Osteoarthritis and Cartilage 31 (2023) 83–95

Osteoarthritis and Cartilage

Harmonising knee pain patient-reported outcomes: a systematic literature review and meta-analysis of Patient Acceptable Symptom State (PASS) and individual participant data (IPD)

V. Georgopoulos † *, S. Smith †, D.F. McWilliams †, M.P.M. Steultjens ‡, A. Williams § ||, A. Price ¶ #, A.M. Valdes †, T.L. Vincent § #, F.E. Watt ‡ § ††, D.A. Walsh †

† *Academic Rheumatology, Pain Centre Versus Arthritis and NIHR Nottingham BRC, School of Medicine, University of Nottingham, UK*
‡ *Centre for Living, School of Health and Life Sciences, Glasgow Caledonian University, UK*
§ *Centre for Osteoarthritis Pathogenesis Versus Arthritis, Kennedy Institute of Rheumatology, University of Oxford, UK*
|| *Fortius Clinic, London, UK*
¶ *Department of Immunology and Inflammation, Imperial College London, UK*
# *Botnar Research Centre, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, UK*
†† *Centre for Sport, Exercise and Osteoarthritis Research Versus Arthritis, UK*

ARTICLE INFO

Article history:
Received 24 March 2022
Accepted 16 August 2022

Keywords:
Knee pain
Patient Acceptable Symptom State
Patient-reported outcome measures
Harmonisation

SUMMARY

Objective: In order to facilitate data pooling between studies, we explored harmonisation of patient-reported outcome measures (PROMs) in people with knee pain due to osteoarthritis or knee trauma, using the Patient Acceptable Symptom State scores (PASS) as a criterion.

Methods: We undertook a systematic literature review (SLR) of PASS scores, and performed individual participant data (IPD) analysis of score distributions from concurrently completed PROM pairs. Numerical rating scales (NRS), visual analogue scales, KOOS and WOMAC pain questionnaires were standardised to 0 to 100 (worst) scales. Meta-regression explored associations of PASS. Bland Altman plots compared PROM scores within individuals using IPD from WebEx, KICK, MenTOR and NEKO studies.

Results: SLR identified 18 studies reporting PASS in people with knee pain. Pooled standardised PASS was 27 (95% CI: 21 to 35; $n = 6{,}339$). PASS was statistically similar for each standardised PROM. Lower PASS was associated with lower baseline pain ($\beta = 0.49$, $P = 0.01$) and longer time from treatment initiation ($Q = 6.35$, $P = 0.04$). PASS scores were lowest in ligament rupture (12, 95% CI: 11 to 13), but similar between knee osteoarthritis (31, 95% CI: 26 to 36) and meniscal tear (27, 95% CI: 20 to 35). In IPD, standardised PROMs each revealed similar group mean scores, but scores within individuals diverged between PROMs (LoA between −7 to −38 and +25 to 52).

Conclusion: Different standardised PROMs give similar PASS thresholds in group data. PASS thresholds may be affected more by patient and treatment characteristics than between PROMs. However, different PROMs give divergent scores within individuals, possibly reflecting different experiences of pain.

Introduction

Knee pain is highly prevalent in people of all ages, and one quarter of the global population over the age of 50 experiences persistent knee pain[1]. Moreover, knee pain can be of significant socioeconomic burden as it limits function, induces disability and distress, and reduces quality of life[2]. In older people, knee pain is most commonly attributed to osteoarthritis (OA), and in younger people it is often associated with internal derangements or external factors such as sporting injury[3,4].

* Address correspondence and reprint requests to: V. Georgopoulos, A26 Academic Rheumatology, Clinical Sciences Building, City Hospital, Nottingham, NG5 1PB, UK. Tel: 44-(0)115-823-1759.

E-mail addresses: vasileios.georgopoulos@nottingham.ac.uk (V. Georgopoulos), stephanie.smith2@nottingham.ac.uk (S. Smith), dan.mcwilliams@nottingham.ac.uk (D.F. McWilliams), martijn.steultjens@gcu.ac.uk (M.P.M. Steultjens), andy.williams@fortiusclinic.com (A. Williams), andrew.price@ndorms.ox.ac.uk (A. Price), ana.valdes@nottingham.ac.uk (A.M. Valdes), tonia.vincent@kennedy.ox.ac.uk (T.L. Vincent), f.watt@imperial.ac.uk (F.E. Watt), david.walsh@nottingham.ac.uk (D.A. Walsh).

https://doi.org/10.1016/j.joca.2022.08.011

Pain is a personal experience validly measured by numerical patient-reported outcome measures (PROMs), which are used to quantify knee pain associated with OA or injury[5]. Each questionnaire measures the participant's experience of pain, but may address different pain characteristics (e.g., weight-bearing or non-weight-bearing), functional impact, or recollection over different time periods. PROMs must meet statistical criteria of validity to enable their interpretation, data pooling and comparisons. Pooling pain data between studies using different PROMs requires that each PROM gives the same interpretation for groups, individual participant and across the full range of pain severities.

Thresholds in PROMs can categorise participants as having a particular pain characteristic. The Patient Acceptable Symptom State (PASS) is the threshold below which patients consider themselves well[6]. Available treatments might relieve but often do not eliminate pain, and PASS is the threshold representing pain below which a patient would accept for the remainder of their life[7]. It may be derived by relating post-treatment outcome status to an external anchor that reflects the patient's perspective using Receiver Operator Curve (ROC) analysis[8]. Alternatively, PASS may be derived based on data distribution within the population[6] or by predictive analysis (logistic regression)[9]. PASS thus indicates a clinical benchmark that might permit comparisons between PROMs.

There is an increasing need for pooling pain data between studies that use different PROMs. Meta-analysis of group data requires data distributions within a population to be similar for different PROMs. Furthermore, an individual's harmonised pain score or categorisation should be similar irrespective of the PROM from which it is derived. Therefore, we explored harmonisation of PROMs in people with knee pain due to osteoarthritis or knee trauma, using the PASS as a criterion. Specifically, we investigated whether harmonised PROMs measuring knee pain severity provide similar estimates of the PASS across groups and individuals with knee pain, irrespective of the host PROM by; (1) systematically reviewing literature deriving PASS thresholds in participant groups, and (2) comparing score distributions in individuals who completed at least 2 different PROMs on the same occasion.

Methods

The predefined and a priori registered systematic literature review (SLR) protocol (PROSPERO: CRD42020203250) followed the guidelines of the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA)[10]. The methodological steps taken for the SLR and subsequent individual participant data (IPD) analysis featured within this manuscript are illustrated in Supplementary Fig. 1.

Systematic literature review of PASS thresholds

Literature search

A systematic online search to identify PASS thresholds from PROMs measuring knee pain was conducted in CENTRAL, MEDLINE, EMBASE, AMED, CINAHL, and SPORTDiscus databases from 1948 until December 2021. A unique search strategy (Supplementary Table 1) incorporated PASS search terms, terms associated with knee pathologies (osteoarthritis, meniscal or ligament tears etc.) and PROMs designed to partially or fully measure pain (e.g., Western Ontario and McMaster Universities Osteoarthritis Index-WOMAC, Knee injury and Osteoarthritis Outcome Score-KOOS, Numerical Rating Scales-NRS, Visual Analogue Scales-VAS)[11]. All study designs were permitted. Citation tracking from identified studies and relevant reviews were also used.

Inclusion/Exclusion criteria

All citations deriving a PASS threshold were considered for inclusion in the systematic review according to the eligibility criteria in Table I. All identified studies were imported into EndNote X9 (Thomson Reuters) reference-handling system, and all duplicates removed. The screening process consisted of two phases, each undertaken by two reviewers (V.G., S.S.) independently. Phase one entailed screening the titles and abstracts for eligibility, and phase two comprised the evaluation of full-text citations clearing phase one. Discrepancies were resolved via discussion, and where consensus was not achieved, a third reviewer (D.A.W.) was consulted.

Data extraction

Two independent reviewers (V.G., S.S.) extracted data from each citation and subsequently validated by a third reviewer (D.F.M.). Missing data were sought from corresponding authors. Data was extracted for; first author, publication year, country of study origin, study design categorised into randomised controlled trial (RCT) or observational study (cohort, case-controlled, case only), number of participants, participant characteristics [age, sex ratio, ethnicity, body mass index (BMI)], clinical details (diagnosis), case and control interventions offered with the length of observation/follow up,

Inclusion criteria
1. Prospective studies of any design that had recruited adult participants experiencing knee pain.
2. Prospective studies exploring the efficacy of any or no intervention with or without a control population.
3. Studies that had derived a PASS cut-off score for pain PROMs or pain-measuring domains of relevant PROMs.
4. Published in English language as an original research article in a peer reviewed journal.
Exclusion criteria
1. Populations diagnosed with an inflammatory rheumatic disease (e.g., rheumatoid arthritis).
2. Studies reporting outcomes which did not include pain or where pain was not a distinguished domain.
3. Duplicate publication of data (follow-up analysis of already published data).
4. Books or book chapters, PhD theses or other dissertations, abstracts of conference presentations.

