**Measurement properties of patient-reported hand function measures in rheumatoid arthritis: a systematic review protocol**

Hisham Arab Alkabeya1, 2, Beatrice E. A. Sankah1, Ann-Marie Hughes1 & Jo Adams1

1 School of Health Sciences, University of Southampton, Southampton, UK

2 Faculty of Allied Medical Sciences, Arab American University, Jenin, Palestine

**CORRESPONDING AUTHOR:**

Hisham Arab Alkabeya,

School of Health Sciences,

Building 67,

University of Southampton,

Highfield Campus,

Southampton,

SO17 1BJ,

United Kingdom,

Email: H.Arab-Alkabeya@soton.ac.uk, Tel: +44 (0) 7599623946

Hisham Arab Alkabeya, Doctoral Research Fellow, MSc BSc,

Beatrice E. A. Sankah, Doctoral Research Fellow, MSc BSc

Ann-Marie Hughes, Associate Professor, PhD MSc PGDip BSc

Jo Adams, Professor of Musculoskeletal Health, PhD MSc DipCOT

Word count

2,529

**Measurement properties of patient-reported hand function measures in rheumatoid arthritis: a systematic review protocol**

 **Abstract**

**Background**: Patient-reported outcome measures (PROMs) are commonly used to evaluate hand function in people with Rheumatoid Arthritis (RA). A decision will always need to be made about which appropriate PROMs to use. The present review therefore aims to describe the available hand function PROMs for use in people with RA by appraising their methodological quality and psychometric properties using a contemporaneous method.

 **Methods/design**: The proposed systematic review will include published studies written in English, which report evidence for psychometric properties and/or practical properties of hand function PROMs in RA. Four major databases (MEDLINE, Embase, PsycINFO, and CINAHL) will be searched from inception to May 2019. A three-staged search strategy will be applied: (1) electronic bibliographic databases for published studies, (2) “named measures” searching approach, and (3) reference lists of studies with included PROMs. The proposed systematic review will be conducted in compliance with the COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) guideline for systematic review of PROMs. Accordingly, the methodological quality of the included studies will be assessed against the updated COSMIN risk of bias checklist, and each study’s results will be assessed for their psychometric quality.

**Conclusion**: The proposed systematic review seeks to provide rigour, and transparent evaluation of PROMs used to evaluate hand function in the RA population. The findings will provide clarity for healthcare professionals and researchers on the appropriate PROMs for hand function assessment. It will also provide a summary of hand function PROM recommendations for RA.

**Keywords**: hand function, rheumatoid arthritis, systematic review, patient reported, measurement properties

# Introduction

Rheumatoid arthritis (RA) is the most prevalent form of inflammatory arthritis and affects approximately 1% of the world’s population [1]. Hand involvement is typically present in about 80-90% of the people with RA [2]. Some common impairments associated with hand RA are pain, stiffness, deformities, limited Range of Motion (ROM) and functional limitations which often present in symmetrical patterns [3]. The importance of hand function in RA is underlined by the fact that RA has a particular predilection for the hands and hand function is an essential component in performing Activities of Daily Living (ADLs). Furthermore, hand function domain is included in the International Classification of Functioning, Disability and Health (ICF) core set for RA [4]. Therefore, it is important to measure, interpret and evaluate hand function in clinical practice and include hand outcomes within the core set of outcomes to be included in trials involving people with hand RA [5]. The assessment of hand function in hand RA is crucial in determining the extent of functional loss and evaluating the outcomes of both surgical and rehabilitative procedures/interventions [6]. The term functional assessment has been used to represent a wide range of assessment techniques and is often used inconsistently [7]. Although the clinical assessment of hand function and disability remains complex and debateable [8], it is essential to measure the progression of hand disability in RA to understand the impact of the disease, determine treatment strategies and evaluate interventions used in hand RA management.

