BMJ Open

Study protocol for the validation of a new pictorial functional scale in patients with knee osteoarthritis: the functional activity scoring tool (FAST)

Zhi Yin Tang [1], Khim Siong Ng [1], Yi Ling Eileen Koh [2], Meredith T Yeung [3]

To cite: Tang ZY, Ng KS, Koh YLE, *et al*. Study protocol for the validation of a new pictorial functional scale in patients with knee osteoarthritis: the functional activity scoring tool (FAST). *BMJ Open* 2024;**14**:e076947. doi:10.1136/bmjopen-2023-076947

► Prepublication history for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2023-076947).

Received 21 June 2023
Accepted 10 December 2023

[1]Department of Physiotherapy, SingHealth Polyclinics, Singapore
[2]SingHealth Polyclinics, Singapore
[3]Health and Social Sciences Cluster, Singapore Institute of Technology, Singapore

Correspondence to
Zhi Yin Tang;
tang.zhi.yin@singhealth.com.sg

ABSTRACT

Background Patient-reported outcome measures (PROMs) are required for patient-centred care. There are limited PROMs with good psychometric properties, and limitations to any language-based scale are often constrained by the written words or numerals used. Therefore, we developed the Functional Activity Scoring Tool (FAST), a self-reporting pictorial scale. FAST measures the impact of knee osteoarthritis on essential activities of daily living (ADL) and the significant changes in the self-perceived functional status over time.

Objectives This study aims to (1) develop FAST with adaptation from the Wong-Baker FACES pain rating scale, (2) validate FAST against the Patient-Specific Functional Scale (PSFS) and Knee Injury and Osteoarthritis Outcome Score (KOOS) and (3) establish the reliability, validity and responsiveness of FAST in individuals with knee osteoarthritis.

Methods and analysis The prospective study protocol investigates the validity, responsiveness and reliability of FAST. The PSFS and KOOS will be gold standard comparisons. Participant recruitment will occur at four public polyclinics that offer physiotherapy outpatient services in Singapore. Onsite physiotherapists familiar with the study eligibilities will refer potential participants to the investigators after the routine physiotherapy assessment. After providing written consent, eligible participants will complete outcome measurements with FAST, the PSFS and KOOS during baseline and follow-up assessments. The Global Rating of Change (GROC) scale will determine how the participant's knee status was changed compared with the beginning of the physiotherapy intervention.

Ethics and dissemination SingHealth Centralised Institutional Review Board approved the study (CIRB reference number: 2022/2602). The final results will be published via scientific publication. FAST will benefit the evaluation and management of those who suffer knee osteoarthritis regardless of English proficiency or language barriers.

Trial registration number NCT05590663

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ The first pictorial patient-specific functional assessment tool—the Functional Activity Scoring Tool (FAST)—is developed.
⇒ This multisite study will validate the novel pictorial scale created and reviewed by an expert panel comprising patients and their families, physiotherapists and family physicians.
⇒ The proposed study aligns with international consensus standards on best practices of instrument development and validation studies—the COnsensus-based Standards for the selection of health Measurement INstruments.
⇒ Validation of patient-reported outcome measures is an iterative process. More testing of its psychometric properties must follow to support its usefulness in patients with other musculoskeletal conditions.
⇒ Although the study protocol will not alter the standardised physiotherapy treatment, we cannot rule out possible confounding variables that may influence the study outcomes.

INTRODUCTION

Healthcare professionals regularly assess the crucial yet trouble-functioning tasks in activities of daily living (ADL). While various condition-specific questionnaires, such as the Roland-Morris Disability Questionnaire[1] or Knee Injury and Osteoarthritis Outcome Score (KOOS),[2] and health status measures, such as the 36-Item Short Form Survey[3] or EuroQol-5D,[4] exist; unfortunately, limited patient-reported outcome measures (PROMs) have been established thus far, especially in the area of osteoarthritis. The application of PROM in orthopaedic is expected to increase.[5] PROMs were initially created for research purposes and eventually adopted for clinical management, seeking to determine the patients' perceptions of their symptoms, functional status and health-related quality of life. PROMs are frequently unfittingly referred to as 'outcome measures', even though they measure health—by comparing a patient's health at different times, the care outcome received can be determined.[6]

PROMs provide additional 'patient-centred' data that is unique in capturing the patients' perspective on the impact of their disease or disorder and its treatment.[5] These self-reported instruments elicit information about a patient's health status directly from the patient without needing interpretation from a healthcare professional.[6] The approach of gathering patient-centred data is integral in informing clinical care and supplementing measurable clinical improvements in the patients as part of the routine practice. Well-validated PROM assessing functional outcomes is required in the era of patient-centred care for holistic management.

