ARTICLE IN PRESS

Journal of Hand Therapy xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of Hand Therapy

journal homepage: www.elsevier.com/locate/jht

Psychometric properties of body structures and functions measures in non-surgical thumb carpometacarpal osteoarthritis: A systematic review

Corey McGee, PhD, OTR/L, CHT [a,*], Kristin Valdes [b], Caitlin Bakker [c], Cindy Ivy [d]

[a] Programs in Occupational Therapy and Rehabilitation Science, University of Minnesota, Minneapolis, MN, USA
[b] Program in Occupational Therapy, Touro University, Henderson, NV, USA
[c] Archer Library, University of Regina, Regina, SK, Canada
[d] Program in Occupational Therapy, Northern Arizona University, Phoenix, AZ, USA

ARTICLE INFO

Article history:
Received 17 January 2023
Revised 19 July 2023
Accepted 20 July 2023
Available online xxxx

Keywords:
Psychometrics
Thumb
Osteoarthritis
Systematic review
Body structures and functions
Measurement
Carpometacarpal

ABSTRACT

Background: Measurement of treatment outcomes and change in health status over time is a critical component of clinical practice and research for people with osteoarthritis. Numerous clinical tools are used to assess the structures and function of the thumb in persons with thumb carpometacarpal osteoarthritis however their psychometrics have not yet been systematically explored.
Purpose: The purpose of this study was to explore the psychometric properties of clinical tools used in persons with non-surgical thumb carpometacarpal osteoarthritis to objectively measure thumb structures and function, evaluate the quality of such studies, and subsequently make clinical and future research recommendations.
Study Design: Systematic review.
Methods: A systematic search and screening was conducted across nine databases. Original research published between 2002 and 2022 that involved the assessment of psychometric properties (validity, reliability, precision, responsiveness, sensitivity, specificity, and minimal clinically important difference) of clinical tools were included. Sample characteristics, methods, and psychometric findings from each study were compiled. The methodological quality of included studies was evaluated using the COnsensus-based Standards for the selection of health Measurement Instruments' checklist. Two independent researchers screened articles and assessed methodological quality and when not in agreement, a third party was consulted.
Results: Eleven studies were included in the review. The mean age of all participants in the studies was 69 years of age. The study designs included prospective case–control, prospective cohort, and cross-sectional to determine the psychometric properties of the measurements and tools. The included studies examined techniques to assess range of motion, strength, and pain-pressure thresholds, and screen for arthritis (ie, provocative tests). The intermetacarpal distance method, Kapandji index, pain-pressure threshold test, and pain-free grip and pinch dynamometry demonstrate excellent reliability and acceptable precision. Metacarpal extension, adduction, and pressure-shear provocative tests have superior sensitivity and specificity and the extension and adduction tests have excellent reliability. Other assessments included in the review yielded less robust psychometric properties. Studies were of variable methodological quality spanning from inadequate to very good.
Conclusions: Based on the available literature on the psychometric properties of assessments of body structures and functions in persons with non-operative thumb carpometacarpal osteoarthritis, we offer a limited set of recommendations for use when screening for arthritis symptomology and measuring hand strength, thumb mobility, and pain thresholds. Additional psychometric research is needed in these domains as well as in dexterity, sensation, and objective measures of hand function. Future research should employ best practices in psychometric research.

* Corresponding author. University of Minnesota, MMC 368 426 Church St SE Minneapolis, MN 55455.
E-mail address: mcge0062@umn.edu (C. McGee).

0894-1130/© 2023 Elsevier Inc. All rights reserved.
https://doi.org/10.1016/j.jht.2023.07.002

Introduction

Thumb carpometacarpal (CMC) osteoarthritis (OA) can be a painful and debilitating condition associated with declines in joint mobility, coordination, and joint receptors, and functional use of the hand.[1–4] The measurement of relevant objective data is vital to determine deficits, note progress, and determine treatment outcomes. Measurement of treatment outcomes and change in health status over time is a critical component of research and clinical practice for people with thumb CMC OA.[5] The use of outcome measures for clinical trials of OA that address the domain of function is promoted by both the Osteoarthritis Research Society International[6] and Outcome Measures in Rheumatology and Clinical Trials.[7] Currently there is no singular gold standard for the assessment of body structures and function[8] (ie, anatomical parts and physiological functions of body systems) in thumb CMC OA.[9]

There are a variety of tools that can be used to assess the body structures and functions (BSF) of the affected thumb CMC joint and there is some variability in how often these are used by hand therapists. The authors of a cross-sectional survey of hand therapy practice patterns of therapists treating patients with thumb CMC OA sought to describe this variability.[10] In this study, the authors reported that over 85% of therapists perform goniometric measurement of the thumb and surrounding joints, approximately 7% of therapists use methods other than goniometry, and 3% of therapists did not measure range of motion (ROM) at all.[10] Most, but not all, therapists reported measuring thumb opposition, but therapists used a variety of measures including verbal description, the use of callipers or a ruler, and a small percentage used the Kapandji opposition scale.[11] Almost all of the respondents reported that they measured grip strength and pinch strength and about a third reported performing manual muscle testing of thumb musculature. Therapists commonly used provocative tests to screen in or out thumb CMC OA symptomology and they reported that the CMC grind test[12] was used more frequently, followed by Finkelstein's test,[13] and ligament laxity tests.[10] The authors concluded that more consistent use of psychometrically-sound BSF outcome measures in thumb CMC OA is needed.

Similarly, the authors of a systematic review that linked the outcome measures used in studies on thumb CMC OA orthotic interventions to the International Classification of Functioning found that the thumb CMC researchers also use a variety of measures that focused on BSF. These measures included grip and pinch strength assessment, range of motion measurements taken with a goniometer, thumb ROM assessed with an infrared camera system, ROM assessed with the Kapandji scale, the O'Conner Dexterity test, and the Sollerman test of hand function.[14] The assessment of grip and pinch strength were the most commonly used outcome measures in eight of the nine studies included in the review.[14]

A scoping review of the clinical measures for thumb CMC OA reported that researchers used 52 different BSF tests for the evaluation of CMC OA.[15] While, as the authors acknowledge, these numbers are notably low, the review may have overestimated the number of BSF tests used in thumb CMC OA as it included numerous articles on measures that (1) could be useful in thumb CMC OA but had actually only been studied in broad "hand OA," and healthy-handed populations and/or (2) were either non-clinical in nature (ie, tools used only for research purposes) or not administered by hand therapists (eg, radiographic assessments). The authors reported gaps in clinical outcome measures that addressed ligamentous structures, biomechanical properties of the CMC joint, neuromuscular structures, and proprioceptive functions and concluded that further research was needed to develop and validate distinct clinical tools to evaluate BSF in thumb CMC OA. This conclusion aligns well with the Osteoarthritis Research Society International recommendations for the use of thumb CMC OA measures that are reliable, valid, responsive to change, feasible, and readily available to clinicians and researchers.[6,15] However, the scope of the review conducted by Normand et al[15] did not yield an exploration of the tests' psychometrics, an assessment of the quality of the psychometric research in non-operative thumb CMC OA, or clinical recommendations.

In conclusion, while there are numerous tools that have or could conceivably be used to assess BSF in persons with thumb CMC OA who are being managed non-operatively, it is best-practice to select tools with measurement properties that are specific to the population being treated and the treatment being used.[16] The objectives of this systematic review are to (1) assess the literature on clinical tools used in the assessment of BSF in persons with non-operative thumb CMC OA and describe their psychometric properties, (2) based on these findings, make recommendations to help guide clinicians and researchers in the selection of instruments to evaluate BSF in patients with non-operative thumb CMC OA, and (3) identify gaps in the literature that might inform future BSF measurement research in persons with non-operative thumb CMC OA.

Methods

Search strategy

In accordance with best practices,[17] we conducted a comprehensive search combining natural language and controlled vocabulary using a combination of terms to reflect the concepts of CMC OA and conservative treatments. A full search strategy included all search terms is available in Appendix A. Search terms included carpometacarpal, thumb, osteoarthritis, orthotic devices, orthopedic equipment, musculoskeletal manipulations, exercise therapy, rehabilitation, occupational therapy, physical therapy, modalities, conservative, non-surgical, intervention, and management. The search was conducted across nine databases: Ovid MEDLINE, Ovid Embase, CINAHL via EBSCO, ClinicalTrials.gov, Global Index Medicus, PubMed, Scopus, SPORTDiscus EBSCO, and Web of Science Core Collection. A medical research librarian trained in conducing systematic review searches performed all searches in August, 2022. No limitations were placed on study design, date of publication, or language of publication. The search protocol was registered with PROSPERO[18] prior to the commencement of screening (CRD42021272694).

Study selection

Screening was completed using Covidence[19] and occurred in two phases: title–abstract screening and full-text screening. Screening at both stages was done by two independent researchers and discrepancies were resolved through consensus or by a third party where necessary. Reasons for exclusion were recorded at the full-text screening phase in accordance with PRISMA guidelines, and are reported in Figure 1.

