Introduction and background

The current best evidence-based treatment method for patellofemoral pain (PFP) is multimodal and may include a mix of exercise therapy, patellar taping and bracing, foot orthoses and surgery (1). However, what constitutes multimodal therapy is not the same across research studies and it is often applied inconsistently in clinical practice (1,2). High quality empirical studies (3,4) confirm that a multimodal approach confers some benefits to patients such as improved pain, function and quality of life, in the short term. However, there is limited evidence to support the longer-term outcomes of a multimodal treatment approach (5-8). In view of the limited benefit and lack of evidence of the long-term success of the multimodal approach, support for the idea of subgrouping patients with PFP has grown in recent years, especially as this approach has proved effective for optimising management in other musculoskeletal conditions, such as, low back pain (9,10).

Strong support for the idea of clinically subgrouping PFP patients and delivering targeted treatment was gained at the First International PFP Research Retreat (11); this was reinforced at the 2nd and 3rd International PFP Research Retreats (12,13), where it was stated that:

“Identification of subgroups remains the ‘holy grail’ for PFP research.” (13)

The concept of identifying subgroups within the PFP population is not actually that new, however, methodological approaches to subgroup identification have advanced considerably. Holmes and Clancy in 1998 (p. 299) (14),
when discussing the management of PFP patients, argue that:

“an adequate classification system should aid in proper diagnosis and treatment of specific problems. If properly devised, it should also aid in the comparison of results between different treatment centres. In addition, it should be a system that is simple and useful in the clinical setting with minimal use of complicated imaging techniques.”

This is a statement with which we wholeheartedly agree, however, it can be seen (Table 1) that early attempts at producing such a system, including that by Holmes and Clancy (1998) (14) themselves, resulted in increasingly complex, multi-layered clinical frameworks.

Merchant (15) proposed a classification system of patellofemoral disorders based on aetiology with five major groups: trauma; patellofemoral dysplasia; idiopathic chondromalacia patellae; osteochondritis dissecans; synovial plicae (Table 1). Thirty-eight subgroups were then described underneath each of these specific pathological conditions. Wilk et al. (16) divided patellofemoral disorders into eight major groups: patellar compression syndromes; patellar instability; biomechanical dysfunction; direct patellar trauma; soft tissue lesions; overuse syndromes; osteochondritis diseases; neurologic disorders (Table 1). Some of these were further subdivided to generate 26 subgroups in total. Treatment suggestions for each of the eight major patellofemoral dysfunction categories were then briefly discussed. Holmes and Clancy (14) suggested that from a clinical perspective, PFP in the skeletally mature adult falls into three broad categories: (I) patellofemoral instability (19 subgroups); (II) PFP with malalignment but no episodes of instability (11 subgroups); and (III) PFP without malalignment (30 subgroups) (Table 1). In total, this sub-classification system yielded 60 PFP subgroups.

In contrast to the increasing complexity of these frameworks, Post (17) presented a paper on the clinical evaluation of patellofemoral disorders which summarised a number of discussions held by the International Patellofemoral Study Group (IPSG). In this paper, a simple two-layer hierarchy was used to initially categorise PFP as “Unstable”, requiring surgical intervention or “Stable”. “Stable” PFP was then categorised into just three subgroups: extremity alignment; soft tissue mobility/flexibility; dynamic control (Table 1). No specific threshold data for allocation to a subgroup was reported, however, treatment advice based on expert opinion for each of these subgroups was presented. Witvrouw et al. (18) presented subgroups based on a consensus of expert opinion reached by the European Rehabilitation Panel (Table 1). Despite moving

<table>
<thead>
<tr>
<th>Author/s (date)</th>
<th>Subgroups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Merchant (1988) (15) (38 subgroups)</td>
<td>Trauma</td>
</tr>
<tr>
<td>Wilk et al. (1998) (16) (26 subgroups)</td>
<td>Patellar compression syndromes</td>
</tr>
<tr>
<td>Post (1999) (17) (4 subgroups)</td>
<td>Unstable (surgery required)</td>
</tr>
<tr>
<td>Witvrouw et al. (2005) (18) (13 subgroups)</td>
<td>Malalignment</td>
</tr>
<tr>
<td></td>
<td>Flexibility</td>
</tr>
</tbody>
</table>