Few osteoarthritis-specific PROMs have been developed and extensively studied. A systematic review[7] identified that these PROMs attempt to measure psychometric properties such as pain, mental functions and moods, physical symptoms such as stiffness and mobility, as well as function in sports and recreation with either the term or subscale level. Overall, the review findings found limited evidence of psychometric properties from these PROMs. Concurrently, straightforward tools to report on self-efficacy were limited. Among these, the Patient-Specific Functional Scale (PSFS)[8] was uniquely developed to enable self-reporting of the impact of musculoskeletal conditions on essential ADL and the significant changes in the self-perceived functional status over time. A recent systematic review[9] of the PSFS against the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) criteria concluded that it is an easy-to-use, reliable and responsive scale in numerous musculoskeletal conditions. The PSFS is applicable across various conditions and body regions as it allows the comparison of functional outcomes across conditions and between studies.[10–12] However, the usefulness of a language-based PROM is restricted if patients and families have limited English proficiency, which is a barrier to healthcare services that is well documented.[13] Singapore is primarily an English-speaking country. However, her geographical location and historical and cultural backgrounds greatly influence the languages used in this city-state. Our anecdotal experience suggests that older patients (especially individuals ≥65 years old) have difficulty understanding and accurately completing the PSFS, as English may not necessarily be their primary spoken language. Difficulty with completing forms can occur for many reasons, such as the written words and numerals not being universally understood, and problems with health literacy are common and underestimated.[14–16] Very often, informal interpretation, such as relying on a family member to translate the communication, has shown to be associated with a more significant number of errors.[17] Pictures or pictorial aids are a useful adjunct to medical information and aid the transfer and comprehension of written and spoken information.[18] One good example is using the pictorial scale to measure pain in the Wong-Baker FACES pain rating scale,[19] with proven benefits such as improving adherence to medications and enhancing the understanding of instructions.[20]

As such, there is a growing need for a new PROM that is simple, reliable and responsive, yet minimises the limitations of any word or language-based outcome measures that are currently in use. The Functional Activity Scoring Tool (FAST) has been developed to address this situation. FAST is a new pictorial-scale PROM measuring function in an individual with osteoarthritis. Several aspects were considered during the conceptualisation of the instrument: the applicability to a broad range of clinical presentations (conditions, limitations and age), simple administration, concise yet effectual for speedy medical documentation and simple interface in electronic medical record systems.

Hypothesis and aim

The confidence of a PROM depends on the psychometric evaluation of its measurement properties, and it must be undertaken to satisfy rigorous criteria.[21] These include validity (to what extent does the instrument measure the construct it purports to measure), reliability (the degree to which measurement is free from error) and responsiveness (the ability of an outcome measure to detect change over time in the construct to be measured).[22] The process to assess these measurement properties must be iterative and studied individually. Thus, we hypothesise that the measure of function, an additional dimension to the quality of life, is possible by the same principle. The new FAST scale can be used to measure function and difficulty in performing ADL in patients with knee osteoarthritis, in an equally valid and reliable manner as the PSFS and KOOS. We aim to provide a standardised tool for gathering and documenting the patients' symptoms. With these considerations, we developed the FAST scale. This study aims to (1) develop the FAST pictorial functional scale with adaptation from the Wong-Baker FACES pain rating scale, (2) validate FAST against the PSFS and KOOS and (3) establish the reliability, validity and responsiveness of FAST in individuals with knee osteoarthritis.

METHOD AND ANALYSIS

Study design and setting

This study will be a prospective validation study to establish the psychometric properties of a newly developed PROM. This study is proposed under the recommendation of Basch *et al*[23] methods for developing patient-reported outcome-based performance measures and uses the procedures that de Vet and colleagues[24] advocated for in developing a PROM. This approach provides evidence for developing a PROM that measures the intended context and its use as an outcome measure in clinical practice and research trials. The study will take place in four physiotherapy outpatient clinics in Singapore over 12 months.