Inclusion criteria

Consensus was required between two reviewers to determine final eligibility. To be included studies must have (1) been conducted on participants with thumb CMC OA (ie, not general hand OA) who had not being treated with surgery or steroid injection, (2) been original, peer-reviewed publications (ie, systematic reviews, meta-analyses, conference proceedings, editorials, book chapters, expert opinion, etc. were excluded), (3) investigated the psychometric properties of clinical tools that objectively measured body structures and function (ie, studies on subjective pain assessments, patient-reported outcomes, mental health, etc. were excluded), and (4) implemented tools only used in clinical settings (ie, not solely for

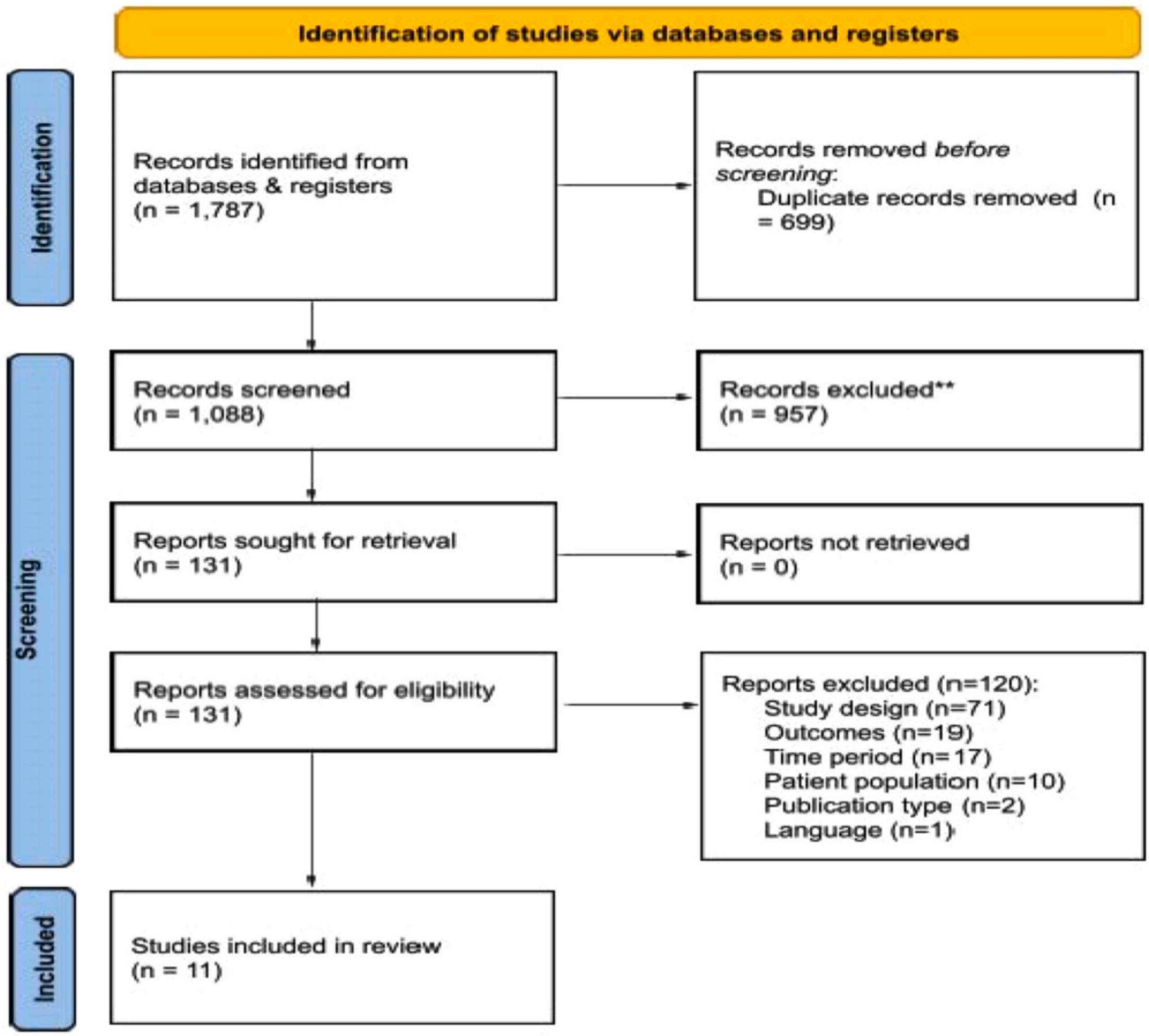

Fig. 1. Flow diagram.

research purposes). Articles more than 20 years old were excluded to ensure that we did not analyze outdated material.[20]

Data extraction

Data on studies' samples, methodology, and psychometric findings were extracted from the included studies. The psychometric findings of interest included reports of tools' properties (ie, reliability, validity, responsiveness, precision, sensitivity, specificity, and minimally importance clinical difference).

Assessment of methodological quality

Prior to formally beginning the review process, several articles were pilot-tested to ensure agreement. The methodological quality of included studies was evaluated by two independent researchers, and consensus arose through discussion. Although not utilized, a third party was available for consultation if consensus was not reached. To assess quality, the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN)[21] was used. This checklist is used to assess methodological quality of research that spans various domains of measurement properties for later use in systematic reviews.[21] The domains assessed in COSMIN include measurement error, validity, reliability, responsiveness, and interpretability with related measurement properties. For each of the measurement properties, the COSMIN checklist consists of five to 18 items to determine methodological quality and each item is rated on a four-point scale (ie, inadequate, doubtful, adequate, and very good).[21] By applying the lowest rating for each item, an overall score is separately generated for each measurement property. A study is rated as inadequate, doubtful, adequate, or very good regarding methodological quality for each of the assessed measurement properties.

Results

Included studies

After deduplication, the search strategy identified 1088 potentially eligible studies. Following a screen of the titles and abstracts, 131 studies were potentially eligible. One hundred and twenty studies were deemed ineligible. Eleven studies were included in the review. A flowchart of this process with additional detail including reasons for exclusion is shown in Figure 1.

Included study characteristics

The characteristics of these studies are summarized in Table 1. A total of 603 subjects with CMC OA were enrolled across all studies with a mean age of 69 years. The psychometric study designs included prospective case–control, prospective cohort, and cross-sectional. The authors of the included studies examined the psychometric properties of the instruments used only in the clinical evaluation of persons with non-operative thumb CMC OA. These psychometric properties are defined in Table 2. The selected studies investigated techniques used to measure ROM, strength, and pain-pressure threshold, and to reproduce arthritic symptomology (ie, provocative tests). Specifically, these authors investigated the criterion validity (ie, sensitivity and specificity) of provocative tests,[22–25] the reliability and precision of tools and techniques used to quantify ROM,[26,27] the reliability, precision, and construct validity (ie, minimal clinically important difference) of several hand strength measures,[28–32] and the reliability and precision of the pain-pressure threshold test in persons with thumb CMC OA.[29] The provocative tests studied included the grind test, traction shift test, metacarpophalangeal (MP) extension test, MP flexion test, and pressure-shear test. Tests of ROM included the intermetacarpal distance (IMD) method, Kapandji index, and goniometric measurements of the thumb. The strength tests included pain-free grip and pinch strength using dynamometry, maximal volitional contraction (MVC) grip strength using dynamometry, and combined thumb abduction/index finger extension strength using myometry. A summary of each study's objectives, participants, methods, and psychometric findings is reported in Table 3.

Measure usability

The researchers of the studies provided sufficient detail regarding the administration of the tests and tools that they used in their research. The cost of the assessments range from no cost to approximately $300.00. All of the procedures used by the researchers can be performed in less than 2 minutes. Additional details, including brief summaries of the testing procedures, can be found in Table B1 in Appendix B.[22–32]

Methodological quality of the included studies

The methodological quality of the included studies varied from inadequate[23,24,28] to very good.[21,25,26] The reliability of the instruments and measurement error were more often provided than the criterion or construct validity. Two of the studies on assessments of thumb mobility were of high methodological quality.[25,26] Studies

Table 1
Included study characteristics

Authors and year of publication	Measure(s)	Study design	Number of participants	Age of participants
Choa et al. 2013[24]	The grind and traction shift tests	Prospective case–control to compare the sensitivity and specificity of the grind and traction shift test in CMC OA	30 healthy subjects 30 subjects with CMC OA	CMC OA mean age 66 Healthy participants mean age 50
Miller and Jerosh-Herold 2013[28]	Maximal pinch strength	Prospective cohort study, repeated measures design to compare the test–retest reliability of the Jamar dynamometer to a digital strain gauge torsion dynamometer (IME)	38 subjects with CMC OA	CMC OA mean age 63
Villafañe and Valdes 2013[29]	Index finger extension and thumb strength, thumb CMC extension, and pain-pressure threshold	Prospective cohort study to measure isometric force of index finger extension and abduction CMC joint, thumb CMC extension, and pain-pressure threshold of the thumb in patients with CMC OA to establish the cutoff value scores for a minimal detectable change	39 subjects with CMC OA 38 healthy subjects	CMC OA mean age 81 Healthy subjects mean age 78
Villafañe and Valdes 2014[30]	Pain-free pinch strength	Prospective cohort study to determine the test–retest reliability of pain-free pinch strength testing	27 subjects with CMC OA	CMC OA mean age 81
Villafañe et al 2015[31]	Pain-free grip strength	Cross-sectional study to determine the test–retest reliability of pain-free grip strength testing in subjects with CMC OA	78 subjects with CMC OA	CMC OA mean age 83
Gelberman et al 2015[22]	Thumb metacarpal adduction and extension tests	Cross-sectional study to determine the diagnostic performance (ie, sensitivity, specificity, inter-rater reliability) of the thumb metacarpal adduction and extension tests	48 with CMC OA 44 with radial sided wrist pain 47 with other wrist pain	CMC OA mean age 62 Radial sided wrist pain mean age 52 Other wrist pain mean age 42
Jha et al 2015[26]	Kapandji index, goniometry	Cross-sectional study to determine the inter-rater reliability of the Kapandji index to goniometric measurement of the thumb	33 patients (54 thumbs) with CMC OA	CMC OA mean age 65
Model et al 2016[23]	Lever, grind, and MP extension tests	Prospective cohort study to compare the effectiveness of the lever test, grind test, and MP extension test	62 subjects with CMC OA	CMC OA mean age 63
Villafañe et al 2017[32]	Pinch and grip strength	Prospective case–control study to determine the MCID in maximal pinch and grip strength in women with CMC OA	57 women subjects with CMC OA 53 healthy subjects	CMC OA mean age 83 Healthy mean age 77
Sela et al 2019[25]	Grind, MP flexion, MP extension, and pressure-shear tests	Prospective cohort study to determine the diagnostic value of the grind, MP flexion, MP extension, and pressure-shear test	104 (127 thumbs) subjects with CMC OA	CMC OA mean age 59
McGee et al 2021[27]	Intermetacarpal distance measure of palmar and radial abduction	Cross-sectional, psychometric study to determine the inter-rater reliability and precision of the intermetacarpal distance	22 subjects (28 thumbs) with CMC OA	CMC OA mean age 59

CMC = carpometacarpal, MP = metacarpophalangeal, MCID = minimal clinically important difference, OA = osteoarthritis.

