Journal of Hand Therapy xxx (2017) 1–10

Contents lists available at ScienceDirect

Journal of Hand Therapy

journal homepage: www.jhandtherapy.org

Scientific/Clinical Article

The reliability and measurement error of protractor-based goniometry of the fingers: A systematic review

Yara E. van Kooij PT, MSc [a,*], Alexandra Fink PT, MSc [a], Maria W. Nijhuis-van der Sanden PT, PhD [b], Caroline M. Speksnijder PT, MSc, PhD [b,c,d]

[a] Handtherapie Nederland, Utrecht, The Netherlands
[b] Radboud University Medical Center, Radboud Institute for Health Sciences, Scientific Institute for Quality of Healthcare (IQ healthcare), Nijmegen, The Netherlands
[c] Physical Therapy Science, University Medical Center Utrecht, Program in Clinical Health Sciences, Utrecht, The Netherlands
[d] Department of Oral-Maxillofacial Surgery, University Medical Center Utrecht, Prosthodontics and Special Dental Care, Utrecht, The Netherlands

ARTICLE INFO

Article history:
Received 30 January 2016
Received in revised form
24 February 2017
Accepted 28 February 2017
Available online xxx

Keywords:
Finger
Goniometry
Range of motion
Reliability
Measurement error

ABSTRACT

Study Design: Systematic review
Purpose of the Study: The purpose was to review the available literature for evidence on the reliability and measurement error of protractor-based goniometry assessment of the finger joints.
Methods: Databases were searched for articles with key words "hand," "goniometry," "reliability," and derivatives of these terms. Assessment of the methodological quality was carried out using the Consensus-Based Standards for the Selection of Health Measurement Instruments checklist. Two independent reviewers performed a best evidence synthesis based on criteria proposed by Terwee et al (2007).
Results: Fifteen articles were included. One article was of fair methodological quality, and 14 articles were of poor methodological quality. An acceptable level for reliability (intraclass correlation coefficient > 0.70 or Pearson's correlation > 0.80) was reported in 1 study of fair methodological quality and in 8 articles of low methodological quality. Because the minimal important change was not calculated in the articles, there was an unknown level of evidence for the measurement error.
Discussion: Further research with adequate sample sizes should focus on reference outcomes for different patient groups. For valid therapy evaluation, it is important to know if the change in range of motion reflects a real change of the patient or if this is due to the measurement error of the goniometer. Until now, there is insufficient evidence to establish this cut-off point (the smallest detectable change).
Conclusion: Following the Consensus-Based Standards for the Selection of Health Measurement Instruments criteria, there was limited level of evidence for an acceptable reliability in the dorsal measurement method and unknown level of evidence for the measurement error.
Level of Evidence: 2a

Introduction

In daily practice, the measurement of active range of motion (ROM) of the fingers is a common procedure in hand therapy. It is used to assess joint restrictions and to evaluate the effectiveness of interventions.[1,2] The ROM of the fingers can be measured with different methods such as wire tracing,[3] visual estimation,[4] composite finger flexion,[5] and goniometery.[6] Hand therapists use a wide selection of goniometric instruments, among these are numerous protractor-based goniometers, pendulum devices, multiangle goniometers, and computerized goniometers.[6,7] The protractor-based goniometer is the most frequently used instrument in clinical practice.[6]

Therapists apply both dorsal and lateral methods and use a diversity of protractor-based goniometers to assess the ROM in the hand.[1,6] In a survey, only 41% of the therapists had implemented a standardized protocol,[1] although it is known that the use of a standardized protocol increases the reliability of ROM assessment.[8-10] Recently, the American Society of Hand Therapists published the third edition of the guideline for hand goniometry in their clinical assessment recommendtions.[7] However, to date, no systematic appraisal of studies has been conducted to evaluate the reliability and measurement error of the goniometer for active ROM

* Corresponding author. Handtherapie Nederland, Wallesteinlaan 45, 3554 HM Utrecht, The Netherlands.
E-mail address: y.vankooij@handencentrum.nl (Y.E. van Kooij).

0894-1130/$ – see front matter © 2017 Hanley & Belfus, an imprint of Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jht.2017.02.012

of the fingers. Knowledge about both the reliability and measurement error is a prerequisite for a valid evaluation of therapy interventions and to qualify joint restrictions.[11] In addition, it is not clear how different types of protractor-based goniometers and placement preferences influence outcomes.[6]

The primary aim of this study was to review the available literature for the reliability and measurement error of protractor-based goniometry assessment of the finger joints and to provide recommendations for standardized evidence-based practice. The secondary aim was to evaluate the influence of design and placement preferences of the protractor-based goniometer on the reliability and measurement error.