Patient and public involvement

Patients and families from physiotherapy outpatient clinics provided input and suggestions to the FAST scale

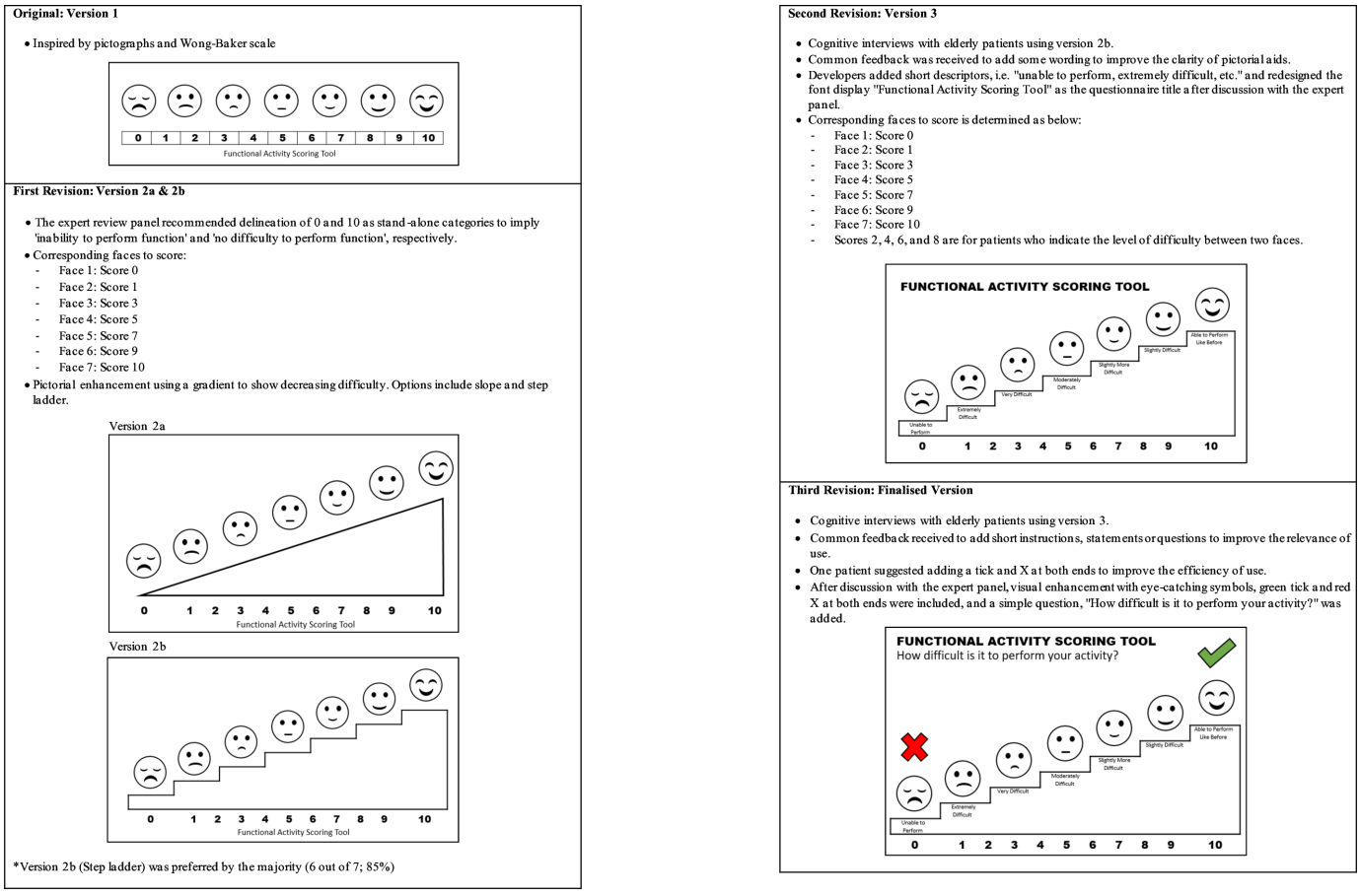

Figure 1 Functional Activity Scoring Tool (FAST) conceptualisation process: versions and revisions.

during its conceptualisation and feasibility stage. Hence, their feedback also shaped the scale design, with the pros and cons of the different versions of the FAST scale discussed with patients and/or their families who will not be recruited as study participants.