Using the ICF terms, RA patients’ hand function limitations are evaluated through measuring impairments (e.g. ROM and grip strength), whilst limitations in hand activity are documented using both performance-based measures and patient-reported outcome questionnaires [9, 10]. Hand impairment measures focus on reflecting the consequences of the disease at the bodily musculoskeletal level. Although impairment measures are relatively simple to obtain, they are reported to be limited in demonstrating how well patients perform their ADLs and often do not capture the full extent of patient disability [11]. Besides, for RA patients, the level of hand disability in daily life may be of greater importance to them than the level of impairment [10, 12]. Impairments of hand function can lead to both physical and psychosocial difficulties in accomplishing daily life activities [12]. Therefore, hand function can be defined as the ability to perform daily life activities [9].

Performance-based measures assess an individual’s ability to perform prescribed tasks. Examples of these measures include Jebsen Taylor Hand Function Test [13] and Grip Ability Test (GAT) [14]. These measures include tasks that simulate real life tasks and might not be meaningful to the patients’ real-life activities. The results obtained are effort dependent and may be influenced by mood[15, 16]. This led to a shift in the focus of methods used to describe health outcomes related to the hand manifestations of RA toward patient-reported outcome measures (PROMs), which goes with a world-wide rheumatology trend for more patient-centred care [17]. PROMs of hand function are standardised and validated self-administrated questionnaires that provide quantitative information about hand use in daily life from the patients’ perspective. A wide variety of hand self-reported questionnaires that assess disease symptoms and activity limitations have been used with RA patients [18]. Although hand self-reported questionnaires have been recommended with RA patients [11], consensus on which of the available instruments should be used in clinical and research setting is currently unavailable [19]. Given the vast number of available PROMs for hand function, the selection of an appropriate measure necessitates the careful consideration of their measurement properties. A scoping search identified two relevant narrative review studies reporting the psychometric properties of PROMs hand function used in arthritis population [18, 20]. However, both reviews presented evidence of the measurement properties of a few PROMs based on data available from the arthritis population in general, which does not provide RA specific evidence. Additionally, due to the nature of narrative reviews, they are not systematic or transparent in their approach to synthesis [21] hence their results are not conclusive. Considering these shortcomings, we concluded that a full systematic review is required. High-quality systematic reviews that employ up-to-date evidence synthesis and recommendations can provide a comprehensive overview of all available measures and support evidence-based recommendations in the selection of the most appropriate PROMs for a given purpose [22]. Therefore, this systematic review aims to critically appraise, compare and summarise the measurement properties (i.e. reliability, criterion validity, construct validity, responsiveness) of all available PROMs for hand function in RA, considering the methodological quality of these studies as well as the quality of evidence. Based on findings of this review, recommendations regarding appropriate PROMs for evaluating hand function in RA will be proposed.

# Methods

## Design

This systematic review will be conducted in compliance with the recently published COSMIN (COnsensus-based Standards for the selection of health Measurement Instruments) guideline for the systematic review of PROMs [22]. Details of the protocol are registered on the PROSPERO (International Prospective Register of Ongoing Systematic Reviews) registry (CRD42019122087). The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [23] and COSMIN [22] checklist will be used to guide the conduct and reporting of this systematic review. Any protocol amendments will be reported in the systematic review publication.

## Search strategy

Prior to the commencement of the present systematic review, the PROSPERO database was searched to avoid duplication of the potential review evidence. From this search, no ongoing or completed systematic reviews were identified. A three-staged search strategy will be used to identify all published studies on the psychometric prosperities and/or interpretability of hand function PROMs in hand RA that meet the inclusion criteria. First, a computerized literature search will be performed in the following databases from inception to May 2019: MEDLINE (Ovid), EMBASE (Ovid), PsycINFO (EBSCO) and CINAHL Plus (EBSCO). The search strategy will be constructed in line with the guidelines published by the COSMIN [22], and guidance from Terwee et al. (2009) [23]. Medical Subject Headings terms and free text searching will be used to reflect three key elements (1) population- RA, (2) construct-hand function, (3) measurement properties (a sensitive and validated search filter published by Terwee et al. (2009) [23]. An example of a search strategy in MEDLINE (Ovid) is presented in appendix 1. Second, following the review of titles and abstracts of included studies, an additional “named measures” search will be developed and applied in each database, with search terms developed as above to reflect: (1) population- RA, (2) construct-hand function, (3) named hand function measure, (4) measure measurement properties. Finally, the reference lists of all selected articles will be searched to achieve a more comprehensive search and identify additional relevant studies.