Table 2
Definitions of the properties of the reviewed tools

Term	Definition
Accuracy	Accuracy of the measurement provided by an instrument is determined by comparing the reading on the device with a standard measure (or known true value).[84]
Inter-rater reliability	The agreement between observers (also known as "interobserver" reliability) when making the same measurement. It is sometimes tested through using Cohen's Kappa Statistic.[85] Reliability, as per the Kappa result, can be interpreted as follows: values ≤0 as indicating no agreement, 0.01-0.20 as none to slight, 0.21-0.40 as fair, 0.41-0.60 as moderate, 0.61-0.80 as substantial, and 0.81-1.00 as almost perfect agreement. Other times it is tested through use of an intraclass correlation coefficient (ICC). Reliability, as per the ICC result, can be interpreted as follows: ≤0.39 = poor, 0.40-0.59 = fair, 0.60-0.74 = good, and ≥0.75 = excellent.[36]
Minimal clinically important difference (MCID)	The smallest improvement considered worthwhile by a patient.[86]
Negative predictive value (NPV)	The percentage of those identified by the test as negative who actually do not have the diagnosis. A value of 1.0 or 100% would indicate 100% of those with a positive test actually having the condition.[87]
Positive predictive value (PPV)	The percentage of individuals identified by the test as positive who actually have the diagnosis. A value of 1.0 or 100% would indicate 100% of those with a negative test do not have the condition.[87]
Precision	Degree to which repeated measurements under unchanged conditions show the same results (ie, measurement error). This is often established through calculating the standard error of the measurement (SEM), and the minimal detectable change (MDC) and MDC%. Both the SEM and MDC indicate the minimal amount of change allowed in a patient's score that is not a result of a measurement error. The MDC is the more stringent of the two analyses. The MDC% is an indicator of how much error (ie, MDC) is present relative to the range of measurements recorded by the tool. This is expressed as a percent and an MDC% of less than 30% is defined as acceptable while one that is less than 10% is excellent.[88]
Test–retest reliability	The agreement between scores of tests administered by the same assessor on two or more occasions (also referred to as "intra-rater" or "intra-observer" reliability).[89]
Sensitivity	The proportion or percentage of individuals with a particular diagnosis who are correctly identified as positive by the test (ie, rate of correct positive diagnoses). A value of 1 indicates that those with the condition will test positive 100% of the time.[90]
Specificity	The proportion or percentage of individuals without a particular diagnosis who are correctly identified as negative by the test (ie, rate of correct negative diagnoses). A value of 1 indicates that those without the condition will test negative 100% of the time.[90]

that examined strength measurements were determine to be of inadequate,[28] doubtful,[29,30] or adequate[27,31] methodological quality. Studies that examined provocative tests were determine to be of inadequate,[23,24] doubtful,[22] or adequate[21] methodological quality. In total, 1/2 of the reliability studies had adequate or better methodological quality, 1/2 of studies reporting on precision were of adequate or very good quality, and 1/4 of the criterion validity (ie, sensitivity and specificity studies) were of adequate quality or better. The one study on construct validity (ie, minimal clinically important difference or "MCID") was of adequate methodological quality. Figures 2 and 3 illustrate these trends. A meta-analysis was not performed given that only three testing approaches were studied on two or more occasions (ie, CMC Grind and MP Extension provocative tests and CMC extension goniometry) and Cochrane stipulates that a minimum of two studies per measurement would be needed.[33] Given the low volume of relevant literature, only one comparison of the weighted estimates of measurement properties of CMC1 provocative tests could possibly be made.

Discussion

In this study, we sought to synthesize the evidence on the psychometric properties of tools used to measure structures of and functions of the thumb in persons with non-operative thumb CMC OA. Further, we intended to explore the methodological rigor of studies in this area. While there is a plethora of clinical assessments that could be used to assess physical function in persons with thumb CMC OA[15] many do not have sound psychometric properties and most have not been studied in persons with non-operative thumb CMC OA. Although studies on physical assessments conducted in healthy, general hand OA, and inflammatory arthritis populations may give some guidance to hand therapists who are seeking tools to measure physical constructs that are perceived to be barriers to occupational performance, if they are not psychometrically sound or do not have established psychometric properties in persons who represent the population being treated, the use of the tool may yield invalid and/or unreliable findings. For these reasons, we sought to study only those assessments of thumb body structures and functions that have been tested in persons with non-operative thumb CMC OA. Given this, and that our review intentionally excluded subjective and patient-reported measures of thumb function, only 11 publications met our inclusion criteria. These tests were limited to those that measured mobility, strength, and pain-pressure threshold, and those intended to provoke arthritis symptoms. Based on the reviewed literature, we have compiled the following list of recommendations for clinical practice that is also summarized in Table 4.

Recommendations for ROM measurement of the thumb

Opposition

When assessing opposition, we recommend the Kapandji index.[11] Jha et al[26] describe this test to have excellent inter-rater reliability (IRR)[34] and acceptable precision.[35]

Palmar and radial abduction

We recommend the IMD method for quantifying radial and palmar abduction because it yields excellent IRR[36] and acceptable precision.[30] While there is evidence to support this method's test–retest reliability in non-clinical and other clinical populations,[37,38] and preliminary evidence to support its test–retest reliability in thumb CMC OA,[39] further research is needed in persons with thumb CMC OA. Other psychometric properties, such as validity, responsiveness, and its minimally clinically important difference, are not yet known. The radial abduction goniometric method described by American Society for Surgery of the Hand[40] appears to yield excellent test–retest reliability[36] and acceptable precision[35] when the evaluator is conducting the assessments within the context of the same therapy session in persons with thumb CMC OA.

Table 3
Results of data extraction

Author/ year/ country	Measure(s)	Objective	Study design	Participants (n, inclusion criteria)	Rater	Results
Choa et al, 2013[23], UK	Grind and traction-shift tests	Compare the sensitivity and specificity of grind and traction-shift tests	Clinical measurement-cross-sectional	n = 60 (30 with CMC OA, 30 healthy), CMC OA diagnosis confirmed by radiography, positive responsiveness to steroid injections, and symptoms persisting for > 6 wk. Healthy had negative X-rays.	Three orthopedic surgeons	**Sensitivity and specificity** *Grind* / *Traction-shift* Sen. 0.3 / 0.67 Spe. 0.97 / 1 PPV 0.9 / 1 NPV 0.57 / 0.75
Miller and Jerosch-Herold, 2013[27], UK	Pain-free maximal tripod pinch with Jamar and MIE grip dynamometers	To compare pain, patient preference, and the test–retest reliability between a single trial of the Jamar dynamometer and the MIE myometer	Clinical measurement-cross-sectional	N = 38, diagnosis of primary OA at the base of the first CMCJ (based on clinical history and provocative testing with or without radiographic images)	Single experienced therapist	**Test–retest (intra-rater) :** *Reliability:* ***Test*** / ***Reliability*** / ***Precision*** Jamar tripod / 0.891(0.799, 0.942) / SEM = 0.48 kg MIE tripod / 0.914 (0.841, 0.954) / SEM = 0.60 kg *Pain:* No difference *Preferences:* ● Jamar: 47% ● MIE: 45% ● Either: 8%
Villafañe and Valdes, 2013,[28] Italy	Combined thumb abduction and index finger extension strength via the Psytech flexion/extension gauge, pain-pressure threshold via algometry, and radial abduction via goniometry	To establish the intra-rater reliability and minimal detectable change score for an isometric test of index finger extension and abduction CMC strength, patient pressure threshold	Clinical measurement-cohort	n = 77, adults with Eaton stage 3 or 4 radiographic unilateral CMC OA in the right, dominant hand	Physical therapist	**Within-session test–retest (intra-rater) reliability:** *Test* / *Reliability* / *Precision* Thumb abduction and index finger extension strength (lbs.) / ICC = 0.54 / MDC = 0.12-0.13 lbs Pain-pressure threshold (kg) / ICC = 0.89 / MDC = 0.23-0.25 kg/cm^2 Radial abduction (deg.) / ICC = 0.91 / MDC = 1.24-1.46°
Villafañe and Valdes, 2014,[29] Italy	Pain-free maximal pinch using Baseline pinch gauge	Examine the test–retest reliability and precision of the pain-free MVC pinch strength test in elderly subjects with unilateral thumb CMC OA	Clinical measurement-cohort	n = 27, radiographically confirmed stage III or IV thumb CMC OA in dominant hand, no ongoing rehabilitation	Therapist with 12 y of experience treating musculoskeletal disorders	**Test–retest reliability and precision:** *Test* / *Reliability* / *Precision* Tip / ICC = 0.93 (0.85, 0.97) / SEM = 0.06 kg Tripod / 0.96 (0.92, 0.98) / SEM = 0.04 kg Key / 0.99 (0.97, 0.99) / SEM = 0.01 kg

(*continued on next page*)

Table 3 (*continued*)

