk factor for future PFP in the military, especially when normalised by BMI (SMD −0.69, CI −1.02, –0.35). Moderate evidence indicated that hip weakness was not a risk factor for future PFP (multiple pooled SMDs range −0.09 to −0.20), but in adolescents, moderate evidence indicated that increased hip abduction strength was a risk factor for future PFP (SMD 0.71, CI 0.39, 1.04).

Conclusions This review identified multiple variables that did not predict future PFP, but quadriceps weakness in military recruits and higher hip strength in adolescents were risk factors for PFP. Identifying modifiable risk factors is an urgent priority to improve prevention and treatment outcomes.

INTRODUCTION
Patellofemoral pain (PFP) is characterised by diffuse retropatellar or periapatellar symptoms throughout activities that load the knee during flexion, such as running, stair descent or squatting.1 PFP is a common pathology in both adolescents2 and adults,3 with prevalence in the general population reported as 22.7%.4 However, the factors associated with PFP development and the incidence of the condition across a variety of populations remains under-evaluated due to limited prospective data and the homogeneity of studied populations.4 5 As PFP is reported to be common across the lifespan and may be the precursor to patellofemoral osteoarthritis,6 7 an improved understanding of the factors associated with the development of PFP and its incidence in differing populations is essential to prevent symptoms.

With the incidence of PFP reported to be high4 and symptoms persisting despite evidence based interventions,8 further investigation is warranted to understand variables that are associated with PFP development and subsequently deliver evidence based preventative strategies. In 1992, Van Mechelen et al presented a theoretical model described as the ‘sequence of prevention’ for sports injury (see figure 1) to guide injury prevention development.9 With the incidence of PFP defined across populations10 (stage 1), an understanding of the aetiology (stage 2) is required to identify the variables associated with the pathology’s development. A variable associated with future pathology should be manipulated as a preventative strategy within a randomised controlled trial (stage 3). The effectiveness of the implemented strategy should then be appraised by re-examining the incidence within a specific population (stage 4).

In 2012, Lankhorst et al completed a systematic review of risk factors for PFP,10 which identified a clear association between low knee extension strength and subsequent risk of PFP, irrespective of measurement method, but no associations with other investigated variables. This is likely due to the low number of included studies (n=7), high data heterogeneity and data pooling being possible for just 13 out of 137 identified variables, but was unexpected given the known cross sectional association between PFP and multiple pathomechanical variables, such as muscle function and lower limb biomechanics.11

Additional risk factors for future PFP have been reported in other systematic reviews using data from single studies. Increased navicular drop in military recruits,12 greater peak hip adduction during running13 and increased forces at foot level during both walking and running14 all increase the risk of future PFP. While these findings are statistically significant, the absence of data pooling and the small to moderate effect sizes limit their impact and clinical applicability. Given the number of subsequently published prospective studies, an updated systematic review on this topic is now appropriate.

The aim of this systematic review was to provide researchers and clinicians with an evidence synthesis concerning predictive variables for PFP, to aid the development of preventative interventions. The review was designed to synthesise the available evidence at stage 2 (aetiology) of the Van Mechelen model (see figure 1), and enable addressing stage 3 (preventative strategies). A secondary aim was to determine the incidence of PFP within the included studies, both as a...
heterogeneous condition and within specific homogenous cohorts. Specific objectives were to (i) establish prospective links between all investigated variables and future PFP, (ii) identify risk factors and PFP incidence specific to individual homogenous cohorts and (iii) inform future studies on PFP prevention.

METHODS
This systematic review was completed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement15 and was registered with PROSPERO prior to completion of the initial search (registration No: CRD42016049327).

Search strategy
The search terms used by Lankhorst et al10 were duplicated for the purpose of this review. The following terms were used for PFP: arthralgia AND knee joint OR anterior knee pain OR (patell* OR femoropatell* OR femoro-patell* OR retropatell*) AND (pain OR syndrome OR dysfunction). Keywords used for risk factors were: risk factor OR association OR relative risk OR RR ≥2.0

Inclusion criteria
A single investigator (ADR) exported all studies identified by the search strategy to Endnote X7 (Thomson Reuters, Philadelphia, USA). Eligibility criteria were adapted from the original review of Lankhorst et al10: (i) studies involving male or female subjects who developed subsequent PFP (synonyms including retropatella pain, chondromalacia or anterior knee pain); (ii) at least one variable investigated as a risk factor for PFP; and (iii) prospective study designs. Studies with <20 PFP subjects were excluded by the review of Lankhorst et al,10 but were included in this review. Two independent authors (BSN and NEL) reviewed all abstracts to determine eligibility. Full texts were screened where eligibility could not be determined by the abstract alone and any discrepancies were resolved at a consensus meeting.

Quality assessment
Methodological quality and risk of bias of included studies were determined by combining the Newcastle–Ottawa Scale16 (NOS) and appraising the number of events per variable.17 Eligible studies were independently rated by two authors blind to the study authors and institutions (SDL and NEL), with discrepancies resolved at a consensus meeting. The NOS contains eight categories relating to methodological quality and each study was given an eventual score out of a maximum of 8 points. A score of 0–3 points equated to a low quality (LQ) study, a score of 4–6 points equated to a moderate quality (MQ) study, with a score of 7–8 points required for a study to be given a score of high quality (HQ). In addition, HQ or MQ studies were reduced to either MQ or LQ respectively, if they were determined to have a high risk of bias as a result of having <10 PFP participants for each investigated variable within their total sample.17 Inter-rater reliability of the NOS was calculated using the percentage agreement method.