Table 1 Continued
Table 1 (continued)

<table>
<thead>
<tr>
<th>Author/s (date)</th>
<th>Subgroups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selhorst et al. (2015) (19) (4 subgroups)</td>
<td>Elevated fear avoidance</td>
</tr>
<tr>
<td></td>
<td>Decreased muscle flexibility</td>
</tr>
<tr>
<td></td>
<td>Functional malalignment</td>
</tr>
<tr>
<td></td>
<td>Decreased muscle strength</td>
</tr>
<tr>
<td>Keays et al. (2015) (20) (4 subgroups)</td>
<td>Hypermobility</td>
</tr>
<tr>
<td></td>
<td>Hypomobility</td>
</tr>
<tr>
<td></td>
<td>Faulty movement pattern</td>
</tr>
<tr>
<td></td>
<td>Osteoarthritis</td>
</tr>
<tr>
<td>Selfe et al. (2016) (21) (3 subgroups)</td>
<td>Strong</td>
</tr>
<tr>
<td></td>
<td>Weak and tight</td>
</tr>
<tr>
<td></td>
<td>Weak and pronated</td>
</tr>
<tr>
<td>Drew et al. (2018) (22) (4 subgroups)</td>
<td>Strong</td>
</tr>
<tr>
<td></td>
<td>Pronation and malalignment</td>
</tr>
<tr>
<td></td>
<td>Weak</td>
</tr>
<tr>
<td></td>
<td>Flexible</td>
</tr>
</tbody>
</table>

back towards increasing complexity, this paper represents a bridge with the more recent efforts to understand subgroups. There were some attempts to define threshold data to guide subgroup allocation and evidence-based treatment recommendations for each of the subgroups were presented. However, the proposed thresholds were based on clinical observation and review of the literature rather than being statistically derived. Other studies have investigated subgroups within the PFP population using specialised high cost equipment not routinely seen in clinical scenarios, e.g., radiographic examination and scintigraphy (23), dynamic magnetic resonance imaging (MRI) (24,25), and six camera three-dimensional motion analysis systems (26). Translation of the results of these types of studies, using complex equipment, into routine clinical practice has been extremely limited. More recently, Selhorst et al. (19) reported on a pilot study of 21 paediatric patients with a mean age of 14 years old and Keays et al. (20) reported on a study of 41 patients that had a very wide age range from 13 to 82 years, with only eight patients in the young adult (20–40 years) age range. Interestingly, both papers described four subgroups of PFP patients, which appear to partially overlap (Table 1).

Few studies in PFP have had a hypothesis-driven approach initially using data to identify clinically important subgroups and then going on to explore the prognostic effect attributed to subgroup membership (27). Selfe et al. (21) and Drew et al. (22) are exceptions to this and both studies have based their approaches on rigorous statistical methods. In the case of Selfe et al. (21), this has led to the development of a robust simple hierarchical algorithm. This algorithm uses objective data generated by low cost clinical assessment tests to categorise patients into 1 of 3 subgroups. Due to the low-cost nature of the clinical assessment tests employed, this approach has high clinical utility. This makes it potentially viable for widespread future roll out into primary care and physiotherapy clinics, both in the UK and internationally, and conforms to the views of Holmes and Clancy (14) discussed earlier. Drew et al. (22) have developed this further and combined known imaging features with other clinical features to explore subgroups using established modifiable clinical, biomechanical and imaging features. The justification for and approaches used by both these studies to derive subgroups are discussed in more detail in the subsequent sections of this review.

The Post (17) and the four more recent papers (19-22) describe just three or four subgroups which significantly improves their clinical utility. Interestingly although they employ differing methodologies and include slightly different populations there are some notable areas of overlap in the proposed subgroups, with all five papers identifying a tight/hypomobile subgroup. Three papers describe separate subgroups where there is (I) decreased strength (19,21,22) or (II) decreased dynamic control/faulty movement patterns (17,19,20). Two papers describe separate subgroups that are (I) strong; or (II) have increased pronation (21,22).