Author/ year/ country	Measure(s)	Objective	Study design	Participants (n, inclusion criteria)	Rater	Results
Villafañe et al, 2015,[30] Italy	Pain-free maximal grip strength using Jamar dynamometer	Examine the test–retest reliability of pain-free MVC grip strength in persons with CMC OA	Clinical measurement-cohort	n = 78, radiographically confirmed stage III or IV thumb CMC OA in dominant hand, no ongoing rehabilitation	Therapist with 12 y of experience treating musculoskeletal disorders.	**Test–retest reliability:** ICC (95% CI): 0.94 (0.91, 0.96) **Precision:** SEM = 0.61 kg
Gelberman et al, 2015[21] US	Thumb metacarpal adduction and extension tests, CMC Grind Test, Point Tenderness of TMC joint	Determine the diagnostic performance sensitivity, specificity, and inter-rater reliability of thumb metacarpal adduction and extension tests against traditional examination maneuvers for thumb CMC arthritis	Clinical measurement-cross-sectional	n = 129 patients (48 with radiographically confirmed CMC OA, 91 with negative X-rays but radial-sided wrist and hand pain or other wrist pain)	Hand surgeon and 4th year medical student evaluated each patient for all tests	**Inter-rater reliability:** • Adduction: k = 0.79 • Extension test: k = 0.84 • CMC Grind: k = 0.31 • Palpation: k = 0.63 **Sensitivity and specificity:** *MP ext.* / *MP add* / *Grind* / *Palpation* Sen. 0.94 (0.82-0.98) / 0.94 (0.82-0.98) / 0.44 (0.30-0.59) / 0.94 (CI, 0.82-0.98) Spe. 0.95 (0.87-0.98) / 0.93 (0.86-0.97) / 0.92 (0.84-0.97) / 0.81 (CI, 0.71-0.88)
Jha et al, 2015,[25] Australia	Thumb CMC goniometry and Kapandji Index Opposition Scale	Examine the inter-rater reliability and precision of thumb CMC goniometric measures and the Kapandji Index	Clinical measurement-cohort	n = 33 (54 thumbs), adults awaiting surgical management without concurrent medical conditions	2 hand therapists (1 physical and 1 occupational therapist) with 9 and 14 years of experience and 1 fellowship-trained orthopedic surgeon	**Inter-rater reliability:** *Movement* / *ICC (95% CI)* CMC extension 0.64 (0.37-0.81) − 0.73 (0.27-0.92) CMC flexion 0.66 (0.22-0.88) − 0.84 (0.50-0.95) Kapandji Index 0.77 (0.55-0.89) − 0.92 (0.72-0.98) **Precision:** *Movement* / *SEM (deg.)* / *MDC (deg.)* CMC extension 9.7 / 23 CMC flexion 10.3 / 24.25 Kapandji index 0.43 / 1
Model et al, 2016,[22] USA	Grind, lever, MP extension, and CMC palpation tests	Compare the characteristics of the grind, lever, MP extension, and CMC palpation tests as a diagnostic tool	Clinical measurement-cross-sectional	n = 62 (100 thumbs), adults with pain localized to the basal joint and positive radiographic reading	Four orthopedic surgeons	**Sensitivity and specificity:** *Test* / *Sen.* / *Spe.* / *PPV* / *NPV* MP ext. 0.91 / 0.76 / 0.95 / 0.49 Lever 0.65 / 0.95 / 1 / 0.26 Grind 0.82 / 0.81 / 0.95 / 0.64 Palpation 0.41 / 1 / 0.98 / 0.36
Villafañe et al, 2017,[31] Italy	MVC grip and pinch strength (kg) using Baseline Grip and Pinch Dynamometers	Establish the minimal clinically important difference (MCID) of MVC grip, tripod pinch, and tip pinch	Clinical measurement-cross-sectional	n = 57, adults with radiographic (stages 1-2) unilateral CMC OA in the right, dominant hand	Therapist with 10 years of experience in musculoskeletal pain disorders	**MCID:** • Grip: 0.84 kg • Tip pinch: 0.33 kg • Tripod pinch 0.35 kg

(*continued on next page*)

Table 3 (*continued*)

Author/ year/ country	Measure(s)	Objective	Study design	Participants (n, inclusion criteria)	Rater	Results					
Sela et al, 2019,[24] USA	Grind, MC flexion, MC extension, and pressure-shear tests	Compare the sensitivity and specificity, of 4 different provocative tests: the grind, metacarpal (MC) flexion, MC extension, and pressure-shear tests	Clinical measurement-cross-sectional	n = 104 (127 thumbs), adults with radial hand/wrist pain and radiographic OA	Board-certified orthopedic surgeon	**Sensitivity and specificity:**					
						Test	*Acc. (95% CI)*	*Sen. (95% CI)*	*Spe. (95% CI)*	*PPV (95% CI)*	*NPV (95% CI)*
						Grind test	0.70 (0.61, 0.78)	0.64 (54, 73)	1.00 (0.78, 1.00)	1.00 (0.92, 1.00)	0.37 (0.25, 0.50)
						MC pressure-shear test	0.98 (0.94, 1.00)	0.99 (0.95, 1.00)	0.95 (0.77, 1.00)	0.99 (0.95, 1.00)	0.95 (0.77, 1.00)
						MC flexion test	0.47 (0.38, 0.56)	0.36 (0.27, 0.46)	1.00 (0.78, 1.00)	1.00 (0.87, 1.00)	0.25 (0.16, 0.35)
						MC extension test	0.55 (0.46, 0.64)	0.46 (0.36, 0.56)	1.00 (0.78, 1.00)	1.00 (0.89, 1.00)	0.28 (0.18, 0.39)
McGee et al., 2021,[26] USA	Intermetacarpal distance (IMD)	Determine inter-rater reliability and precision of the IMD method for measuring palmar and radial abduction	Clinical measurement-cross-sectional	n = 22 (28 thumbs), adults with radiographic OA or positive provocative test	Two OT CHTs with 9 and 40 years of clinical experience	**Inter-rater reliability:**					
						Movement	ICC (95% CI)				
						P. Abd	0.85 (0.69, 93)				
						R. Abd	0.76 (0.54,0.88)				
						Precision:					
						Movement	*SEM (deg.)*	*MDC/MDC% (deg.)*			
						P. Abd	2.4	5.6/8.4			
						R. Abd	3.8	8.8/13.7			

Acc. = accuracy; CI = confidence interval; deg. = degrees; ICC = intraclass correlation coefficient; lbs. = pounds; k = Cohen's kappa coefficient; kg = kilograms; MCID = minimally clinically important difference; MDC = minimal detectable change; MVC = maximal voluntary contraction; PPV = positive predictive value; SEM = standard error of the measurement; Sen. = sensitivity; Spe. = specificity.

Carpometacarpal flexion, MP flexion/extension, and interphalangeal flexion/extension

While IMD and goniometric assessments of CMC radial abduction (ie, extension), respectively, yield excellent inter and IRR in persons with thumb CMC OA, goniometric assessments of thumb CMC, MP, and interphalangeal (IP) flexion and extension, have variable IRR and low precision. Interphalangeal flexion measurements appear to have poor to good IRR whereas IP extension, and MP and CMC extension and flexion measurements appear to have moderate-to-good IRR.[26] These findings are primarily in agreement with those of McGee et al[41] who reported good-to-excellent IRR for MP and IP flexion measurements but poor IRR for CMC flexion measurements in healthy adults.

Because IRR for goniometric assessment of CMC and MP flexion and extension and IP extension is good-to-excellent, therapists should anticipate having comparable findings yet may want to proceed by either having a consistent therapist take these measurements for the same client or interpreting these measurements with caution by ensuring that change exceeds the published standard error of the measurement (SEM). We recommend that only the same therapist take IP flexion measurements. Should only one therapist be involved in assessing a client's radial abduction, one might argue that there is evidence to support that either IMD method or American Society for Surgery of the Hand goniometric method could be used. However, further evidence is needed to evaluate the test–retest reliability of all of the measures when the retest is occurring at a time point that is more in alignment with clinical practice (ie, 1-2 weeks after the initial assessment).

While others haven reported on the psychometrics for various approaches for quantifying thumb CMC palmar and radial abduction/adduction (ie, Pollexograph, radius-metacarpal goniometry, intermetacarpal goniometry, and thumb-distal-interphalangeal distance),[37,42] these studies were not carried out in clinical populations and thus should be used with caution until further validated for use in persons with thumb CMC OA. Additionally, more evidence is needed on the validity, test–retest reliability, responsiveness, and minimally clinically important difference for the aforementioned IMD and goniometric measurements in persons with thumb CMC OA.

Recommendations for strength measurement of the thumb

Pinch

We recommend the use of the Baseline pinch gauge and a modified version of the Mathiowetz et al[43] procedures (ie, three maximal pain-free trials). According to Villafane and Valdes,[30] the average of three pain-free maximal pinch measurements using the Baseline Pinch Gauge yields excellent testCHTsretest reliability[36] at 1-week follow-up for tip, three-point, and lateral pinch. Based on the data provided, we were able to estimate the minimal detectable change (MDC) and subsequently calculate the MDC%. From this, we were able to determine the precision for each measurement using the Baseline to be excellent.[35] Miller and Jerosch-Herold[28] also explored the test–retest reliability of pain-free maximal pinch strength measurements but did so using the Jamar and MIE (MIE Medical Research Ltd) pinch gauges. Their methods were distinct from those used by Villafane and Valdes,[30] in that they only included 1 trial, only involved the assessment of three-point pinch, and the retest occurred within the same measurement session. Both tools demonstrated excellent test–retest reliability[36] but only the MIE demonstrated acceptable precision.[35] Additionally, no significant differences were found between the tools in terms of pain intensity associated with testing or patient preference.

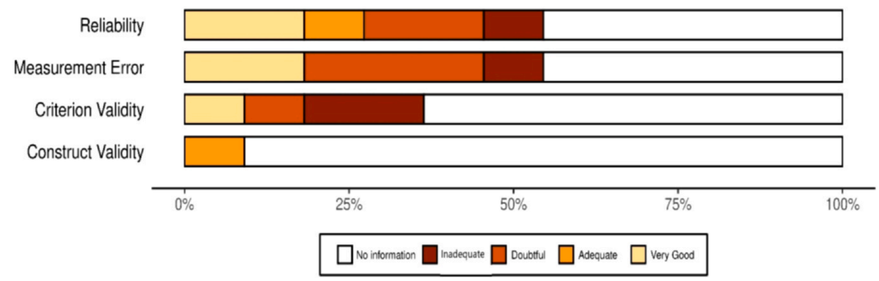

Fig. 2. Methodological quality of the included studies.

