Is Diagrammatic Goniometry Feasible for Finger ROM Evaluation and Self-evaluation?

Valdas Macionis MD

Received: 30 July 2012 / Accepted: 20 December 2012 / Published online: 5 January 2013
© The Association of Bone and Joint Surgeons® 2013

Abstract

Background While “diagrammatic” evaluation of finger joint angles using two folded paper strips as goniometric arms has been proposed and could be an alternative to standard goniometry and a means for self-evaluation, the measurement differences and reliability are unknown.

Questions/purposes This study assessed the standard and diagrammatic finger goniometry performed by an experienced examiner on patients in terms of (1) intragoniometer and intergoniometer (ie, intrarater) differences and reliability; (2) interrater differences and reliability relative to patients’ diagrammatic self-evaluation; and (3) the interrater differences related to patient’s hand dominance.

Each author certifies that he or she, or a member of his or her immediate family, has no funding or commercial associations (eg, consultancies, stock ownership, equity interest, patent/licensing arrangements, etc) that might pose a conflict of interest in connection with the submitted article.

All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research* editors and board members are on file with the publication and can be viewed on request.

Each author certifies that his or her institution approved the human protocol for this investigation, that all investigations were conducted in conformity with ethical principles of research, and that informed consent for participation in the study was obtained.

V. Macionis

Clinic of Rheumatology, Orthopedics and Traumatology,
and Reconstructive Surgery, Vilnius University Faculty
of Medicine, Vilnius, Lithuania

V. Macionis (✉)

Department of Plastic and Reconstructive Surgery, Centre
of Plastic and Reconstructive Surgery, Vilnius University
Hospital Santariskiu Klinikos, Zygimantu 3/1, LT-01102
Vilnius, Lithuania
e-mail: valdas.macionis@mf.vu.lt

Methods Sixty-one patients without previous training self-evaluated active extension of all joints of the fifth finger of one hand once using two rectangular strips of paper. A practitioner used a goniometer and a diagram to perform parallel evaluations once in 12 patients and three times in 49 patients. The diagrams were scanned and measured. All evaluations and proportions of differences between the paired measurements of 5° or less were combined for analysis.

Results Intrarater intraclass correlation coefficients (ICC) based on the second and third practitioner’s trials for the proximal interphalangeal joint were greater than 0.99. Reliability was poor when calculations involved the first measurement of the practitioner (ICCs < 0.38). Interrater reliability was poor regardless of the practitioner’s trial (ICCs < 0.033). The proportions of the absolute differences of 5° or less between all paired practitioner’s measurements were similar. The proportions of the acceptable differences between paired practitioner’s and patients’ measurements were nonequivalent for the interphalangeal joints. The interrater differences did not depend on patients’ handedness.

Conclusions In experienced hands both techniques produce clinically comparable reliability, but patients’ performance in extempore diagrammatic self-evaluation is inadequate. Further studies are necessary to explore whether appropriate training of patients can improve consistency of diagrammatic self-evaluation.

Level of Evidence Level III, diagnostic study. See Guidelines for Authors for a complete description of levels of evidence.

Introduction

The clinical importance of reliable and universally interpretable measurements of ROM has been acknowledged

[24]. Reliability is of particular importance when assessing multiarticular finger motion prone to errors owing to various anatomic, biomechanical, and clinical factors [22].

Recording finger ROM is more easily understood if expressed graphically [28]. Diagrams have been used to explain notation of ROM of the other joints [18, 24]. Clinicians and patients can read diagrams showing actual positions of the appropriate segments of the joints. One also can measure the drawings to obtain the traditional expression of joint angles in degrees. Drawing the finger dorsal silhouette by means of wire tracing [17] is probably the only method of graphic finger motion recording that has gained any substantial use [2, 19, 21, 30, 35]. However, the wire method is unreliable [13] and technically complex. The traditional protractor can be replaced by other newer and reasonably reliable techniques of finger ROM measurement but these new methods involve high technology [7, 8, 15, 32, 33, 37, 38] or require mathematical transformation of measurements [6]. Goniometry thus remains the standard technique. In everyday practice, clinicians need simple and readily available evaluation techniques. The simplest assessment of finger ROM by visual estimation is of low accuracy (median percentage error, 22%) [31] and may lead to as much as 51° errors [4]. Similar limitations apply to another quick technique of linear measurement of flexion deficit [26] in which the distance between the pulp and distal palm crease is determined.

I recently showed that finger ROM could be evaluated by using a diagrammatic technique in which the joint angles are obtained by appropriately aligning two folded paper strips as goniometric arms and then tracing their edges [27]. The key advantages of this method include ease and rapidity of making a substitute for a standard goniometer of custom dimensions and shapes. Another characteristic of the paper strip technique is the possibility of obtaining the measurement with one hand.

Patient self-care is a major part of the overall treatment of hand disorders accompanied by limited ROM of finger joints. Frequent followups are necessary to assess effectiveness of patients' efforts. It would be helpful if such patients were able to self-evaluate finger motion at home with some readily available instrument. Self-evaluation of motion would be especially useful if used as a means of enhancing patients' motivation to participate in the prescribed treatment. Routine ROM measurements can be regarded as a contribution to the overall treatment [1, 29]. Features of the diagrammatic technique could enable patients to self-evaluate their finger motion [27].