Development of FAST

During the feasibility stage, surveys were conducted on patients, families and healthcare professionals to gather feedback on the application of the PSFS. The most prevalent verbatim was 'difficulty comprehending PSFS due to its being too lengthy and the lack of pictorial aid to assist patient's comprehension of the scale'. Therefore, a prototype of the FAST scale was created and reviewed by an expert panel of academics, researchers and clinicians (n=7) and a series of cognitive interviews with a purposive sample of patients older than 65 years (n=12) to elicit feedback on its relevance, clarity and acceptability. The final version of FAST was developed after three revisions. Figure 1 presents the conceptualisation and revision process of the FAST development. The final version of the FAST scale from this revision process will be used to test for reliability and validity in this study protocol. It consists of a pictorial diagram with seven expression faces corresponding to an 11-point Likert scale, with Face 1 (the saddest expression) on the left of the scale paired with a score '0' and a verbal description of 'unable to perform', Face 4 (neutral expression) to agree with score '5' with a description of 'moderately difficult' and Face 7 (the happiest expression) on the right of the scale matching with score '10' and a descriptor of 'able to perform like before'. The red 'cross' on the left and green 'tick' on the right accentuate the effects of the facial expression and association with the verbal descriptors.

Sample size

The size of the retest sample was estimated based on a method developed to calculate the required number of participants in a reliability study.[25] The probabilities of type I and type II error were α=0.05 and β=0.20, respectively. An intraclass correlation coefficient (ICC) value of <0.50 indicated poor reliability, whereas values between 0.50 and 0.75 indicated fair to good reliability; an ICC value >0.9 showed excellent reliability.[26] We hypothesised that our findings would be consistent with a minimum coefficient of 0.75. This level of reliability is at least appropriate for person-level comparisons. Following these assumptions, a minimum of 50 participants will be necessary for the test-retest analysis for this study. According to COSMIN guidelines, validity calculations are considered good-excellent if the sample size exceeds 100 (n=100).[27] To allow for a possible attrition rate of 20%, a minimum sample size of 120 will be needed.

Participant recruitment and selection criteria

Participant recruitment will occur at four public polyclinics that offer physiotherapy outpatient services in different districts of Singapore. Onsite physiotherapists familiar with the study protocol will identify eligible participants during the routine initial physiotherapy assessment. Inclusion criteria based on the National Institute for Health and Care Excellence criteria[28] will be individuals diagnosed with knee osteoarthritis and referred for physiotherapy care at the polyclinics, age 45 years and above and proficient in colloquial/conversational English. Potential participants will be excluded if there are additional underlying medical or trauma conditions of the knees (eg, trauma, fracture, infection, inflammatory disease and tumour), history of knee surgery within the last 3 months or clinically recognisable cognitive impairment that inhibits the comprehension and completion of the questionnaires. Participation in the study is strictly voluntary and will not impact the type or quality of the individual's physiotherapy treatments based on prevailing evidence.

Instruments

The self-administered KOOS is a knee-specific instrument developed to assess the patients' opinions about their knees and associated short-term and long-term problems.[2] It is a validated tool in Singapore for knee osteoarthritis patients.[29] It consists of 42 items in 5 subscales, that is, pain (9 questions), symptoms (7 questions), activities in daily living (17 questions), sports and recreation function (4 questions) and knee-related quality of life (4 questions). The 5-point Likert scale scoring system ranges from '0' (no problem) to '2' (moderate problem) to '4' (extreme problem), and the score for each domain is calculated by summing the questions. Scores will be converted to a 0 to 100 scale, with 0 representing extreme knee problems and 100 representing no knee problems. The use of the 0 to 100 score is practical as it projects a direct reference to the percentage concept.[2]

The PSFS is a self-reported, patient-specific measure that assesses patients' functional status.[8] Patients are asked to identify three activities most affected by their conditions and then rate their ability on an 11-point Likert 0 to 10 scale for each activity, where '0' is unable to perform the activity and '10' being able to perform the activity at the same level as before the onset of symptoms. The total score is computed by dividing the sum of the activity scores by the number of activities listed.