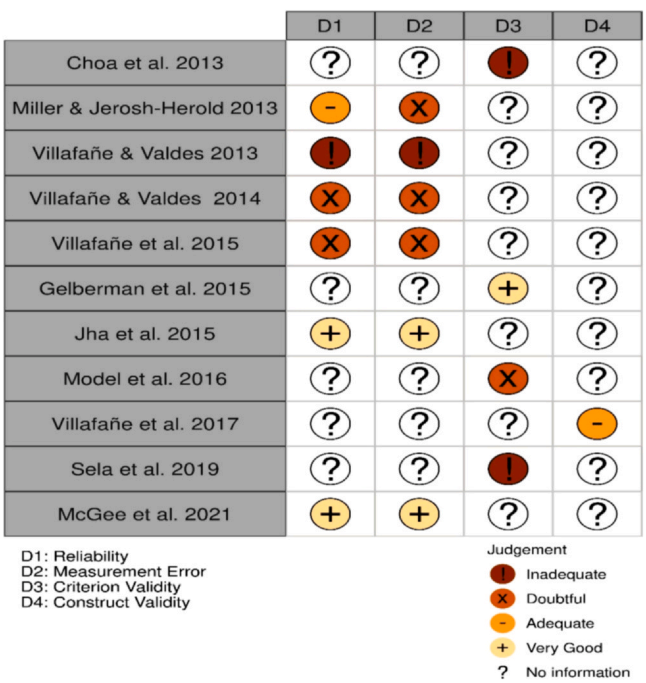

Fig. 3. Characteristic and psychometric reporting of individual studies.

Efforts have also been made to establish the MCID of pinch strength measures. Through the use of a distribution-based approach, Villifane et al[32] determined the MCID of maximal tip and three-point pinch measurements gathered with the Baseline pinch gauge as per the procedures described by Mathiowetz et al.[43] These values (0.33 kg for tip and 0.35 kg for three-point) exceed the error (ie, precision) estimates described by Villafane and Valdes,[30] and should be surpassed in order for a therapist to be confident that the change in pinch strength will have an impact on the client's daily experiences.

Grip

We recommend the use of the Jamar Grip Dynamometer using a modified version of the Mathiowetz et al[43] procedures (ie, three maximal pain-free trials). In a study design similar to that of Villafane and Valdes,[30] Villafane et al[31] reported that recording the average of three pain-free maximal measurements using the Jamar grip dynamometer yields excellent test–retest reliability[36] and acceptable precision[35] at 1-week follow-up.

Villifane et al[32] also determined the MCID of maximal grip strength measurements gathered with the Baseline dynamometer as per Mathiowetz et al.[43] According to these authors, for a therapist to be confident that the change in grip strength will have an impact on the client's daily experiences, it should exceed 0.84 kg.

Other measures of hand strength

Based on our review, we cannot recommend any additional measures of hand strength in thumb CMC OA. Villafane and Valdes[11] reported the intrasession reliability of a combined measure of thumb abduction and index finger extension strength via the Psytech Flexion/Extension gauge to only be "fair." A pinch-collapse test,[44] where maximal pinch at the time of thumb MP collapse is assessed via dynamometry, is also described in the literature but not yet psychometrically tested. There were also several myometers and research aparati that did not meet our inclusion criteria. The Rotterdam Intrinsic Hand Myometer (RIHM),[45] a clinical tool for assessing isolated measures of hand strength, has well-established psychometric properties in numerous non-clinical and clinical populations as well as reference values however has not yet been studied in thumb CMC OA.[46–50] A myometer developed for research purposes[51] was used to assess thumb abduction and adduction strength in persons with thumb CMC OA in response to exercise; however, its psychometrics were not described. Other tools developed for research purposes include force sensing jar tools used to measure cylindrical grasp strength in persons with thumb CMC OA[52]and to quantify the effects of joint protection strategies on hand forces.[53] While both tools are described to have sound psychometric properties,[54,55] they were not designed for clinical use.

Recommendations for fine motor/dexterity

Our review did not reveal any studies designed to test the psychometrics of dexterity and fine motor assessments in persons with non-operative thumb CMC OA. However, there are several studies that did not meet our inclusion criteria but involved tests of fine motor and dexterity skills in persons with non-operative thumb CMC OA. Carreira et al[56] evaluated the effects of a short opponens orthosis vs a no-treatment control on dexterity via the O'Connor Tweezer Dexterity Test.[57] Loyley et al[58] used the Nine Hole Peg Test[59] to comparatively evaluate the effects of three orthosis designs as well as no-treatment group on fine motor skills. The Strength-Dexterity Test has been used in various descriptive studies of hand function in persons with and without conditions affecting the hand, including persons with thumb CMC OA.[60–62] The test kit is comprised of numerous springs with variable tensile strengths that, when successfully compressed, are combined indicators of strength and dexterity. However, to date there are no published psychometric studies on these tests or others such as the Functional Dexterity Test (FDT),[63] Purdue Pegboard,[64] Box and Blocks,[65] or the Minnesota Rate of Manipulation (MRM).[66]

Recommendations for hand function

Although conceptually measures of body functions, our review did not yield any studies on the psychometrics of objective assessments of activity performance. While there are instruments such as the Arthritis Hand Function Test (AHFT),[67] the Jebsen-Taylor Hand Function Test (JTHFT),[68] or the Sollerman Hand Function Test (SHFT)[69] whose psychometrics have been studied in hand OA and gout populations,[70–73] their psychometrics are not yet known in the thumb CMC OA subset.

Recommendations for sensory function

Based on our review, we cannot recommend any tests for sensory function in persons with thumb CMC OA.

Joint position sense

While there were not any studies on clinical assessments of thumb CMC proprioception that met our inclusion criteria, new approaches to measuring joint position sense have emerged in thumb CMC OA. A goniometric method introduced by Ouegnin and Valdes[74] has been used to describe differences in joint position sense (JPS) between those with and without thumb CMC OA and has been used to evaluate the outcomes of proprioceptive training in those with thumb CMC OA[75] however, its psychometrics are not yet known. Additionally, in a conference paper by McGee et al,[39] the authors report strong preliminary psychometrics of a joint position sense measurement that involves the use of the aforementioned IMD method; however, there are not yet any peer-reviewed publications on this approach.

Cutaneous sensation

There are currently no published psychometric studies on clinical measures of cutaneous sensation like two-point discrimination (2PD) or cutaneous sensory threshold testing (eg, Semmes Weinstein Monofilaments)[76] in thumb CMC OA however, recent evidence suggests there is a link between CMC OA synovitis and altered function of the superficial branch of the radial nerve.[77]

Right-left discrimination

Some recent evidence suggests that, like in persons with other chronic pain conditions,[78] persons with hand OA may have altered body schemas. While not specifically in thumb CMC OA, a 2018 case–control study[79] revealed that persons with hand OA have altered right-left discrimination sense relative to persons without.

Force matching

No studies on this topic met our inclusion criteria; however, evidence suggests that persons with thumb CMC OA have impaired pinch and grip force matching accuracy relative to healthy controls.[80] Although causality cannot be inferred, these findings align well with known alterations in conscious proprioception in this population.[74] While this study did not involve dynamometers commonly used in clinical examination, the procedures could easily be adapted for use with grip and pinch gauges that are more often used in clinical environments. Further study is needed.

Recommendations for pain-pressure threshold

We recommend that pain-pressure threshold, as described by Villafañe and Valdes,[11] be used as an objective measure of pain tolerance in thumb CMC OA. In this exam, pressure is applied to the base of the anatomical snuffbox with an algometer with the highest tolerated pressure being indicated of the patient's threshold. This approach has excellent intra-session reliability,[29] known precision,[29] and has been used in several interventions studies on the effects of nerve and joint mobilizations on pain in persons with thumb CMC OA.[1,81]

Methodological quality

Only 3/11 studies were of "very good" methodological quality which compounds the issue of the already acknowledged low volume of BSF assessment research in non-operative thumb CMC OA. Common issues with the methodological quality of the reviewed reliability studies included uncertainty about procedures for keeping evaluators blinded to previous test scores, and uncertainty about the appropriateness of time intervals between initial and follow-up tests. Inadequate statistical analysis and design flaws (eg, provocative testing not being conducted on healthy hands) were the most common methodological issues in criterion validity (ie, provocative test) studies.

Future research

The results of this review suggest that there is a shortfall of psychometrically tested tools for assessing thumb body structures and functions in persons with non-operative thumb CMC OA. Future research should explore the validity, MCID, and responsiveness of mobility, strength, dexterity, sensory, and hand function measures in persons with non-operative thumb CMC OA. Reliability studies are needed for specific strength measurements, including the RIHM and the Pinch Collapse test. Reliability studies are also needed in joint position sense (eg, goniometry and IMD), cutaneous sensation (eg, 2PD and sensory threshold testing), force matching (eg, pinch dynamometry), dexterity (eg, NHPT, FDT, MRM), and hand function tests (eg, AHFT, JTHFT, SHFT). To prevent future issues with methodological rigor, recommend the use of COSMIN criteria[21] when planning and reporting future measurement research. Table 5 summarizes these recommendations.

Limitations

Our practice recommendations are limited due to the small volume of literature meeting inclusion criteria. Although we believe our inclusion criteria were justified, early psychometric studies and studies on patients who underwent injection and arthroplasty may have expanded our recommendations.

