An exploration of normative values in New Zealand to inform the Targeted Interventions for Patellofemoral Pain approach

Original research using quantitative data

Kim Hébert-Losier[a]*, Ivana Hanzlíková[a], Sana Oladi Ghadikolaei[a], Jessie Janssen[b], James Selfe[c], Jim Richards[d]

[a] Division of Health, Engineering, Computing and Science
Te Huataki Waiora School of Health
University of Waikato, New Zealand

[b] Department of Health Sciences
Institute Therapeutic Sciences
IMC University of Applied Sciences Krems, Austria

[c] Department of Health Professions
Faculty of Health, Psychology and Social Care
Manchester Metropolitan University, United Kingdom

[d] Allied Health Research unit,
Faculty of Allied Health and Well-being
University of Central Lancashire, United Kingdom

*Corresponding author: Kim Hébert-Losier

E-mail: kim.hebert-losier@waikato.ac.nz

ORCiD: https://orcid.org/0000-0003-1087-4986

Twitter: @KimHebertLosier

ResearcherID: P-6671-2016

Research Gate: https://www.researchgate.net/profile/Kim_Hebert-Losier

LinkedIn: https://www.linkedin.com/in/kim-hébert-losier-854794ab/

University of Waikato Adams Centre for High Performance

52 Miro Street, Mount Maunganui

3116 Tauranga, New Zealand

Telephone: +64 7 837 9476

Fax: +64 7 838 4504

Ivana Hanzlíková

E-mail: ih27@students.waikato.ac.nz

ORCiD: https://orcid.org/0000-0002-2259-9312

Twitter: @HanzlikovaIvana

Sana Oladi Ghadikolaei

E-mail: ro59@students.waikato.ac.nz

ORCiD: https://orcid.org/0000-0003-3844-3838

Twitter: @OladiSana

Jessie Janssen
E-mail: jessie.janssen@fh-krems.ac.at
ORCiD: https://orcid.org/0000-0002-5961-2736
Twitter: @DrJessieJanssen
LinkedIn: https://www.linkedin.com/in/jessie-janssen-2469909/

James Selfe
E-mail: j.selfe@mmu.ac.uk
ORCiD: https://orcid.org/0000-0001-9931-4998

Jim Richards
E-mail: jrichards@uclan.ac.uk
ORCiD: https://orcid.org/0000-0002-4004-3115
Twitter: @ProfJimRichards

Word count: 3708 words

ABSTRACT

Background: The Targeted Interventions for Patellofemoral Pain studies (TIPPs) have identified three subgroups exist in UK and Turkish patellofemoral pain (PFP) populations: Strong; Weak and Tight; and Weak and Pronated, based on six clinical assessments. The thresholds used to develop the subgrouping algorithms were based on normative values sourced from various populations and countries.

Objectives: Explore normative scores from the clinical assessments in a singular non-PFP population whilst considering potential differences between ethnicities and sex (primary aim). Revisit inter-rater reliability of each assessment (secondary aim).

Design: Cross-sectional and test-retest.

Method: The six assessments; rectus femoris length, gastrocnemius length, patellar mobility, hip abductor strength, quadriceps strength, and Foot Posture index (FPI) were measured in 89 New Zealanders (34% Māori, 45% female). Two raters independently assessed 17 participants to examine inter-rater reliability.

Results: Significant interactions between ethnic group and sex were noted for rectus femoris length and patella mobility. Māori versus European males exhibited greater rectus femoris tightness (p = 0.001). Māori versus European females demonstrated greater patellar mobility (p = 0.002). Females were significantly weaker than males in normalised strength measures (p < 0.001), and had lower FPIs. Mean differences between testers for all measures were small and not significant, except for FPI which had a 2.0 point median difference (*p* = 0.021).

Conclusions: Our results indicate that sex is an important factor worth considering within the TIPPs subgrouping approach, more than ethnicity, especially for the normalised strength measures. The sub-optimal reliability of FPI warrant reconsideration of its inclusion within TIPPs.

Keywords: ethnicity, knee, physiotherapy, normative data.

HIGHLIGHTS

- The Targeted Interventions for Patellofemoral Pain algorithm uses 6 clinical tests
- The algorithm identifies three patellofemoral pain (PFP) subgroups
- There was no main effect of ethnicity (Māori vs NZ European) on clinical scores
- Sex differences in the 2 strength tests warrant a sex-based stratification approach
- The suboptimal Foot Posture Index reliability queries its inclusion in the algorithm

INTRODUCTION

Patellofemoral pain (PFP) is a challenging clinical condition with a wide variety of theories postulated to explain casual mechanisms (Janssen 2017). Despite the publication of clinical practice guidelines (Willy et al. 2019) and consensus statements (Collins et al. 2018), no standardised treatment approach for PFP existed. Given that it takes on average 17 years for research to be translated into practice (Morris, Wooding, and Grant 2011), current physiotherapy management of PFP varies and frequently relies on a trial-and-error multimodal approach that includes exercise therapy, patellar taping, bracing, and foot orthoses (Smith et al. 2017; Collins et al. 2018). An international consensus has highlighted the need for musculoskeletal studies to adopt subgrouping approaches to improve our understanding of the underlying mechanisms and optimise patient management (Foster et al. 2009). This consensus highlighted that the heterogeneity of patient samples in previous studies has led to findings of small or no treatment effect, with inferences that non-pharmacological interventions in musculoskeletal conditions might lead to little patient benefit. Indeed, the majority of randomised control trials indicate no clinical benefit of conservative treatment over placebo, sham, or other approaches in reducing PFP symptomatology (Saltychev et al. 2018). However, such heterogeneity may be masking a range of individual responses and true treatment effects.

Subgrouping approaches have proved effective for optimising management in other musculoskeletal conditions, such as low back pain (Brennan et al. 2006; Hill et al. 2011), and there is growing agreement from academics and clinicians on the potential benefit of delivering tailored interventions to improve PFP outcomes (Lack et al. 2018). A recent review summarises the range of PFP subgrouping approaches, and highlights that no consensus yet exists on the number or best classification approach to PFP subgroups (Selfe et al. 2018). The majority of classification systems are of limited use because they do not include clear diagnostic criteria for each subgroup, or they rely on imaging or surgical findings that may not always be available to clinicians (Willy et al. 2019).

Our research team commenced the **T**argeted **I**ntervention for **P**atellofemoral **P**ain **s**tudies (TIPPs) programme to investigate potential PFP subgroups that were: (1) identified by simple evidence-based clinical tests; (2) based on tests that clinicians could use routinely in a variety of settings, e.g., from primary care facilities to teaching hospitals; (3) based on assessments that required minimal expertise and training for competent performance, and involved no or low-cost equipment; (4) based on published thresholds for potentially important factors that could be used to assign patients to specific subgroups; and (5) matched to a specific and credible treatment intervention for each identified subgroup (Selfe et al. 2013). The thresholds used to develop the TIPPs algorithms (Selfe et al. 2013) were based on published normative values sourced from the literature, and were hence derived from a range of studies and populations from various countries (Maffiuletti 2010; Witvrouw et al. 2000; Redmond, Crane, and Menz 2008; Herrington, Malloy, and Richards 2005; Youdas et al. 2005). Normative data derive from a reference population and establish a baseline distribution for a score or measurement, and against which the score or measurement can be compared (Campbell 2013).