The Global Rating of Change (GROC) scale is an outcome measure that assesses the patients' self-perception of change in their condition between sessions.[30] The GROC scale is quantified on a 15-point Likert scale from '-7' (a very great deal worse) to '0' (about the same) to '7' (a very great deal better). The scale is easy to administer as it requires minimal skills or training, has good reproducibility and is sensitive to changes.[31]

Procedure

Eligible individuals will be informed of the study's purpose and data collection procedures. Written informed consent will be obtained from every participant before data collection commences. The participants' confidentiality and anonymity will be maintained throughout the study process with a unique identifier, and only the study researchers will have access to the data. Participants will receive standardised care, and their participation status will not be shared with the attending physiotherapists apart from the initial identifications for eligibility. All data collection forms will be coded with the same unique identifier, and the study team will not retain any identifiable information. Only anonymised data will be used for data analysis. The project investigators will perform all data collection. Demographic data, clinical characteristics and primary outcome measurements with FAST, the PSFS and KOOS will be collected during baseline assessment (week 0). Follow-up assessment with FAST, the PSFS and KOOS will be scheduled 2–3 weeks post initial assessment together with the administration of GROC to evaluate the efficacy of the standard physiotherapy treatment that the participants will be receiving regardless of the participation status in this study. With reference to a prior study,[10] the 2-week to 3-week period is chosen as it is also the typical duration between the initial and follow-up physiotherapy session in the local setting. Figure 2 depicts the workflow of the data collection procedures. This study will not require any alteration or deviation from the standard protocol for knee osteoarthritis physiotherapy management.

Statistical analyses

All statistical analysis will be conducted using IBM SPSS 29.0 with the statistical significance set as $p<0.05$. Descriptive statistics will be used to describe the demographic variables using mean and SD or median and IQR for continuous variables and frequencies and percentages for categorical variables. To determine the profile of the subjects with the FAST scoring, Mann-Whitney U test or Kruskal-Wallis test can be used for the continuous FAST score and the categorical demographics (ie, gender, ethnicity, marital status and education level), while Spearman's correlation can be used to compare the continuous FAST and demographics (ie, age).

Validation

Face validity

The qualitative methods used to determine the face validity of FAST involved face-to-face meetings with an expert panel of academics, researchers and clinicians (n=7) and a series of cognitive interviews with patients (n=12). Three essential criteria were determined in establishing face validity: clarity (the extent to which an item is open to more than one possible interpretation), relevancy (the extent to which an item will be relevant to its component) and acceptability (the extent to which readers would easily understand an item).