Therefore, to explore clinical feasibility of the paper strip diagrammatic goniometry of finger joints, this study assessed the standard and diagrammatic finger goniometry performed by an experienced examiner on patients in terms

of (1) intragoniometer and intergoniometer (ie, intrarater) measurement differences and reliability; (2) interrater measurement differences and reliability relative to patients' diagrammatic self-evaluation; and (3) the interrater measurement differences related to patient's hand dominance.

Patients and Methods

Sixty-one patients (12 women, 49 men) with a deficit of active extension of the fifth finger were included in the study, as a sample of convenience, from January 2011 to November 2011. Thirteen other patients who were asked to participate were not included in the study because of a language barrier (3), inability to understand the task (3), a vision disorder (1), and undisclosed reasons (6). All but one evaluated subject were inpatients hospitalized for planned surgery as a result of various hand disorders. The hand conditions included Dupuytren's contracture ($n = 50$), claw finger resulting from an ulnar nerve lesion ($n = 5$), post-traumatic contracture ($n = 4$), postburn contracture ($n = 1$), and contracture resulting from osteoarthritis ($n = 1$). Bilateral hand involvement was observed in 36 patients, 32 of whom had Dupuytren's disease. Thirty-two patients performed self-evaluation with their left hands, of which one was dominant and one ambidextrous. Twenty-nine patients self-evaluated with their right hands, which were dominant in 28 participants. Patients' ages ranged from 23 to 81 years (mean, 55.7 years; SD, 12.5 years). None of the patients had self-evaluated their joints before the study. I have 4 years experience in paper strip goniometry and performed parallel evaluations of the patients' joints.

A baseline flexion-hyperextension plastic finger goniometer with a scale graduated by 1°-increments (Fig. 1) and two rectangular 3.8 × 10.7-cm folded paper strips were used to obtain finger joint angles (Fig. 2).

Fig. 1 The standard finger goniometer is used to measure the PIP joint angle.

Fig. 2A–B The (A) initial position and (B) proper alignment of the paper strips during the patient's self-evaluation of the PIP joint are shown.

All inpatients underwent evaluations in their wards at bedside before surgery. Two patients were evaluated in an auditorium at a table. All joints of the fifth finger of one hand were evaluated in their simultaneous maximal active extension. Single individual evaluation sessions included brief technical familiarization of the patient with the diagrammatic technique, patients' self-evaluation with the paper goniometer, and evaluation of the same joints by me (hereafter referred to as the practitioner) with the paper goniometer and with a standard finger goniometer. The order of the measurements was as follows: Trial 1, practitioner with the paper goniometer, patient with the paper goniometer, practitioner with the standard goniometer; and Trials 2 and 3, practitioner with the paper goniometer, practitioner with the standard goniometer. In each of the subsessions of instrument use, the joints were evaluated in the following order: the metacarpophalangeal (MCP), proximal interphalangeal (PIP), and distal interphalangeal (DIP). The first 12 patients participated only in the first trial. One of them underwent only professional evaluation and self-evaluation of the MCP and PIP joints with the paper goniometer and then discontinued participation because of the scheduled operation. For the remaining 49 patients, the evaluation protocol was enhanced by

Table 1. Number of evaluated patients by joint, tool, rater, and trial

Joint	P first trial; P(Pt)	G first trial	G second trial; G third trial; P second trial; P third trial
MCP	61	60	49
PIP	61	60	49
DIP	60	60	49

MCP = metacarpophalangeal joint; PIP = proximal interphalangeal joint; DIP = distal interphalangeal joint; P = practitioner's measurements with paper goniometer; P(Pt) = patients' measurements with the paper strip goniometer; G = practitioner's measurements with the standard finger goniometer.

Fig. 3 Diagrammatic self-evaluation of hyperextension is easier performed with vertically oriented paper strips.

appending the second and third practitioner's trials. The distribution of the evaluated patients is summarized (Table 1).

The patients were instructed to crisscross the paper strips side by side at an appropriate fulcrum point to obtain the lengths of the paper goniometer arms according to the dimensions of the segments of the joint of interest. In cases of limited extension, the initial angle between the paper strips had to be made smaller than that formed by the digital segments. The patients then were required to manipulate the prepared paper goniometer with one hand in a fanning-cards manner to align the appropriate edges of the strips with the dorsal middle line of the appropriate joint segments (Fig. 2). In cases of hyperextension, manipulation was similar; however, the paper strips had to be oriented with their longer edges perpendicular to the longitudinal axes of the appropriate joint segments (Fig. 3). Having achieved adequate alignment, the patients laid the paper strips flat onto a blank and traced the edges of the strips to obtain a diagram of the joint angle (Fig. 4). The standard goniometer was applied dorsally, except when the position of the neighboring joints interfered with the length of the arms of the instrument. In the latter situations, lateral alignment was used. Before self-evaluating for the study, the patients

Fig. 4A–B The process of obtaining a joint diagram is shown. Care is taken not to displace the strips when (A) laying them flat and (B) tracing their edges.

performed trial paper strip manipulations and self-evaluations of the finger of interest until they felt comfortable with the technique. The practitioner performed diagrammatic recordings in a similar fashion by manipulating the paper strips with both hands. The obtained angle values were recorded by the same practitioner. The measurement session lasted approximately 15 minutes. The diagrams were scanned and then measured with the ImageJ program (National Institutes of Health, Bethesda, MD, USA) by the practitioner. Differences between patient and practitioner or between the practitioner trials of 5° or less were considered acceptable and not clinically meaningful [9].