PASS: Patient Acceptable Symptom State, PhD: Doctorate of Philosophy, PROMs: Patient Reported Outcome Measures.

Table I

Study eligibility criteria

position on treatment pathway (pre/post-treatment delivery), pain PROMs, outcome score distributions within the studied populations, calculated PASS thresholds with estimates of precision, PASS derivation methods along with anchor question.

Quality and content assessment

Two reviewers (V.G., S.S.) independently conducted risk of bias and quality assessments on each included citation using a modified version of the Quality In Prognosis Studies (QUIPS) Tool[12] as it evaluates study domains that can influence individual study quality as well as overall SLR results. Disagreements in methodological quality were resolved by consensus or by consultation with a third reviewer (D.A.W.).

Data synthesis and analysis

PROMs data were harmonised by linear transformation to a 0–100 scale (0 = no pain, 100 = extreme pain). Summary statistics are presented as mean (standard deviation; SD or 95% confidence intervals; 95% CI), or median (interquartile range; IQR). SDs were imputed for ROC-derived PASS thresholds using the mean of SDs reported for distribution-derived PASS thresholds, weighted for each study sample size. The median, lower and higher available SDs, were also imputed for the purposes of sensitivity analyses. Citations were categorised based on the reported pain PROM, and by PASS derivation method (ROC, 75th percentile in the distribution of pain intensity for satisfied participants, and predictive modelling), diagnosis (osteoarthritis, meniscal or ligament tears), intervention (surgical, non-surgical), participants' position on the treatment pathway (baseline, follow-up), and level of bias (high, moderate, low).

Forest plots of pooled data from included studies used random-effects models in R (meta package, R Core Team 2020, Austria). When undertaking subgroup analyses, PASS thresholds reported in a study population at more than one timepoint have been used in

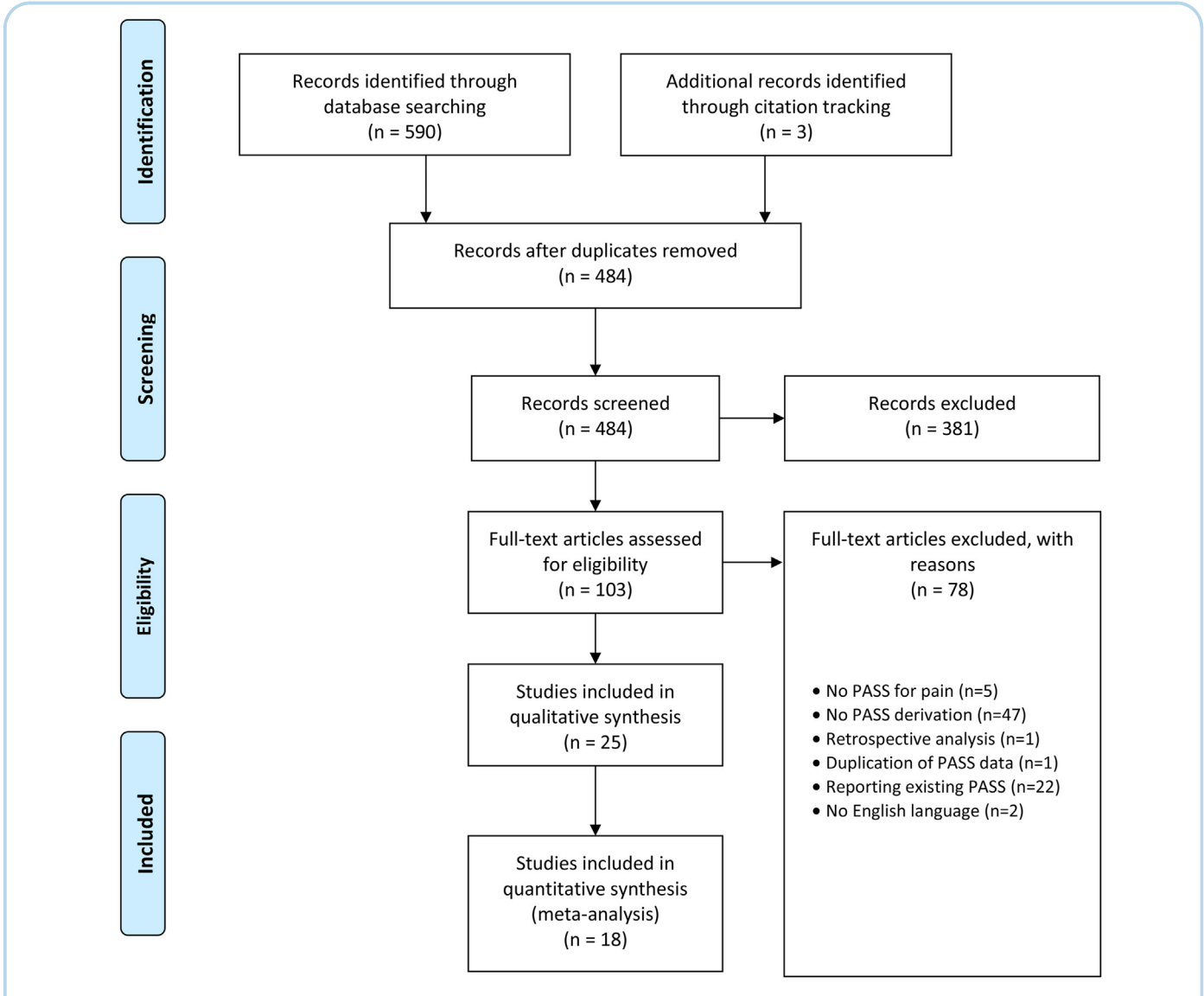

Fig. 1

Osteoarthritis and Cartilage

PRISMA flow diagram.

separate meta-analysis models for each timepoint. PASS thresholds were compared between subgroups using Cochran's Q-test. Heterogeneity measured using I^2 was considered low (I^2 < 25%), moderate (25% ≥ I^2 < 50%) or high (I^2 ≥ 50%), and significant if P < 0.10. Heterogeneity above 75% prompted subgroup analysis based on methodological quality. Publication bias was assessed by a funnel plot with its associated Egger's test, and overall risk of bias for each study determined using modified QUIPS. Post-hoc meta-regressions explored in separate models the relationships between PASS thresholds and baseline pain, the percentage of female study participants, diagnosis or treatment. For implementation purposes, overall PASS thresholds were rounded to their closest 10/100 to reflect the ordinal nature of NRS scores.

Individual participant data distribution analysis

In order to extend our group level data findings from SLR and apply to individual data, secondary data analysis from four existing knee pain studies explored data distributions for identified PROMs. Web-Based Exercises for Treating Knee Osteoarthritis (WebEx) (NCT03545048) was an RCT exploring the effectiveness of internet-based exercises in individuals with a diagnosis of knee osteoarthritis, which reported NRS and WOMAC pain scales[13]. Knee Injury Cohort at the Kennedy (KICK) (NCT02667756) was a longitudinal cohort study aiming to identify biomarkers predicting outcomes in individuals with a clinically significant acute knee injury, which collected NRS and KOOS pain scales[14]. Meniscal Tear and Osteoarthritis Risk (MenTOR) (NCT02684864/REC15/SC/0551) is an active observational cohort study exploring clinical outcome prediction by knee synovial fluid biomarkers in individuals with degenerative meniscal tear, which also collected NRS and KOOS pain scales. Neuromuscular control in Knee Osteoarthritis (NEKO) (NCT02314715) was a cross-sectional study exploring muscle performance and biomechanics in individuals with knee osteoarthritis and reported VAS and KOOS pain scales[15]. All four studies received UK research ethics committee's approvals, details of which can be found in prior publications or else are given here. WOMAC pain scores were derived from the corresponding five KOOS items when not otherwise available[16]. VAS pain scores were transformed into NRS pain scores[17]. PROM data were harmonised into a 0–100 scale before analysis (0 = no pain, 100 = extreme pain). All IPD were merged to analyse paired PROMs completed by individuals simultaneously (NRS/VAS vs KOOS, NRS/VAS vs WOMAC, KOOS vs WOMAC). Data were categorised around the pooled harmonised PASS threshold derived in the SLR, rounded to its closest 10/100 to reflect the ordinal nature of NRS scores. Linear regression models explored association and variance between PROM pairs. Source study was coded as a categorical variable and included as an interaction term in multivariable models exploring whether discrete studies influenced the strength of association between PROMs. Skewness, kurtosis, deviation from normality (Shapiro–Wilk test) and percentage of participants below the indicative harmonised PASS threshold were measured for each PROM. Cohen's kappa measured agreement and Chi^2 test differences between PROMs for categorising people below or above the pooled harmonised PASS threshold determined by SLR. Cohen's Kappa agreement was considered none (0–4%), minimal (5–15%), weak (16–35%), moderate (36–63%), strong (64–81%) or near perfect (82–100%)[18]. Bland–Altman plots and analysis visually evaluated agreement between PROMs, and established 95% limits of agreement (LoA, the range within which 95% of the differences between two separate means are expected to lie). PROMs inter-correlation was measured using concordance correlation coefficient (CCC), and considered little or zero (ρ = 0.00 to 0.25), fair (ρ > 0.25 to 0.50), moderate to good (ρ > 0.50 to 0.75), or good to excellent (ρ > 0.75)[19].