From this work, three hypothesis-driven subgroups emerged: (1) Strong; (2) Weak and Tight; and (3) Weak and Pronated (Selfe et al. 2016). Continued work in this area has led to the development of a subgrouping algorithm based on objective data generated by six low-cost clinical tests to categorise PFP patients into one of these three subgroups. The algorithm and subgroup allocation are delivered to clinicians via the Appatella™ mobile application (DigitalLabs@MMU 2019). This approach to PFP subgrouping suggests that it is ecologically valid as it has high clinical value due to its low-cost and accessibility, making it potentially viable for widespread implementation in primary care and physiotherapy clinics internationally.

The Prognosis Research Strategy (PROGRESS) partnership (Hingorani et al. 2013a) and the Medical Research Council (2018) provide recommendations on development, design, and analysis in stratification research. Both frameworks suggest an initial hypothesis setting stage, which defines the problem and population. Thereafter, recommendations are to progress to identifying the variables that define subgroups and then gain an understanding of the properties of the tests. Researchers are encouraged to continue considering implementation of tests from both a patient and health professional perspective. These considerations help direct the choice of tests, number of subgroups, analytical approaches, and thresholds for patient subgroup allocation. Approaches based on targeted interventions, such as the TIPPs framework, require thorough evaluation and exploration of the potential mechanisms and biological reasons underpinning individualised treatment responses (Hingorani et al. 2013b).

In line with these recommendations, we embarked on the present study to understand further the properties of the clinical assessments used in the TIPPs algorithm, and notably to verify the normative values in a given population. Our primary aim was to explore the clinical assessment scores from a non-PFP population outside of the United Kingdom (Selfe et al. 2016) and Turkey (Yosmaoğlu et al. 2020), i.e., countries where previous work on TIPPs has been undertaken in PFP populations, whilst considering ethnicity and sex. Since the original TIPPs thresholds stem from various normative population groups, this exploration may help refine the approach by verifying the normative values of all the clinical assessments included in the TIPPs within a singular population of non-injured individuals. Given that reliability of measures depends on sample characteristics (Matheson 2019), our secondary aim was to revisit the inter-rater reliability of the individual clinical assessments of the TIPPs algorithm within this singular cohort of individuals.

MATERIALS AND METHODS

Participants

Normative data characterise what is typical in a defined population. In the context of PFP, young active adults are a suitable population to define normative values for TIPPs benchmarking considering that PFP typically manifests in young physically active populations (Smith et al. 2018). Previous TIPPs research has involved 95 to 127 individuals (Yosmaoğlu et al. 2020; Selfe et al. 2016), with normative reference data from clinical gait analysis services derived from 81 participants (Pinzone et al. 2014). With this in mind, we sought to recruit a minimum of 80 individuals to establish normative values from a university-aged population to reflect young active individuals.

Of the 196 undergraduate university students in the Health, Sport and Human Performance programme at the University of Waikato who were invited to participate via an online forum, 89 voluntarily accepted to participate and completed all TIPPs assessments. All 89 participants were invited to participate in the inter-rater reliability component of this study, with a subgroup of 17 individuals agreeing to participate. All recruitment and data collection processes were undertaken within a 4-week period in the month of May 2019. For inclusion, participants needed to be free from injury, pain, or a medical contraindication to physical activity. All participants signed an informed consent document prior to participation. The study protocol was approved by the local University human ethics committee [UoW HREC(Health)#2017-54], which was conducted in accordance with international ethical standards (Harriss, Macsween, and Atkinson 2017), adhered to The Code of Ethics of the World Medical Association (*Declaration of Helsinki*), and prepared in accordance with the Strengthening the Reporting of Observational studies in Epidemiology (STROBE) studies guidelines for cross-sectional studies (von Elm et al. 2007).

Study design

This cross-sectional study aimed to establish a dataset from a New Zealand (NZ) population and allowed the exploration of potential differences between the two main ethnic groups (Māori and NZ European) and sex. Data were collected across two sessions, 7 days apart, due to the availability of the participants and physiotherapists. We used a standard test-retest design to establish inter-rater reliability from n = 17 participants, with two qualified physiotherapists with more than 3 years of experience assessing each participant 60 minutes apart. The physiotherapists were blinded to each other's measures to reduce bias. Assessments were conducted in a movement laboratory using medical plinths, standard universal goniometers, cloth metric tapes, stabilisation straps, and a Lafayette Hand-Held Dynamometer (Model 01165, range: 0 – 136.1 kg, resolution: 0.1 kg, Lafayette Instrument Company, IN, USA). We assessed the dominant leg only, defined as the preferred leg to kick a ball. Four of the 89 individuals were left-leg dominant. All test-retest participants were right-leg dominant, with test conducted in a random order between the two therapists.

Similar to previously conducted TIPPs studies (Selfe et al. 2013; Selfe et al. 2016), the two qualified physiotherapists collecting the clinical measures undertook a series of comprehensive personalised training sessions. These sessions involved training on the research processes and on how to undertake the standardised clinical tests. Experts in the field led and observed the physiotherapists during these sessions and provided peer feedback. The physiotherapists were given a comprehensive manual outlining standard operating procedures and standardised data recording forms.

Clinical assessments

The TIPPs algorithm includes six clinical assessments (Selfe et al. 2016): (1) passive prone knee flexion (rectus femoris length) (Witvrouw et al. 2000); (2) calf flexibility standing method (gastrocnemius length) (Witvrouw et al. 2000); (3) hip abductor strength(Maffiuletti 2010); (4) quadriceps (knee extension) strength (Maffiuletti 2010);

(5) patellar mobility (medial plus lateral glide) using the patellar glide test (Witvrouw et al. 2000; Janssen et al. 2019); and (6) foot pronation using the Foot Posture Index (FPI) (Redmond, Crane, and Menz 2008). The two physiotherapists conducting the clinical assessments attended three training sessions, conducted a series of practice assessments using the Appatella™ mobile application (DigitalLabs@MMU 2019), and were given a manual outlining the assessment procedures. A brief summary of each clinical assessment is provided (**Table 1**) given that these procedures are presented in detail elsewhere (Selfe et al. 2013).

Statistical analysis

All data were examined using Shapiro Wilk tests and found to be suitable for parametric analysis, except the FPI. For the muscle length, strength, and patellar mobility assessments, univariate analyses were performed to explore the effect of ethnicity (Māori, NZ European), sex (males, females), and their interaction on outcomes. Post-hoc pairwise comparisons were performed where significant main effects were observed, and mean differences with 95% confidence intervals [lower, upper] were calculated. For the FPI, Mann-Whitney U tests were performed between ethnic groups and sexes, and the median value with 25^{th} and 75^{th} percentiles was calculated. Data from all 89 and 17 participants were available for analysis from the normative and reliability samples (i.e., no missing data).

Table 1. Overview of the procedures for each of the six clinical assessments used in the TIPPs algorithm.