Normality of the measurements and their differences was assessed with the Shapiro-Wilk test. Only the PIP joint angle measurements could be used for parametric analysis, as not all other measurements and their differences were normally distributed. Means and standard deviations of the intragoniometer, intergoniometer, and interrater measurement differences (Table 2), and the 95th percentiles of the absolute values of those differences were obtained. Proportions of the intragoniometer, intergoniometer, and interrater absolute measurement differences falling within the conventionally acceptable 5° were identified as reported earlier [9]. These proportions were appropriately combined

for analysis of intragoniometer, intergoniometer, and interrater differences (Table 2). Combination of the measurements and their 5° or less differences into multiple comparison pairs was based on the logic that replacing one value in the pair with another one (related to another trial or the other instrument) will change the result of the appropriate computation proportionally to the degree of differences between the interchanged values. The degree of intragoniometer, intergoniometer, and interrater differences was assessed by comparing the appropriately paired proportions using the two-sided McNemar test. To find whether the degree of the interrater differences depended on dominance of the hand with which the patients performed self-evaluation, a two-sided Fisher's exact test was used for patients' handedness-related proportions of 5° or less differences derived from the patients' self-evaluation and the practitioner's evaluations with each instrument in each trial (Table 2). Intragoniometer, intergoniometer, and interrater reliabilities were determined by calculating intraclass correlation coefficients (ICC) (two-way mixed, single-measure model) for the appropriately paired groups of measurements (Table 2). In the paired analyses of the groups of unequal sizes (Table 1), the missing data were handled by pairwise deletion. Bonferroni correction was used for multiple comparisons. The data were analyzed with Microsoft Excel (Microsoft Inc, Redmond, WA, USA) and IBM SPSS Versions 19 and 20 (SPSS Inc, Chicago, IL, USA).

Results

Professional examiner's evaluations with the standard goniometer and paper goniometer produced similar descriptive statistics of intragoniometer measurement differences (Table 3). The descriptors of the practitioner's intragoniometer measurement differences were comparable to those of intergoniometer differences (Table 3). Parity of the goniometers could be well seen from the comparison of the descriptors of differences between the measurements of the second and third trials. When one of the measurements in the formula for calculation of intragoniometer differences was replaced with a reading of the other instrument, no notable change in the descriptive statistics was observed (Table 4). However, when the practitioner's first trial measurements were used in the calculations, the descriptive statistics of intragoniometer and intergoniometer differences tended to be slightly poorer relative to the corresponding descriptors based on the second and third trials (Table 5). Similarity of intragoniometer and intergoniometer differences was further seen from the equivalence of the appropriate proportions of absolute measurement differences falling within the conventional 5° (the lowest $p > 0.06$) (Table 6). All readings obtained with

Table 2. The statistical approaches and corresponding combinations of measurements

Statistical approaches	Combinations of measurements				
	Intrarater (study question 1)			Interrater (study question 2)	Interrater (related to patient's hand dominance) (study question 3)
	Intragoniometer		Intergoniometer		
	Standard goniometer	Paper strips			
Descriptive statistics of measurement differences	G – G (N = 3)	P – P (N = 3)	G – P; G – G; G – P (N = 9)	P(Pt) – P; P(Pt) – G (N = 6)	n/a
Comparison of proportions of 5° or less absolute differences by McNemar test	$[G - G]^{\%} \vee [G - G]^{\%}$ (N = 3)	$[P - P]^{\%} \vee [P - P]^{\%}$ (N = 3)	$[G - G]^{\%} \vee [P - P]^{\%};$ $[G - G]^{\%} \vee [G - P]^{\%};$ $[G - P]^{\%} \vee [G - P]^{\%}$ (N = 99)	$[P(Pt) - P]^{\%} \vee [G - G]^{\%};$ $[P(Pt) - P]^{\%} \vee [P - P]^{\%};$ $[P(Pt) - P]^{\%} \vee [G - P]^{\%};$ $[P(Pt) - G]^{\%} \vee [G - G]^{\%};$ $[P(Pt) - G]^{\%} \vee P - P]^{\%};$ $[P(Pt) - G]^{\%} \vee [G - P]^{\%}$ (N = 99)	n/a
Comparison of proportions of 5° or less absolute differences by Fisher's test	n/a	n/a	n/a	n/a	$[P(Pt \text{ Dm}) - P]^{\%} \vee [P(Pt \text{ Ndm}) - P]^{\%};$ $[P(Pt \text{ Dm}) - G]^{\%} \vee [P(Pt \text{ Ndm}) - G]^{\%}$ (N = 6)
Calculation of ICC	G $\vee$ G (N = 3)	P $\vee$ P (N = 3)	G $\vee$ P (N = 9)	P(Pt) $\vee$ P; P(Pt) $\vee$ G (N = 6)	n/a