Data extraction
Data related to study characteristics were initially extracted from all included studies by one author (ADR) and subsequently reviewed by a second author (BSN). This included participant numbers (separating those who developed PFP and those who did not), characteristics of those groups (such as population), study duration and publication details (author and year). A second author (BSN) extracted all data pertaining to potential risk variables to be included in the meta-analysis. Means and SD were extracted for variables of interest, which included (but were not limited to) anthropometrics and demographics (such as sex, body mass index (BMI)), biomechanical variables (such as kinematics and kinetics) and muscle function (such as strength or onset timing).

Statistical methods
Statistical analyses were undertaken using Review Manager 5.0 (The Cochrane Collaboration, Copenhagen, Denmark). Analyses were completed initially by one author (SDL) and subsequently reviewed by a second author (BSN). Means and SDs were extracted for continuous scaled variables and used to calculate a standardised mean difference (SMD) with 95% confidence intervals (CIs). Calculated individual or pooled SMDs were categorised as small (0.59), medium (0.60–1.19) or large (1.20).18 For nominal scaled variables, raw counts of injured and uninjured participants (eg, PFP incidence in men and women) were extracted and used to calculate risk ratios (RR) with 95% CIs, with a small effect indicated by an RR ≥2.0 and a large effect by an RR ≥4.0.18

Data were pooled and have been presented as both a heterogeneous PFP cohort and further pooled by specific homogeneous subgroup where possible. Where methodological approaches between studies were deemed to be adequately comparable, a meta-analysis was performed and the level of statistical heterogeneity for pooled data determined using I² statistics (heterogeneity defined as I² >50%, P<0.05). Random effects were used due to the variation in study methods and populations, and the typically low number of studies, therefore reducing the possibility of a type 1 error.19

Only outcomes incorporating data from a minimum of two studies are presented in the main body of the review, due to the risk of reporting inappropriate levels of evidence where data pooling was not possible.

Evidence based recommendations
Levels of evidence were assigned to each calculated variable (pooled or otherwise), as described by Van Tulder et al,20 which...
incorporate both assigned methodological quality of included studies and statistical outcomes.

Strong evidence
Pooled results derived from three or more studies, including a minimum of two high quality studies that are statistically homogenous.

Moderate evidence
Pooled results derived from multiple studies, including at least one high quality study, that are statistically heterogeneous; or from multiple moderate or low quality studies which are statistically homogenous.

Limited evidence
Results from one high quality study or multiple moderate or low quality studies that are statistically heterogeneous.

Very limited evidence
Results from one moderate or low quality study.

RESULTS
Search results
The search resulted in 3044 titles and abstracts being identified for screening. Following the removal of duplicates and studies that did not meet the inclusion criteria of the review, 18 studies involving a total of 4818 participants were included (see figure 2), 21–38 483 of whom went on to develop symptoms consistent with PFP. This is indicative of a heterogeneous PFP incidence of 10%. Extracted data relating to study characteristics are presented in table 1.

Subgroups and PFP incidence
Three distinct subgroups were identified during the data extraction process. There were a total of seven studies involving military recruits, 21 23–25 28 36 37 six studies involving adolescents 22 26 27 29 31 38 and five studies involving recreational runners. 30 32–35 Studies involving military recruits involved a total of 2435 participants, 280 of whom went on to develop PFP; reflective of an incidence of 11% (range 3–43%). Studies involving adolescents involved a total of 1118 participants, 128 of whom went on to develop PFP, reflective of an incidence of 11% (range 5–15%). Studies involving recreational runners

Results
The search resulted in 3044 titles and abstracts being identified for screening. Following the removal of duplicates and studies that did not meet the inclusion criteria of the review, 18 studies involving a total of 4818 participants were included (see figure 2), 21–38 483 of whom went on to develop symptoms consistent with PFP. This is indicative of a heterogeneous PFP incidence of 10%. Extracted data relating to study characteristics are presented in table 1.

Subgroups and PFP incidence
Three distinct subgroups were identified during the data extraction process. There were a total of seven studies involving military recruits, 21 23–25 28 36 37 six studies involving adolescents 22 26 27 29 31 38 and five studies involving recreational runners. 30 32–35 Studies involving military recruits involved a total of 2435 participants, 280 of whom went on to develop PFP; reflective of an incidence of 11% (range 3–43%). Studies involving adolescents involved a total of 1118 participants, 128 of whom went on to develop PFP, reflective of an incidence of 11% (range 5–15%). Studies involving recreational runners