Table 4
Synopsis of recommendations

Domain	Practice recommendations
Mobility	Kapandji index[26] for thumb opposition, IMD test for radial and palmar abduction;[27] goniometry for thumb MP and IP extension and MP flexion[26]
Strength	Three trials of pain-free maximal hand strength using the Jamar dynamometer[31] and baseline pinch dynamometer[30]
Fine motor/dexterity	Insufficient evidence
Sensation/proprioception/perception	Insufficient evidence
Provocative tests	Thumb metacarpal adduction stress test[22] or extension stress test[22]; do not use grind compression
Pain threshold	Pain-pressure threshold test via algometry[29]

IMD = intermetacarpal distance; IP = interphalangeal; MP = metacarpophalangeal

Table 5
Suggestions for future research

Domain	Future research recommendations
Mobility	• Studies on the validity, test–retest, MCID, responsiveness of IMD measures for radial and palmar abduction • Studies on test–retest validity, reliability, responsiveness, and MCID for Kapandji index • Studies on validity, test–retest reliability, inter-rater reliability, precision, MCID, and responsiveness of the following measure of thumb CMC mobility: Pollexograph, radius-metacarpal goniometry, and thumb-distal-interphalangeal distance
Strength	• Studies on validity and responsiveness for all strength measurements • Studies on test–retest, inter-rater reliability, precision, and MCID for the RIHM and Pinch Collapse Test[44]
Sensory	• Studies on test–retest reliability, inter-rater reliability, precision, MCID, and responsiveness of goniometer-based JPS method • Studies on validity, test–retest (larger sample needed) inter-rater reliability, precision, MCID, and responsiveness of IMD JPS method • Studies on validity, test–retest reliability, inter-rater reliability, precision, MCID, and responsiveness for sensory threshold testing, 2PD, and pinch/grip force matching
Dexterity	• Studies on validity, test–retest reliability, inter-rater reliability, precision, MCID, and responsiveness of NHPT, FDT, and box and blocks
Hand function	• Studies on validity, test–retest reliability, inter-rater reliability, precision, MCID, and responsiveness of AHFT, JTHF, and SHFT

AHFT = Arthritis Hand Function Test; FDT = Functional Dexterity Test; JTHT = Jebsen-Taylor Hand Function Test; JPS = Joint Position Sense; IMD = intermetacarpal distance, MCID = minimal clinically important difference; NHPT = Nine Hole Peg Test; SHFT = Sollerman Hand Function Test; 2PD = Two Point discrimination.

Conclusions

We recommend that consistent outcome measures with sound psychometric properties be used in clinical evaluation and believe our findings will help to support this practice. We hope that our findings will also help to guide consensus groups such as Wouters et al[5] with formulating future measurement recommendations. Furthermore, as is suggested by earlier research, we recommend that future studies on the effectiveness of interventions with thumb CMC OA use uniform outcome measures.[82,83] Future psychometric research for the purpose of growing our library of clinical measures of body structures and functions for persons with non-operative thumb CMC OA is needed.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Search strategy

MEDLINE (Ovid) search.

1. exp Carpometacarpal Joints/
2. exp Metacarpus/
3. (cmc or carpometacarpal or "carpal metacarpal" or (thumb adj1 base)).tw,kw.
4. or/1-3
5. exp Osteoarthritis/
6. (osteoarthrit* or OA).tw,kw.
7. or/5-6
8. 4 and 7
9. exp Orthotic Devices/
10. exp Orthopedic Equipment/
11. exp Musculoskeletal Manipulations/
12. exp Exercise Therapy/
13. exp Occupational Therapy/
14. rehabilitation.fs.
15. exp Physical Therapy Modalities/
16. ("physical therap*" or physiotherap* or orthotic* or orthosis or orthoses or exercis* or kinesiotherap* or rehabilitat* or "occupational therap*").tw,kw.
17. ((conservative or "non-surgical" or nonsurgical or "non surgical" or "non-operative" or nonoperative) adj2 (therap* or treat* or intervention* or management)).tw,kw.
18. or/9-17
19. 8 and 18
20. exp Animals/ not exp Humans/
21. 19 not 20
22. (exp Child/ or exp Infant/ or exp Adolescent/) not exp Adult/
23. 21 not 22
24. remove duplicates from 23

Appendix B

Table B1
Usability of the measurement

Test domain	Measure	Description	Administration	Time to administer	Scoring	Cost*
Provocative Tests	Grind Test	Manual test applied by the examiner	Examiner applies axial compression along the plane of the metacarpal bone and rotates the thumb metacarpal base.[23]	< 1 min	Test is positive if it reproduces pain at the joint.	N/A
	Traction Shift Test	Manual test applied by the examiner	Longitudinal traction to the thumb CMC joint, alternate volar and dorsal pressure over the base of the metacarpal to provoke subluxation and relocation of the joint.[24]	< 1 min	Test is positive if it reproduces pain at the joint.	N/A
	Lever Test	Manual test applied by the examiner	Examiner puts their thumb and index on both sides of the thumb CMC joint and levers the first MC joint radially and ulnarly to the endpoints at the CMC[23]	< 1 min	Test is positive if it reproduces pain at the joint.	N/A
	MP extension	Manual test applied by the examiner	Examiner provides resistance to active thumb MP extension by placing index finger on the thumb IP joint.[23]	< 1 min	Test is positive if it reproduces pain at the joint.	N/A
	Thumb metacarpal adduction and extension test	Manual test applied by the examiner	Firm adduction force downward to metacarpal head; firm extension force to end range with examiner's thumb along radial aspect of distal thumb metacarpal 5-10 mm proximal to MP joint.[22]	< 2 min	Test positive with pain at CMC	N/A
Strength	Grip strength testing	Use of dynamometer to measure strength	Needle placed at "0" position Dynamometer in position #2 Inform client they will feel no resistance Right hand first followed by left hand Measurer supports dynamometer without restricting movement Say "squeeze" "harder" To limits of pain	< 2 min	Measured in kg Can be scored as change over time, comparison to uninvolved side and comparison to norms	$300.00-$384.00
	Pinch strength test	Use of pinch meter to measure strength	Apply force to pinch groove while holding pinch gauge between thumb and two fingers in 3-point pinch or lateral pinch. Use index to thumb for tip to tip To limits of pain	< 2 min	Measured in kg or lbs.	$268.00325.00
	Digital strain gauge torsion dynamometer (MIE)	Sensor whose resistance varies with applied force; torque value is displayed with torsion dynamometer	Pinch using maximal force for 5 s or one time to pain limits.	< 2 min	Measured in kg. or lbs	$1500.00
	Index extension and CMC abduction with Psytech Finger Flexion/Extension gauge (Psytech; Fabrication Enterprises, Inc, Irvington, NY)	Use of device that tests isometric abduction of the thumb CMC joint and extension of index finger	Standardized testing position of seated with shoulder adducted and neutrally rotated, elbow 90 degrees flexion, forearm neutral pronation/supination[29]	< 2 min	Opening strength measured in lbs	$94.75

(continued on next page)

Table B1 (*continued*)

Test domain	Measure	Description	Administration	Time to administer	Scoring	Cost*
ROM	Kapandji index	Active range of motion test	The patient is asked to oppose the thumb to 10 locations moving radial to ulnar: radial side of proximal phalanx of index finger, radial side of middle phalanx of index finger, tip of index, tip of middle, tip of ring, tip of little finger (digit V), DIP crease of digit V, PIP joint crease of digit V, MP joint crease of digit V and distal palmar crease[11]	< 2 min	Each location is numbered (1-10) and examiner records the highest number that the patient can touch with the tip of their thumb	N/A
	Intermetacarpal distance	Measurement of distance between metacarpals	Dorsal midpoints of first and secnd metacarpal heads are located and marked during active palmar and radial abduction; measure distance using digital calipers[27]	< 2 min	Measurement in mm	$2.50-$34.99
	Goniometry	ROM of CMC, MCP	CMC extension/flexion MCP extension/flexion and IP extension by consistent therapist CMC radial abduction as described by ASHT[26,29,91]	< 2 min	Measurement in degrees	$9.99-$18.99
Pain-pressure threshold	Pressure algometer	Measures pressure pain threshold	Progressive pressure applied at the base of the anatomical snuffbox with algometer, with the highest tolerated pressure being recorded	< 1 min	Measurement read in grams of force at highest tolerated pressure	$114.85-$395.00

lbs. = pounds; kg = kilograms; MP =metacarpophalangeal; cm = centimeter.
*Costs determined through range found online at time of manuscript submission.