[%] Proportion of 5° or less absolute differences between the measurements within the brackets; G and P = practitioner's measurements obtained with standard finger goniometer and paper goniometer, respectively; P(Pt) = patients' measurements with the paper strip goniometer; N = total number of combinations due to three combinations of practitioner's trials (1 versus 2, 1 versus 3, and 2 versus 3); n/a = not applicable; P(Pt Dm) and P(Pt Ndm) = patients' measurement with the paper strip goniometer obtained with dominant and nondominant hands, respectively; v = versus; ICC = intraclass correlation coefficient.

both instruments by a professional examiner could be used interchangeably without notable influence on the results of comparisons of the appropriate proportions of 5° or less intrarater differences. Intragoniometer and intergoniometer reliabilities for the PIP joint varied depending on the practitioner's trial (Table 7). Reliability was almost perfect when calculation of ICC was based on the practitioner's measurements obtained in the second and third trials (the lowest ICC > 0.99). Reliability was always poor when the first trial measurements were involved (the highest ICC < –0.4). The practitioner evaluated the DIP joints more consistently than the MCP and PIP joints across both instruments and trials.

Although the practitioner's and patients' paired measurements looked similar (Fig. 5), interrater differences

were substantial. The descriptive statistics of measurement differences of the paired patients' and practitioner's evaluations were notably poorer than those based exclusively on practitioner's measurements (Table 3). The corresponding proportions of the absolute measurement differences of 5° or less were considerably different for PIP and DIP joints (the lowest $p < 0.001$) (Table 6). Interrater reliability for the PIP joint was poor (the highest ICC < –0.04). Interrater differences were more notable for the DIP joint than for the MCP and PIP joints. Interrater differences did not depend on patients' handedness. Patients' hand dominance-related proportions of 5° or less differences between patients' and practitioner's measurements across both tools and all trials were equivalent (Table 6).

Table 3. Descriptive statistics of intragoniometer (G – G and P – P), intergoniometer (G – P), and interrater (P(Pt) – G and P(Pt) – P) measurement differences

Joint, subgroup of differences/ trials	Mean	SD	95% within*	Within 5°*	Number of patients	Number of measurement pairs
MCP						
G – G (Trials 1–2; 1–3; 2–3)	0.6°	4.0°	9°	83%	49	147
P – P (Trials 1–2; 1–3; 2–3)	0°	5.2°	10°	75%	49	147
G – P (Trials 1–2; 1–3; 2–3)	0.6°	4.9°	11°	76%	49	452
P(Pt) – P Trial 1	–1.1°	7.3°	18°	62%	61	61
P(Pt) – G Trial 1	–1.2°	6.6°	14°	63%	60	60
PIP						
G – G (Trials 1–2; 1–3; 2–3)	0.4°	3.7°	8°	89%	49	147
P – P (Trials 1–2; 1–3; 2–3)	0°	3.5°	7°	87%	49	147
G – P (Trials 1–2; 1–3; 2–3)	1.1°	4.0°	9°	85%	49	452
P(Pt) – P Trial 1	–0.8°	5.7°	11°	69%	61	61
P(Pt) – G Trial 1	–2.5°	6.5°	15°	60%	60	60
DIP						
G – G (Trials 1–2; 1–3; 2–3)	0.2°	2.8°	6°	94%	49	147
P – P (Trials 1–2; 1–3; 2–3)	0.2°	3.0°	5°	95%	49	147
G – P (Trials 1–2; 1–3; 2–3)	0.8°	3.4°	7°	90%	49	452
P(Pt) – P Trial 1	–4.4°	7.5°	20°	60%	60	60
P(Pt) – G Trial 1	–5.4°	7.4°	19°	58%	60	60

* Absolute differences; G and P = paper goniometer and standard finger goniometer, respectively; P(Pt) = patients' measurements with the paper strip goniometer; MCP = metacarpophalangeal joint; PIP = proximal interphalangeal joint; DIP = distal interphalangeal joint.

Table 4. Descriptive statistics reflecting similarity of intragoniometer (P – P and G – G) and intergoniometer (P – G and G – P) measurement differences

Joint, subgroup of differences/trial number	Mean	SD	95% within*	Within 5°*	Number of measurement pairs
MCP					
P2 – P3	–0.1°	4.7°	7°	88%	49
P2 – G3	1.2°	5.0°	8°	71%	49
G2 – G3	0.1°	3.8°	7°	82%	49
G2 – P3	–1.1°	4.1°	9°	82%	49
PIP					
P2 – P3	0.2°	2.9°	6°	92%	49
P2 – G3	1.2°	3.2°	7°	92%	49
G2 – G3	–0.5°	3.3°	7°	90%	49
G2 – P3	–1.5°	3.9°	7°	86%	49
DIP					
P2 – P3	0.1°	2.6°	5°	98%	49
P2 – G3	1.1°	3.3°	6°	88%	49
G2 – G3	0.1°	2.4°	5°	96%	49
G2 – P3	–0.9°	2.7°	5°	96%	49

* Absolute differences; MCP = metacarpophalangeal joint; PIP = proximal interphalangeal joint; DIP = distal interphalangeal joint; P and G = standard finger goniometer and paper goniometer, respectively.