Name	Method	Equipment	Outcome	Summary of procedure
Rectus femoris length	Passive prone knee bend	Universal goniometer	Knee angle (°)	Patient prone on a plinth. Foot of non-tested leg on the floor with hip at 90°. Knee of tested leg passively flexed maximally (Norkin and White 2016; Witvrouw et al. 2000).
Gastrocnemius length	Calf flexibility standing method	Universal goniometer, tape measure	Ankle angle (°)	Patient faces wall with toes of tested leg 60 cm away and knee straight. Tested leg parallel with and behind non-tested leg (i.e., toes of tested leg level with heel of non-tested leg). Hands on wall for support. Keeping heel on floor, ankle of tested leg flexed maximally (Norkin and White 2016; Witvrouw et al. 2000).
Patellar mobility	Patellar glide (medial – lateral)	Tape measure, felt pen	Total displacement (cm)	Patient supine with quadriceps relaxed and knees straight. Medial force applied to lateral border of patella. Maximal displacement of patella pole marked. Lateral force then applied to medial border of patella. Maximal displacement of patella pole marked. Distance between maximal displacements recorded (Witvrouw et al. 2000; Janssen et al. 2019).
Hip abductor strength (normalised to body weight)	Maximum voluntary isometric strength test	HHD, tape measure	Moment (Nm/kg)	Patient side lying with legs in neutral position. Test leg on top with stabilising strap around test leg and plinth, adjusted to ensure neutral position during effort. HHD under strap. Patient abducts test leg towards ceiling, applying maximum force against HHD for 3 s. Foot remains parallel to ceiling (i.e., no hip rotation). Maximum from 3 tests done 20 s apart used. Moment arm measured as distance from hip-joint axis (i.e., proximal part of greater trochanter) to HHD (Selfe et al. 2016).
Quadriceps strength (normalised to body weight)	Maximum voluntary isometric strength test	HHD, tape measure	Moment (Nm/kg)	Patient seated with knee at 90° over edge of plinth. HHD under strap, perpendicular to tibia, proximal to the malleoli. Patient extends test knee, applying maximum force against HHD, holding sides of plinth for stability. Maximum from 3 tests done 20 s apart used. Moment arm measured as distance from knee-joint axis (i.e., lateral femoral epicondyle) to HHD (Maffiuletti 2010).
Foot pronation	Foot Posture Index	None	Number between -12 to +12	Patient relaxed standing, double limb support. Six items assessed: (1) talar head palpation; (2) lateral malleolar curvature; (3) calcaneal frontal plane position; (4) prominence in the region of the talonavicular joint; (5) congruence of the medial longitudinal arch; (6) abduction/adduction of forefoot on rearfoot. Each item scored from -2 to +2, with positive values indicating greater pronation (Redmond, Crosbie, and Ouvrier 2006).

Notes. Abbreviations: HHD, handheld dynamometer. TIPPs, Targeted Intervention for Patellofemoral Pain studies.

RESULTS

Participants

The demographic characteristics of the 89 participants included in the normative database (35% Māori, 45% female) and subgroup of 17 participants included in the inter-rater reliability analysis (29% Māori, 29% female) are presented in **Table 2**.

Table 2. Demographic characteristics of participants contributing to the TIPPs normative database (*n* = 89) and inter-rater reliability analysis (*n* = 17).

Characteristic	Normative database		Inter-rater reliability	
	Male (*n* = 49)	Female (*n* = 40)	Male (*n* = 12)	Female (*n* = 5)
Age (years)	19.6 (2.1)	19.4 (1.7)	21.7 (4.3)	22.0 (4.0)
Height (cm)	180.7 (9.4)	167.9 (5.7)	173.3 (12.4)	166.8 (4.3)
Mass (kg)	84.7 (15.4)	67.7 (7.6)	78.4 (14.5)	69.6 (9.2)
Māori (*n*)	37% (18)	30% (12)	33% (4)	20% (1)

Notes. Values are mean (SD). Abbreviations: SD, standard deviation; TIPPs, Targeted Intervention for Patellofemoral Pain studies.

Normative database

The univariate analyses showed significant interactions between sex and ethnic group for rectus femoris length and patellar mobility (**Table 3**). Post-hoc pairwise comparisons indicated significant differences in rectus femoris length in males only, with Māori showing greater muscle tightness (mean difference: 7.0° [2.9, 11.1], p = 0.001). For patellar mobility, a significant difference was seen within females only, with the Māori group showing greater patellar mobility (mean difference: 4.4 mm [1.6, 7.3], p = 0.002). No other differences were seen between ethnic groups for any of the clinical assessment measures. In addition, females were significantly weaker than males in both normalised quadriceps (mean difference: -0.36 Nm/kg [-0.52, -0.20], $p < 0.001$) and hip abductor (mean difference: -0.425 Nm/kg [-0.62, -0.23], $p < 0.001$) strength measures, and demonstrated less pronated foot postures on the FPI (median difference: -2.0 points, z-score -2.277, $p = 0.023$), **Table 3**.

Inter-rater reliability

Descriptive statistics from the inter-rater reliability assessments are presented in **Table 4**. There were no significant differences between the testers for the clinical measures, except for the FPI (median difference: 2.0 points, **Table 4**). The muscle strength and rectus femoris length measures showed ‘excellent’ reliability between the two testers, whereas the gastrocnemius length and patellar mobility correlation metrics indicated ‘fair to good’ reliability (**Table 4**). The Bland-Altman plots showed that in all measures, excluding FPI, mean differences between the two testers were small (**Figure 1**). Noteworthy are the relatively wide limits of agreement for all measures, especially the patellar mobility (width of 2.1 cm), rectus femoris length (width of 28°), and gastrocnemius length (width of 14°).

Table 3. Descriptive statistics for each of the six clinical assessments used in the TIPPs algorithm presented by sex and ethic group. Results from univariate analyses, Mann-Whitney U tests, and post-hoc comparisons indicated.