Results

Systematic literature review of PASS thresholds

Characteristics of included studies

The study selection process is shown in Fig. 1, included study characteristics in Table II, and an overview of study data in Supplementary Table 2. Twenty-five studies met the inclusion criteria[6,20–41]. All were prospective observational studies of people undergoing treatment for knee pain. No RCTs were identified.

Study characteristics	
Studies; *n*	25
Participants; *n*	11,550
Female; *n* (%)	7,035 (61%)
Age; mean (SD)	60 (17)
BMI; mean (SD)	28.5 (4.8)
Pain; 0–100, mean (SD)	47.8 (22.3)
Study location	
Europe	15
North America	10
Asia	4
Oceania	3
Africa	2
Study setting	
Clinical	23
University	2
Diagnosis	
Knee OA	15
Meniscal tear	6
ACL tear	3
Cartilage defect	1
Intervention	
Surgical	19
Non-surgical	6
Affected site*	
Knee	25
Hip	2
Pain outcome measures*	
KOOS	13
WOMAC	5
NRS	4
VAS	1
Follow-up assessment*	
>12 months	9
12 months	11
6 months	8
3 months	4
1 month	3
7 days	1

ACL: Anterior Cruciate Ligament, BMI: Body Mass Index, KOOS: Knee Injury and Osteoarthritis Outcome Score, NRS: Numerical Rating Scale, OA: Osteoarthritis, SD: Standard Deviation, VAS: Visual Analogue Scale, WOMAC: Western Ontario and McMaster Universities Arthritis Index.

* One study may include more than one site, outcome measure and follow-up assessment.

Table II Osteoarthritis and Cartilage

Summary of study characteristics

Weighted mean harmonised pain score (0–100; 0 = no pain, 100 = extreme pain) was 56.9 (±20.4) for the 11,550 included participants. Average participant age and BMI weighted for number of participants was 60 (±17) years and 28.5 (±4.8) respectively, while 61% of participants were women (Supplementary Table 2).

Knee was the joint of primary interest in all 25 studies. OA was the most commonly studied diagnosis (15 studies), with others focused on traumatic or degenerative meniscal (6/25) or anterior cruciate ligament (ACL) tears (3/25), or cartilage defects (1/25). Interventions were total[28,29,36–38,40,42] or partial[43] knee replacement, pharmacological[6,34,39,41], arthroscopic partial meniscectomy[26,27,30], arthroscopic meniscal repair[22–24], ACL reconstruction[23,33,35], meniscal allograft transplantation (MAT)[31], primary microfracture procedure[25], platelet-rich plasma injections[21], and multimodal (education, physical therapy, pharmacological analgesia, and lifestyle and weight reduction advice)[32]. KOOS was the most commonly used pain PROM (13/25) followed by WOMAC (5/25), 0–100 NRS (4/25) or VAS (1/25). A single study reported PASS threshold for both KOOS and NRS[28].

PASS derivation

Time to follow-up at which PASS was determined varied between studies. Almost half of the studies reported PASS threshold at 6 or 12-months post-intervention (8 and 11/25 respectively). PASS thresholds >12 months after treatment were reported in 9/25 studies, and at 3-months, 1-month or 7-days in 4, 3 and 1 studies, respectively. All studies used an anchoring question. Twenty-four established symptom acceptability after treatment finished, and one asked participants whether, with hindsight having experienced the entire process, they would still proceed with the same treatment. Sixteen studies asked whether participants would be satisfied with or find acceptable their current state for the rest of their lives (yes/no). Six used a predefined satisfaction threshold using a 0–10 NRS (3 studies) or a 5-point Likert scale (3 studies). Two studies did not report the anchor question. Most (17/25) studies used ROC analysis to derive the PASS threshold, whereas 7 used the value corresponding to the 75th percentile of the distribution of participants who reported acceptable pain levels. A single (1/25) study used predictive modelling (logistic regression) to identify the PASS cut-off point.

Harmonised PASS threshold for people with knee pain

Data from 18 studies could be included in the meta-analysis. Seven studies were excluded because either pain outcome was not reported discrete from function[29,36,38,42,43] or PASS for knee pain was not reported discrete from hip pain[34,41]. Nine of the 18 studies were rated as low, five as moderate and four as high risk of bias (Table III). Data from KOOS, WOMAC, NRS and VAS pain PROMs

Modified Quality in Prognosis Studies (QUIPS)

Studies & Risk of bias tool criteria			Study participation	Study attrition	Outcome measurement	Study confounding	Statistical analysis and reporting	Overall risk of bias
Observational PASS Derivation Studies	Included in Meta-analyses	Agarwalla *et al.*, 2019	Low	N/A	High	N/A	Low	**Moderate**
		Beletsky *et al.*, 2021	Low	N/A	Low	Moderate	Low	**Low**
		Boffa *et al.*, 2021	Low	N/A	Low	N/A	Low	**Low**
		Chahal *et al.*, 2021	Moderate	Moderate	Low	Low	Low	**Low**
		Chahla *et al.*, 2020	Low	N/A	Low	Low	Low	**Low**
		Connelly *et al.*, 2019	High	High	Low	N/A	Low	**High**
		Dwyer *et al.*, 2020	Low	High	Low	Low	Low	**Moderate**
		Escobar & Riddle, 2014	Low	High	Low	N/A	Low	**High**
		Escobar *et al.*, 2012	Low	High	Low	N/A	Low	**High**
		Gowd *et al.*, 2019	Low	Low	Low	Low	Low	**Low**
		Ingelsrud *et al.*, 2015	Low	Low	Low	N/A	Low	**Low**
		Liu *et al.*, 2019	Low	Low	High	Low	Low	**High**
		Maheshwer *et al.*, 2021	Low	Moderate	Low	Low	Low	**Low**
		Mahler *et al.*, 2018	Low	Moderate	Low	Low	Low	**Low**
		Muller *et al.*, 2016	Low	Low	Low	N/A	Low	**Low**
		Pedersen *et al.*, 2021	Low	Moderate	Low	Moderate	Low	**Moderate**
		Perrot & Bertin, 2013	Low	Moderate	Low	Moderate	Low	**Moderate**
		Tubach *et al.*, 2005	Low	Moderate	Low	Low	Moderate	**Moderate**
	Excluded from Meta-analyses	Bellamy *et al.*, 2015	Low	Low	Low	N/A	Low	**Low**
		Felix *et al.*, 2019	Low	Low	Low	Low	Low	**Low**
		Goh *et al.*, 2021	Low	Low	Low	Low	Low	**Low**
		Ingelsrud *et al.*, 2021	High	Moderate	Low	N/A	Low	**Moderate**
		Keurentjes *et al.*, 2014	Moderate	Moderate	Moderate	Low	Low	**Moderate**
		Naal *et al.*, 2015	Low	Moderate	Low	N/A	Moderate	**Moderate**
		Tubach *et al.*, 2012	Low	Low	Low	N/A	Low	**Low**

Studies have been marked within each domain as high, medium and low risk of bias and an overall mark was given based on their performance on 'study attrition', 'outcome measurement' and 'statistical analysis and reporting', which were *a priori* set as the most important domains. Overall grades about study risk of bias were based on study performance in the aforementioned domains and were given after consensus between the two reviewers. 'Study confounding' was excluded except where explicitly stated as a methodological step of a study. The 'prognostic factor measurement' domain featured in the original QUIPS was removed as it was not applicable to the design and purpose of the individual studies. Studies reporting pain outcomes not discrete from function or PASS scores for knee pain not discrete from hip pain were excluded from meta-analysis.

Table III

Assessment of risk of bias

were harmonised to a 0–100 scale (0 = no pain, 100 = extreme pain). The pooled estimate of harmonised PASS was 26.70 (95% CI: 22.23 to 31.17, $n = 7,485$) (Fig. 2). A funnel plot displayed symmetry (Egger's test: 1.80, $P = 0.09$) suggesting little or no bias, but high heterogeneity between studies ($I^2 = 95\%$, $P < 0.01$, Fig. 3). Sensitivity analyses for a PASS threshold with imputed median, lowest or highest SDs observed amongst the 13 studies for which data were available gave pooled PASS thresholds of 26.63 (95% CI: 22.31 to 30.96), 26.58 (95% CI: 22.40 to 30.77) and 27.18 (95% CI: 21.88 to 32.48) respectively, and were therefore similar to the primary analysis in Fig. 2.