Table 5. Descriptive statistics reflecting intragoniometer (ie, intrarater) trial-related differences

Joint, subgroup of differences/trial number	Mean	SD	95% within*	Within 5°*	Number of measurement pairs
MCP					
P1 – P3	0°	5.6°	11°	67%	49
P2 – P3	0°	4.7°	7°	88%	49
G1 – G3	0.8°	4.3°	9°	88%	49
G2 – G3	0.1°	3.7°	7°	82%	49
PIP					
P1 – P3	0°	3.7°	7°	86%	49
P2 – P3	0.2°	2.9°	6°	92%	49
G1 – G3	–0.6°	4.0°	9°	88%	49
G2 – G3	–0.5°	3.3°	7°	90%	49
DIP					
P1 – P3	0.1°	2.6°	6°	94%	49
P2 – P3	1.1°	3.3°	5°	98%	49
G1 – G3	0.1°	2.4°	8°	92%	49
G2 – G3	–0.9°	2.7°	5°	96%	49

* Absolute differences; MCP = metacarpophalangeal joint; PIP = proximal interphalangeal joint; DIP = distal interphalangeal joint; P and G = standard finger goniometer and paper goniometer, respectively.

Table 6. Summary of the results of multiple comparisons of the proportions of absolute measurement differences of 5° or less

Joint	Lowest p values (number of the proportion pairs different at p < 0.05)				
	Intrarater			Interrater	Interrater (related to patient's hand dominance)
	Intragoniometer		Intergoniometer		
	Standard goniometer	Paper strips			
	[G – G] [%] v	[P – P] [%] v	[G – G] [%] v		
	[G – G] [%]	[P – P] [%]	[G – P] [%] v	[P(Pt) – P] [%] v [G – G] [%] ; [P(Pt) – P] [%] v [P – P] [%] ; [P(Pt) – P] [%] v [G – P] [%] ; [P(Pt) – G] [%] v [G – G] [%] ; [P(Pt) – G] [%] v [P – P] [%] ; [P(Pt) – G] [%] v [G – P] [%] (N = 99)	[P(Pt Dm) – P] [%] v [P(Pt Ndm) – P] [%] ; [P(Pt Dm) – G] [%] v [P(Pt Ndm) – G] [%] (N = 6)
	(N = 3)	(N = 3)	[G – P] [%] [G – P] [%] (N = 99)		
MCP	1	0.063	0.099	0.09	1
PIP	1	1	0.4	< 0.001 (8)	1
DIP	1	1	1	< 0.001 (65)	0.67

[%] Proportion of 5° or less absolute differences between the measurements in the brackets; G and P = practitioner's measurements obtained with paper goniometer and standard finger goniometer, respectively; P(Pt) = patients' measurements with the paper strip goniometer; v = versus; N = total number of combinations due to three combinations of practitioner's trials (1 versus 2, 1 versus 3, and 2 versus 3); MCP = metacarpophalangeal joint; PIP = proximal interphalangeal joint; DIP = distal interphalangeal joint.

Table 7. Summary of the results of reliability analysis for the PIP joint

Measurement combinations according to practitioner's trial	Intraclass correlation coefficient (95% CI)			
	Intrarater		Interrater P(Pt) vs P; P(Pt) vs G (N = 6)	
	Intragoniometer	Intergoniometer		
	Standard goniometer G vs G (N = 3)	Paper strips P vs P (N = 3)	G vs P (N = 9)	
Trial 1 vs Trial 2; Trial 1 vs Trial 3	0.015 ^H (–0.271 to 0.296)	0.212 ^H (–0.169 to 0.390)	0.373 ^H (0.132 to 0.572)	0.032 ^H (–0.225 to 0.283)
Trial 2 vs Trial 3	0.994 (0.99 to 0.997)	0.996 (0.992 to 0.998)	0.995 ^H (0.990 to 0.997) 0.991 ^L (0.982 to 0.995)	–0.409 ^H (–0.629 to –0.137) 0.413 ^L (–0.632 to –0.141)

^HHighest value; ^Llowest value; G and P = practitioner's measurements obtained with paper goniometer and standard finger goniometer, respectively; P(Pt) = patients' measurements with the paper strip goniometer; N = total number of combinations due to three combinations of practitioner's trials (1 versus 2, 1 versus 3, and 2 versus 3).

Discussion

Numerous reliable techniques have been proposed to facilitate evaluation of multiarticular finger motion [6–8, 15, 32, 33, 37, 38] as alternatives to standard goniometry. Most techniques other than visual recording or goniometry, however, have not found widespread clinical use owing to their complexity. Finger ROM recording and interpretation can be simplified by diagrammatic presentation [28]. The most commonly used graphic technique of readily interpretable silhouette diagrams of the dorsal digital surface, obtained by means of wire tracing [17], is reportedly not reliable [13]. A recent paper [27] suggested finger joint angle diagrams can be readily obtained by using an improvised goniometer assembled of two folded paper strips. Manipulation of the

finger-held paper strips resembles the standard goniometry procedure and can be used for self-evaluation, which could be beneficial for augmenting patient's participation in the treatment process. To explore clinical feasibility of such uses of the diagrammatic goniometry, I assessed the standard and diagrammatic finger goniometry performed by an experienced examiner on patients in terms of (1) intragoniometer and intergoniometer differences and reliability; (2) interrater differences and reliability relative to patients' diagrammatic self-evaluation; and (3) the interrater differences related to patient's hand dominance.