Table 1 Study characteristics

<table>
<thead>
<tr>
<th>Study</th>
<th>NOS score</th>
<th>Risk of bias</th>
<th>Cohort</th>
<th>PFP</th>
<th>Sample size</th>
<th>Incidence (%)</th>
<th>Study duration (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boiling et al</td>
<td>H</td>
<td>L</td>
<td>Military (USA)</td>
<td>40 (M=16, F=24)</td>
<td>1319</td>
<td>3</td>
<td>30</td>
</tr>
<tr>
<td>Duvigneaud et al</td>
<td>H</td>
<td>L</td>
<td>Military (Belgium)</td>
<td>26 (F=26)</td>
<td>62</td>
<td>42</td>
<td>1.5</td>
</tr>
<tr>
<td>Finnoff et al</td>
<td>MD</td>
<td>H</td>
<td>Adolescents (USA)</td>
<td>5 (M=2, F=3)</td>
<td>98</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Foss et al</td>
<td>H</td>
<td>L</td>
<td>Adolescents (USA)</td>
<td>39 (F=39)</td>
<td>262</td>
<td>15</td>
<td>24</td>
</tr>
<tr>
<td>Herbst et al</td>
<td>H</td>
<td>L</td>
<td>Adolescents (USA)</td>
<td>38 (F=38)</td>
<td>255</td>
<td>15</td>
<td>12</td>
</tr>
<tr>
<td>Hetsroni et al</td>
<td>H</td>
<td>L</td>
<td>Military (Israel)</td>
<td>61 (MF=7)</td>
<td>405</td>
<td>15</td>
<td>4</td>
</tr>
<tr>
<td>Holden et al</td>
<td>MD</td>
<td>H</td>
<td>Adolescents (Ireland)</td>
<td>8 (F=8)</td>
<td>76</td>
<td>11</td>
<td>24</td>
</tr>
<tr>
<td>Luedeke et al</td>
<td>MD</td>
<td>H</td>
<td>Recreational runners (USA)</td>
<td>3 (M=1, F=2)</td>
<td>57</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>Milgrom et al</td>
<td>MD</td>
<td>L</td>
<td>Military (Israel)</td>
<td>60 (M=60)</td>
<td>390</td>
<td>15</td>
<td>3.5</td>
</tr>
<tr>
<td>Myer et al</td>
<td>MD</td>
<td>H</td>
<td>Adolescents (USA)</td>
<td>14</td>
<td>145</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>Noehren et al</td>
<td>MD</td>
<td>H</td>
<td>Recreational runners (USA)</td>
<td>15 (F=15)</td>
<td>400</td>
<td>3</td>
<td>24</td>
</tr>
<tr>
<td>Ramskov et al</td>
<td>H</td>
<td>L</td>
<td>Recreational runners (Denmark)</td>
<td>24 (M=10, F=14)</td>
<td>629</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>Thijssen et al</td>
<td>H</td>
<td>L</td>
<td>Military (Belgium)</td>
<td>36 (M=25, F=11)</td>
<td>84</td>
<td>43</td>
<td>1.5</td>
</tr>
<tr>
<td>Thijssen et al</td>
<td>MD</td>
<td>H</td>
<td>Recreational runners (Belgium)</td>
<td>17 (M=1, F=16)</td>
<td>102</td>
<td>17</td>
<td>2.5</td>
</tr>
<tr>
<td>Thijssen et al</td>
<td>MD</td>
<td>H</td>
<td>Recreational runners (Belgium)</td>
<td>16 (F=16)</td>
<td>77</td>
<td>21</td>
<td>2.5</td>
</tr>
<tr>
<td>Van Tiggelen et al</td>
<td>H</td>
<td>L</td>
<td>Military (Belgium)</td>
<td>31 (M=31)</td>
<td>96</td>
<td>32</td>
<td>1.5</td>
</tr>
<tr>
<td>Van Tiggelen et al</td>
<td>H</td>
<td>L</td>
<td>Military (Belgium)</td>
<td>26 (M=26)</td>
<td>79</td>
<td>33</td>
<td>1.5</td>
</tr>
<tr>
<td>Witrout et al</td>
<td>MD</td>
<td>H</td>
<td>Adolescents (Belgium)</td>
<td>24 (M=11, F=13)</td>
<td>282</td>
<td>9</td>
<td>24</td>
</tr>
</tbody>
</table>

F, female; H, high; L, low; M, male; MD, moderate; NOS, Newcastle–Ottawa Scale; PFP, patellofemoral pain.
involved a total of 1265 participants, 75 of whom went on to develop PFP, reflective of an incidence of 6% (range 4–21%).

**Quality assessment**

After evaluation of study quality and the risk of bias, a total of nine HQ studies and a further nine MQ studies were identified. The questions with the lowest percentage agreement were question No 5 (does the study control for any confounding variables) and question No 7 (was follow-up time clearly defined).

**Anthropometrics and demographics**

Data pooling was possible for seven individual variables (sex, height, weight, BMI, body fat percentage, age and limb length).

**Sex**

There is moderate evidence from three HQ and four MQ studies that sex is not a risk factor for future PFP ($I^2=73\%$, RR 1.33, CI 0.76, 2.34) (see figure 3). This outcome does not change when pooling data only for military subjects (moderate evidence, $I^2=91\%$, RR 0.82, CI 0.25, 2.74), adolescents (moderate evidence, $I^2=0\%$, RR 1.23, CI 0.38, 2.07) or recreational runners (moderate evidence, $I^2=76\%$, RR 3.08, CI 0.59, 15.99). While subgroup data pooling was non-significant, six of the seven included studies that reported data on sex had a greater proportion of women in their PFP cohort. The highest of which was observed in the recreational runner subgroup.