Effects of study and participant characteristics on knee pain PASS

No significant differences in harmonised PASS threshold were demonstrated between studies reporting different pain PROMs ($Q = 3.64$, $P = 0.16$) (Supplementary Fig. 2). Subgroup pooled PASS thresholds were; KOOS (23.14, 95% CI: 18.65 to 27.63, $n = 3,371$), WOMAC (27.50, 95% CI: 22.60 to 32.40, $n = 1,433$), and NRS/VAS (34.81, 95% CI: 21.75 to 47.87, $n = 3,064$). High methodological heterogeneity was found between studies reporting KOOS or NRS/VAS ($I^2 = 94\%$ and 90% respectively, each $P < 0.01$), and moderate heterogeneity for WOMAC ($I^2 = 48\%$, $P = 0.17$).

Meta-regression revealed that PASS thresholds were significantly lower with lower baseline pain severity ($\beta = 0.49$, 95% CI: 0.08 to 0.90, $P = 0.01$) (Supplementary Fig. 3). A 30/100 higher baseline pain was associated with approximately mean 15/100 higher PASS.

Differences in PASS were found between diagnostic subgroups ($Q = 188.63$, $P < 0.001$) (Supplementary Fig. 4). Pooled PASS thresholds for knee OA (30.52, 95% CI: 24.58 to 36.45, $n = 4,918$) were similar to meniscal tear (26.03, 95% CI: 24.58 to 27.47, $n = 1,575$, $Q = 2.07$, $P < 0.15$), but significantly higher than for ACL tear (12.07, 95% CI: 10.59 to 13.54, $n = 992$, $Q = 46.81$, $P < 0.0001$). Significant differences were observed also between meniscal tear and ACL tear subgroups ($Q = 175.03$, $P < 0.0001$). Heterogeneity was high within the knee OA subgroup ($I^2 = 81\%$, $P < 0.01$), but not detected within meniscal and ACL tear subgroups (each $I^2 = 0\%$, each $P > 0.56$).

Studies of surgical interventions reported lower pooled PASS thresholds (23.33, 95% CI: 19.45 to 27.21, $n = 4,589$) than those of non-surgical interventions (37.15, 95% CI: 27.54 to 46.76, $n = 2,896$, $Q = 5.32$, $P = 0.02$) (Supplementary Fig. 5). Studies within each surgical and non-surgical intervention subgroup displayed high heterogeneity ($I^2 = 93\%$ and 82% respectively, each $P < 0.01$).

Studies with PASS determined after longer follow-up reported lower PASS thresholds ($Q = 9.84$, $P = 0.01$) (Supplementary Fig. 6). Pooled PASS thresholds were; up to 6-months follow-up (28.29, 95% CI: 22.17 to 34.41, $n = 5,569$), 12-month follow up (23.83, 95% CI: 18.55 to 29.11, $n = 3,303$), and beyond 12-months (15.75, 95% CI: 10.45 to 21.06, $n = 1,880$). Studies within 6-month, 12-month and beyond 12-month follow up subgroups each displayed high heterogeneity ($I^2 = 97\%$, 87% and 78% respectively, each $P < 0.01$).

No significant differences were demonstrated between studies with different risks of bias ($Q = 2.43$, $P = 0.30$) (Supplementary Fig. 7), or using different PASS derivation methods ($Q = 0.07$, $P = 0.78$) (Supplementary Fig. 8). Pooled PASS thresholds were; studies with low (22.94, 95% CI: 16.12 to 29.76, $n = 2,016$), moderate (32.07, 95% CI: 22.83 to 41.31, $n = 3,555$), high risk of bias (26.34, 95% CI: 21.12 to 31.56, 1,914), ROC approach (25.35, 95% CI: 20.68 to 30.02, $n = 2,320$), logistic approach (26.00, 95% CI: 24.50 to 27.50, $n = 614$), and distribution approach (29.45, 95% CI: 17.70 to 41.20, $n = 4,552$). Studies with low, moderate or high risk of bias each displayed high heterogeneity ($I^2 = 84\%$, 91% and 63% respectively, each $P < 0.05$), as did studies using ROC, distribution or logistic PASS derivation methods ($I^2 = 52\%$, 99% and 0% respectively, each $P < 0.05$). No significant association was demonstrated between

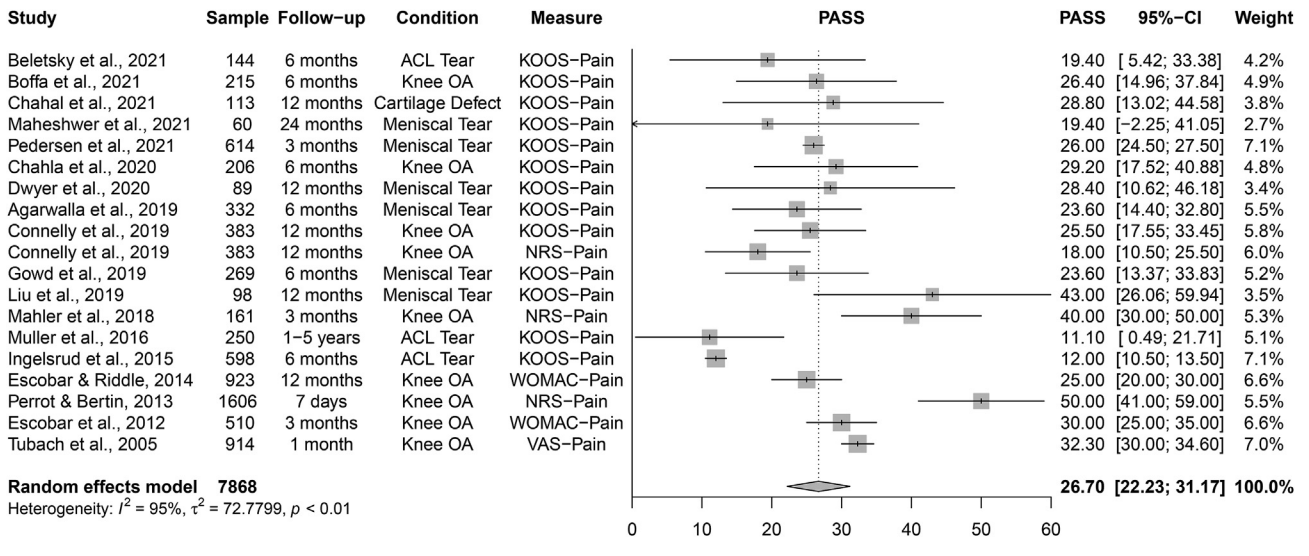

Fig. 2

Osteoarthritis and Cartilage

Forest plots showing the PASS score from studies of knee pain. ACL: Anterior Cruciate Ligament, BMI: Body Mass Index, CI: Confidence Interval, KOOS: Knee Injury and Osteoarthritis Outcome Score, NRS: Numerical Rating Scale, OA: Osteoarthritis, SD: Standard Deviation, VAS: Visual Analogue Scale, WOMAC: Western Ontario and McMaster Universities Arthritis Index. The study by Connelly *et al.*, 2020 has been included twice in the model as the authors reported PASS threshold scores for two distinct PROMs on the same population. The study's sample size has been included only once in the model total sample size ($n = 7,485$).