There are several limitations of this study. First, the patients as raters were at a comparative disadvantage owing to the absence of previous training. This disparity may be difficult to eliminate completely since it does not

Fig. 5 These box plots summarize the joint angle measurements obtained by the patients and by the practitioner in Trial 1. P(Pt) = patients' measurements with the paper strip goniometer; G1 and P1 = practitioner's measurements in Trial 1 with the paper strip goniometer and with the standard finger goniometer, respectively; MCP = metacarpophalangeal joint ($n = 61$); PIP = proximal interphalangeal joint ($n = 61$); DIP = distal interphalangeal joint ($n = 60$). The horizontal lines in the boxes represent the medians. The ends of the whiskers represent the maximum and minimum values. The circle indicates an outlier.

seem possible to train patients to the level matching a practitioner's goniometric experience. Thus, relatively large differences between professional ROM measurements and patients' self-evaluations should be expected. Second, the study did not explore patient's intrarater differences. This was for the patients' convenience, given their first-time involvement and multiple measurements of the practitioner. Third, the study included patients with various hand disorders. Some joints had no extension deficit and even showed hyperextension (Fig. 5), perhaps accounting for the considerable variability of the obtained data. However, the wide range of joint angles were those that would be encountered in normal clinical practice. Third, the sample sizes could be larger to achieve data normality and enable more parametric analyses. I found no finger goniometric reliability studies using more subjects than the current study, and only one study involved more than 50 subjects [32]. The presented means and standard deviations of the measurement differences can be used for power analysis for subsequent studies directed at determination of SEs of measurement [23, 36] and limits of agreement [23]. Fourth, the assessment of interrater differences and reliability was based on the assumption that multiple patients represent a theoretical single rater performing self-evaluation, (ie, one professional rater was compared against multiple patients who apparently had various degrees of dexterity to perform self-evaluation). This limitation is inherent to self-measurement studies.

The data suggest parity of the diagrammatic paper strip technique and standard goniometry in the hands of the experienced evaluator. The 95th percentiles of the absolute values of differences between measurements made with different instruments were between 5° and 11° (ie, above the conventional 5° limit). Percentages of the intergoniometer differences within 5° , however, were between 75% and 95%. More importantly, the proportions of paired intragoniometer differences of 5° or less using different tools were not different, which I believe would justify clinical use of the technique considering that professional traditional goniometry was the reference standard. Dijkstra et al. [11], in a study of three raters measuring the MCP joint of the fifth finger in hyperextension in 30 healthy subjects, found that intrarater residual standard deviations of the repeated measurements (ie, the SEs of measurement [9, 12, 23, 36]) of the MCP joint of the fifth finger in hyperextension ranged from 2.12° to 3.88° . After appropriate transformation of these values [12, 36], measurements of the individual raters were within more than 5.8° 95% of the time. A study by Ellis et al. [13] used 40 professional examiners to evaluate all joints of two splinted fingers in one healthy subject three times with a standard finger goniometer. The repeatability coefficients derived from the variances of the repeated measurements ranged from 3.8° to 9.9° . Burr et al. [5] used the same design features for comparative evaluation of interphalangeal finger joints with three different goniometers and determined that the 95th percentiles of the differences between maximum and minimum range from 5° to 16° . The above-mentioned study by Ellis et al. [13] was the only source that could be identified to report reliability of diagrammatic finger motion recording. Along with the goniometric measurements, Ellis et al. performed parallel evaluations with malleable wire. They concluded that wire tracing is less reliable than goniometry. The intrarater repeatability coefficients derived from the appropriate measurements of the wire diagrams ranged from 8.0° to 10.9° . The corresponding values of the diagrammatic technique in the current study were 5° and 10° . Again, these findings are difficult to compare, because in an investigation by Ellis et al. [13], evaluations were obtained by multiple professional examiners and only two standard static finger positions of only one subject were used, which seems to contribute to lower variability of the measurements.

The obtained negative ICC values, although theoretically not possible, are observable in practice and indicate low reliability [16, 34]. When explaining the low intrarater (ie, intragoniometer and intergoniometer) ICCs related to the practitioner's first trial, several factors related to the rater and subject should be taken into account. Given the practitioner's experience in goniometry, it is doubtful that a rater's ability to perform consistent repeated measurements increased owing to the evaluations in the first trial. It is

more likely that a learning effect occurred because the patients, being novice to the procedure, improved consistency of presentation of their fingers in active extension. The low intrarater reliability related to the first trial, however, was not of clinical concern, as the appropriate proportions of the clinically acceptable 5° differences were equivalent.