The search strategy will be constructed and modified for each of the databases in collaboration with an experienced librarian. All searches will be performed by two independent researchers (HA and BS) on the same day. Results will be exported to a single EndNote® (version X8) library. Then duplicate records will be deleted using the EndNote automatic de-duplication option.

## Eligibility criteria

To be eligible for selection, a study must focus on a PROM(s) used for measuring hand function (the construct of interest). In addition, the following inclusion and exclusion criteria will be applied:

### Inclusion criteria

1. Studies reporting the assessment of one or more psychometric properties or evaluating the interpretability (e.g. determination of clinically important change) of hand function PROM(s).
2. The study population should be adults (≥18) with RA making up at least 50% of the total study sample.
3. Studies reporting cross-cultural validation of hand function PROM(s) in the RA population.
4. Full-text articles published in peer reviewed English language journals.

### Exclusion criteria

1. Studies reporting the assessment of psychometric properties of performance based, or clinician-rated, instruments for hand function.
2. Questionnaires developed for evaluating hand function in the juvenile rheumatoid arthritis (JRA) population.
3. Studies or trials evaluating the effectiveness of interventions where a hand function PROM is used as an end point.
4. Editorial letters, reviews, conference abstracts and dissertations.

## Selection of articles

Following the step of removing duplicates, the study selection process will be completed in two stages. The first stage includes examining only the titles and abstracts of the search results to eliminate all clearly ineligible publications. The second stage encompasses a full-text review of articles that appear to meet the inclusion criteria or in cases when a decision cannot be made based on the title and abstract alone. All stages will be completed by two reviewers (HA and BS) independently. At any stage, if the reviewers are unable to reach a consensus, reviewers (AMH or JA) within this review will be consulted, and disagreement will be resolved through dialogue. Reasons for exclusion at the full-text stage of screening will be recorded. The process of selection will be summarized using a PRISMA flow diagram.

## Data extraction

Data from the selected studies will be extracted independently by two reviewers (HA and BS), using data collection forms adopted from the COSMIN methodology guideline [24]. The COSMIN taxonomy will be used to decide which property has been evaluated in the study [25]. Discrepancies between the two reviewers will be resolved through discussion and, if necessary, through consultation with a third reviewer (AMH or JA). Data on the following will be extracted:

1. Characteristic of the study population including the age, gender, ethnicity, disease duration, and questionnaire administration.
2. Questionnaire characteristics including name, language, response options, scoring method, domains, number of items and recall period.
3. Evidence regarding the measurement properties of the questionnaire including data analysis, and results of the measurement properties.
4. Operational characteristics such as interpretability, patient acceptability, and feasibility of administration for staff will be reported.

The interpretability and feasibility related to the measure being evaluated will be extracted and reported using the COSMIN data extraction forms designed for this purpose [24].