**Height**

There is strong evidence from five HQ and seven MQ studies that height is not a risk factor for future PFP ($I^2=0\%$, SMD $−0.08$, CI $−0.21$, 0.05) (see figure 4). This outcome does not change when pooling data for only military recruits (strong evidence, $I^2=41\%$, SMD $−0.15$, CI $−0.42$, 0.12), adolescents (moderate evidence, $I^2=0\%$, SMD $0.06$, CI $−0.23$, 0.35) or recreational runners (moderate evidence, $I^2=0\%$, SMD $−0.15$, CI $−0.43$, 0.13).

**Weight**

There is strong evidence from five HQ and seven MQ studies that weight is not a risk factor for future PFP ($I^2=0\%$, SMD $0.02$, CI $−0.11$, 0.16) (see figure 5). This outcome does not change when pooling data for only military recruits (strong evidence, $I^2=0\%$, SMD $0.05$, CI $−0.12$, 0.23), adolescents (moderate evidence, $I^2=28\%$, SMD $−0.10$, CI $−0.46$, 0.25) or recreational runners (moderate evidence, $I^2=0\%$, SMD $0.10$, CI $−0.18$, 0.37).

**Body mass index**

There is strong evidence from four HQ and three MQ studies that BMI is not a risk factor for future PFP ($I^2=33\%$, SMD $0.10$, CI $−0.12$ to 0.32) (see figure 6). This outcome does not change when pooling data for only military recruits (moderate evidence, $I^2=65\%$, SMD $0.09$, CI $−0.48$ to 0.65), adolescents (moderate evidence, $I^2=75\%$, SMD $0.23$, CI $−0.72$ to 1.18) or recreational runners (moderate evidence, $I^2=0\%$, SMD $0.15$, CI $−0.13$ to 0.43).

**Body fat percentage**

There is moderate evidence from one HQ study and one MQ study that body fat percentage is not a risk factor for future PFP in adolescents ($I^2=0\%$, SMD $−0.13$, CI $−0.40$, 0.13) (see figure 7). This variable was not investigated in either military recruits or recreational runners.

**Age**

There is strong evidence from three HQ and five MQ studies that age is not a risk factor for future PFP ($I^2=13\%$, SMD $0.06$, CI $−0.13$, 0.25) (see figure 8). This outcome does not change when pooling data for only military recruits (moderate evidence, $I^2=0\%$, SMD $−0.05$, CI $−0.36$, 0.27), adolescents (limited evidence, $I^2=80\%$, SMD $0.04$, CI $−0.98$, 1.07) or recreational runners (moderate evidence, $I^2=0\%$, SMD $0.16$, CI $−0.09$, 0.40).

**Limb length**

There is limited evidence from two MQ studies that limb length is not a risk factor for future PFP ($I^2=0\%$, SMD $0.06$, CI $−0.28$, 0.25). This variable was not investigated in recreational runners and no data pooling was possible within any individual subgroups.

**Lower limb alignment**

Data pooling was only possible for static Q angle. Limited evidence from one HQ and one MQ study indicates that Q angle is not a risk factor for future PFP ($I^2=0\%$, SMD $−0.22$, 0.33) (see figure 9). No data pooling was possible for any identified subgroup.

**Strength measures**

**Quadriceps strength**

When pooling all available data for quadriceps strength, regardless of cohort or measurement method, there is strong evidence that quadriceps weakness is a risk factor for future PFP (moderate evidence, $I^2=65\%$, small SMD $−0.32$, CI $−0.42$ to $−0.22$).

Data pooling was only possible for the military subgroup for all quadriceps strength measures. There is moderate evidence from two HQ studies that quadriceps weakness is a risk factor for future PFP when measured with an isokinetic dynamometer concentrically at 60˚/s ($I^2=0\%$, moderate SMD $−0.66$, CI $−0.99$, $−0.32$) (see figure 10A) or concentrically at 240˚/s ($I^2=17\%$, small SMD $−0.49$, CI $−0.85$ to $−0.12$) (see figure 10B).

For normalised quadriceps strength measured with an isokinetic dynamometer, there is moderate evidence from two HQ studies that quadriceps weakness is a risk factor for future PFP when normalised by body mass at 60˚/s ($I^2=0\%$, moderate SMD $−0.61$, CI $−0.94$ to $−0.28$) (see figure 10C) or at 240˚/s ($I^2=0\%$, small SMD $−0.53$, CI $−0.87$ to $−0.20$) (see figure 10D). When normalised by BMI, moderate evidence remains that quadriceps weakness is a risk factor for future PFP when measured at both 60˚/s ($I^2=0\%$, moderate SMD $−0.69$, CI $−1.02$ to $−0.35$) (see figure 10E) and 240˚/s ($I^2=0\%$, small SMD $−0.51$, CI $−0.84$ to $−0.18$) (see figure 10F).