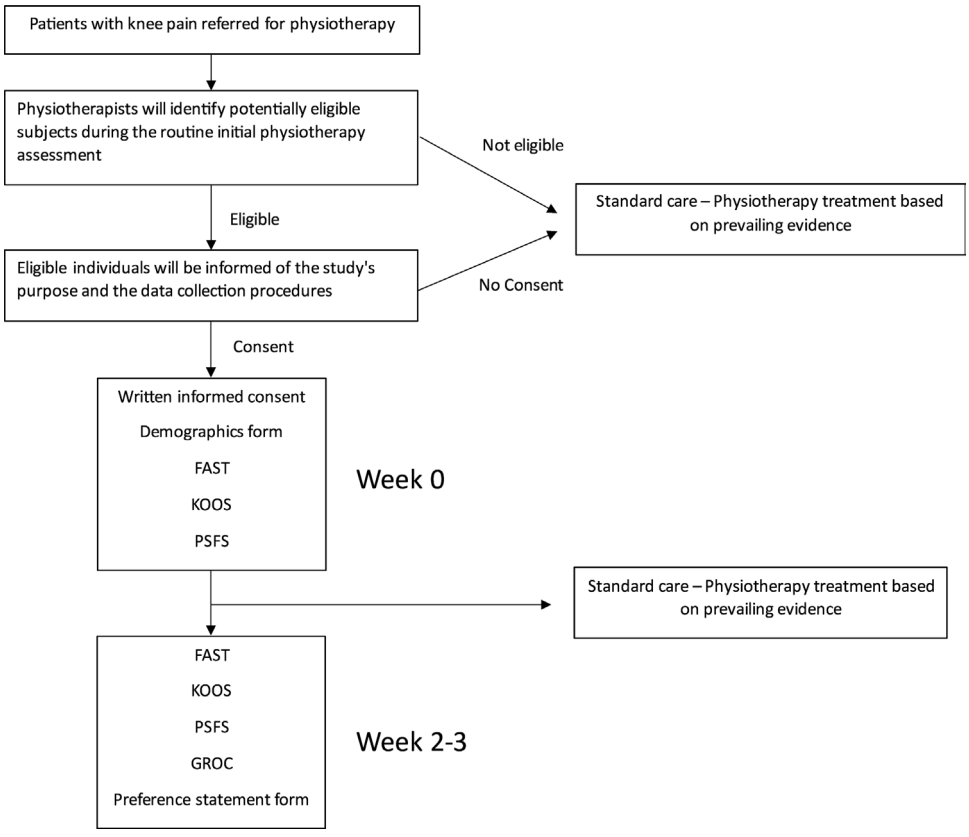

Figure 2 Workflow of the data collection procedures.

Content validity

The content validity index (CVI) and context validity ratio (CVR) will determine the content validity.[32] CVI is the most widely reported method for determining content validity in instrument development and assessing its relevance and clarity. There are two methods of calculation, namely, item CVI (I-CVI) and scale-level CVI (S-CVI).[33] This study will use a 4-point Likert scale, '1', unacceptable; '2', needs some revision; '3', needs minor revision; and '4', acceptable, for the calculation of the I-CVI from the total rating scores from all panel members. Where I-CVI is >0.79, the item is acceptable; between 0.70 and 0.79, the item will require revision; and when it is <0.70, the item will be eliminated.[34] Similarly, the S-CVI will be determined by the number of items in an instrument that receives a 'highly acceptable' grade. The universal agreement (UA) among the panel members (S-CVI/UA) and the average CVI (S-CVI/Ave) are two ways of determining S-CVI.[33] S-CVI/UA will be calculated by the sum of all items with I-CVI equal to 1 divided by the total number of items, and S-CVI/Ave is equal to the sum of all the I-CVI divided by the number of items. Content validity is excellent when the S-CVI/UA is >0.8 and the S-CVI/Ave is >0.90.[34]

CVR quantifies the essentiality of an item.[35] CVR ranges from −1 to 1; a higher score represents a greater agreement between panel members. CVR = (Ne − N/2)/(N/2), where Ne is the number of panel members who rated an item as 'essential', and N is the total number of panel members.[33] Each element of the FAST scale will be evaluated on a 3-point Likert scale (1, not essential; 2, useful but not essential; and 3, essential).

Criterion validity

The KOOS Singapore English version and PSFS will serve as the criterion for disability in the knee osteoarthritis population. The two validated self-administered questionnaires are specific and sensitive to change over time. The correlations between FAST, KOOS and the PSFS will assess the criterion validity of the FAST scale. Spearman's correlation will investigate the criterion validity against the PSFS, KOOS and GROC and the measurement of agreement according to the following criteria: high (rho ≥0.60), moderate (rho <0.60 to ≥0.30) or low (rho <0.30).[36] The higher the rho, the higher the agreement between the two instruments.

Responsiveness

Responsiveness is defined as the ability to measure and recognise change when a change has occurred. Similarly, Spearman's correlation coefficients (rho) can be used to determine strong (rho ≥0.60), moderate (rho <0.60 to ≥0.30) or weak (rho <0.30) correlations.[36]

Reliability

The test-retest reliability of FAST will be calculated via intraclass correlation coefficient (ICC) for absolute agreement using a two-way mixed-effect analysis of the variance

model between the scores of two stable assessment periods (ie, GROC <3). ICC values >0.75 are indicative of good-excellent reliability.[26] Participants who scored between −3 and +3 on GROC were included in the test-retest analysis and were assumed that they did not demonstrate any clinically relevant changes during this interval period.[37]

Cronbach alpha measures the internal consistency of the instrument; a value of >0.7 is considered to be acceptable. For good internal consistency, the value should be >0.8, and for excellent internal consistency, the value should be >0.9.[38]

Measurement errors were determined by calculating the standard error of measurement (SEM) and the minimal detectable change (MDC). MDC is calculated using the formula $MDC = z \approx \sqrt{2} \times \sqrt{MSE}$, where z=1.64 and is the score associated with a 90% confidence interval, $\sqrt{2}$ reflects the uncertainty introduced by using scores at two different points in time and the square root of the mean square error (MSE) term represents SEM.[39] MSE was found by constructing a one-way analysis of variance (ANOVA) table of the baseline and follow-up scores of the stable group.[39]