## Assessment of the methodological quality of the included studies

The methodological quality of the included studies on a measurement property will be evaluated using the updated COSMIN risk of bias checklist [22, 26]. The updated COSMIN risk of bias checklist contains 10 domains, two for content validity and the remaining for cross-cultural validity/measurement invariance, reliability, internal consistency, measurement error, criterion validity, hypothesis testing for construct validity and responsiveness. The quality of each study on a measurement property will be assessed individually using the pertinent COSMIN box. A four-point rating scale (i.e., very good, adequate, doubtful, inadequate) of the COSMIN checklist will be used to rate each study, with the lowest score in each category considered the final overall rating for the methodological quality in that category for the study assessed. For instance, if the measurement error was rated “very good” in one question, but “doubtful” on another, the overall score for measurement error would be “doubtful”. The assessment of the included studies will be evaluated against this criterion, and a summary score will be presented. The COSMIN steering committee reached a consensus that no “gold standard” exists for PROMs [24]. Therefore, included studies will be categorized as measuring criterion validity only when a shortened version of a PROM is compared to the original long version. Measurement properties regarding PROMs content validity require the inclusion of studies of the development of the measures as well as studies focusing on content validity and expert opinion. However, for this proposed systematic review, content validity will not be evaluated. This is because the majority of hand function PROMs commonly used with persons with rheumatic diseases were not developed to evaluate hand RA [18]. Therefore, a comprehensive evaluation of content validity of all available RA hand function PROMs should be conducted in a future systematic review. The overall interrater agreement percentages of the original COSMIN risk of bias checklist was reported to be high and interrater reliability was reasonable [27]. The new version of the COSMIN risk of bias checklist has been used in a recent systematic review of PROMs for soft-tissue facial reconstruction and demonstrated very good interrater reliability (ICC=.81-87) [28].

## Assessment of the study results against criteria for good measurement properties

Findings from each study on the measurement properties will be evaluated using the recently updated good measurement properties criteria [22], wherein the quality of each study’s results will be rated as either “sufficient”, “insufficient” or “indeterminate”. The assessment of the methodological quality and psychometric properties of included studies will be performed by four reviewers who will work in pairs; any discrepancies encountered will be resolved through discussion.

## Evidence synthesis

The results obtained from the two assessments detailed above will be pooled and used to produce a global score for each measurement property of each PROM as outlined by the COSMIN guideline [22, 24]. Results can be sufficient (positive), insufficient (negative), inconsistent, or indeterminate by employing a “75 percent in agreement” rule (i.e., for a positive outcome on internal consistency, 75 percent or more of the studies reporting internal consistency must be positive) [24]. If the results per study are all sufficient (or all insufficient), the overall rating will also be sufficient (or insufficient). If the results are inconsistent, then exploration is warranted to determine the reasons for inconsistency. Where reasons are identified, overall ratings will be provided for relevant subgroups. Where no reasons could be found, then the overall rating will be inconsistent and in the case where inadequate information is available, the overall rating will be indeterminate.

Following the above stage, all evidence per measurement property of PROMs will be graded using the modified Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach [22, 24]. Using this approach, the quality of evidence will be graded as high, moderate, low or very low evidence. Four out of the five factors of the modified GRADE approach namely risk of bias, inconsistency, indirectness and imprecision will be used to determine the quality of the evidence. The fifth factor (i.e. publication bias) according to Mokkink et al [24] is difficult to assess in studies on measurements properties, hence not considered in the methodology for systematic reviews of PROMs. Premised on the above, the choice to exclude publication bias in the evaluation of the overall quality of evidence of the proposed systematic review was made. If the overall rating for a specific measurement property is indeterminate, then no grading of the quality will be given since the quality of the measure cannot be decided. Detailed methods for using the modified GRADE approach is described in the COSMIN manual for systematic reviews for PROMs [24]. Two independent reviewers (HA & BS) will perform the evidence synthesis, and a third reviewer (AMH or JA) will be consulted in case of unresolved disagreement. Finally, based on the combined results of each measurement category and modified GRADE evaluation, recommendations will be formulated on the appropriateness of each identified PROM.

# Conclusion

This proposed systematic review seeks to systematically identify and critically appraise PROMs used in assessing hand function in people with hand RA. It will provide a comprehensive summary of available hand function PROMs and the quality of their measurement properties. In addition, findings of this proposed systematic review will inform the selection of PROMs evaluating hand function for both clinical and research purposes with an aim to enhance hand function assessment in routine clinical practice, service evaluation and research.

**Disclosure statement**

Reviewers report no conflict of interest.

# References

1. Gibofsky A. Overview of epidemiology, pathophysiology, and diagnosis of rheumatoid arthritis. Am J Manag Care. 2012;18(13 Suppl):S295-302.

2. Durmus D, Uzuner B, Durmaz Y, et al. Michigan Hand Outcomes Questionnaire in rheumatoid arthritis patients: relationship with disease activity, quality of life, and handgrip strength. J Back Musculoskelet Rehabil. 2013;26(4):467-473.