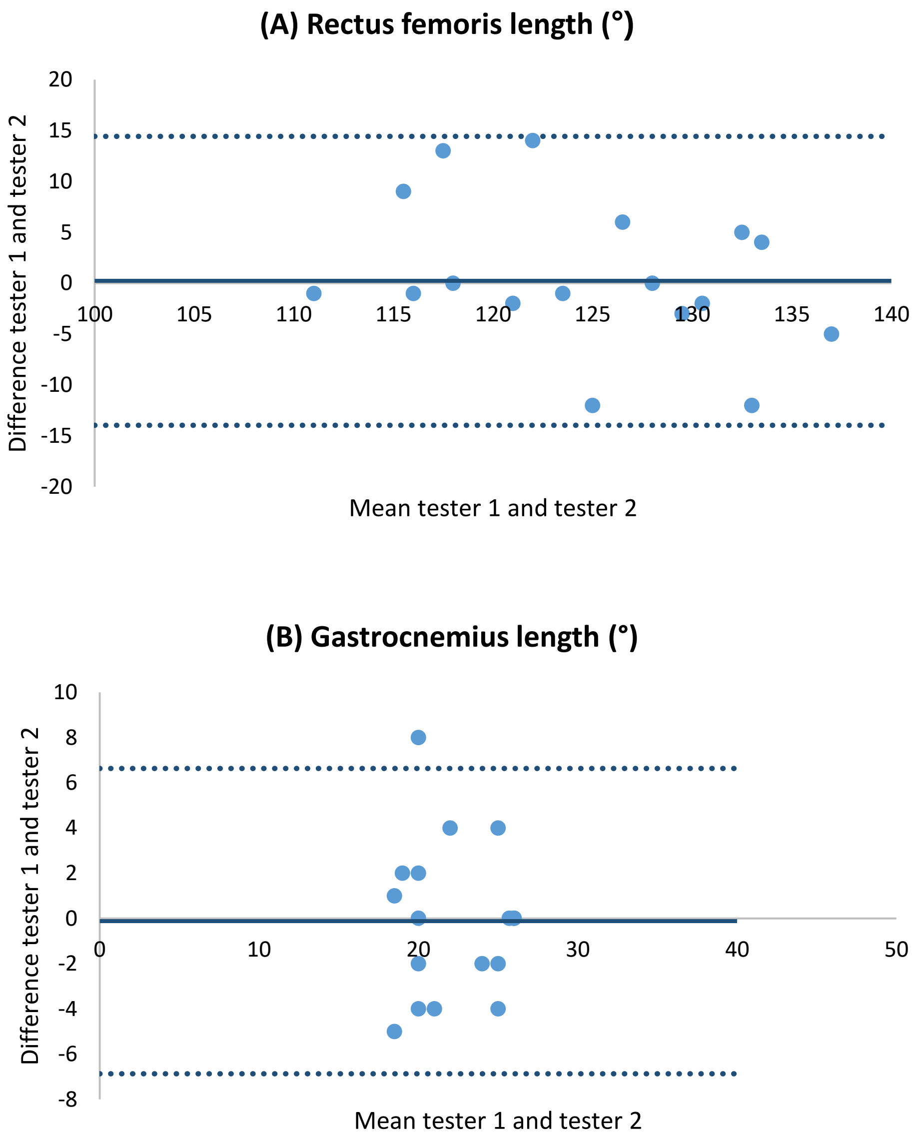
	Females			Males		
	Māori (n = 12)	NZ European (n = 28)	All (n = 40)	Māori (n = 18)	NZ European (n = 31)	All (n = 49)
Rectus femoris length (°)Ϯ	132.2 (10.7)	130.0 (6.8)	130.6 (8.1)	123.6 (7.2)†	130.6 (8.1)†	128.0 (7.2)
Gastrocnemius length (°)	19.7 (2.9)	20.5 (4.7)	20.3 (4.2)	21.4 (4.3)	21.3 (3.9)	20.9 (4.1)
Patellar mobility (mm)Ϯ	24.1 (3.4)†	19.6 (4.4)†	21.0 (4.6)	19.1 (6.3)	21.7 (4.4)	21.0 (5.0)
Hip abductor strength (Nm/kg)*	1.75 (0.51)	1.49 (0.40)	1.57 (0.44)	2.07 (0.38)	2.03 (0.44)	2.04 (0.41)
Quadriceps strength (Nm/kg)*	1.95 (0.42)	1.99 (0.33)	1.98 (0.36)	2.27 (0.35)	2.34 (0.34)	2.40 (0.36)
Foot Posture Index*	2.00 (0.25, 4.75)	1.00 (-2.00, 4.00)	2.00 (0.00, 4.00)	5.00 (3.00, 6.00)	3.00 (1.00, 6.00)	4.00 (1.00, 6.00)

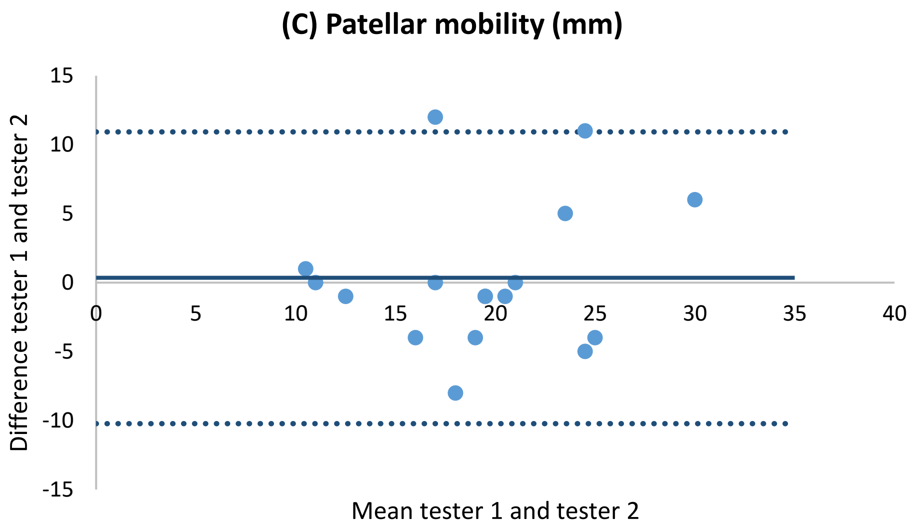
Notes. Values are mean (SD), except for the Foot Posture Index where values are displayed as median (25^{th}, 75^{th} percentiles). Ϯ Significant interaction between ethnic group and sex. * Significant difference between males and females. † Significant difference between Māori and NZ European during post-hoc comparisons. Significance set at $p \leq 0.05$. Abbreviations: NZ, New Zealand; SD, standard deviation.

Table 4. Inter-rater reliability descriptive statistics from 17 participants for each of the six clinical assessments used in the TIPPs algorithm presented for the two testers. Results from paired t-test and related samples Wilcoxon Signed Rank test comparisons between testers indicated.

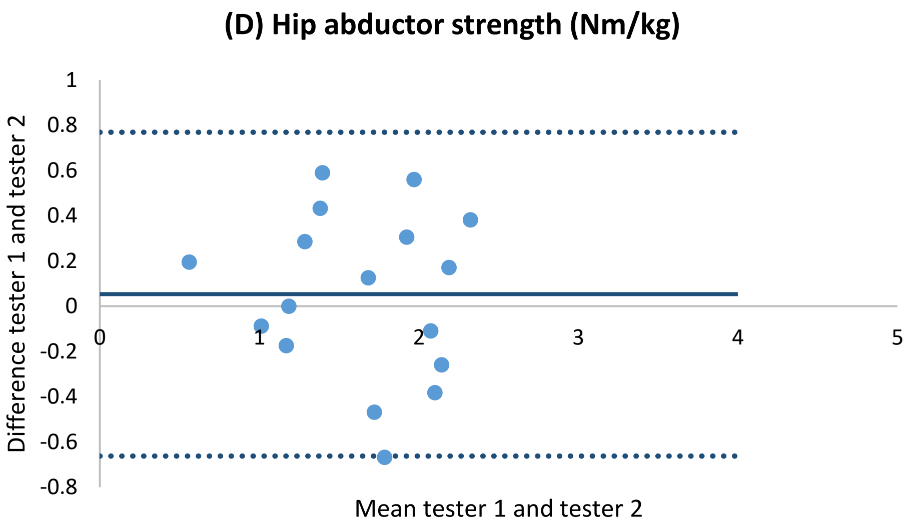
Variable	Tester 1	Tester 2	Difference	*t*-statistic	*p*-value	ICC [lower, upper]
Rectus femoris length (°)	124.8 (7.5)	124.6 (9.1)	0.24 (7.24)	0.134	0.895	0.777 [0.368, 0.920]
Gastrocnemius length (°)	22.0 (3.3)	22.2 (3.4)	-0.12 (4.39)	-0.141	0.890	0.652 [0.001, 0.875]
Patella mobility (mm)	19.5 (6.3)	19.2 (5.4)	0.4 (5.4)	0.270	0.791	0.745 [0.280, 0.909]
Hip abductor strength (Nm/kg)	1.67 (0.51)	1.62 (0.54)	0.05 (0.37)	0.600	0.557	0.867 [0.635, 0.952]
Quadriceps strength (Nm/kg)	2.22 (0.46)	2.14 (0.51)	0.07 (0.36)	0.834	0.416	0.846 [0.581, 0.944]
Foot Posture Index †	5.00 (2.50, 7.50)	3.00 (-1.50, 5.50)	-2.0 (-1.0, 4.0)	-2.299	0.021*	nc

Notes. Values are mean (SD), 95% confidence intervals [lower, upper]. † For the Foot Posture Index, related samples Wilcoxon Signed Rank test was used for comparisons, values are displayed as median (25^{th}, 75^{th} percentiles). * Significant difference between testers. Abbreviations: ICC, intra-class correlation coefficient; nc, not calculated.