The intrarater ICCs based on the second and third trial measurements correspond to the appropriate earlier observations. Three studies [14, 20, 25] reported ICC values for active motion of separate finger joints evaluated with standard goniometry. In a study by Groth et al. [20], 39 raters evaluated active flexion and extension of interphalangeal joints of two fingers in one patient. Intrarater ICCs across all joints and motions for six randomly chosen raters were 0.99 and 0.86 for lateral and dorsal alignment methods, respectively. Engstrand et al. [14], with eight raters measuring active flexion and extension of all joints of four fingers in 13 patients, reported that interrater ICCs ranged from 0.832 to 0.973 depending on the joint and motion. The range of ICC values was much wider in a study by Lewis et al. [25] who had seven raters measure angles of active extension of one finger in 20 healthy subjects. The resultant interrater and intrarater ICCs varied from 0.55 to 0.88 and from 0.64 – 0.99, respectively.

A literature search did not reveal studies that explored self-evaluation of finger joint motion. The obtained relatively low interrater reliability and high differences indicate poor patient performance in self-measurement if compared with the appropriate assessments of the experienced evaluator. The reasons for inadequate interrater reliability in the current study include the patients' minimal training, difficulty in manipulation with only one hand, functional deficit of the manipulating hand, and inconvenient position of the fifth finger in terms of access for self-evaluation. Thus, clinical feasibility of the paper goniometer for self-evaluation requires further studies involving control of those factors, especially that of training. In addition, the issue of patients' intrarater reliability needs to be addressed. If patients' intrarater reliability is clinically acceptable, patients could use the paper goniometer for self-control of ROM dynamics even if interrater reliability of the patient's and practitioner's measurements is poor. A good prerequisite for clinical use of the paper goniometer is that hand dominance seems to have no influence on the consistency of self-evaluation. Reliable patient self-evaluation at home would open a perspective for beneficial cost reduction of ambulatory consultations for monitoring finger ROM dynamics. In situations in which a possible surgical intervention is determined by appropriate dynamics of ROM, patients' visits would be necessary only when a certain angle of the involved joint is self-recorded or if there is no considerable progress in ROM during a certain

time. This approach could be enhanced by readily using the patients' drawn diagrams in the telecommunication-based virtual visits. Patients could compare their diagrams obtained before surgery with those made postoperatively to show the benefits of the procedure, which should improve patient and surgeon rapport. Diagrammatic joint self-evaluation could make assessment of patient satisfaction more objective. The parity of the current hand outcome questionnaires including the Patient Evaluation Measure, DASH, and Michigan Hand Outcome Measure [10] may need to be explored in respect to patients' involvement in self-evaluation. An important aspect of reliability of the diagrammatic technique concerns conversion of the diagrams to the angle degrees. In terms of patient's self-monitoring, however, numerical angle expression is not so meaningful since diagrams can be compared by superimposing them on each other [2, 17, 21, 30]. Reliability of translating diagrams into angle degrees with a standard goniometer or simple protractor needs to be explored, especially in terms of varying diagram size. Routine use of diagram scanning and computerized measurement, which was used in the current investigation, would damage the simplicity of the technique. Another possibility for further research could be to study use of the diagrammatic ROM self-evaluation as a simple instrument for goniometric biofeedback suggested earlier [3].

The observations reported here suggest that in the hands of an experienced evaluator, graphic recording of finger ROM by means of a paper strip diagrammatic technique produces similar results to those obtained by standard finger goniometry, but a patients' ability to self-evaluate their ROM is clinically inadequate. Further studies are necessary to explore whether appropriate training of patients can improve consistency of diagrammatic finger joint self-evaluation.

References

1. Albee FH, Gilliland AR. Metrotherapy, or the measurement of voluntary movement: its value in surgical reconstruction. *JAMA*. 1920;75:983–986.
2. Bexon C, Salter M. Assessment. In Salter M, Cheshire L, eds. *Hand Therapy: Principles and Practice*. Oxford, UK: Butterworth Heinemann; 2000:23–24.
3. Brown DM, DeBacher GA, Basmajian JV. Feedback goniometers for hand rehabilitation. *Am J Occup Ther*. 1979;33:458–463.
4. Bruton A, Ellis B, Goddard J. Comparison of visual estimation and goniometry for assessment of metacarpophalangeal joint angle. *Physiotherapy*. 1999;85:2011–2018.
5. Burr N, Pratt AL, Stott D. Inter-rater and intra-rater reliability when measuring interphalangeal joints: comparison between three hand-held goniometers. *Physiotherapy*. 2003;89:641–652.
6. Catalano LW 3rd, Browne RH, Carter PR, Frobish AC, Ezaki M, Littler JW. The Littler line method and the area under a Gaussian curve: a new method of assessing digital range of motion. *J Hand Surg Am*. 2001;26:23–30.