For quadriceps strength measured isometrically with a handheld dynamometer (HHD), there is moderate evidence from one HQ and one MQ study that quadriceps strength is not a risk factor for future PFP ($I^2=82\%$, small SMD $−0.25$, CI $−0.74$, 0.25).
<table>
<thead>
<tr>
<th>Question</th>
<th>R1</th>
<th>R2</th>
<th>R1</th>
<th>R2</th>
<th>R1</th>
<th>R2</th>
<th>R1</th>
<th>R2</th>
<th>R1</th>
<th>R2</th>
<th>R1</th>
<th>R2</th>
<th>R1</th>
<th>R2</th>
<th>R1</th>
<th>R2</th>
<th>R1</th>
<th>R2</th>
<th>R1</th>
<th>R2</th>
<th>R1</th>
<th>R2</th>
<th>R1</th>
<th>R2</th>
<th>R1</th>
<th>R2</th>
<th>R1</th>
<th>R2</th>
<th>R1</th>
<th>R2</th>
<th>R1</th>
<th>R2</th>
<th>R1</th>
<th>R2</th>
<th>R1</th>
<th>R2</th>
</tr>
</thead>
<tbody>
<tr>
<td>The cohort was truly or somewhat representative of a typical PFP cohort</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Selection of the non-PFP cohort was from the same community as the PFP cohort</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Ascertainment of PFP was made via secure record or structured interview</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Demonstration that outcome of interest was not present at start of study</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Cohorts were comparable on the basis of the design or confounders controlled for</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Assessment of outcome was independent OR linked to medical records</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>Follow-up time was clearly defined</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>Follow-up was adequate (all subjects accounted for or ≤20% attrition)</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td></td>
</tr>
</tbody>
</table>

| Percentage agreement (%) | 94 | 94 | 100 | 100 | 89 | 100 | 89 | 94 |

? = unable to determine.
N, no; PFP, patellofemoral pain; R, rater; Y, yes.
Hamstrings strength
There is moderate evidence from two HQ studies\(^{25, 37}\) that hamstrings strength is not a risk factor for future PFP in the military when measured with an isokinetic dynamometer concentrically at 60°/s (\(I^2=0\%\), SMD –0.09, CI –0.42, 0.24) or 240°/s (\(I^2=0\%\), SMD –0.10, CI –0.43, 0.22). This variable was not investigated in either adolescents or recreational runners.

Hip strength
There is moderate evidence from one HQ \(^{21}\) and two MQ \(^{26, 35}\) studies that hip extension (\(I^2=0\%\), SMD –0.18, CI –0.44, 0.09) (see figure 11A), hip internal rotation (\(I^2=20\%\), SMD –0.17, CI –0.43, 0.10) (see figure 11B) and hip flexion strength (\(I^2=0\%\), SMD –0.17, CI –0.43, 0.10) (see figure 11C), measured isometrically with a HHD, are not risk factors for future PFP. There is also limited evidence from two MQ studies \(^{26, 35}\) that both hip adduction strength (\(I^2=0\%\), SMD –0.20, CI –0.67, 0.28) (see figure 11D) and hip flexion strength (\(I^2=52\%\), SMD –0.08, CI –0.82, 0.67) (see figure 11E) are not risk factors for future PFP when measured with a HHD. No data pooling was possible for any identified subgroup for these strength measures.

There is moderate evidence from two HQ \(^{21, 27}\) and two MQ \(^{26, 35}\) studies that hip abduction strength is not a risk factor for future PFP (\(I^2=86\%\), SMD 0.25, CI 0.38, 0.88) (see figure 12A) when measured isometrically with a HHD. When the data were pooled for the adolescent cohort, there is moderate evidence from one HQ \(^{27}\) and one MQ study \(^{26}\) that increased hip abduction strength is a risk factor for future PFP (\(I^2=0\%\), SMD 0.71, CI 0.39, 1.04) (see figure 12B) when measured isometrically with a HHD. Data pooling was not possible for the military or recreational runner subgroups.

Biomechanics
Dynamic knee valgus angle
Moderate evidence from one HQ study \(^{21}\) and one MQ study \(^{29}\) indicates that knee valgus angle during a jump land task is not a risk factor for future PFP (\(I^2=99\%\), SMD 4.17, CI –4.19, 12.53). No data pooling was possible for any identified subgroup.

Foot kinetics
One HQ study \(^{36}\) and one MQ study \(^{34}\) investigated foot kinetics during walking and running, respectively. When these data were pooled, moderate evidence indicates no significant associations between time to peak force at any investigated region of the foot, which included the hallux, the metatarsal heads and the medial/lateral heel.