(A) Rectus femoris length (°)
Difference tester 1 and tester 2
Mean tester 1 and tester 2
(B) Gastrocnemius length (°)
Difference tester 1 and tester 2
Mean tester 1 and tester 2


(C) Patellar mobility (mm)
Difference tester 1 and tester 2
15
10
5
0
-5
-10
-15
0
5
10
15
20
25
30
35
40
Mean tester 1 and tester 2

C


(D) Hip abductor strength (Nm/kg)
Difference tester 1 and tester 2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
0
1
2
3
4
5
Mean tester 1 and tester 2

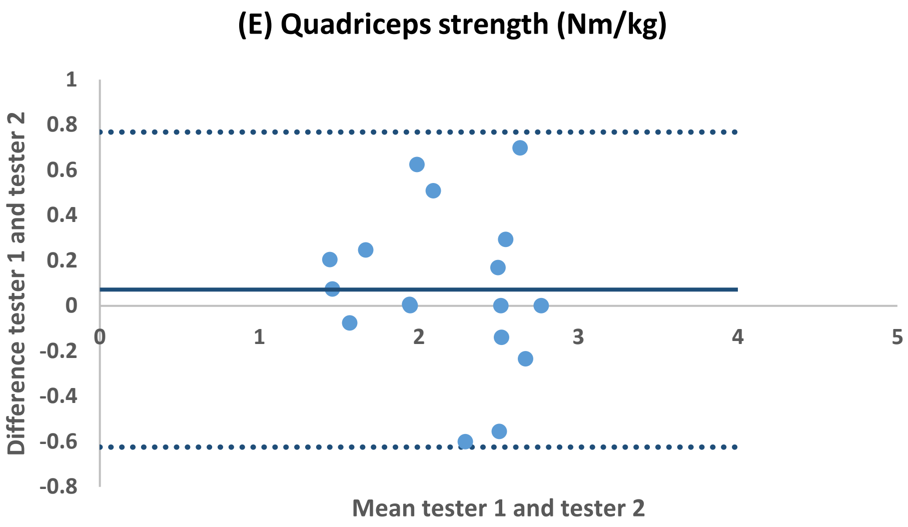


Figure 1. Bland and Altman plots with limits of agreements for: (A) rectus femoris length (°), (B) gastrocnemius length (°), (C) patellar mobility (mm), (D) hip abductor strength (Nm/kg), and (E) quadriceps strength (Nm/kg).

DISCUSSION

New Zealand university students are culturally and ethnically diverse (Taylor et al. 2017). It was hypothesised that ethnic differences might be seen within this population in the TIPPs subscales, warranting population-specific cut-off scores within the TIPPs algorithm. In multi-ethnic countries, studies have identified differences in knee pain between ethnic groups and sex, highlighting the potential for population-specific risk factors (Veerapen, Wigley, and Valkenburg 2007). Māori versus NZ European males exhibited greater rectus femoris tightness, with Māori versus NZ European females demonstrating greater patellar mobility. Else, there were no main effects of ethnic group in the TIPPs subscales, which would agree with the similar articular mobility (Klemp, Williams, and Stansfield 2002) and grip strength (Teh et al. 2014) levels observed between these two ethnic groups. Using the same TIPPs assessment technique, Gichuru et al. (2020) found differences between UK and Turkish population groups related to foot pronation attributed to a highly supinated Turkish sample. Despite suggestions that healthy Māori may be predisposed to more pronated and flatter feet than NZ Europeans (Gurney et al. 2012) and different plantar loading patterns (Gurney, Kersting, and Rosenbaum 2009) based on Harris mat measures (static foot morphology) and pressure distribution during walking (dynamic foot function), FPI values were similar between groups. The similarities in FPI outcomes between groups might be in part due to the relatively poor reliability we observed in this measure between testers. These results suggest an alternative assessment method to identify foot pronation in the TIPPs and allocate individuals to the “Weak and Pronated” subgroup should be considered, such as Harris mat (Gurney et al. 2012) or midfoot mobility (McPoil et al. 2009) methods.

There are relatively few studies comparing physical fitness attributes between ethnic groups within a NZ context (Quarrie and Williams 2002; Rush et al. 2007). Māori and NZ European demonstrate different body composition and anthropometric characteristics, with greater fat-free mass and lesser appendicular fat mass in Māori (Rush et al. 2007). The normalisation of isometric strength measures to body mass accounted for some of these distinct anthropometric characteristics, with no difference between ethnic groups seen in normalised quadriceps and hip abductor strength. Females, however, remained weaker compared to males despite normalisation. Males are overall stronger (Courtright et al. 2013) and have proportionally larger muscles (Janssen et al. 2000) than females. A previous study on quadriceps muscle strength using a similar methodology to the one presented here found ~20% lower strength in females than males (Weng et al. 2015), consistent with the difference seen here. These findings suggest that normalisation of lower body strength measures to body mass might be insufficient to account for the established sex strength differences (Courtright et al. 2013), warranting stratified algorithms based on sex within the TIPPs algorithm. The separate consideration of sex is important given that the prevalence of PFP is greater in females than males across a wide range of populations, from adolescents to military (Smith et al. 2018).

There is a paucity of reliability and validity data for a number of tests to assess passive accessory motion of the patella relative to the femur (Willy et al. 2019). The data generated in this study help to address this issue. For the patellar mobility test, the ICC for inter-tester reliability was ‘fair to good’, and the Bland-Altman plots showed that the mean difference between the two testers was small. The two testers recorded a mean

(SD) patellar mobility value of 19.5 (6.3) and 19.2 (5.4) mm. These values are consistent with a previous UK report using identical methodology where the mean patellar mobility of 44 knees in 22 healthy participants was 16.4 (5.3) mm (Janssen et al. 2019). Together, these findings indicate that measuring patellar mobility using the total medial-lateral patellar glide test is reliable and support its inclusion in the TIPPs subgrouping algorithm and use to guide clinical management.

One of the identified clinical challenges in the management of PFP is the identification of hip and thigh muscle weakness using accurate strength measures (Willy et al. 2019). Maximum voluntary isometric quadriceps strength quantified using an isokinetic dynamometer is highly reliable between testers (ICC = 0.97 to 0.98) (Chmielewski et al. 2004), but such a tool is not accessible to the majority of clinicians and researchers. Isometric muscle strength testing using a handheld dynamometer is less expensive and more accessible to practitioners, and has been the most widely used assessment tool for this purpose (Van Cant et al. 2014). The reliability and validity of measures from HHD rely on proper testing methods and stabilisation. HHD with belt stabilisation to assess knee and hip maximal isometric strength has been shown as reliable (Florencio et al. 2019; Hansen et al. 2015; Ishøi, Hölmich, and Thorborg 2019) and valid in comparison to isokinetic (Hansen et al. 2015) HHD without belt stabilisation (Florencio et al. 2019), albeit providing corresponding lower (Hansen et al. 2015) and higher (Florencio et al. 2019) values. Our reliability data align with previous reports, with ICC values of 0.85 to 0.87 and 'excellent' correlation between testers. These data suggest that muscle strength testing with an HHD with belt stabilisation is valid and reliable and supports its inclusion in the TIPPs subgrouping algorithm, although this is not necessarily interchangeable with values collected using isokinetic dynamometers or HHD without stabilisation (Weng et al. 2015).