Methods

Data sources and searches

The following computerized bibliographic databases were searched up to April 2016: PubMed (1966-2015), Embase (1974-2015), and CINAHL (EBSCOhost; 1981-2015). The databases were searched with the keywords "hand," "goniometry," "reliability," and derivatives of these terms (Appendix I).[12] Reference lists were screened to identify additional relevant studies.

Study selection

A study was included if it was a full-text original article, primarily concerning the evaluation of the reliability and measurement error of finger goniometry using protractor-based goniometers during active ROM assessment. Movements should consist of flexion and/or extension of the fingers. We also included studies in which the fingers were fixed (in an orthosis) during the measurements. This fixed position is used in healthy participants to reproduce common clinical limitation of flexion and/or extension of the fingers. The participants in the study had to be adults (>18 years). Articles in all languages were included.

Studies were excluded if no clear description of placement method, design, and type of the goniometer was described; if passive ROM measurements were described because passive measurements are strongly affected by the variation of force applied by the therapist[13]; and if computerized goniometers were used because therapists prefer protractor-based goniometers rather than computerized measurements.[6] Two reviewers (Y.K. and A.F.) independently assessed titles and abstracts of the studies retrieved by the literature search. In case of disagreement between the 2 reviewers, a third reviewer (C.M.S.) made the decision regarding inclusion of the article. Authors were contacted if the full text was not available.

Data extraction

Two reviewers (Y.K. and C.M.S.) extracted relevant characteristics of the studies. These characteristics were country and setting of the study, characteristics of the participants and raters, used instrumentation, placement preferences, joint movement, reliability, and measurement error scores.

Measurement properties

Reliability is the extent to which scores for patients who have not changed are the same for repeated measurements. Interrater reliability means that the repeated measurements are performed by 2 or more persons at the same moment.[11] Intrarater reliability means that both measurements are done by the same person(s) on different occasions with the assumption that the patient is stable in this specific interval.[11] Measurement error is the systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured.[11] A detailed discussion and definition on these measurement properties have been published elsewhere.[14] Based on criteria proposed by Terwee et al, the possible overall rating for the reliability and measurement error of the included studies was "positive," "indeterminate," or "negative" (Table 1).[15] A positive rating for the reliability is an intraclass correlation coefficient (ICC) score of at least 0.70 or a Pearson's correlation (Pearson's r) of 0.80. A positive rating for measurement error is given when (1) the smallest detectable change (SDC) is smaller than the minimal important change (MIC); (2) the MIC is outside the limits of agreement (LoA); and (3) there are convincing arguments that the measurement error is acceptable (Table 1).[15]

Quality assessment

To determine the methodological quality of the included studies the Consensus-Based Standards for the Selection of Health Status Measurement Instruments (COSMIN) checklist was used. The COSMIN checklist was developed to investigate the methodological quality of health questionnaires and consists of 9 boxes with standards for the optimal methodological designs.[16] The methodological quality of each measurement property is evaluated in 1 of the 9 boxes of the checklist. For this review, the assessment of the methodological quality of the included studies was carried out using the boxes' reliability and measurement error.[16] Each item can be scored on a 4-point scale (ie, "poor," "fair," "good," or "excellent").[17] An overall score for the methodological quality of a study was determined by taking the lowest rating of any of the items. The validity and reliability of the COSMIN checklist is appropriate.[18] However, the COSMIN checklist was developed for studies that investigate health questionnaires. Regarding sample size, the COSMIN recommendation of 50 participants may differ for a measurement instrument such as the goniometer. Therefore, if a study described an appropriate power calculation and reached the needed number of participants, the study was scored excellent on the sample size parameter.

Two reviewers (C.M.S. and Y.K.) independently performed data extraction and assessment of the methodological quality of the included studies. In case of disagreement, a third reviewer (A.F.) made the decision. C.M.S. is a senior physical therapist and senior epidemiologist and trained by the COSMIN team on quality assessment and data extraction. Both Y.K. and A.F. are senior physical therapists and were trained by C.M.S. on the COSMIN checklist.

Best evidence synthesis—levels of evidence

In the best evidence synthesis, the results of the different included studies were combined adjusted for their methodological

Table 1
Quality criteria for reliability and measurement error (based on Terwee et al)[15]

Property	Rating[a]	Quality criteria
Reliability	+	ICC/weighted kappa $\geq$0.70 OR Pearson's $r \geq 0.80$
	?	Neither ICC/weighted kappa nor Pearson's r determined
	−	ICC/weighted kappa < 0.70 OR Pearson's $r < 0.80$
Measurement error	+	MIC > SDC OR MIC outside the LoA
	?	MIC not defined
	−	MIC $\leq$ SDC OR MIC equals or inside LoA

MIC = minimal important change; SDC = smallest detectable change; LoA = limits of agreement; ICC = intraclass correlation coefficient.
[a] + = positive rating; ? = indeterminate rating; − = negative rating.