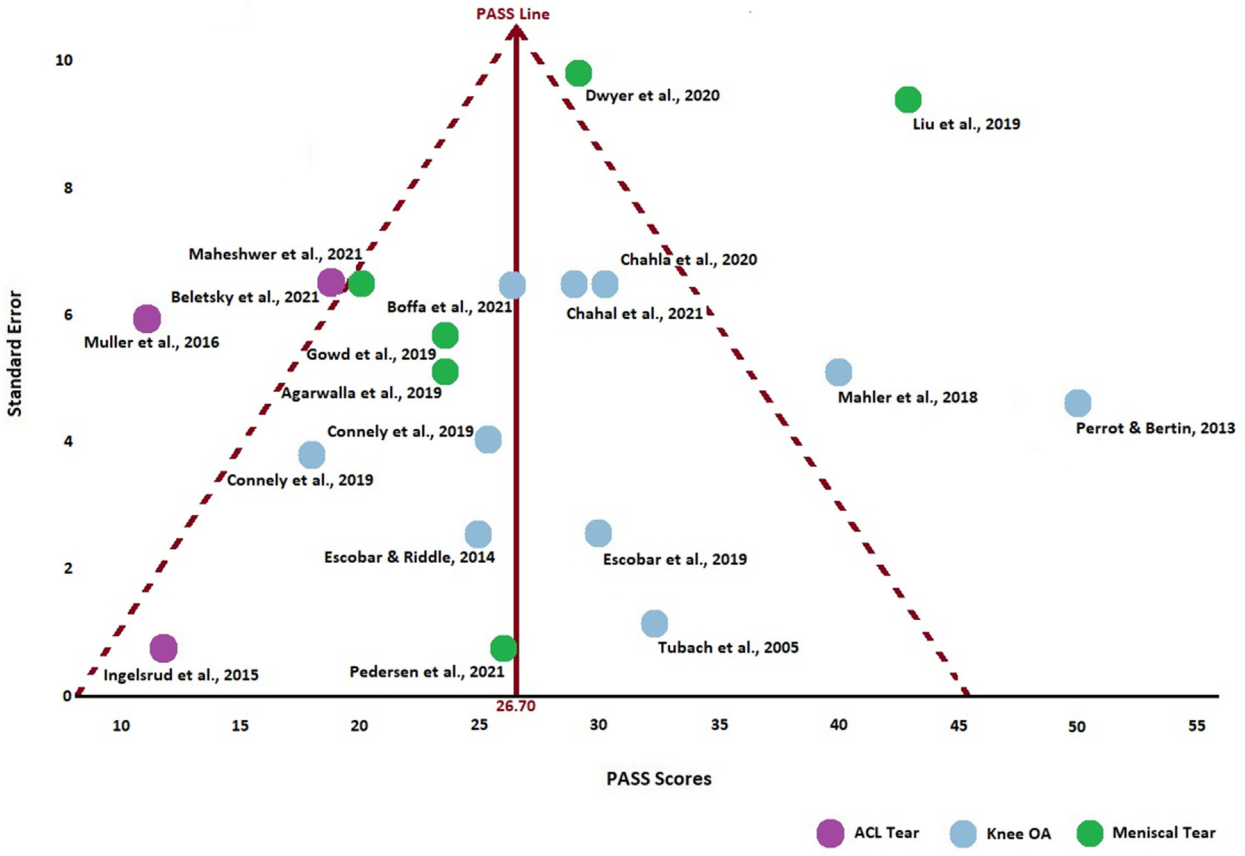

Fig. 3

Osteoarthritis and Cartilage

Funnel plot with PASS line for knee pain studies (***n*** = 18) examining PASS scores measured via different tools and derived via different methods depicting a symmetrical presentation but high variability due to the range of PASS scores (12–50/100).

PASS thresholds and the percentage of females reported in each study ($\beta = 0.02$, 95% CI: −0.01 to 0.03, $P = 0.10$).

Individual participant data comparisons between knee pain PROMs

We used IPD to extend our findings from SLR of group data, and applied the derived harmonised PASS thresholds to individual data. Demographics and data distributions are presented in Table IV for each group for which paired PROM data were available. Paired NRS and KOOS data were available for 325, NRS and WOMAC for 430, and KOOS and WOMAC for 325 participants. Each of the three harmonised PROMs displayed similar data distributions (Table IV), except that NRS data displayed more negative kurtosis than KOOS or WOMAC. Each harmonised PROM was positively associated with each other PROM, with small mean score differences (Fig. 4, Table IV). Taking part in any particular study did not significantly influence the relationship between PROMs. Differences were found between harmonised PROMs in the proportion of participants scoring ≤30/100 (NRS: 40–43%, KOOS: 31%, WOMAC: 51–56%, $Chi^2 = 54.40$ to 109.79, $P < 0.0001$, Table IV). Bland–Altman plots (Fig. 4) showed wide LoA. Commonality, which refers to the agreement between PROMs in categorising people as having achieved (PASS ≤ 30/100) or not achieved (PASS > 30/100) an acceptable pain state post-intervention was 23–31%. Cohen's agreement was 68–75%.

Discussion

Our SLR and meta-analysis of published studies indicates that harmonised PROMs (0–100) produce similar PASS thresholds, converging at approximately 30/100. We found that acceptable pain levels were influenced by baseline pain severity, time since starting treatment, and diagnostic/treatment group. IPD analysis supports the findings of this SLR, indicating similar group mean scores based on harmonised scores from different PROMs in different studies. However, different PROMs may give widely different individual patient level scores. This might result from the complexity and diversity of pain, such that PROMs might reflect different aspects that are more or less relevant to different individuals. Overall, our data support how pain PROMs might be harmonised for analysis of group data by applying a threshold such as a harmonised PASS, but suggest caution in pain PROMs harmonisation for investigating individuals with knee pain.

We found that PASS thresholds were similar irrespective of the method used to determine PASS, or the questionnaire from which it was derived. Overall, people with knee pain might consider that a pain score of 30/100 or less after a given intervention will be acceptable for the rest of their lives or it might be that people set their expectations for future health based on what they perceive as normal within the population, which might correspond to the 75th

Variables (Value, Range, %)		Descriptives and Normality Testing								
		NRS/VAS vs KOOS			NRS/VAS vs WOMAC			KOOS vs WOMAC		
Demographics and Descriptives	No. participants	325			430			325		
	Female (*n* (%))	107 (33%)			178 (41%)			107 (33%)		
	Age (y)	42 ± 16			48 ± 18			42 ± 16		
	BMI (kg/m^2)	28 ± 6			29 ± 6			28 ± 6		
	NRS (0–100)	40 (20–60)			40 (20–60)			–		
	KOOS (0–100)	39 (25–53)			–			39 (25–53)		
	WOMAC (0–100)	–			30 (20–45)			30 (15–45)		
Comparisons and Associations	NRS − KOOS mean diff. (95% CI)	2.58 (0.34–4.82)			–			–		
	NRS − WOMAC mean diff. (95% CI)	–			10.24 (8.33–12.24)			–		
	KOOS − WOMAC mean diff. (95% CI)	–			–			9.03 (8.16–9.90)		
	NRS ≤30 (*n* (%))	138 (43%)			173 (40%)			–		
	KOOS ≤30 (*n* (%))	100 (31%)			–			100 (31%)		
	WOMAC ≤30 (*n* (%))	–			220 (51%)			181 (56%)		
	≤30 Commonality (*n* (%))	73 (23%)			126 (29%)			99 (31%)		
	Cohen's kappa (95% CI)	0.44 (0.34–0.54)			0.35 (0.26–0.44)			0.50 (0.40–0.60)		
	Cohen's agreement (%)	72%			68%			75%		
	Chi^2 (x (*P*))	58.81 (<0.0001)			54.40 (<0.0001)			109.79 (<0.0001)		
	CCC (*P*)	0.81 (<0.0001)			0.77 (<0.0001)			0.86 (<0.0001)		
Distribution	PROMs (Range)	Shapiro–Wilk (*P*-value)	Skewness (95% CI)	Kurtosis (95% CI)	Shapiro–Wilk (*P*-value)	Skewness (95% CI)	Kurtosis (95% CI)	Shapiro–Wilk (*P*-value)	Skewness (95% CI)	Kurtosis (95% CI)
	NRS/VAS (0–100)	0.97 (<0.0001)	0.30 (0.03–0.56)	−0.73 (−1.27 to −0.20)	0.97 (<0.0001)	0.24 (0.01–0.48)	−0.72 (−1.18 to −0.25)	–	–	–
	KOOS (0–100)	0.99 (0.02)	0.41 (0.14–0.68)	−0.05 (−0.58 to 0.49)	–	–	–	0.99 (0.02)	0.41 (0.14–0.68)	−0.05 (−0.58 to 0.49)
	WOMAC (0–100)	–	–	–	0.97 (<0.0001)	0.51 (0.28–0.74)	−0.02 (−0.48 to 0.45)	0.96 (<0.0001)	0.61 (0.34–0.88)	0.14 (−0.40 to 0.67)

BMI: Body mass index, NRS: Numerical rating scale. KOOS: Knee injury and Osteoarthritis Score pain subscale, WOMAC: Western Ontario and McMaster Universities Arthritis Index pain subscale. Data are test statistics and probabilities for pooled individual patient data for cases included in each paired comparison (NRS-KOOS, NRS-WOMAC and KOOS-WOMAC). Data are from WebEx, KICK, MenTOR and NEKO studies. All scales were standardised by linear transformation to 0 to 100, with 100 indicating worst pain. Data are *n* (%), mean (±SD), median (interquartile range) or k (95% CI).