The rectus femoris (ICC 0.777) and gastrocnemius (ICC 0.652) muscle length tests demonstrated 'fair to good' inter-rater reliability based on ICCs. Our inter-tester reliability for rectus femoris appears low compared to other work using similar methodology (ICC 0.91) (Piva et al. 2006), and in the mid-range of the breadth of ICCs values reported for gastrocnemius muscle length testing (ICC 0.29 to 0.92) (Willy et al. 2019; Barton et al. 2010). The variability and differences in results seen here may be related to a number of factors, such as how much therapists 'push' at the end of range during rectus femoris length assessment, and on how participants interpret the instructions for both the gastrocnemius and rectus femoris length tests. For example, the perception of 'pain' or 'discomfort' may vary between individuals, and affect the outcomes of length assessments. These are issues that future studies using these assessments need to be aware of and determine if these can be mitigated through training. For the TIPPs algorithm, the relatively lower reliability of the gastrocnemius muscle length test with a 'fair to good' inter-rater reliability compared to the strength measures with 'excellent' correlations indicates that the sensitivity to detect between the "Weak and Tight" and "Weak and Pronated" is not as robust as detecting between the "Strong" and both "Weak" groups.

The FPI is the only test that relies on subjective clinical observations in the TIPPs, with the remainder resulting in a specific objective measurement. The FPI is an aggregate of 6 separate observations, which is reported to help overcome this subjectivity. Despite demonstrated validity against three-dimensional (Patel et al. 2020) and radiographic

(Hegazy et al. 2020) imaging, our results indicate a mean difference in FPI scores of around 2 points between testers, which suggest that the difference we observed in FPI between sex is within tester error. Cornwall et al. (2008) found similar between-tester differences and cautioned against using the FPI. In clinical practice, this variation in FPI might influence foot posture categorisation, as well as orthotic prescription and management. Although our two examiners were qualified physiotherapists with over 3 years of experience and were suitably trained in the use of the FPI, their relative inexperience with the FPI likely had a negative impact on the inter-tester reliability outcome (Cornwall et al. 2008). Preliminary evidence had suggested that midfoot mobility measures may predict PFP patients who respond favourably to foot orthoses (Vicenzino et al. 2010; Matthews et al. 2017; Mills et al. 2012), which could be a potential alternative to the FPI in the TIPPs. One method to calculate midfoot mobility involves recording the change in midfoot width measured at 50% of the total foot length between non-weight bearing and weight bearing. 'Excellent' inter-rater (ICC > 0.83) and intra-rater (ICC > 0.97) reliability has been reported for this measure (McPoil et al. 2009). A recent randomised clinical trial, however, reported no association between midfoot mobility and treatment outcomes from using foot orthoses versus hip exercise (Matthews et al. 2020), indicating that perhaps foot width measure might not necessarily be a better alternative to the FPI for subgroup allocation via the TIPPs algorithm.

We acknowledge that our sample examined a targeted subset of a NZ university population. These undergraduate students were all studying in the area of Health, Sport and Human Performance, and limit the external validity of our findings to other populations. It is reasonable to assume that these students may have additional knowledge and a vested interest in maintaining a healthy lifestyle, potentially biasing their ‘normative’ values. We purposefully excluded individuals with PFP as the goal was to establish a normative baseline in individuals who not presented with pain. It is possible that differences between Māori and NZ European patients with PFP exist, but requires further research. In addition, the inter-rater reliability of measures was conducted 60-minutes apart in a non-injured population. It is possible that familiarisation or learning might have influenced the reliability outcomes, and worth highlighting that the reliability of measures in non-injured populations does not ensure that tests are reliable in presence of injury.

CONCLUSIONS

Overall, our study suggests that sex is an important factor worthy of further consideration within the TIPPs algorithm and subgrouping approaches, more so than ethnicity, although it is noteworthy that Māori males showed tighter rectus femoris and Māori females demonstrated greater patellar mobility. Based on this information, we recommend that a lower strength threshold be applied for the classification of females into the “strong” group for the allocation of individuals with PFP into subgroups. Overall, the different tests used to classify subgroups within TIPPs showed acceptable reliability, except for the FPI. This clinical measure relies on subjective interpretations to a greater extent than the others, and correspondingly was less reliable. This study provides direction to improve TIPPs, and warrants a sex-based stratified approach to PFP subgrouping. Ethnic differences have been reported to exist in patients presenting with musculoskeletal pain, including to the knee (Veerapen, Wigley, and Valkenburg 2007); therefore studies examining ethnic differences within PFP cohorts should be further explored.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the participants for their time and voluntary participation in this study. The authors would also like to acknowledge Versus Arthritis for supporting the Appatella™ development in previous grant funding applications (POC21427).

DECLARATIONS OF INTEREST

None.

REFERENCES

Barton, C. J., D. Bonanno, P. Levinger, and H. B. Menz. 2010. 'Foot and ankle characteristics in patellofemoral pain syndrome: A case control and reliability study', *Journal of Orthopaedic & Sports Physical Therapy*, 40: 286-96.

Brennan, G. P., J. M. Fritz, S. J. Hunter, A. Thackeray, A. Delitto, and R. E. Erhard. 2006. 'Identifying subgroups of patients with acute/subacute "nonspecific" low back pain: Results of a randomized clinical trial', *Spine*, 31: 623-31.

Campbell, Daniel. 2013. 'Normative Data.' in Fred R. Volkmar (ed.), *Encyclopedia of Autism Spectrum Disorders* (Springer New York: New York, NY).

Chmielewski, T. L., S. Stackhouse, M. J. Axe, and L. Snyder-Mackler. 2004. 'A prospective analysis of incidence and severity of quadriceps inhibition in a consecutive sample of 100 patients with complete acute anterior cruciate ligament rupture', *Journal of Orthopaedic Research*, 22: 925-30.

Collins, Natalie J, Christian J Barton, Marienke van Middelkoop, Michael J Callaghan, Michael Skovdal Rathleff, Bill T Vicenzino, Irene S Davis, Christopher M Powers, Erin M Macri, Harvi F Hart, Danilo de Oliveira Silva, and Kay M Crossley. 2018. '2018 Consensus statement on exercise therapy and physical interventions (orthoses, taping and manual therapy) to treat patellofemoral pain: recommendations from the 5th International Patellofemoral Pain Research Retreat, Gold Coast, Australia, 2017', *British Journal of Sports Medicine*, 52: 1170-78.

Cornwall, Mark W., Thomas G. McPoil, Michael Lebec, Bill Vicenzino, and Jodi Wilson. 2008. 'Reliability of the Modified Foot Posture Index', *Journal of the American Podiatric Medical Association*, 98: 7-13.

Council, Medical Research. 2018. "The MRC framework for the development, design and analysis of stratified medicine research." In, 1-57. Swindon, UK.