Recent frameworks for subgrouping studies demonstrate why many of the attempts to subgroup patients in PFP have not translated well into clinical practice. The PROGRESS partnership provides some broad recommendations and the Medical Research Council provides a framework on development, design and analysis in stratification research (28,29). Both suggest a similar pathway from an initial hypothesis setting stage, which defines the problem and population. This then progresses to identifying the variables to define subgroups, understanding of the properties of the test, through to studies to identify the subgroups. Once subgroups are identified, this is followed by verification and validation and then robust evaluation of the effectiveness of subgrouping on outcome in clinical practice. As shown above, few previous studies have been hypothesis and data driven and, as yet, those that have applied this approach are not mature enough along the pathway to be tested in practice (21,22).

While these frameworks outline considerations for
the researcher at each stage, they do not provide clear recommendations on which statistical approach, for model prediction to identify subgroups is best in different circumstances (30). This is clearly an issue for the researcher when different methods can give different results with the same dataset and, thus, may identify different subgroups. This problem is compounded in this field as some techniques, such as regression methods, require large datasets, which are uncommon in PFP research. Furthermore, their outputs may be difficult to interpret clinically, particularly, for “theragnostic” markers (29), i.e., those that aim to identify which patients will respond to different treatments. Latent profile analysis approaches are increasingly considered as better analytically than more traditional hierarchical clustering models (31). This is because they are based directly on the distributional properties of the relevant variables. However, hierarchical models reflect better the clinical decision-making process around which treatments to choose for which patients.

An important issue stressed in both the PROGRESS recommendations and the MRC framework is the consideration throughout development, design and analysis of the clinical relevance and appropriateness of the marker, especially if the purpose of the identification of subgroups is to optimise current treatment (28,29). Researchers need to ensure early and continuing consideration of the feasibility and acceptability of implementing both the test and the treatment for both patients and for health professionals. This might help direct the choice of tests, number of subgroups to identify, analytical approaches, thresholds for allocation of patients to subgroups and evaluation methods and outcomes.

**Subgroup derived targeted intervention for patellofemoral pain (TIPPs)**

The TIPPs programme of work (21,32) has to date consisted of three phases in order to identify and validate potential subgroups within the PFP population using readily available, low cost, easy to use tools found in clinical practice:

(I) Literature search to identify appropriate low-cost clinical assessments, linked to reported thresholds to identify clinically relevant subgroups; mapped to credible evidence-based treatment interventions;

(II) Feasibility study to investigate if these assessments could be performed in routine clinical practice, if they could identify clinically relevant subgroups and what the optimum test thresholds for subgroup allocation might be within a UK population;

(III) Validation study of the subgroups in a Turkish population using the same assessment protocol.

Future work will aim to identify if these subgroups of patients with PFP respond better to specifically targeted exercises compared to best evidence usual care.

Phase 1 included an in-depth literature search to identify assessments that were, or could be, used in clinical practice and that had the potential to identify possible subgroups. One of the key documents guiding this phase of our work was the First International PFP Research Retreat (11). This consensus proposed three subgroups based on the anatomical region thought to be responsible for the problem, i.e., proximal, local and distal. In order to facilitate implementation into clinical settings, assessments were deemed appropriate when they were: based on evidence of diagnostic performance; applicable to be used in a wide range of clinical settings; easy to learn and administer; free to use or available at a low cost; linked to reported thresholds; linked to a credible evidence-based treatment intervention. Through this literature review, seven assessments were identified, which were all applied in the next phase of the programme (Table 2).

In the phase 2 feasibility study, four National Health Service (NHS) physiotherapy clinics, serving the general population, in the UK recruited 130 people with PFP over a one-and-a-half-year period. This was to investigate if the assessments could be performed in routine clinical practice, if they could identify clinically relevant subgroups and to establish what the optimum test thresholds for subgroup allocation might be within a UK population. Participants were between 18 and 40 years old, experienced uni- or bilateral PFP for at least 3 months, and had not yet started physiotherapy treatment. Additional study details and eligibility criteria are presented in (21,32).