Table IV

Participant demographics based on comparison data between pain PROMs taken from 4 discrete studies of people with knee pain

Osteoarthritis and Cartilage

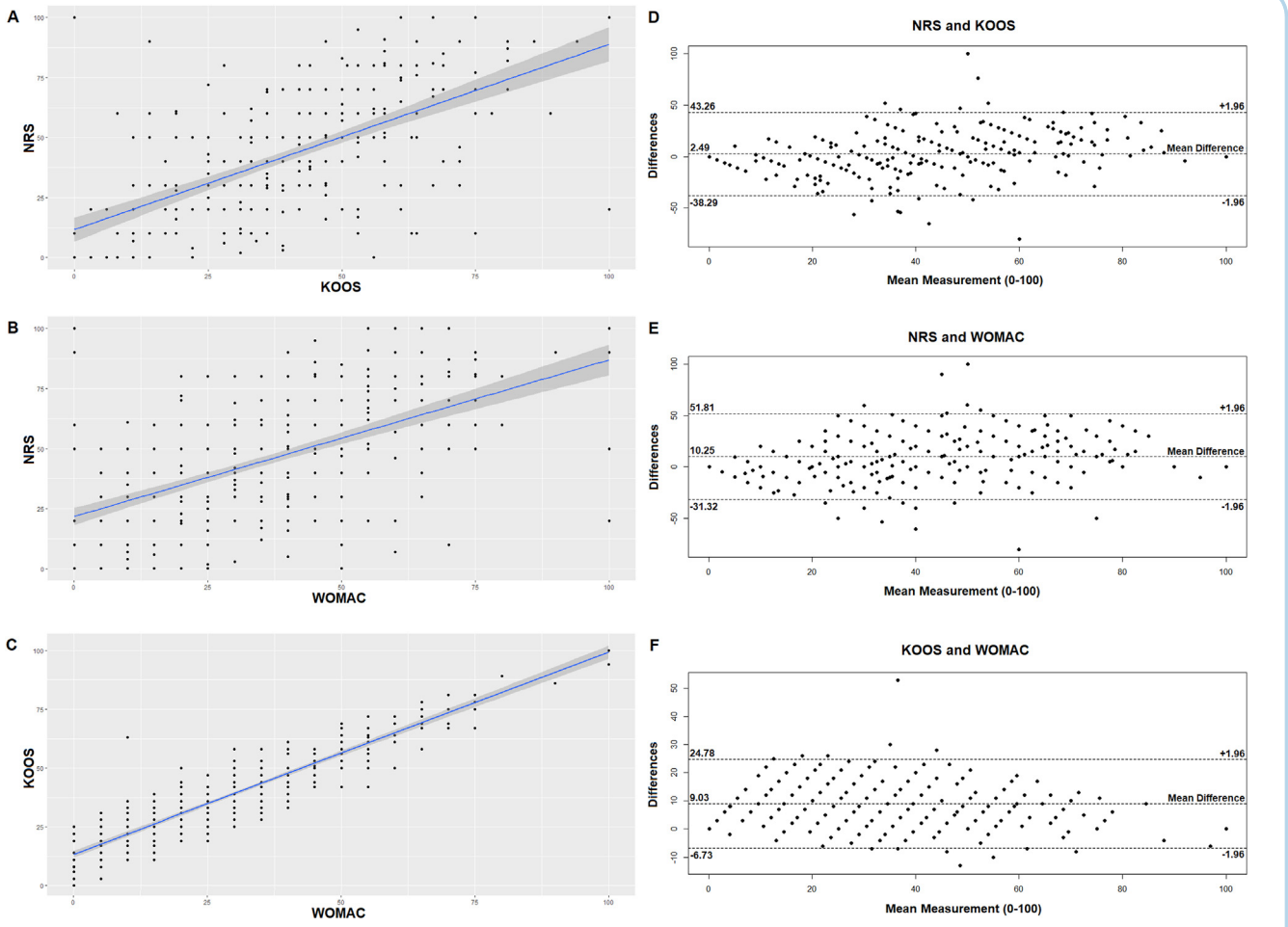

Fig. 4

Osteoarthritis and Cartilage

Relationships and Bland–Altman plots between standardised scores of pain PROMs completed in pairs at the same time in individuals with knee pain. A: NRS/VAS and KOOS: $n = 325$, $\beta = 0.76$ (95% CI: 0.65 to 0.88), Intercept: 12.10 (95% CI: 7.01 to 17.19), $R^2 = 0.35$ ($P < 0.0001$); B: NRS/VAS and WOMAC: $n = 430$, $\beta = 0.65$ (95% CI: 0.56 to 0.74), Intercept: 21.77 (95% CI: 18.20 to 25.33), $R^2 = 0.31$ ($P < 0.0001$), C: KOOS and WOMAC: $n = 325$, $\beta = 0.86$ (95% CI: 0.82 to 0.90), Intercept: 13.44 (95% CI: 11.98 to 14.90), $R^2 = 0.85$ ($P < 0.0001$). D: NRS/VAS and KOOS: $n = 325$, mean difference 2.49, −LoA: −38.29 (95% CI: −42.22 to −34.36); +LoA: 43.26 (95% CI: 39.33 to 47.20); E: NRS/VAS and WOMAC: $n = 430$, mean difference 10.25, −LoA: −31.32 (95% CI: −34.80 to −27.84); +LoA: 51.81 (95% CI: 48.33 to 55.29), F: KOOS and WOMAC: $n = 325$, mean difference 9.03, −LoA: −6.73 (95% CI: −8.24 to −5.21); +LoA: 24.78 (95% CI: 23.26 to 26.30). NRS/VAS: Numerical rating scale/Visual Analogue Scale. KOOS: Knee injury and Osteoarthritis Score pain subscale, WOMAC: Western Ontario and McMaster Universities Arthritis Index pain subscale. Data are from WebEx, KICK, MenTOR and NEKO studies. All scales were standardised by linear transformation to 0 to 100, with 100 indicating worst pain.

percentile used in various different distribution-based methods for PASS derivation, possibly explaining why we did not observe a difference between the methods used. Despite this apparent consistency between PROMs, we found that PASS might be significantly influenced by a range of person and treatment factors; baseline pain, time from treatment initiation, diagnosis and type of treatment. In meta-regression, PASS thresholds were associated with baseline pain severity, indicating that those with more severe pain at baseline might be more willing to accept more severe residual pain after treatment. This might represent a realistic acceptance, to the extent that pain severity, pain chronicity and the number of painful sites have all been associated with worse future pain in people with knee osteoarthritis following treatment[44]. However, our findings indicate that the pain severity that people can find acceptable may wane over time after treatment. Shortly after treatment people may feel that they can accept more pain for the rest of their life than subsequently proves to be the case. PASS thresholds were lower in studies of ligament tears. This might reflect high health, treatment and physical activity expectations of young individuals with soft tissue injuries associated with sports. Other explanations might include differences between pathologies in their associated sensory and emotional experience of pain and its impact on the individual. Surgical interventions yielded lower PASS thresholds than non-surgical interventions, which might suggest that people undergoing surgical form of treatment have different expectations in regards to treatment efficacy and future pain.

Although other clinical benchmarks (e.g., minimum clinically important difference-MCID), might legitimately vary between PROMs, our data indicate consistency between harmonised PROMs in the PASS benchmark. Consistency in PASS thresholds has previously been reported across rheumatological diagnoses[7]. Other factors, such as sex, considered to influence pain reporting were not significantly associated with PASS in our study. Heterogeneity has previously raised concerns about the universal validity of PASS and its readiness to be applied with confidence in clinical practice[45]. Our study suggests that PASS as a harmonised threshold may be more validly applied to group data rather than to individual patients or research participants. PASS threshold selection should consider patient and treatment characteristics appropriate to the research or clinical question being addressed.

Our IPD analysis found that even though mean harmonised PROMs scores were similar, different PROMs produced different scores for individuals. KOOS was originally developed from the WOMAC questionnaire in order to be targeted to populations with knee injury as well as those with osteoarthritis[46]. KOOS pain scale incorporates the WOMAC pain scale items, and both scales demonstrate high content validity (>90%) and similar responsiveness to arthroplasty treatment[47]. However, KOOS and WOMAC might display different construct validity (discrimination) in people with OA[48], or provide different scores depending on individual age, sex, and BMI in healthy participants[49]. We similarly show similar data distributions and high correlation between KOOS and WOMAC in our study. The slightly lower mean harmonised scores derived from WOMAC compared to KOOS, although statistically significant, did not reach the threshold of 15/100 proposed to be an MCID[50]. However, at -individual participant level, agreement between harmonised WOMAC and KOOS scores was low (+LoA up to 25/100). This likely reflects the inclusion in KOOS of items that target people with different pain phenotypes. We similarly found comparable data distributions, prevalence of acceptable pain levels and strong correlation between NRS and other harmonised pain PROMs, but even lower agreement than observed between KOOS and WOMAC scores at an individual level.