Courtright, S. H., B. W. McCormick, B. E. Postlethwaite, C. J. Reeves, and M. K. Mount. 2013. 'A meta-analysis of sex differences in physical ability: revised estimates and strategies for reducing differences in selection contexts', *Journal of Applied Psychology*, 98: 623-41.

DigitalLabs@MMU. 2019. 'Appatella', Accessed 28 June 2019. https://digitallabs.mmu.ac.uk/what-we-do/products/appatella/.

Florencio, L. L., J. Martins, M. R. B. da Silva, J. R. da Silva, G. L. Bellizzi, and D. Bevilaqua-Grossi. 2019. 'Knee and hip strength measurements obtained by a hand-held dynamometer stabilized by a belt and an examiner demonstrate parallel reliability but not agreement', *Physical Therapy in Sport*, 38: 115-22.

Foster, Nadine E., Krysia S. Dziedzic, Danielle A. W. M. van der Windt, Julie M. Fritz, and Elaine M. Hay. 2009. 'Research priorities for non-pharmacological therapies for common musculoskeletal problems: nationally and internationally agreed recommendations', *BMC Musculoskeletal Disorders*, 10: 3.

Gichuru, Phillip , Jessie Janssen, Hayri Yosmaoglu, James Selfe, Paola Dey, and Emel Sonmezer. 2020. 'Validation of presence of three main subgroups identified within a UK patellofemoral pain population (PFP) using a Turkish data set', *Physiotherapy*, 107: e140-e41.

Gurney, J. K., C. Kuch, D. Rosenbaum, and U. G. Kersting. 2012. 'The Māori foot exhibits differences in plantar loading and midfoot morphology to the Caucasian foot', *Gait & Posture*, 36: 157-9.

Gurney, Jason K., Uwe G. Kersting, and Dieter Rosenbaum. 2009. 'Dynamic foot function and morphology in elite rugby league athletes of different ethnicity', *Applied Ergonomics*, 40: 554-59.

Hansen, E. M., C. N. McCartney, R. S. Sweeney, M. R. Palimenio, and T. L. Grindstaff. 2015. 'Hand-held dynamometer positioning impacts discomfort during quadriceps strength testing: A validity and reliability study', *International Journal of Sports Physical Therapy*, 10: 62-8.

Harriss, D. J., A. Macsween, and G. Atkinson. 2017. 'Standards for ethics in sport and exercise science research: 2018 update', *International Journal of Sports Medicine*, 38: 1126-31.

Hegazy, F. A., E. A. Aboelnasr, Y. Salem, and A. A. Zaghloul. 2020. 'Validity and diagnostic accuracy of foot posture Index-6 using radiographic findings as the gold standard to determine paediatric flexible flatfoot between ages of 6-18 years: A cross-sectional study', *Musculoskeletal Science and Practice*, 46: 102107.

Herrington, L., S. Malloy, and J. Richards. 2005. 'The effect of patella taping on vastus medialis oblique and vastus laterialis EMG activity and knee kinematic variables during stair descent', *Journal of Electromyography and Kinesiology*, 15: 604-7.

Hill, J. C., D. G. Whitehurst, M. Lewis, S. Bryan, K. M. Dunn, N. E. Foster, K. Konstantinou, C. J. Main, E. Mason, S. Somerville, G. Sowden, K. Vohora, and E. M. Hay. 2011. 'Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial', *Lancet*, 378: 1560-71.

Hingorani, A. D., D. A. Windt, R. D. Riley, K. Abrams, K. G. Moons, E. W. Steyerberg, S. Schroter, W. Sauerbrei, D. G. Altman, and H. Hemingway. 2013a. 'Prognosis research strategy (PROGRESS) 4: Stratified medicine research', *British Medical Journal*, 346: e5793.

Hingorani, Aroon D, Daniëlle A van der Windt, Richard D Riley, Keith Abrams, Karel G M Moons, Ewout W Steyerberg, Sara Schroter, Willi Sauerbrei, Douglas G Altman, and Harry Hemingway. 2013b. 'Prognosis research strategy

(PROGRESS) 4: Stratified medicine research', *BMJ : British Medical Journal*, 346: e5793.

Ishøi, L., P. Hölmich, and K. Thorborg. 2019. 'Measures of hip muscle strength and rate of force development using a fixated handheld dynamometer: Intra-tester intra-day reliability of a clinical set-up', *International Journal of Sports Physical Therapy*, 14: 715-23.

Janssen, Ian, Steven B. Heymsfield, ZiMian Wang, and Robert Ross. 2000. 'Skeletal muscle mass and distribution in 468 men and women aged 18-88 yr', *Journal of Applied Physiology*, 89: 81-8.

Janssen, J. 2017. 'Concepts of Patellofemoral Pain.' in J. Selfe, J. Janssen and M. Callaghan (eds.), *Patellofemoral Pain: An Evidence-Based Clinical Guide* (Nova Science Publishers: Hauppauge, NY).

Janssen, Jessie, Paola Dey, Canpolat Celik, Jim Richards, and James Selfe. 2019. 'How useful is a single measurement of patellar mobility in the assessment of patients with patellofemoral pain?', *Physiotherapy Practice and Research*, 40: 29-35.

Klemp, P, SM Williams, and SA Stansfield. 2002. 'Articular mobility in Maori and European New Zealanders', *Rheumatology*, 41: 554-57.

Lack, S., B. Neal, D. De Oliveira Silva, and C. Barton. 2018. 'How to manage patellofemoral pain - Understanding the multifactorial nature and treatment options', *Physical Therapy in Sport*, 32: 155-66.

Maffiuletti, N. A. 2010. 'Assessment of hip and knee muscle function in orthopaedic practice and research', *The Journal of Bone and Joint Surgery. American Volume*, 92: 220-9.

Matheson, G. J. 2019. 'We need to talk about reliability: making better use of test-retest studies for study design and interpretation', *PeerJ*, 7: e6918.

Matthews, M., M. S. Rathleff, A. Claus, T. McPoil, R. Nee, K. Crossley, and B. Vicenzino. 2017. 'Can we predict the outcome for people with patellofemoral pain? A systematic review on prognostic factors and treatment effect modifiers', *British Journal of Sports Medicine*, 51: 1650-60.

Matthews, Mark, Michael Skovdal Rathleff, Andrew Claus, Tom McPoil, Robert Nee, Kay M Crossley, Jessica Kasza, and Bill T Vicenzino. 2020. 'Does foot mobility affect the outcome in the management of patellofemoral pain with foot orthoses versus hip exercises? A randomised clinical trial', *British Journal of Sports Medicine*, Online First

McPoil, T. G., B. Vicenzino, M. W. Cornwall, N. Collins, and M. Warren. 2009. 'Reliability and normative values for the foot mobility magnitude: a composite measure of vertical and medial-lateral mobility of the midfoot', *Journal of Foot and Ankle Research*, 2: 6.

Mills, K., P. Blanch, P. Dev, M. Martin, and B. Vicenzino. 2012. 'A randomised control trial of short term efficacy of in-shoe foot orthoses compared with a wait and see policy for anterior knee pain and the role of foot mobility', *British Journal of Sports Medicine*, 46: 247-52.

Morris, Z. S., S. Wooding, and J. Grant. 2011. 'The answer is 17 years, what is the question: understanding time lags in translational research', *Journal of the Royal Society of Medicine*, 104: 510-20.