Discussion
This systematic review aimed to provide a synthesis of the evidence concerning predictive variables for PFP development. Despite the inclusion of 11 additional prospective studies and 55 additional variables when compared with the previous review of Lankhorst et al.,\(^{10}\) high data heterogeneity and a limited ability to pool data remained. Just two predictive variables, lower isokinetic quadriceps strength in the military and higher isometric hip abduction strength in adolescents, were identified. Heterogeneous incidence of PFP was 10%, with incidence also identified

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Females</th>
<th>Males</th>
<th>Control Females</th>
<th>Control Males</th>
<th>Std. Mean Difference</th>
<th>Std. Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Event</td>
<td>Total</td>
<td>Event</td>
<td>Total</td>
<td>IV, Random, 95% CI</td>
<td>IV, Random, 95% CI</td>
</tr>
<tr>
<td>Duvigneaud 2008 (M/HQ)</td>
<td>16.68</td>
<td>7.58</td>
<td>26</td>
<td>167.1</td>
<td>6.2</td>
<td>36</td>
</tr>
<tr>
<td>Finnis 2011 (A/MQ)</td>
<td>147</td>
<td>5.10</td>
<td>10</td>
<td>172.10</td>
<td>92</td>
<td>2.2%</td>
</tr>
<tr>
<td>Holden 2015 (A/MQ)</td>
<td>161</td>
<td>8.50</td>
<td>159</td>
<td>7.68</td>
<td>3.2%</td>
<td></td>
</tr>
<tr>
<td>Milgrom 1991 (M/MQ)</td>
<td>177.8</td>
<td>7.30</td>
<td>60</td>
<td>177.61</td>
<td>330</td>
<td>23.0%</td>
</tr>
<tr>
<td>Myer 2010 (A/MQ)</td>
<td>160.7</td>
<td>9.20</td>
<td>14</td>
<td>160.8</td>
<td>131</td>
<td>5.7%</td>
</tr>
<tr>
<td>Ramskov 2015 (R/HQ)</td>
<td>174</td>
<td>10.24</td>
<td>24</td>
<td>175.10</td>
<td>605</td>
<td>10.5%</td>
</tr>
<tr>
<td>Thijs 2007 (M/HQ)</td>
<td>175.94</td>
<td>7.54</td>
<td>36</td>
<td>179.28</td>
<td>7.73</td>
<td>48</td>
</tr>
<tr>
<td>Thijs 2008 (M/MQ)</td>
<td>164.5</td>
<td>26.88</td>
<td>17</td>
<td>167.4</td>
<td>7.58</td>
<td>85</td>
</tr>
<tr>
<td>Thijs 2011 (R/MQ)</td>
<td>166.9</td>
<td>4.90</td>
<td>16</td>
<td>167.65</td>
<td>61</td>
<td>5.7%</td>
</tr>
<tr>
<td>Van Tiggelen 2004 (M/HQ)</td>
<td>178.4</td>
<td>5.30</td>
<td>31</td>
<td>181.5</td>
<td>6.4</td>
<td>65</td>
</tr>
<tr>
<td>Van Tiggelen 2009 (M/MQ)</td>
<td>180.6</td>
<td>6.12</td>
<td>26</td>
<td>180.5</td>
<td>6.22</td>
<td>53</td>
</tr>
<tr>
<td>Wijtvrouw 2000 (A/MQ)</td>
<td>179.3</td>
<td>5.30</td>
<td>24</td>
<td>180.16</td>
<td>25</td>
<td>258</td>
</tr>
</tbody>
</table>

Total (95% CI)

<table>
<thead>
<tr>
<th>Females</th>
<th>Males</th>
<th>Control Females</th>
<th>Control Males</th>
<th>Std. Mean Difference</th>
<th>Std. Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>287</td>
<td>1832</td>
<td>100%</td>
<td></td>
<td>–0.08 (–0.21, 0.05)</td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: Tau\(^2\) = 0.00; Chi\(^2\) = 10.46; df = 11 (P = 0.49); I\(^2\) = 0% Test for overall effect: Z = 1.20 (P = 0.23)
within the specific homogenous cohorts of military recruits (11%), adolescents (11%) and recreational runners (6%).

Isokinetic quadriceps weakness is predictive of future PFP in a military cohort and this is in agreement with the previous review of Lankhorst et al. Unfortunately, the strength of this evidence has not changed, as no new prospective studies using an isokinetic dynamometer to measure quadriceps strength have been published since 2011. New data from two studies investigating isometric quadriceps strengths in military cohorts, not included by Lankhorst et al, demonstrated no significant association with future PFP. While this could be interpreted as conflicting evidence, it could be that isometric muscle testing may not be sensitive enough to identify military recruits at risk of PFP. This limits the clinical applicability of these results, as isometric testing with a HHD is a more accessible tool for clinicians to use when measuring muscle strength.

Quadriceps weakness was not identified as a risk factor for future PFP in an adolescent group. While this further validates the importance of investigating risk factors within homogenous groups, it is also important to consider the implications of these findings in relation to risk modification interventions within differing populations. A similar disparity between adult and adolescent populations has been observed cross-sectionally, with no differences in hip or knee strength between adolescents with PFP and a group of asymptomatic controls matched for both age and sex, but significant strength deficits in adults with PFP compared with control groups. These findings offer indications as to why rehabilitation programmes have been shown to be of significant benefit in adults with PFP but are of only limited additional benefit to education alone in adolescents.

In contrast with the quadriceps data, hip muscle strength, regardless of test direction, was not a risk factor for future PFP in military recruits or recreational runners. However, increased baseline isometric hip abduction strength predicts future PFP in adolescents. Herbst et al make the suggestion that increased hip abduction strength could be the result of greater eccentric hip abductor demands due to higher peak hip abduction during dynamic tasks. When pooling data from both military and adolescent cohorts, dynamic knee valgus angle was also not a risk factor for future PFP. However, Holden et al reported higher knee valgus displacement in adolescent women who develop PFP (mean difference 7.19°). Despite these reported kinematic deficits, hip strength and altered kinematics during dynamic tasks are consistently negatively correlated, contradicting this hypothesis.