Ethics and dissemination

The SingHealth Centralised Institutional Review Board (CIRB) approved this research protocol (CIRB reference number: 2022/2602). There are no potential risks for participants taking part in this study. All participants will provide written consent to participate and have the right to withdraw from participation in the project at any time without any compromise or disadvantage to them in any form. All participants will be assigned a unique de-identified code to protect the confidentiality of the participants. Access to the data is restricted to the project investigators, and only anonymised data will be used during data analysis. All investigators declare no financial or other competing interests at all study sites. This study will validate the new pictorial functional scale (FAST) in patients with knee osteoarthritis and hope to investigate if the new scale correlates with similar existing PROM with good validity and reliability. The final results and establishment of the new PROM will be published via scientific publication. This will be advantageous to healthcare professionals in evaluating functional status changes in individuals with osteoarthritis regardless of English proficiency or language barriers.

Trial status

The study is at its pilot trial stage at the time of submission of this study protocol.

Acknowledgements The authors thank and acknowledge SingHealth Polyclinics and the Singapore Institute of Technology for supporting this study.

Contributors ZYT and KSN drafted the manuscript; YLEK and MTY critically revised the manuscript. All authors approved the final version of the article.

Funding The publication cost is supported by SingHealth Polyclinics – Centre Grant (CG21APR3006 (NMRC/CG3/001/2022-SHP) - Research Seed Fund SHPCG-RSF22-01.

Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Zhi Yin Tang http://orcid.org/0000-0002-8641-8146
Khim Siong Ng http://orcid.org/0000-0002-0374-326X
Yi Ling Eileen Koh http://orcid.org/0000-0001-9596-7158
Meredith T Yeung http://orcid.org/0000-0002-4294-4187

REFERENCES

1 Roland M, Morris R. A study of the natural history of back pain. part I: development of a reliable and sensitive measure of disability in low-back pain. *Spine (Phila Pa 1976)* 1983;8:141–4.
2 Roos EM, Roos HP, Lohmander LS, *et al*. Knee injury and osteoarthritis outcome score (KOOS)—Development of a self-administered outcome measure. *J Orthop Sports Phys Ther* 1998;28:88–96.
3 Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. *Medical Care* 1992;30:473–83.
4 EuroQol Group. Euroqol--a new facility for the measurement of health-related quality of life. *Health Policy* 1990;16:199–208.
5 Kyte DG, Calvert M, van der Wees PJ, *et al*. An introduction to patient-reported outcome measures (Proms) in Physiotherapy. *Physiotherapy* 2015;101:119–25.
6 Black N. Patient reported outcome measures could help transform Healthcare. *BMJ* 2013;346:bmj.f167.
7 Lundgren-Nilsson Å, Dencker A, Palstam A, *et al*. Patient-reported outcome measures in osteoarthritis: a systematic search and review of their use and Psychometric properties. *RMD Open* 2018;4:e000715.
8 Stratford P, Gill C, Westaway M. Assessing disability and change on individual patients: A report of a patient specific measure. *Physiotherapy Canada* 1995;47:258–63.
9 Pathak A, Wilson R, Sharma S, *et al*. Measurement properties of the patient-specific functional scale and its current uses: an updated systematic review of 57 studies using COSMIN guidelines. *Journal of Orthopaedic & Sports Physical Therapy* 2022;52:262–75.
10 Chatman AB, Hyams SP, Neel JM, *et al*. The patient-specific functional scale: measurement properties in patients with knee dysfunction. *Phys Ther* 1997;77:820–9.
11 Hefford C, Abbott JH, Arnold R, *et al*. The patient-specific functional scale: validity, reliability, and responsiveness in patients with upper extremity musculoskeletal problems. *J Orthop Sports Phys Ther* 2012;42:56–65.
12 Westaway MD, Stratford PW, Binkley JM. The patient-specific functional scale: validation of its use in persons with neck dysfunction. *J Orthop Sports Phys Ther* 1998;27:331–8.
13 Espinoza J, Derrington S. How should Clinicians respond to language barriers that exacerbate health inequity. *AMA J Ethics* 2021;23:E109–116.
14 Williams MV, Parker RM, Baker DW, *et al*. Inadequate functional health literacy among patients at two public hospitals. *JAMA* 1995;274:1677–82.
15 Baker DW, Gazmararian JA, Sudano J, *et al*. The association between age and health literacy among elderly persons. *J Gerontol B Psychol Sci Soc Sci* 2000;55:S368–74.
16 Williams MV, Davis T, Parker RM, *et al*. The role of health literacy in patient-physician communication. *Fam Med* 2002;34:383–9.
17 Whitaker KL, Krystallidou D, Williams ED, *et al*. Addressing language as a barrier to Healthcare access and quality. *Br J Gen Pract* 2022;72:4–5.
18 Ghiassi R, Murphy K, Cummin AR, *et al*. Developing a pictorial Epworth Sleepiness scale. *Thorax* 2011;66:97–100.
19 Garra G, Singer AJ, Taira BR, *et al*. Validation of the Wong-Baker FACES pain rating scale in pediatric emergency Department patients. *Acad Emerg Med* 2010;17:50–4.