Participants completed demographic, clinical, and psychosocial questionnaires related to aspects of PFP and were clinically assessed using the seven tests. Baseline demographics, such as, gender distribution and age, were in line with those reported by others (4,37). This study also identified an average wait time of almost 4 years for people with PFP symptoms before they consulted a physiotherapist. A causal pathway diagram, based on the broader literature review, specific consensus documents and expert opinion around the proximal, local and distal subgroups was drawn up to inform the analytical approach. Both hierarchical agglomerative cluster analysis and latent
profile analysis were used to explore the existence of subgroups within the sample. Surprisingly, the Hamstrings length test mean scores (36) were similar across all three subgroups identified by preliminary analyses and was excluded from further analysis. Three subgroups were found: “weak and tight” (39% of participants), “weak and pronated” (39%), and “strong” (22%). The two largest subgroups were both classified as having weak quadriceps and hip abductor muscles; these subgroups might benefit from strengthening exercises. In addition to being weak, the people with PFP in the “weak and pronated” subgroup had a significantly higher mean Foot Posture Index (FPI) than the other two groups. This weak and pronated subgroup had a FPI with a mean greater than 6; we would set this as the threshold for subgroup allocation as a FPI of 6 or more is clinically relevant for treatment needs. Therefore, in addition to strengthening exercises prescribing a correcting foot orthotic to people meeting this criteria in the “weak and pronated group” might also be beneficial.

Both weak subgroups were consistent with current treatment practices for PFP (1,2). The third identified subgroup (“strong”) is a novel previously unrecognised group that falls outside the current treatment recommendations as no weakness in strength or shortening in muscle length was identified. The people in this subgroup also experienced higher levels of function and quality-of-life. It is currently our hypothesis that this group is overloading their patellofemoral joint due to reduced motor control therefore perhaps proprioceptive training is the answer for improving their PFP. There is evidence to suggest motor control of the quadriceps may be problematic in some PFP patients (4,18,38,39), therefore neuromuscular retraining could be the focus of any rehabilitation strategy rather than strengthening exercises for this subgroup of patients. Recently Greuel et al. (40) have independently confirmed the existence of a strong group of PFP patients. They reported that there were no differences in strength between healthy subjects and a strong group of PFP patients. However, they reported an increased level of muscle inhibition in the strong PFP patients, suggestive of a motor control problem. The efficacy of these proposed treatments need to be demonstrated in a randomised controlled trial (RCT).

Phase 3 aimed to validate the findings of phase 2 in a different population. In Turkey, an identical TIPPs study was set up exploring subgroups in the PFP population. Forty-six participants took part and underwent the six assessments, which were demonstrated as useful in phase 2 of the UK study. Publication of the findings of this study will strengthen the evidence for the three subgroups. An interesting consideration is the potential for different distributions of subgroups in different populations; this raises the possibility of environmental or genetic influencing factors in some subgroups and/or different norms, which might have an impact on subgroup thresholds.

Currently, it is still unknown if targeted treatment of the three subgroups will lead to improved patient outcomes. The TIPPs programme of research presented here used a one-off assessment and therefore no outcome data are available. Future research should investigate the prognostic implications of these subgroups and establish the level of efficacy of more targeted intervention.

### PFP subgroups derived from imaging

There have been numerous attempts at classifying and subgrouping PFP using imaging (41). However, no consensus exists on which imaging modalities should be used or which patellofemoral joint features are associated with PFP compared to asymptomatic individuals (42). There is, in addition, in the UK a pressure to reduce imaging in clinical practice due to