A more plausible explanation for the association between increased isometric hip abduction strength and future PFP in young adolescents is a high level of physical activity, common within this age group. The mean age of the Herbst et al cohort is 12.7 years and may therefore have increased lower limb muscle strength as a consequence of high physical activity levels. It may be that lower limb muscle strength correlates with duration of symptoms in adolescents, with strength deficits presenting later in life when symptoms persist and if activity levels subsequently reduce. As a result, it is sensible to question the role of increasing muscle strength in adolescent cohorts.

Figure 5 Weight. Forest plot detailing standardised mean differences for weight when comparing participants who developed patellofemoral pain (PFP) with controls. A, adolescents; HQ, high quality; IV, inverse variance; M, military recruits; MQ, medium quality; R, recreational runners.

Figure 6 Body mass index (BMI). Forrest plot detailing standardised mean differences for BMI when comparing participants who developed patellofemoral pain (PFP) with controls. A, adolescents; HQ, high quality; IV, inverse variance; M, military recruits; MQ, medium quality; R, recreational runners.
as a preventative measure. It is advised that future prospective studies both report and stratify for activity levels when investigating the association between strength variables and future PFP. In 2011, Coppack et al.\(^4\) reported a significant reduction in PFP risk after completion of a quadriceps and gluteal strengthening programme when compared with a non-specific control group. It is surprising that an exercise intervention designed to increase quadriceps strength has not been investigated further, given both level\(^1\) and level\(^2\) evidence identifying quadriceps weakness as a preventative treatment target. While increased baseline hip abduction strength increases the risk for future PFP in adolescents, this is potentially a surrogate indicator of activity level. Given the positive results of education and load management interventions in this population,\(^44\) we suggest prioritising these interventions in future in order to reduce the incidence of PFP in adolescents.

Multiple variables often described as risk factors for future PFP, perhaps due to strong associations in cross sectional studies, were not found to be so in this meta-analysis. Participant height, weight, BMI, body fat percentage, age and knee angle did not predict future PFP in any cohort. The recent systematic review of Hart et al.\(^48\) reports a cross sectional association between high BMI and both PFP and patellofemoral osteoarthritis in adults, again perhaps due to a reduction in activity levels after symptom development.\(^49\) Higher BMI was not a risk factor for future PFP in either adults or adolescents, nor was a high BMI linked to intervention outcomes in participants with PFP.\(^48\) While these data question the biologically plausible suggestion that a high BMI may influence treatment outcomes and this suggestion may not be necessary.

Using data from the work of Boling et al.\(^21\) the previous 2012 review of Lankhorst et al.\(^10\) reported that women are at a higher risk of developing PFP within the military (OR 2.23, 95% CI 1.16, 4.10). Our results are in conflict with this, with pooled data from seven studies\(^21\) identifying no significant links between female sex and PFP development. Pooling data for the identified individual subgroups is also non-significant, but a greater proportion of women developing subsequent PFP was reported in six of the seven studies. The largest proportion of women occurs among the recreational runner subgroup and this is in fact statistically significant when a fixed effects model is used for the meta-analysis, meaning observed results are most likely a result of selection bias in source studies. However, given the low number of studies (n=3) and high heterogeneity, a fixed effects model is inappropriate and increases the chance of sustaining a type I error.\(^39\) Given the absence of a causal association between female sex and future PFP, the frequent bias towards women in trial sampling and the need to control for sex as a confounder may not be necessary.

The heterogeneous incidence of PFP was 10% in this review, demonstrating that PFP affects up to 1 in 10 persons across multiple populations. The recent systematic review of PFP incidence and prevalence by Smith et al.\(^4\) identified a wide range of PFP incidence among military recruits (9.7–571.4 cases per 1000 person-years), which is similar to the incidence range identified by this review (3–43%). The variance is likely explained by the four studies\(^24\) included in this review not included by Smith et al.\(^4\) and the three studies included by Smith et al.\(^4\) not eligible for inclusion within this review.\(^47\)\(^51\) The incidence range for PFP within adolescent cohorts are identical (5–15%), despite two studies from this review\(^22\)\(^29\) not being included by Smith et al.\(^4\).

**Limitations and future research directions**

This review is not without limitations, which must be considered when interpreting the results. There is currently no accepted method for determining study methodological quality or ascertaining risk of bias. While the NOS is advocated by the Cochrane Group, it is possible that using a different quality appraisal tool may have yielded different levels of eventual evidence. It should also be considered that the NOS does not have a component pertaining to the...
Figure 9  Q angle. Forrest plot detailing standardised mean differences for Q angle when comparing participants who developed patellofemoral pain (PFP) with controls. HQ, high quality; IV, inverse variance; M, military recruits; MQ, medium quality; R, recreational runners.