7. Chiu HY, Su FC, Wang ST, Hsu HY. The motion analysis system and goniometry of the finger joints. *J Hand Surg Br.* 1998;23:788–791.
8. Cook JR, Baker NA, Cham R, Hale E, Redfern MS. Measurements of wrist and finger postures: a comparison of goniometric and motion capture techniques. *J Appl Biomech.* 2007;23:70–78.
9. de Vet HC, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. *J Clin Epidemiol.* 2006;59:1033–1039.
10. Dias JJ, Rajan RA, Thompson JR. Which questionnaire is best? The reliability, validity and ease of use of the Patient Evaluation Measure, the Disabilities of the Arm, Shoulder and Hand and the Michigan Hand Outcome Measure. *J Hand Surg Eur Vol.* 2008;33:9–17.
11. Dijkstra PU, de Bont LG, van der Weele LT, Boering G. Joint mobility measurements: reliability of a standardized method. *Cranio.* 1994;12:52–57.
12. Eliasziw M, Young SL, Woodbury MG, Fryday-Field K. Statistical methodology for the concurrent assessment of interrater and intrarater reliability: using goniometric measurements as an example. *Phys Ther.* 1994;74:777–788.
13. Ellis B, Bruton A, Goddard JR. Joint angle measurement: a comparative study of the reliability of goniometry and wire tracing for the hand. *Clin Rehabil.* 1997;11:314–320.
14. Engstrand C, Krevers B, Kvist J. Interrater reliability in finger joint goniometer measurement in Dupuytren's disease. *Am J Occup Ther.* 2012; 66:98–103.
15. Georgeu GA, Mayfield S, Logan AM. Lateral digital photography with computer-aided goniometry versus standard goniometry for recording finger joint angles. *J Hand Surg Br.* 2002;27: 184–186.
16. Giraudeau B. Negative values of the intraclass correlation coefficient are not theoretically possible. *J Clin Epidemiol.* 1996;49:1205.
17. Glanville H. Objective assessment of treatments of hand injuries: a new method of measurement. *Ann Phys Med.* 1964;8:304–306.
18. Greene WB, Heckman JD, eds. *The Clinical Measurement of Joint Motion.* Rosemont, IL, USA: American Academy of Orthopaedic Surgeons; 1994:6–8.
19. Greenstock MH. Use of the Odstock wire as a simple method of evaluating multi-articular function in digits. *J Hand Surg Br.* 1997;22:681.
20. Groth GN, VanDeven KM, Phillips EC, Ehretzman R. Goniometry of the proximal and distal interphalangeal joints, Part II: placement preferences, interrater reliability, and concurrent validity. *J Hand Ther.* 2001;14:23–29.
21. Hage JJ, de Groot PJ. A method of two-dimensional measurement for evaluating finger motion impairment. *J Hand Surg Br.* 1996;21:564–565.
22. Hamilton GF, Lachenbruch PA. Reliability of goniometers in assessing finger joint angle. *Phys Ther.* 1969;49:465–469.
23. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med.* 2000; 30:1–15.
24. Lea RD, Gerhardt JJ. Range of motion measurements. *J Bone Joint Surg Am.* 1995;77:784–798.
25. Lewis E, Fors L, Tharion WJ. Interrater and intrarater reliability of finger goniometric measurements. *Am J Occup Ther.* 2010;64:555–561.
26. Macdermid JC, Fox E, Richards RS, Roth JH. Validity of pulp-to-palm distance as a measure of finger flexion. *J Hand Surg Br.* 2001;26:432–435.
27. Macionis V. A technique for graphical recording of range of motion using an improvised paper goniometer. *J Hand Ther.* 2011;24:374–377.
28. Nutter JA. Reconstructive surgery: the problem of records. *JAMA.* 1919;72:410–411.
29. Palmer ML, Epler ME. *Fundamentals of Musculoskeletal Assessment Techniques.* 2nd ed. Philadelphia, PA, USA: Lippincott Williams & Wilkins; 1998:12.
30. Robins RH. Hand assessment charts. *J Hand Surg Br.* 1986;11: 287–298.
31. Rose V, Nduka CC, Pereira JA, Pickford MA, Belcher HJ. Visual estimation of finger angles: do we need goniometers? *J Hand Surg Br.* 2002;27:382–384.
32. Smith RP, Dias JJ, Ullah A, Bhowal B. Visual and computer software-aided estimates of Dupuytren's contractures: correlation with clinical goniometric measurements. *Ann R Coll Surg Engl.* 2009;91:296–300.
33. Stam HJ, Ardon MS, den Ouden AC, Schreuders TA, Roebroek ME. The compangle: a new goniometer for joint angle measurements of the hand. A technical note. *Eura Medicophys.* 2006;42:37–40.
34. Tammemagi M, Frank J, Streiner D. Negative values of the intraclass correlation coefficient are not theoretically possible: response. *J Clin Epidemiol.* 1996;49:1206.
35. Thomine JM. The clinical examination of the hand. In: Tubiana R, ed. *The Hand.* Philadelphia, PA, USA: Saunders Company; 1981:616–647.
36. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res.* 2005;19:231–240.
37. Weiss PL, August S, Peters G, Sampalis J. Using the Exos Handmaster to measure digital range of motion: reliability and validity. *Med Eng Phys.* 1994;16:323–328.
38. Williams NW, Penrose JM, Caddy CM, Barnes E, Hose DR, Harley P. A goniometric glove for clinical hand assessment: construction, calibration and validation. *J Hand Surg Br.* 2000;25:200–207.