References

1. Villafañe JH, Silva GB, Diaz-Parreño SA, Fernandez-Carnero J. Hypoalgesic and motor effects of kaltenborn mobilization on elderly patients with secondary thumb carpometacarpal osteoarthritis: a randomized controlled trial. *J Manipulative Physiol Ther.* 2011;34(8):547–556. https://doi.org/10.1016/j.jmpt.2011.08.005
2. Haara MM, Heliövaara M, Kröger H, et al. Osteoarthritis in the carpometacarpal joint of the thumb. Prevalence and associations with disability and mortality. *J Bone Joint Surg.* 2004;86(7):1452–1457. https://doi.org/10.2106/00004623-200407000-00013
3. Mobargha N, Ludwig C, Ladd AL, Hagert E. Ultrastructure and innervation of thumb carpometacarpal ligaments in surgical patients with osteoarthritis. *Clin Orthop Relat Res.* 2014;472(4):1146–1154. https://doi.org/10.1007/s11999-013-3083-7
4. Lawrence EL, Fassola I, Werner I, Leclercq C, Valero-Cuevas FJ. Quantification of dexterity as the dynamical regulation of instabilities: comparisons across gender, age, and disease. *Front Neurol.* 2014;5:53. https://doi.org/10.3389/fneur.2014.00053
5. Wouters RM, Jobi-Odeneye AO, de la Torre A, Joseph A, Hovius SER. A standard set for outcome measurement in patients with hand and wrist conditions: consensus by the International Consortium for Health Outcomes Measurement Hand and Wrist Working Group. *J Hand Surg.* 2021;46(10):841–855.e7. https://doi.org/10.1016/j.jhsa.2021.06.004
6. Kloppenburg M, Maheu E, Kraus VB, et al. OARSI Clinical Trials Recommendations: design and conduct of clinical trials for hand osteoarthritis. *Osteoarthr Cartil.* 2015;23(5):772–786. https://doi.org/10.1016/j.joca.2015.03.007
7. Kroon FPB, van der Heijde D, Maxwell LJ, et al. Core outcome measurement instrument selection for physical function in hand osteoarthritis using the OMERACT Filter 2.1 process. *Semin Arthritis Rheum.* 2021;51(6):1311–1319. https://doi.org/10.1016/j.semarthrit.2021.08.014
8. World Health Organization. International classification of functioning, disability and health: ICF; 2001.
9. Pham T, van der Heijde D, Altman RD, et al. OMERACT-OARSI initiative: Osteoarthritis Research Society International set of responder criteria for osteoarthritis clinical trials revisited. *Osteoarth Cartil.* 2004;12(5):389–399. https://doi.org/10.1016/j.joca.2004.02.001
10. O'Brien VH, McGaha JL. Current practice patterns in conservative thumb CMC joint care: survey results. *J Hand Ther.* 2014;27(1):14–22. https://doi.org/10.1016/j.jht.2013.09.001
11. Kapandji A. Clinical test of apposition and counter-apposition of the thumb. *Ann Chir Main.* 1986;5(1):67–73. https://doi.org/10.1016/s0753-9053(86)80053-9
12. Merritt MM, Roddey TS, Costello C, Olson S. Diagnostic value of clinical grind test for carpometacarpal osteoarthritis of the thumb. *J Hand Ther.* 2010;23(3):261–267. https://doi.org/10.1016/j.jht.2010.02.001
13. Keyes HB. Stenosing tendovaginitis at the radial styloid process. *Ann Surg.* 1938;107(4):602–606. https://doi.org/10.1097/00000658-193804000-00015
14. Valdes K, Naughton N, Algar L. Linking ICF components to outcome measures for orthotic intervention for CMC OA: a systematic review. *J Hand Ther.* 2016;29(4):396–404. https://doi.org/10.1016/j.jht.2016.06.001
15. Normand M, Tang TS, Brismée J-M, Sobczak S. Clinical evaluation of thumb base osteoarthritis: a scoping review. 2021;26(2):63–78. https://doi.org/10.1177/17589983211002560
16. Draak THP, de Greef BTA, Faber CG, Merkies ISJ. The minimum clinically important difference: which direction to take. *Eur J Neurol.* 2019;26(6):850–855. https://doi.org/10.1111/ene.13941
17. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ.* 2021;372:n71. https://doi.org/10.1136/bmj.n71
18. McGee C., Valdes K., Ivy C., Bakker C.Clinical measures used to assess body structures and functions in patients with carpometacarpal osteoarthritis patent application CRD42021272694; 2021.
19. Innovation V.H. Covidence systematic review software. Melbourne, Australia: ⟨www.covidence.org⟩. accessed November 22, 2022.
20. Helfer B, Prosser A, Samara MT, et al. Recent meta-analyses neglect previous systematic reviews and meta-analyses about the same topic: a systematic examination. *BMC Med.* 2015;13:82. https://doi.org/10.1186/s12916-015-0317-4
21. Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LM, de Vet HCW. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res.* 2012;21(4):651–657. https://doi.org/10.1007/s11136-011-9960-1
22. Gelberman RH, Boone S, Osei DA, Cherney S, Calfee RP. Trapeziometacarpal arthritis: a prospective clinical evaluation of the thumb adduction and extension provocative tests. *J Hand Surg.* 2015;40(7):1285–1291. https://doi.org/10.1016/j.jhsa.2015.04.012
23. Model Z, Liu AY, Kang L, Wolfe SW, Burket JC, Lee SK. Evaluation of physical examination tests for thumb basal joint osteoarthritis. *Hand.* 2016;11(1):108–112. https://doi.org/10.1177/1558944715616951
24. Choa RM, Parvizi N, Giele HP. A prospective case-control study to compare the sensitivity and specificity of the grind and traction-shift (subluxation-relocation) clinical tests in osteoarthritis of the thumb carpometacarpal joint. *J Hand Surg.* 2014;39(3):282–285. https://doi.org/10.1177/1753193413508714
25. Sela Y, Seftchick J, Wang WL, Baratz ME. The diagnostic clinical value of thumb metacarpal grind, pressure-shear, flexion, and extension tests for