Figure 10  (A–F) Quadriceps strength. Forrest plot detailing standardised mean differences for quadriceps strength when comparing military recruits who developed patellofemoral pain (PFP) with controls. HQ, high quality; IV, inverse variance; M, military recruits.

reliability of exposure data collection, focusing more on the validity of outcome data. As per the PRISMA guidelines,15 three databases were searched, but it is also possible that increasing the number of databases searched may have yielded additional studies for inclusion. An attempt was made to mitigate this risk by completing a citing reference search in Google Scholar in addition to hand searching the reference lists of included studies. It must be stressed that
incidence data have been calculated only from included studies, and the addition of other epidemiology studies that do not fit the inclusion criteria of this review would have affected the values reported. It was also not possible to express incidence data relative to a timeframe given the high heterogeneity observed between included studies.

Some included studies provided data that were not suitable for inclusion in a meta-analysis (ie, no mean/SD or raw counts) and efforts to obtain raw data directly from study authors were unsuccessful. Despite the addition of 11 new studies, the ability to pool data was limited, which is partly attributable to the 116 individual variables investigated across the 18 included studies that could not be pooled. A total of eight studies26 29–32 34 35 38 failed to adhere to the rule of 1017 (ie, ensuring a minimum of 10 PFP events for each variable of interest), resulting in a high risk of bias and reduced methodological quality.

Given the lack of associations identified by this review (pooled data or otherwise), it is sensible to suggest that perhaps the current body of research is not placing appropriate focus on variables of interest. Altered hip and knee kinematics during running are known to have moderate to strong cross-sectional associations with PFP,13 yet there remains just one prospective investigation of these variables in women runners.32 There is also an emerging evidence base surrounding the association between psychological variables and PFP, with levels of anxiety, depression, catastrophising and fear of movement reported to be elevated in persons with PFP by a recent systematic review.52

The prospective studies included within this review have sought to detect an association between single variables and risk of PFP. The inherent limitation of this approach is the inability to consider interactions between multiple variables. Consequently, Given the lack of associations identified by this review (pooled data or otherwise), it is sensible to suggest that perhaps the current body of research is not placing appropriate focus on variables of interest. Altered hip and knee kinematics during running are known to have moderate to strong cross-sectional associations with PFP,13 yet there remains just one prospective investigation of these variables in women runners.32 There is also an emerging evidence base surrounding the association between psychological variables and PFP, with levels of anxiety, depression, catastrophising and fear of movement reported to be elevated in persons with PFP by a recent systematic review.52

The prospective studies included within this review have sought to detect an association between single variables and risk of PFP. The inherent limitation of this approach is the inability to consider interactions between multiple variables. Consequently,
Figure 12  (A, B) Hip abduction strength. Forrest plot detailing standardised mean differences for hip abduction strength when comparing participants who developed patellofemoral pain (PFP) with controls. A, adolescents; HQ, high quality; IV, inverse variance; M, military recruits; MQ, medium quality; R, recreational runners.

Figure 13  Potential causal inference diagram for patellofemoral pain (PFP). BMI, body mass index.

research needs to move towards a complex systems approach to better understand injury aetiology.53 Rather than endeavouring to identify a singular causal factor, studies should be designed to investigate the interactions between a ‘web of determinants’ that are likely to be non-linear in nature.54 This approach has significant methodological challenges and requires the use of a statistical learning approach such as a Bayesian network.55 Examples of variables that could fit into a web of determinants for PFP from the published literature include muscle strength (quadriceps and gluteal), hip/knee kinematics, activity levels/sporting workload and psychosocial measures (see figure 13).

No variable included within this systematic review identified a link with future PFP in recreational runners. High peak hip adduction is known to be associated with future PFP in female runners,32 and future studies should further explore the causal associations between lower limb kinematics and PFP. While not presented in a fashion that allowed for data pooling, an increased eccentric hip abduction strength reduces the risk of future PFP in recreational runners.33 The distinct limitation of this study design is that no guidance was given to the included runners regarding training frequency or intensity, which is likely to be a significant confounder, as more aggressive run volume progressions increase the risk of injury development.36
CONCLUSION

Quadriiceps weakness, measured using an isokinetic dynamometer and whether or not normalised to either bodyweight or BMI, is a risk factor for future PFP in military recruits and should be investigated as a preventative strategy in a future randomised controlled trial. While increased hip abduction strength is a risk factor for future PFP in adolescents, this may simply be a composite of activity level. Overall, our understanding of what contributes to the development of PFP is inadequate and requires further scientific exploration, although the relationship between given variables and PFP risk is likely to be both complex and individual.

What are the new findings

- Heterogeneous incidence of patellofemoral pain (PFP) is 10%.
- Higher baseline hip abduction strength predicts future PFP development in adolescents.
- Multiple variables (including sex, body mass index and Q angle) were not found to be risk factors for future PFP development.

Contributors

ADR and BSN completed the search. BSN and NEL determined eligible papers for inclusion. SDL and NEL completed the quality appraisal of included papers. BSN and ADR completed the data extraction. BSN, SDL and DHH completed the meta-analysis. All authors contributed to the writing of the final manuscript.

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

None declared.

Patient consent

Not required.

Provenance and peer review

Not commissioned; externally peer reviewed.

REFERENCES