Interpretation of our findings is subject to a number of limitations. PASS thresholds for PROMs that were not explored in the current study may display different characteristics. Meta-analysis demonstrated no significant differences between identified PROMs, although high variability can be observed in the wide confidence intervals. NRS and VAS scales use a range of lead questions, relating to recall period or specific pain characteristics, but heterogeneity between these different scales could not be explored with the data available in this study. Other characteristics might differentiate between PROMs other than those studied here. Although our SLR search strategy incorporated a wide range of diagnostic terms and pain measurement tools, it remains possible that not all relevant studies were identified. Imputation of variance data in meta-analyses may introduce bias, despite following existing guidelines, and despite evidence from our sensitivity analyses that imputation was adequate. Significant heterogeneity between studies has been identified, suggesting that factors additional to those explored in subgroup analyses could also influence PASS thresholds. The subgroup analyses were undertaken using separate statistical models, and we cannot exclude confounding by other variables. Factors not examined in this review include additional group demographics such as ethnicity, age, country of study origin, depression as well as expectations about study participation, selected treatment pathway or outcome. Each might influence outcomes in people with knee pain. Not all studies of meniscal tear excluded people with OA and studies might have been blurred these two diagnostic subgroups. We recognise we might have missed smaller associations due to limitations of power as well as our statistical approaches. Limited study quality reduces certainty in our findings, although most studies included in our SLR were categorised as of low (13/25) or moderate (8/25) risk of bias, and we found no significant effect of study quality on our main findings. We used PASS as a clinical threshold, and, although similar data distributions were observed between scales, we cannot assume that consistency would be demonstrated for a choice of other benchmarks. Such research awaits clinical validation of benchmarks such as the threshold between severe and moderate pain. Alternative approaches to data harmonisation have been proposed, such as development and validation of crosswalk tables from large datasets in which multiple PROMs have been completed[51]. Our findings suggest that harmonisation might be affected by a number of factors such as diagnosis or treatment, more than by which PROM has been used and further studies might explore whether crosswalk tables might differ for different patient groups or be applied to every setting.

In conclusion, our SLR and meta-analysis of published studies and our IPD data analysis converge to suggest that harmonised PROMs (0–100) might display similar data distributions, and might produce similar PASS thresholds at approximately 30/100, and population mean values. However, acceptable pain levels may depend on baseline pain severity, time since starting treatment, and diagnostic/treatment group, and different PROMs may give widely different scores at the individual patient level. Future work should extend our findings to other PROMs, and to other clinical benchmarks such as MCIDs. In prospective research, pain PROMs should be selected to meet the scientific objectives, including responsiveness to treatment effects. KOOS may be of particular value in research on knee Injury, WOMAC for osteoarthritis, and NRS may best support confirmatory analyses. We show how pain PROMs can be harmonised for analysis of group data, but suggest caution in PROMs harmonisation for investigating individuals with knee pain.

Data availability

Primary data availability for the systematic literature review are in the referenced publications. Primary data for the 4 host studies are available in the primary publications from those studies.

Author contributions

Conception and design: David Walsh, Fiona Watt, Tonia Vincent.

Data acquisition: Vasileios Georgopoulos, Stephanie Smith, Fiona Watt, Andy Williams, Ana Valdes, Martijn Steultjens.

Analysis and interpretation of the data: Vasileios Georgopoulos, Stephanie Smith, Daniel McWilliams, Tonia Vincent, Fiona Watt, David Walsh.

Drafting of the article: Vasileios Georgopoulos.

Critical revision of the article for important intellectual content: All co-authors.

Final approval of the article: All co-authors.

Dr. Vasileios Georgopoulos takes responsibility for the integrity of the work as a whole.

Fiona Watt is Chief Investigator of the KICK and MenTOR studies, James Woodburn is Chief Investigator of the NEKO study and Ana Valdes is Chief Investigator of WebEx.

Conflict of interest

This work was in part carried out to inform activities within the STEp-UP OA consortium. Fiona Watt reports consultancy fees from Pfizer unrelated to this work. Tonia Vincent, Fiona Watt and David Walsh are members of the STEp-UP OA consortium through which this work was supported. STEpUP OA receives some of its research funding from Galapagos, Fidia, Novartis, Pfizer, UCB and Biosplice.

Daniel McWilliams has grants from Pfizer and Eli Lilly. Daniel McWilliams' salary is supported by the NIHR Nottingham Biomedical Research Centre.

Role of the funding source

The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Acknowledgements

This project was funded by the STEpUP OA Consortium grant from the Kennedy Trust for Rheumatology Research (KENN171806) and Versus Arthritis. Fiona Watt is a UKRI Future Leaders Fellow (MR/S016538/1 and MR/S016538/2). The KICK study has received funding from Versus Arthritis (21509) and the Kennedy Trust for Rheumatology Research. The MenTOR study has been supported in part by the NIHR Oxford Biomedical Research Centre and in part by the Centre for Osteoarthritis Pathogenesis Versus Arthritis (grants 20205 and 21621). WebEx was funded in part by grants 21960 and 18769 from the Versus Arthritis UK Pain Centre (paid to the University of Nottingham) and by the National Institute for Health Research Nottingham Biomedical Research Centre. The Joint Academy provided their software for free for this study. WebEx was also supported by Pfizer Global medical grant No. WI243608, Global Awards for Advancing Chronic Pain Research (ADVANCE) 2018. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. The authors are grateful to the contributions and support from patients and staff at Imperial College Healthcare NHS Trust and Oxford University Hospitals NHS Trust. The authors thank James Woodburn for his contribution to the NEKO study. The procedures followed were in accordance with the ethical standards of the responsible committees on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2000.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.joca.2022.08.011.

References

1. Hunter DJ, Schofield D, Callander E. The individual and socioeconomic impact of osteoarthritis. Nat Rev Rheumatol 2014;10:437–41.
2. Hootman JM, Helmick CG. Projections of US prevalence of arthritis and associated activity limitations. Arthritis Rheum 2006;54:226–9.
3. Ridley T, McCarthy MA, Bollier MJ, Wolf BR, Amendola A. Age differences in the prevalence of isolated medial and lateral meniscal tears in surgically treated patients. Iowa Orthop J 2017;37:91.
4. Brelin AM, Rue J. Return to play following meniscus surgery. Clin Sports Med 2016;35:669–78.
5. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: visual analog scale for pain (vas pain), numeric rating scale for pain (nrs pain), mcgill pain questionnaire (mpq), short-form mcgill pain questionnaire (sf-mpq), chronic pain grade scale (cpgs), short form-36 bodily pain scale (sf-36 bps), and measure of intermittent and constant osteoarthritis pain (icoap). Arthritis Care Res 2011;63:S240–52.
6. Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, *et al*. Evaluation of clinically relevant states in patient reported outcomes in knee and hip osteoarthritis: the patient acceptable symptom state. Ann Rheum Dis 2005;64:34–7.
7. Kvien TK, Heiberg T, Hagen KB. Minimal clinically important improvement/difference (MCII/MCID) and patient acceptable symptom state (PASS): what do these concepts mean? Ann Rheum Dis 2007;66:iii40–1.
8. Copay AG, Subach BR, Glassman SD, Polly Jr DW, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. Spine J 2007;7:541–6.
9. Terluin B, Eekhout I, Terwee CB, de Vet HC. Minimal important change (MIC) based on a predictive modeling approach was more precise than MIC based on ROC analysis. J Clin Epidemiol 2015;68:1388–96.
10. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JPA, *et al*. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Ann Intern Med 2009;151:W65–94.
11. Watt F, Corp N, Kingsbury S, Frobell R, Englund M, Felson D, *et al*. Towards prevention of post-traumatic osteoarthritis: report from an international expert working group on considerations for the design and conduct of interventional studies following acute knee injury. Osteoarthritis Cartilage 2019;27:23–33.
12. Hayden JA, Côté P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. Ann Intern Med 2006;144:427–37.
13. Gohir SA, Eek F, Kelly A, Abhishek A, Valdes AM. Effectiveness of internet-based exercises aimed at treating knee osteoarthritis: the iBEAT-OA randomized clinical trial. JAMA Netw Open 2021;4:e210012.
14. Garriga C, Goff M, Paterson E, Hrusecka R, Hamid B, Alderson J, *et al*. Clinical and molecular associations with outcomes at 2 years after acute knee injury: a longitudinal study in the Knee Injury Cohort at the Kennedy (KICK). Lancet Rheumatol 2021;3:e648–58.
15. Smith SL, Allan R, Marreiros SP, Woodburn J, Steultjens MP. Muscle co-activation across activities of daily living in individuals with knee osteoarthritis. Arthritis Care Res 2019;71:651–60.
16. van de Graaf VA, Wolterbeek N, Scholtes VA, Mutsaerts EL, Poolman RW. Reliability and validity of the IKDC, KOOS, and WOMAC for patients with meniscal injuries. Am J Sports Med 2014;42:1408–16.
17. Breivik EK, Björnsson GA, Skovlund E. A comparison of pain rating scales by sampling from clinical trial data. Clin J Pain 2000;16:22–8.
18. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med 2012;22:276–82.
19. Portney LG, Watkins MP. Foundations of Clinical Research: Applications to Practice 2009.
20. Beletsky A, Naami E, Lu Y, Polce EM, Nwachukwu BU, Okoroha KR, *et al*. The patient acceptable symptomatic state in primary anterior cruciate ligament reconstruction: predictors of achievement. Arthroscopy 2021;37:600–5.
21. Boffa A, Andriolo L, Franceschini M, Di Martino A, Asunis E, Grassi A, *et al*. Minimal clinically important difference and patient acceptable symptom state in patients with knee osteoarthritis treated with PRP injection. Orthop J Sports Med 2021;9, 23259671211026242.
22. Chahal J, Lansdown DA, Davey A, Davis AM, Cole BJ. The clinically important difference and patient acceptable symptomatic state