carpometacarpal osteoarthritis. *J Hand Ther.* 2019;32(1):35–40. https://doi.org/10.1016/j.jht.2017.09.005
26. Jha B, Ross M, Reeves SWSL, Couzens GB, Peters SE. Measuring thumb range of motion in first carpometacarpal joint arthritis: the inter-rater reliability of the Kapandji Index versus goniometry. 2016;21(2):45–53. https://doi.org/10.1177/1758998315616399
27. McGee C, O'Brien V, Johnson J, Wall K. Thumb carpometacarpal palmar and radial abduction in adults with thumb carpometacarpal joint pain: Inter-rater reliability and precision of the inter-metacarpal distance method. *J Hand Ther.* 2022;35(3):454–460.
28. Miller L, Jerosch-Herold C. Intra-tester reliability of a single trial of pinch strength in patients with osteoarthritis of the first carpometacarpal joint. 2013;18(1):17–23. https://doi.org/10.1177/1758998313484672
29. Villafañe JH, Valdes K. Combined thumb abduction and index finger extension strength: a comparison of older adults with and without thumb carpometacarpal osteoarthritis. *J Manipulative Physiol Ther.* 2013;36(4):238–244. https://doi.org/10.1016/j.jmpt.2013.05.004
30. Villafañe JH, Valdes K. Reliability of pinch strength testing in elderly subjects with unilateral thumb carpometacarpal osteoarthritis. *J Phys Ther Sci.* 2014;26(7):993–995. https://doi.org/10.1589/jpts.26.993
31. Villafañe JH, Valdes K, Vanti C, Pillastrini P, Borboni A. Reliability of handgrip strength test in elderly subjects with unilateral thumb carpometacarpal osteoarthritis. *Hand.* 2015;10(2):205–209. https://doi.org/10.1007/s11552-014-9678-y
32. Villafañe JH, Valdes K, Bertozzi L, Negrini S. Minimal clinically important difference of grip and pinch strength in women with thumb carpometacarpal osteoarthritis when compared to healthy subjects. *Rehabil Nurs.* 2017;42(3):139–145. https://doi.org/10.1002/rnj.196
33. Cumpston M, Li T, Page MJ, et al. Updated guidance for trusted systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions. *Cochrane Database Syst Rev.* 2019;10:ED000142. https://doi.org/10.1002/14651858.ED000142
34. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull.* 1979;86(2):420–428. https://doi.org/10.1037/0033-2909.86.2.420
35. Smidt N, van der Windt DA, Assendelft WJ, et al. Interobserver reproducibility of the assessment of severity of complaints, grip strength, and pressure pain threshold in patients with lateral epicondylitis. *Arch Phys Med Rehabil.* 2002;83(8):1145–1150. https://doi.org/10.1053/apmr.2002.33728
36. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol Assess.* 1994;6(4):284–290. https://doi.org/10.1037/1040-3590.6.4.284
37. de Kraker M, Selles RW, Schreuders TAR, Stam HJ, Hovius SER. Palmar abduction: reliability of 6 measurement methods in healthy adults. *J Hand Surg.* 2009;34(3):523–530. https://doi.org/10.1016/j.jhsa.2008.10.028
38. Murugkar PM, Brandsma JW, Anderson AM, Gurung K, Pun Y. Reliability of thumb web measurements. *J Hand Ther.* 2004;17(7):58–63. https://doi.org/10.1197/j.jht.2003.10.008
39. McGee C., Adducci K., Bonte E., Francen K., Olson D., Roe H. Test-retest reliability and precision of thumb position sense testing using the intermetacarpal distance method in individuals with thumb carpometacarpal osteoarthritis. Presented at: FESSH-EFSHT Congress 2020; Basel, Switzerland. ⟨https://fessh2020.com/down/program_eow_2020.pdf⟩.
40. ASSH The Hand. *Examination and Diagnosis.* Ann Arbor: Churchill; 1990.
41. McGee C, Carlson K, Koethe A, Mathiowetz V. Inter-rater and inter-instrument reliability of goniometric thumb active and passive flexion range of motion measurements in healthy hands. 2017;22(3):110–117. https://doi.org/10.1177/1758998317690754
42. Holzbauer M, Hopfner M, Haslhofer D, Kwasny O, Duscher D, Froschauer SM. Radial and palmar active range of motion measurement: reliability of six methods in healthy adults. *J Plast Surg Hand Surg.* 2021;55(1):41–47. https://doi.org/10.1080/2000656X.2020.1828899
43. Mathiowetz V, Weber K, Volland G, Kashman N. Reliability and validity of grip and pinch strength evaluations. *J Hand Surg.* 1984;9(2):222–226. https://doi.org/10.1016/s0363-5023(84)80146-x
44. Duong V, Robbins SR, Deveza LA, et al. Carpometacarpal and metacarpophalangeal joint collapse is associated with increased pain but not functional impairment in persons with thumb carpometacarpal osteoarthritis. *J Hand Therapy.* 2021;34(4):561–566. https://doi.org/10.1016/j.jht.2020.07.003
45. Med.engineers. Rotterdam Intrinsic Hand Myometer. ⟨http://med-engineers.com/index.php/project/rihm/?lang=en⟩, accessed October 27, 2020.
46. McGee C. Measuring intrinsic hand strength in healthy adults: the accuracy intrarater and inter-rater reliability of the Rotterdam Intrinsic Hand Myometer. *J Hand Ther.* 2018;31(4):530–537. https://doi.org/10.1016/j.jht.2017.03.002
47. Schreuders TAR, Selles RW, Roebroeck ME, Stam HJ. Strength measurements of the intrinsic hand muscles: a review of the development and evaluation of the Rotterdam intrinsic hand myometer. *J Hand Ther.* 2006;19(4):393–401. https://doi.org/10.1197/j.jht.2006.07.024
48. Schreuders TAR, Roebroeck ME, Jaquet J-B, Hovius SER, Stam HJ. Measuring the strength of the intrinsic muscles of the hand in patients with ulnar and median nerve injuries: reliability of the Rotterdam Intrinsic Hand Myometer (RIHM). *J Hand Surg.* 2004;29(2):318–324. https://doi.org/10.1016/j.jhsa.2003.10.024
49. McGee C, Johnson L, Casper J, Gregg K. Reliability and precision of measuring strength of extrinsic muscles of the hand with the Rotterdam Intrinsic Hand Myometer. *J Hand Surg.* 2019;44(7):754–756. https://doi.org/10.1177/1753193419850125
50. McGee C, Hoehn A, Hoenshell C, McIlrath S, Sterling H, Swan H. Age- and gender-stratified adult myometric reference values of isometric intrinsic hand strength. *J Hand Ther.* 2020;33(3):402–410.e2. https://doi.org/10.1016/j.jht.2019.03.005
51. Vocelle. Determining isolated thumb forces in osteoarthritic and healthy persons. *J Biomech Eng.* 2021;143(3):031008.
52. Coughlan MJ, Bourdillon A, Crisco JJ, Kenney D, Weiss A-P, Ladd AL. Reduction in cylindrical grasp strength is associated with early thumb carpometacarpal osteoarthritis. *Clin Orthop Relat Res.* 2017;475(2):522–528. https://doi.org/10.1007/s11999-016-5151-2
53. McGee C, Mathiowetz V. Evaluation of hand forces during a joint-protection strategy for women with hand osteoarthritis. *Am J Occup Ther.* 2017;71(1) https://doi.org/10.5014/ajot.2017.022921 7101190020p1-7101190020p8.
54. Ellingson AM, McGee C, Nuckley DJ, Ferkul M, Mathiowetz VG. Development, construct validity, and reproducibility of a mimetic sealed jar measuring the dynamics of opening. *J Rehabil Assist Technol Eng.* 2017;4:2055668317692222.
55. McGee Ca.M.V.Reliability and validity of a novel instrument for the quantification of hand forces during a jar opening task. J Hand Ther.
56. Carreira ACG, Jones A, Natour J. Assessment of the effectiveness of a functional splint for osteoarthritis of the trapeziometacarpal joint on the dominant hand: a randomized controlled study. *J Rehabil Med.* 2010;42(5):469–474. https://doi.org/10.2340/16501977-0542
57. Hines M, O'Connor J. A measure of finger dexterity. *J Pers Res.* 1926;4(1):379–382.
58. Loyley MST, Davis L, Worsley P, Adams J. E054 Comparison of the functional impact of verum and placebo thumb base orthoses: a proof of concept study. *Rheumatology.* 2019;58(Supp_3):kez110.053. https://doi.org/10.1093/rheumatology/kez110.053 (Oxford, England).
59. Kellor M, Frost J, Silberberg N, Iversen I, Cummings R. Hand strength and dexterity. *Am J Occup Ther.* 1971;25(2):77–83.
60. Valero-Cuevas FJ, Smaby N, Venkadesan M, Peterson M, Wright T. The strength-dexterity test as a measure of dynamic pinch performance. *J Biomech.* 2003;36(2):265–270. https://doi.org/10.1016/s0021-9290(02)00340-8
61. Lawrence EL, Fassola I, Dayanidhi S, Leclercq C, Valero-Cuevas FJ. An evaluation of clustering techniques to classify dexterous manipulation of individuals with and without dysfunction. *IEEE.* 2013:1254–1257. https://doi.org/10.1109/NER.2013.6696168
62. Lawrence EL, Fassola I, Werner I, Leclercq C, Valero-Cuevas FJ. Quantification of dexterity as the dynamical regulation of instabilities: comparisons across gender, age, and disease. *Front Neurol.* 2014;5:53. https://doi.org/10.3389/fneur.2014.00053
63. Aaron DH, Jansen CWS. Development of the Functional Dexterity Test (FDT): construction, validity, reliability, and normative data. *J Hand Ther.* 2003;16(1):12–21. https://doi.org/10.1016/S0894-1130(03)80019-4
64. Tiffin J, Asher EJ. The Purdue pegboard norms and studies of reliability and validity. *J Appl Psychol.* 1948;32(3):234–247. https://doi.org/10.1037/h0061266
65. Mathiowetz V, Volland G, Kashman N, Weber K. Adult norms for the Box and Block Test of manual dexterity. *Am J Occup Ther.* 1985;39(6):386–391. https://doi.org/10.5014/ajot.39.6.386
66. Desrosiers J, Rochette A, Hébert R, Bravo G. The Minnesota manual dexterity test: reliability, validity and reference values studies with healthy elderly people. *Can J Occup Ther (1939).* 1997;64(5):270–276. https://doi.org/10.1177/000841749706400504
67. Backman C, Gibson SCD, Parsons J. Assessment of hand function: the relationship between Pegboard dexterity and applied dexterity. *Can J Occup Ther (1939).* 1992;59(4):208–213. https://doi.org/10.1177/000841749205900406
68. Agnew PJ, Maas F. Hand function related to age and sex. *Arch Phys Med Rehabil.* 1982;63(6):269–271.
69. Sollerman C, Ejeskär A. Sollerman hand function test. A standardised method and its use in tetraplegic patients. *Scand J Plast Reconst Surg Hand Surg.* 1995;29(2):167–176. https://doi.org/10.3109/02844319509034334
70. Backman C, Mackie H. Reliability and validity of the arthritis hand function test in adults with osteoarthritis. *Occup Ther J Res.* 1997;17(1):55–66. https://doi.org/10.1177/153944929701700104
71. Backman C, Mackie H. Arthritis hand function test: Inter-rater reliability among self-trained raters. *Arthritis Rheum.* 1995;8(1):10–15. https://doi.org/10.1002/art.1790080105
72. Sears ED, Chung KC. Validity and responsiveness of the Jebsen-Taylor Hand Function Test. *J Hand Surg.* 2010;35(1):30–37. https://doi.org/10.1016/j.jhsa.2009.09.008
73. Dalbeth N, Collis J, Gregory K, Clark B, Robinson E, McQueen FM. Tophaceous joint disease strongly predicts hand function in patients with gout. *Rheumatology.* 2007;46(12):1804–1807. https://doi.org/10.1093/rheumatology/kem246
74. Ouegnin A., Valdes K. Joint position sense impairments in older adults with carpometacarpal osteoarthritis: a descriptive comparative study. J Hand Ther 33(4):547–552. ⟨doi:10.1016/j.jht.2019.01.006⟩.
75. Cantero-Téllez Ra.A.L.Aa.V.K.Aa.N.N.. Clinical effects of proprioceptive thumb exercise for individuals with carpometacarpal joint osteoarthritis: a randomized controlled trial. J Hand Ther.
76. Bell-Krotoski JA, Fess EE, Figarola JH, Hiltz D. Threshold detection and Semmes-Weinstein monofilaments. *J Hand Ther.* 1995;8(2):155–162. https://doi.org/10.1016/s0894-1130(12)80314-0
77. Umay Ea.G.Ea.S.Aa.G.Ia.U.C.. Is superficial radial nerve affected in patients with hand osteoarthritis? J Hand Ther.
78. Pelletier R, Bourbonnais D, Higgins J, Mireault M, Danino MA, Harris PG. Left right judgement task and sensory, motor, and cognitive assessment in participants with wrist/hand pain. *Rehabil Res Pract.* 2018;2018:1530245. https://doi.org/10.1155/2018/1530245

79. Magni NE, McNair PJ, Rice DA. Sensorimotor performance and function in people with osteoarthritis of the hand: a case-control comparison. *Semin Arthritis Rheumat.* 2018;47(5):676–682. https://doi.org/10.1016/j.semarthrit.2017.09.008
80. Magni NE, McNair PJ, Rice DA. Impairments in grip and pinch force accuracy and steadiness in people with osteoarthritis of the hand: a case-control comparison. *Musculoskelet Sci Pract.* 2021;55:102432https://doi.org/10.1016/j.msksp.2021.102432
81. Villafañe JH, Silva GB, Bishop MD, Fernandez-Carnero J. Radial nerve mobilization decreases pain sensitivity and improves motor performance in patients with thumb carpometacarpal osteoarthritis: a randomized controlled trial. *Arch Phys Med Rehabil.* 2012;93(3):396–403. https://doi.org/10.1016/j.apmr.2011.08.045
82. Wajon A, Vinycomb T, Carr E, Edmunds I, Ada L. Surgery for thumb (trapeziometacarpal joint) osteoarthritis. *Cochrane Database Syst Rev.* 2015; 2015(2):CD004631. https://doi.org/10.1002/14651858.CD004631.pub4
83. Vermeulen GM, Slijper H, Feitz R, Hovius SER, Moojen TM, Selles RW. Surgical management of primary thumb carpometacarpal osteoarthritis: a systematic review. *J Hand Surg.* 2011;36(1):157–169. https://doi.org/10.1016/j.jhsa.2010.10.028
84. Balthazar C, Vendrely A. *Rehabilitation Research: Principles and Applications.* St. Louis: Elsevier; 2022.
85. Cohen J. A coefficient of agreement for nominal scales. *Edu Psychol Measure.* 1960;20(1):37–46. https://doi.org/10.1177/001316446002000104
86. Chiu E-C, Wu W-C, Chou C-X, Yu M-Y, Hung J-W. Test-retest reliability and minimal detectable change of the test of visual perceptual skills-third edition in patients with stroke. *Arch Phys Med Rehabil.* 2016;97(11):1917–1923. https://doi.org/10.1016/j.apmr.2016.04.023
87. Safari S, Baratloo A, Elfil M, Negida A. Evidence based emergency medicine part 2: positive and negative predictive values of diagnostic tests. *Emergency.* 2015;3(3):87–88.
88. Furlan L, Sterr A. The applicability of standard error of measurement and minimal detectable change to motor learning research-a behavioral study. *Front Huma Neurosci.* 2018;12:95. https://doi.org/10.3389/fnhum.2018.00095
89. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. *Tutor Quant Methods Psychol.* 2012;8(1):23–34. https://doi.org/10.20982/tqmp.08.1.p023
90. Lalkhen AG, McCluskey A. Clinical tests: sensitivity and specificity. *Continuing Edu Anaesth Crit Care Pain.* 2008;8(6):221–223. https://doi.org/10.1093/bjaceaccp/mkn041
91. Therapists A, So H. *Clinical Assessment Guidelines.* 3rd ed., New Jersey: American Society of Hand Therapists; 2015.