Norkin, C.C., and D.J. White. 2016. *Measurement Of Joint Motion: A Guide To Goniometry* (F.A. Davis Company: Philadelphia, PA).

Patel, S., A. Bernasconi, J. Thornton, O. Buraimoh, N. P. Cullen, M. J. Welck, and D. Singh. 2020. 'Relationship between foot posture index and weight bearing computed tomography 3D biometrics to define foot alignment', *Gait & Posture*, 80: 143-47.

Pinzone, O., M. H. Schwartz, P. Thomason, and R. Baker. 2014. 'The comparison of normative reference data from different gait analysis services', *Gait & Posture*, 40: 286-90.

Piva, Sara, Kelley Fitzgerald, James Irrgang, Scott Jones, Benjamin Hando, David Browder, and John Childs. 2006. 'Reliability of measures of impairments associated with patellofemoral pain syndrome', *BMC Musculoskeletal Disorders*, 7: 33.

Quarrie, K, and S Williams. 2002. 'Factors associated with pre-season fitness attributes of rugby players.' in W Spinks, T Reilly and A Murphy (eds.), *Science and Football IV* (Routledge: New York, NY).

Redmond, A. C., J. Crosbie, and R. A. Ouvrier. 2006. 'Development and validation of a novel rating system for scoring standing foot posture: The Foot Posture Index', *Clinical Biomechanics*, 21: 89 - 98.

Redmond, Anthony C., Yvonne Z. Crane, and Hylton B. Menz. 2008. 'Normative values for the Foot Posture Index', *Journal of Foot and Ankle Research*, 1: 6.

Rush, E. C., J. H. Goedecke, C. Jennings, L. Micklesfield, L. Dugas, E. V. Lambert, and L. D. Plank. 2007. 'BMI, fat and muscle differences in urban women of five ethnicities from two countries', *International Journal of Obesity*, 31: 1232-39.

Saltychev, M., R. A. Dutton, K. Laimi, G. S. Beaupre, P. Virolainen, and M. Fredericson. 2018. 'Effectiveness of conservative treatment for patellofemoral pain syndrome: A systematic review and meta-analysis', *Journal of Rehabilitation Medicine*, 8: 393-401.

Selfe, J., J. Janssen, M. Callaghan, E. Witvrouw, C. Sutton, J. Richards, M. Stokes, D. Martin, J. Dixon, R. Hogarth, V. Baltzopoulos, E. Ritchie, N. Arden, and P. Dey. 2016. 'Are there three main subgroups within the patellofemoral pain population? A detailed characterisation study of 127 patients to help develop targeted intervention (TIPPs)', *British Journal of Sports Medicine*, 50: 873-80.

Selfe, James, Michael Callaghan, Erik Witvrouw, James Richards, Maria Paola Dey, Chris Sutton, John Dixon, Denis Martin, Maria Stokes, Jessie Janssen, Elizabeth Ritchie, and David Turner. 2013. 'Targeted interventions for patellofemoral pain syndrome (TIPPS): Classification of clinical subgroups', *BMJ Open*, 3: e003795.

Selfe, James, Jessie Janssen, Benjamin Drew, and Paola Dey. 2018. 'Anterior knee pain subgroups: The first step towards a personalized treatment', *Annals of Joint*, 3: 1-9.

Smith, B. E., J. Selfe, D. Thacker, P. Hendrick, M. Bateman, F. Moffatt, M. S. Rathleff, T. O. Smith, and P. Logan. 2018. 'Incidence and prevalence of patellofemoral pain: A systematic review and meta-analysis', *PLOS ONE*, 13: e0190892.

Smith, Benjamin E., Paul Hendrick, Marcus Bateman, Fiona Moffatt, Michael Skovdal Rathleff, James Selfe, Toby O. Smith, and Pip Logan. 2017. 'Current management strategies for patellofemoral pain: An online survey of 99 practising UK physiotherapists', *BMC Musculoskeletal Disorders*, 18: 181.

Taylor, Lynne, Erik Brogt, Ursula Cheer, Natalie Baird, John Caldwell, and Debra Wilson. 2017. 'Ethnicity and engagement in first-year New Zealand law programmes', *Higher Education Research and Development*, 36: 1047-60.

Teh, R., N. Kerse, M. Kepa, R. N. Doughty, S. Moyes, J. Wiles, C. Wham, K. Hayman, T. Wilkinson, M. Connolly, C. Mace, and L. Dyall. 2014. 'Self-rated health, health-related behaviours and medical conditions of Maori and non-Maori in advanced age: LiLACS NZ', *The New Zealand medical journal*, 127: 13-29.

Van Cant, J., C. Pineux, L. Pitance, and V. Feipel. 2014. 'Hip muscle strength and endurance in females with patellofemoral pain: A systematic review with meta-analysis', *International Journal of Sports Physical Therapy*, 9: 564-82.

Veerapen, K., R. D. Wigley, and H. Valkenburg. 2007. 'Musculoskeletal pain in Malaysia: a COPCORD survey', *The Journal of rheumatology*, 34: 207-13.

Vicenzino, B., N. Collins, J. Cleland, and T. McPoil. 2010. 'A clinical prediction rule for identifying patients with patellofemoral pain who are likely to benefit from foot orthoses: a preliminary determination', *British Journal of Sports Medicine*, 44: 862-6.

von Elm, Erik, Douglas G. Altman, Matthias Egger, Stuart J. Pocock, Peter C. Gøtzsche, Jan P. Vandenbroucke, and Strobe Initiative for the. 2007. 'The strengthening therReporting of observational studies in epidemiology (STROBE) statement: Guidelines for reporting observational studies', *PLoS Medicine*, 4: e296.

Weng, Philipp, Jessie Janssen, Jim D. Richards, and James Selfe. 2015. 'Validity of two clinical knee strength assessments compared to the reference standard', *International Journal of Physiotherapy and Research*, 3: 1264-70.

Willy, Richard W., Lisa T. Hoglund, Christian J. Barton, Lori A. Bolgla, David A. Scalzitti, David S. Logerstedt, Andrew D. Lynch, Lynn Snyder-Mackler, and Christine M. McDonough. 2019. 'Clinical practice guidelines linked to the international classification of functioning, disability and health from the Academy of Orthopaedic Physical Therapy of the American Physical Therapy Association', *Journal of Orthopaedic & Sports Physical Therapy*, 49: CPG1-CPG95.

Witvrouw, E., R. Lysens, J. Bellemans, D. Cambier, and G. Vanderstraeten. 2000. 'Intrinsic risk factors for the development of anterior knee pain in an athletic population. A two-year prospective study', *The American Journal of Sports Medicine*, 28: 480-9.

Yosmaoğlu, H. B., J. Selfe, E. Sonmezer, E. Sahin İ, SÇ Duygu, M. Acar Ozkoslu, J. Richards, and J. Janssen. 2020. 'Targeted treatment protocol in patellofemoral Pain: Does treatment designed according to subgroups improve clinical outcomes in patients unresponsive to multimodal treatment?', *Sports Health*, 12: 170-80.

Youdas, J. W., D. A. Krause, J. H. Hollman, W. S. Harmsen, and E. Laskowski. 2005. 'The influence of gender and age on hamstring muscle length in healthy adults', *Journal of Orthopaedic & Sports Physical Therapy*, 35: 246-52.