---

**Table 2 Seven assessments mapped to the appropriate evidence-based treatment option**

<table>
<thead>
<tr>
<th>Assessments</th>
<th>Treatment intervention option</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hand held dynamometry for hip abductor strength (Nm/kg) (33)</td>
<td>Hip abductor strengthening</td>
</tr>
<tr>
<td>Hand held dynamometry for quadriceps strength (Nm/kg) (33)</td>
<td>Quadriceps strengthening</td>
</tr>
<tr>
<td>Medial-lateral patellar mobility test (mm) (34)</td>
<td>Patella stabilisation or mobilisation</td>
</tr>
<tr>
<td>Foot Posture Index (FPI) (35)</td>
<td>Foot orthotics</td>
</tr>
<tr>
<td>Rectus femoris length test (degrees) (34)</td>
<td>Muscle stretching</td>
</tr>
<tr>
<td>Hamstrings length test (degrees) (36)</td>
<td>Muscle stretching</td>
</tr>
<tr>
<td>Gastrocnemius length test (degrees) (34)</td>
<td>Muscle stretching</td>
</tr>
</tbody>
</table>
resource constraints, this risks the biological component of the biopsychosocial model being overlooked and potentially stifles research that can improve our understanding of how local joint pathology influences clinical presentation (43). A recent systematic review has demonstrated that a number of MRI features are associated with PFP, i.e., MRI bisect offset and CT congruence angle analysed at 0 knee flexion and 15 knee flexion respectively (42).

Early imaging research into subgroups was based on observed pathological changes within the patella with no demonstrable link to clinical symptoms or their potential to be modified through clinical interventions. One of the earliest examples of this type of classification is the five sub-types of chondromalacia patellae identified by Ficat et al. (44): lateral facet chondromalacia, medial facet chondromalacia, central chondromalacia, bipolar chondromalacia and total chondromalacia.

The term malalignment can be misleading as some authors use this term to describe differences among both static and dynamic observations (45). For the sake of clarity, here malalignment will be used to describe a static observation and maltracking will refer to dynamic assessment. Sheehan et al. (24) classified their PFP group into maltrackers and non-maltrackers by classifying all non-maltrackers with a patellofemoral lateral-medial displacement of ≥0.45 mm and a patellofemoral varus angle slope ≤0.25 mm/°. Using discriminatory analysis this maltracking criteria yielded a 90% agreement reinforcing the existence of these two sub-groups (24). Employing the same maltracking criteria, Harbaugh et al. (25) explored the relationship of these maltracking and non-maltracking groups with quantifiable femoral and patella shape. They showed that compared to non PFP maltrackers, the maltracking subgroup showed a 20% smaller lateral trochlear inclination (LTI) (25). Linking these subgroups with femoral shape proposes an anatomical explanation for the observed differences in subgroups with the increased LTI in the non-lateral maltrackers acting as an osseous constraint to lateral displacement (25). This idea is supported by an in vitro study which showed using a simulated trochleoplasty (and increasing the LTI) that lateral patella displacement is reduced by ~2.5 mm (46). It is worth noting that these studies selected patients based on PFP plus the presence of at least one maltracking sign including large static Q-angle, positive apprehension test, positive J-sign or clinical lateral patella hypermobility (24, 25), which may affect the generalisability of these findings when compared to a typical group of individuals with PFP. The MRI scans were also acquired in non-weight bearing.

Evolving these imaging subgroups concept further, a series of papers by Pal and colleagues (47-49) explored this idea of maltracking using full weight bearing MRI. In contrast to the previous studies, they classified their maltracking PFP subgroup as being greater than the 75th percentile of a non-Gaussian two-parameter Weibull distribution model. Using gender-specific thresholds they showed that compared to a non-maltracking PFP group, a maltracking subgroup is significantly associated with a delay in vastus medialis (VM) activation during the normal gait cycle (R2 =0.89) and an increased patella height when measured using both the Caton-Deschamps and Blackburne-Peel techniques.