20 Houts PS, Doak CC, Doak LG, *et al*. The role of pictures in improving health communication: A review of research on attention, comprehension, recall, and adherence. *Patient Educ Couns* 2006;61:173–90.
21 Mokkink LB, Terwee CB, Patrick DL, *et al*. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010;19:539–49.
22 Mokkink LB, de Vet HCW, Prinsen CAC, *et al*. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27:1171–9.
23 Basch E, Spertus J, Dudley RA, *et al*. Methods for developing patient-reported outcome-based performance measures (PRO-PMs). *Value Health* 2015;18:493–504.
24 de Vet HCW, Terwee CB, Mokkink LB, *et al*. In: *Measurement in Medicine: A Practical Guide*. Cambridge: Cambridge University Press, 11 August 2011.
25 Bonett DG. Sample size requirements for estimating Intraclass correlations with desired precision. *Stat Med* 2002;21:1331–5.
26 Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice*. Upper Saddle River (NJ): Pearson/Prentice Hall, 2009.
27 Gagnier JJ, Lai J, Mokkink LB, *et al*. COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. *Qual Life Res* 2021;30:2197–218.
28 Skou ST, Koes BW, Grønne DT, *et al*. Comparison of three sets of clinical classification criteria for knee osteoarthritis: a cross-sectional study of 13,459 patients treated in primary care. *Osteoarthritis and Cartilage* 2020;28:167–72.
29 Xie F, Li S-C, Roos EM, *et al*. Cross-cultural adaptation and validation of Singapore English and Chinese versions of the knee injury and osteoarthritis outcome score (KOOS) in Asians with knee osteoarthritis in Singapore. *Osteoarthritis and Cartilage* 2006;14:1098–103.
30 Jaeschke R, Singer J, Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. *Control Clin Trials* 1989;10:407–15.
31 Kamper SJ, Maher CG, Mackay G. Global rating of change scales: A review of strengths and weaknesses and considerations for design. *Journal of Manual & Manipulative Therapy* 2009;17:163–70.
32 Waltz CF, Bausell BR. Nursing research: design statistics and computer analysis. F.A. Davis Company, 1981.
33 Zamanzadeh V, Ghahramanian A, Rassouli M, *et al*. Design and implementation content validity study: development of an instrument for measuring patient-centered communication. *J Caring Sci* 2015;4:165–78.
34 Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? appraisal and recommendations. *Res Nurs Health* 2007;30:459–67.
35 Yamada J, Stevens B, Sidani S, *et al*. Content validity of a process evaluation checklist to measure intervention implementation fidelity of the EPIC intervention. *Worldviews Evid Based Nurs* 2010;7:158–64.
36 Hinkle DE, Wiersma W, Jurs SG. *Applied statistics for the behavioral sciences*. Boston (MA): Houghton Mifflin, 2003.
37 Lim HHR. Cross-cultural adaptation, reliability, validity, and responsiveness of the simplified-Chinese version of neck disability index *Spine (Phila Pa 1976)* 2020;45:E477–8.
38 Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297–334.
39 Schmitt JS, Di Fabio RP. Reliable change and minimum important difference (MID) proportions facilitated group responsiveness comparisons using individual threshold criteria. *J Clin Epidemiol* 2004;57:1008–18.