3. Horsten NC, Ursum J, Roorda LD, et al. Prevalence of hand symptoms, impairments and activity limitations in rheumatoid arthritis in relation to disease duration. J Rehabil Med. 2010;42(10):916-21.

4. Stucki G, Cieza A, Geyh S, et al. ICF Core Sets for rheumatoid arthritis. J Rehabil Med. 2004;(44 Suppl):87-93.

5. Eberhardt K, Sandqvist G, Geborek P. Hand function tests are important and sensitive tools for assessment of treatment response in patients with rheumatoid arthritis. Scand J Rheumatol. 2008;37(2):109-112.

6. Jones LA, Lederman SJ. Human hand function. Oxford: Oxford University Press; 2006.

7. Kimmerle M, Mainwaring L, Borenstein M. The functional repertoire of the hand and its application to assessment. Am J Occup Ther. 2003;57(5):489-98.

8. Goodson A, McGregor AH, Douglas J, et al. Direct, quantitative clinical assessment of hand function: Usefulness and reproducibility. Man Ther. 2007;12(2):144-152.

9. Fowler NK, Nicol AC. Functional and biomechanical assessment of the normal and rheumatoid hand. Clin Biomech (Bristol, Avon). 2001;16(8):660-6.

10. van Lankveld WG, van 't Pad Bosch P, van de Putte L. Predictors of changes in observed dexterity during one year in patients with rheumatoid arthritis. Br J Rheumatol. 1998;37(7):733-9.

11. Waljee JF, Chung KC, Kim HM, et al. Validity and responsiveness of the Michigan Hand Questionnaire in patients with rheumatoid arthritis: a multicenter, international study. Arthritis Care Res (Hoboken). 2010;62(11):1569-77.

12. Nicklasson M, Jonsson H. Experience of participation as described by people with hand deformity caused by rheumatic disease. Br J Occup Ther. 2012;75(1):29.

13. Jebsen RH, Taylor N, Trieschmann RB, et al. An objective and standardized test of hand function. Arch Phys Med Rehabil. 1969;50(6):311-9.

14. Dellhag B, Bjelle A. A Grip Ability Test for use in rheumatology practice. J Rheumatol. 1995;22(8):1559-65.

15. Hochberg M, Silman A, Smolen J, et al. Rheumatoid arthritis Philadelphia: Mosby/Elsevier; 2009.

16. Louie GH, Reveille JD, Ward MM. Challenges comparing functional limitations in rheumatoid arthritis and ankylosing spondylitis. Clin Exp Rheumatol. 2009;27(4 Suppl 55):S83-91.

17. Gossec L, Dougados M, Dixon W. Patient-reported outcomes as end points in clinical trials in rheumatoid arthritis. RMD Open. 2015;1(1):e000019.

18. Poole JL. Measures of hand function: Arthritis Hand Function Test (AHFT), Australian Canadian Osteoarthritis Hand Index (AUSCAN), Cochin Hand Function Scale, Functional Index for Hand Osteoarthritis (FIHOA), Grip Ability Test (GAT), Jebsen Hand Function Test (JHFT), and Michigan Hand Outcomes Questionnaire (MHQ). Arthritis Care Res (Hoboken). 2011;63 Suppl 11:S189-99.

19. Klokker L, Terwee CB, Waehrens EE, et al. Hand-related physical function in rheumatic hand conditions: a protocol for developing a patient-reported outcome measurement instrument. BMJ Open. 2016;6(12):e011174.

20. Poole JL. Measures of adult hand function: Arthritis Hand Function Test (AHFT), Grip Ability Test (GAT), Jebsen Test of Hand Function, and The Rheumatoid Hand Functional Disability Scale (The Duruöz Hand Index [DHI]). Arthritis Rheum. 2003;49(S5):S59-S66.

21. Popay J, Roberts H, Sowden A, et al. Guidance on the conduct of narrative synthesis in systematic reviews. A product from the ESRC methods programme Version. 2006;1:b92.

22. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res. 2018;27(5):1147-1157.

23. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med. 2009;6(7):e1000100.

24. Mokkink LB, Prinsen C, Patrick DL, et al. COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs). User manual. 2018.

25. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010;63(7):737-45.

26. 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(5):1171-1179.

27. Mokkink LB, Terwee CB, Gibbons E, et al. Inter-rater agreement and reliability of the COSMIN (COnsensus-based Standards for the selection of health status Measurement Instruments) checklist. BMC Med Res Methodol. 2010;10:82.

28. Dobbs TD, Gibson JAG, Hughes S, et al. Patient-Reported Outcome Measures for Soft-Tissue Facial Reconstruction: A Systematic Review and Evaluation of the Quality of Their Measurement Properties. Plast Reconstr Surg. 2019;143(1):255-268.

**Appendix 1**

Sample search strategy for MEDLINE (Ovid)

|  |  |
| --- | --- |
| # | Search terms |
| 1 | HAND/ OR ((hand or hands) adj3 (activit\* or abilit\* or function\* or perform\* or skill\* or impair\* or disabilit\*)).ti,ab. |
| 2 | rheumatoid arthritis/ or RA.ti,ab. or "Rheumatoid Arthritis".ti,ab. |
| 3 | (instrumentation or methods).sh. OR (Validation Studies or Comparative Study).pt. OR exp Psychometrics/ OR psychometr\*.ti,ab. OR (clinimetr\* or clinometr\*).tw. OR exp “Outcome Assessment (Health Care)”/ OR outcome assessment.ti,ab. OR outcome measure\*.tw. OR exp Observer Variation/ OR observer variation.ti,ab. OR exp Health Status Indicators/ OR exp “Reproducibility of Results”/ OR reproducib\*.ti,ab. OR exp Discriminant Analysis/ OR (reliab\* or unreliab\* or valid\* or coefficient or homogeneity or homogeneous or “internal consistency”).ti,ab. OR (cronbach\* and (alpha or alphas)).ti,ab. OR (item and (correlation\* or selection\* or reduction\*)).ti,ab. OR (agreement or precision or imprecision or “precise values” or test-retest).ti,ab. OR (test and retest).ti,ab. OR (reliab\* and (test or retest)).ti,ab. OR (stability or interrater or inter-rater or intrarater or intra-rater or intertester or inter-tester or intratester or intra-tester or interobserver or inter-observer or intraobserver or intraobserver or intertechnician or inter-technician or intratechnician or intra-technician or interexaminer or inter-examiner or intraexaminer or intra-examiner or interassay or interassay or intraassay or intra-assay or interindividual or inter-individual or intraindividual or intra-individual or interparticipant or inter-participant or intraparticipant or intra-participant or kappa or kappa’s or kappas or repeatab\*).ti,ab. OR ((replicab\* or repeated) and (measure or measures or findings or result or results or test or tests)).ti,ab. OR (generaliza\* or generalisa\* or concordance).ti,ab. OR (intraclass and correlation\*).ti,ab. OR (discriminative or “known group” or factor analysis or factor analyses or dimension\* or subscale\*).ti,ab. OR (multitrait and scaling and (analysis or analyses)).ti,ab. OR (item discriminant or interscale correlation\* or error or errors or “individual variability”).ti,ab. OR (variability and (analysis or values)).ti,ab. OR (uncertainty and (measurement or measuring)).ti,ab. OR (“standard error of measurement” or sensitiv\* or responsive\*).ti,ab. OR ((minimal or minimally or clinical or clinically) and (important or significant or detectable) and (change or difference)).ti,ab. OR (small\* and (real or detectable) and (change or difference)).ti,ab. OR (meaningful change or “ceiling effect” or “floor effect” or “Item response model” or IRT or asch or “Differential item functioning” or DIF or “computer adaptive testing” or “item bank” or “cross-cultural equivalence”).ti,ab. |
| 4 |  1 AND 2 AND 3  |