for commonly used patient-reported outcomes after knee cartilage repair. Am J Sports Med 2021;49:193–9.
23. Maheshwer B, Wong SE, Polce EM, Paul K, Forsythe B, Bush-Joseph C, *et al.* Establishing the minimal clinically important difference and patient-acceptable symptomatic state after arthroscopic meniscal repair and associated variables for achievement. Arthroscopy 2021;37:3479–86.
24. Pedersen JR, Roos EM, Thorlund JB, Terluin B, Ingelsrud LH. Cutoff values to interpret short-term treatment outcomes after arthroscopic meniscal surgery, measured with the knee injury and osteoarthritis outcome score. J Orthop Sports Phys Ther 2021;51:281–8.
25. Chahla J, Kunze KN, Tauro T, Wright-Chisem J, Williams BT, Beletsky A, *et al.* Defining the minimal clinically important difference and patient acceptable symptom state for microfracture of the knee: a psychometric analysis at short-term follow-up. Am J Sports Med 2020;48:876–83.
26. Dwyer T, Zochowski T, Ogilvie-Harris D, Theodoropoulos J, Whelan D, Chahal J. Determining the patient acceptable symptomatic state for patients undergoing arthroscopic partial meniscectomy in the knee. Am J Sports Med 2020;48:847–52.
27. Agarwalla A, Gowd AK, Liu JN, Lalehzarian SP, Christian DR, Cole BJ, *et al.* Predictive factors and duration to return to sport after isolated meniscectomy. Orthopaedic Journal of Sports Medicine 2019;7.
28. Connelly JW, Galea VP, Rojanasopondist P, Matuszak SJ, Ingelsrud LH, Nielsen CS, *et al.* Patient acceptable symptom state at 1 and 3 years after total knee arthroplasty: thresholds for the Knee Injury and Osteoarthritis Outcome Score (KOOS). J Bone Joint Surg Am 2019;101:995–1003.
29. Felix J, Becker C, Vogl M, Buschner P, Plötz W, Leidl R. Patient characteristics and valuation changes impact quality of life and satisfaction in total knee arthroplasty - results from a German prospective cohort study. Health Qual Life Outcomes 2019;17:180.
30. Gowd AK, Lalehzarian SP, Liu JN, Agarwalla A, Christian DR, Forsythe B, *et al.* Factors associated with clinically significant patient-reported outcomes after primary arthroscopic partial meniscectomy. Arthroscopy 2019;35:1567.
31. Liu JN, Gowd AK, Redondo ML, Christian DR, Cabarcas BC, Yanke AB, *et al.* Establishing clinically significant outcomes after meniscal allograft transplantation. Orthop J Sports Med 2019;7.
32. Mahler EAM, Boers N, Bijlsma JWJ, Van Den Hoogen FHJ, Den Broeder AA, Van Den Ende CHM. Patient acceptable symptom state in knee osteoarthritis patients succeeds across different patient-reported outcome measures assessing physical function, but fails across other dimensions and rheumatic diseases. J Rheumatol 2018;45:122–7.
33. Muller B, Yabroudi MA, Lynch A, Lai C-L, van Dijk CN, Fu FH, *et al.* Defining thresholds for the patient acceptable symptom state for the IKDC Subjective Knee Form and KOOS for patients who underwent ACL reconstruction. Am J Sports Med 2016;44:2820–6.
34. Bellamy N, Hochberg M, Tubach F, Martin-Mola E, Awada H, Bombardier C, *et al.* Development of multinational definitions of minimal clinically important improvement and patient acceptable symptomatic state in osteoarthritis. Arthritis Care Res 2015;67:972–80.
35. Ingelsrud LH, Granan L-P, Terwee CB, Engebretsen L, Roos EM. Proportion of patients reporting acceptable symptoms or treatment failure and their associated KOOS values at 6 to 24 months after anterior cruciate ligament reconstruction. Am J Sports Med 2015;43:1902–7.
36. Naal FD, Impellizzeri FM, Lenze U, Wellauer V, von Eisenhart-Rothe R, Leunig M. Clinical improvement and satisfaction after total joint replacement: a prospective 12-month evaluation on the patients' perspective. Qual Life Res 2015;24:2917–25.
37. Escobar A, Riddle DL. Concordance between important change and acceptable symptom state following knee arthroplasty: the role of baseline scores. Osteoarthritis Cartilage 2014;22:1107–10.
38. Keurentjes JC, Van Tol FR, Fiocco M, So-Osman C, Onstenk R, Koopman-Van Gemert AWMM, *et al.* Patient acceptable symptom states after total hip or knee replacement at mid-term follow-up: thresholds of the Oxford hip and knee scores. Bone Joint Res 2014;3:7–13.
39. Perrot S, Bertin P. "Feeling better" or "feeling well" in usual care of hip and knee osteoarthritis pain: determination of cutoff points for patient acceptable symptom state (PASS) and minimal clinically important improvement (MCII) at rest and on movement in a national multicenter cohort study of 2414 patients with painful osteoarthritis. Pain 2013;154:248–56.
40. Escobar A, Gonzalez M, Quintana JM, Vrotsou K, Bilbao A, Herrera-Espiñeira C, *et al.* Patient acceptable symptom state and OMERACT-OARSI set of responder criteria in joint replacement. Identification of cut-off values. Osteoarthritis Cartilage 2012;20:87–92.
41. Tubach F, Ravaud P, Martin-Mola E, Awada H, Bellamy N, Bombardier C, *et al.* Minimum clinically important improvement and patient acceptable symptom state in pain and function in rheumatoid arthritis, ankylosing spondylitis, chronic back pain, hand osteoarthritis, and hip and knee osteoarthritis: results from a prospective multinational study. Arthritis Care Res 2012;64:1699–707.
42. Ingelsrud LH, Terluin B, Gromov K, Price A, Beard D, Troelsen A. Which Oxford Knee Score level represents a satisfactory symptom state after undergoing a total knee replacement? Acta Orthop 2021;92:85–90.
43. Goh GS, Liow MHL, Chen JY, Tay DK-J, Lo N-N, Yeo S-J. The patient acceptable symptom state for the knee society score, oxford knee score and short form-36 following unicompartmental knee arthroplasty. Knee Surg Sports Traumatol Arthrosc 2021:1–10.
44. Wylde V, Sayers A, Lenguerrand E, Gooberman-Hill R, Pyke M, Beswick A, *et al.* Preoperative widespread pain sensitization and chronic pain after hip and knee replacement: a cohort analysis. Pain 2015;156:47–54.
45. MacKay C, Clements N, Wong R, Davis A. A systematic review of estimates of the minimal clinically important difference and patient acceptable symptom state of the Western Ontario and McMaster Universities Osteoarthritis Index in patients who underwent total hip and total knee replacement. Osteoarthritis Cartilage 2019;27:1408–19.
46. Roos EM, Lohmander LS. The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. Health Qual Life Outcomes 2003;1:1–8.
47. Roos EM, Toksvig-Larsen S. Knee injury and Osteoarthritis Outcome Score (KOOS)–validation and comparison to the WOMAC in total knee replacement. Health Qual Life Outcomes 2003;1:1–10.
48. Gandek B, Ware Jr JE. Validity and responsiveness of the knee injury and osteoarthritis outcome score: a comparative study among total knee replacement patients. Arthritis Care Res 2017;69:817–25.
49. Marot V, Murgier J, Carrozzo A, Reina N, Monaco E, Chiron P, *et al.* Determination of normal KOOS and WOMAC values in a healthy population. Knee Surg Sports Traumatol Arthrosc 2019;27:541–8.

50. Singh JA, Luo R, Landon GC, Suarez-Almazor M. Reliability and clinically important improvement thresholds for osteoarthritis pain and function scales: a multicenter study. J Rheumatol 2014;41:509–15.

51. Soh S, Harris I, Cashman K, Graves S, Ackerman I. Crosswalks between the Oxford hip and knee scores and the HOOS-12 and KOOS-12 instruments. Osteoarthritis Cartilage 2022;30:570–7.