The growing support for identifying PFP subgroups using imaging and increased understanding of how these groups link to other clinical features, such as, VM activation, offers potential treatment strategies moving forwards. Patella tilt has been shown to be modifiable with patella bracing (50) and patellofemoral bisect offset/lateral displacement modifiable with both patella bracing (50) and patella taping (51). Recently, expanding on these efforts to combine known imaging features with other clinical features, Drew et al. (22) explored subgroups using established modifiable clinical (hamstring length, quadriceps length, gastrocnemius length and foot posture), biomechanical (knee extension strength, hip abduction strength, peak knee flexion angle and peak hip internal rotation angle) and imaging features (MRI bisect offset and MRI patella tilt). They identified “Strong”, “Pronation & Malalignment”, “Weak” and “Flexible” subgroups. Furthermore, the natural prognosis of these subgroups was established. By adjusting for known covariates, they showed, compared to a “Strong” subgroup, a substantive directional trend that the “Weak” subgroup was the least likely [32% (7/22); odds ratio (OR): 0.30; 95% CI, 0.07–1.36] and the “Flexible” subgroup most likely [64% (7/11); OR 1.24; 95% CI, 0.20–7.51] to report a favourable outcome at 12 months follow-up.

**Conclusions**

There have been many attempts at defining subgroups within the PFP population over the years. Post (17); Selhorst et al. (19); Keays et al. (20); Selfe et al. (21); Drew et al. (22); using quite different approaches, describe just three or four subgroups with some notable areas of overlap,
all five papers refer to a tight/hypomobile subgroup. Three papers describe separate subgroups where there is (I) decreased strength; or (II) decreased dynamic control/faulty movement patterns. Two papers describe separate subgroups that are (I) strong; or (II) have increased pronation. Although yet to reach a consensus on the optimal approach, the development of robust frameworks to guide stratification research, sophisticated statistical modelling techniques and the drive towards personalised medicine have stimulated new efforts in subgrouping research for PFP, which is gathering momentum. Our experience has highlighted some of the challenges and opportunities in undertaking such subgrouping research in PFP. One is small sample size, which precludes many of the more complex, statistical methods for classifying subgroups and/or optimising thresholds. In the Selfe et al. (21) study, it also precluded cross-validation studies for internal verification requiring reliance on using two different statistical methods instead. Given sample size is a difficulty in many PFP studies, consideration should be given to establishing large prospective datasets, which may require collaboration across institutions and countries. Such an initiative requires a core dataset of putative markers, such as the tests above, but also others for which there may be emerging evidence of their prognostic impact, e.g., psychosocial factors (52) and a core set of outcome measures. While progress is being made on the latter with the development of the KOOS-PF (53) there remains a bewildering variety of different tests used to measure the same clinical phenomenon; some are more practical to use than others. Finally, we also need carefully collected normative data on key measures to allow for appropriate interpretation of comparative test data in PFP patients. To date no definitive RCTs have been conducted to evaluate the potential benefits of targeted interventions for PFP subgroups in terms of improved patient outcomes so this warrants further research (21,54).

Acknowledgements

The authors would like to thank all the patients who kindly volunteered to take part in these studies.

The authors thank the wider TIPPs team and specifically the following physiotherapists for performing the research assessments Steve Hill, Stephen Kirk, Gary McCall, Christine Dewsbury, Kim Patterson and Sophie Chatwin. The authors would also like to thank the following service managers for their support Keith Mills, Elaine Nicholls, Barbara Sharp, Chantel Ostler and Kim Patterson. The authors would also like to thank Brian Francis for his advice on latent profile analysis and its application. The TIPPs team acknowledge the support of the National Institute for Health Research, through the Comprehensive Clinical Research Network.

This work was supported by National Institute for Health Research (NIHR) Clinical Doctoral Research Fellowship (CDRF -2013-04-044); Arthritis Research UK (grant number 19950) and involves collaboration with the Arthritis Research UK Centre for Sport, Exercise and Osteoarthritis. Arthritis Research UK Musculoskeletal Pain CSG, also funded a Think Tank meeting.

This work was also supported in part by funding from the Arthritis Research UK Experimental Osteoarthritis Treatment Centre (Ref 20083) and the Arthritis Research UK Centre for Sport, Exercise and Osteoarthritis (Ref 20194).

Footnote

Conflicts of Interest: The author has no conflicts of interest to declare.

References

2004;34:504-10.